Reliability of GPS based traffic data: an experimental evaluation

Author 1: Xiaoyun Zhao
Author 2: Kenneth Carling
Author 3: Johan Håkansson
Editor: Hasan Fleyeh

Nr: 2014:17
Reliability of GPS based traffic data: an experimental evaluation
Xiaoyun Zhao*, Kenneth Carling, Johan Håkansson

Abstract
GPS tracking of mobile objects provides spatial and temporal data for a broad range of applications including traffic management and control as well as transportation routing and planning. Previous transportation research has focused on GPS tracking data as an appealing alternative to travel diaries. Yet, the GPS based data are gradually becoming a cornerstone for real-time traffic management. Tracking data of vehicles from GPS devices are however susceptible to measurement errors – a neglected issue in transportation research. By conducting a randomized experiment, we assess the reliability of GPS based traffic data on geographical position, velocity, and altitude for three types of vehicles; bike, car, and bus. We find the geographical positioning reliable, but with an error greater than postulated by the manufacturer and a non-negligible risk for aberrant positioning. Velocity is slightly underestimated, whereas altitude measurements are unreliable.

Key words: GPS tracking device, reliability, transportation, road network

1. Introduction
Global Positioning System (GPS) is a Global Navigation Satellite System (GNSS) for geopositioning. The availability and usability of GPS devices in geo-positioning and tracking mobile objects has grown enormously in the past decades and is still increasing. The GPS has emerged for civilian use in the 1990s as the space geodetic technique being accurate and affordable (Zumberge et al., 1995). In their review, Theiss et al. (2005) identified a wide range of applications of GPS tracking data including timing, logistics, traffic management, and weather forecasting and concluded that it will change the way companies and organizations run their business.

GPS tracking technologies have extensively been applied in transportation studies, in particular for studying the routes of motorized vehicles (Zito et al., 1995; Quiroga and Bullock, 1998; Murakami and Wagner, 1999). For instance, Schönfelder (2002) presented an approach to collect GPS longitudinal travel behavior data on humans and described the complexity of their daily life with the interaction between periodicity and variability.

GPS is also applied to study the travel pattern and prediction of human mobility (Ashbrook et al. 2002, 2003). For instance, Jia et al. (2012) confirmed the scaling property and identified the Levy flight characteristic of human mobility by using the GPS tracking data of car movements. GPS data is also applied in environment control. For instance, Carling et al. (2013) and Jia et al. (2013)

* Xiaoyun Zhao is a PhD-student in Micro-data analysis and corresponding author: xzh@du.se, phone: +46 23-778509. Kenneth Carling is a professor in Statistics and Johan Håkansson is a professor in Geography. All are at the School of Technology and Business Studies at Dalarna University, Sweden
studied the induced pollutant emissions of CO2 from car movements by using a GPS tracking data of car movements.

Even though GPS tracking data opens up for interesting applications, gathering information of spatial-temporal mobility by GPS is subject to critical reflections. Leduc (2008) examined recent developments in road traffic data collection and discussed the potentials and bottlenecks related to new GPS technologies. Moreover, Van der Spek et al. (2009) concluded that GPS offers a widely useable instrument to collect invaluable spatial-temporal data on different scales and in different settings adding new layers of knowledge to urban studies, but the use of GPS-technology and deployment of GPS-devices still offers significant challenges for future research. Besides, the enormous use of GPS tracking technologies hinges critically on the functioning of the device.

Nowadays, the internal system of a portable, inexpensive GPS tracking device is designed in a complex way due to the desire for precision and accuracy. Configuration of a GPS device when conducting field tracking is becoming more complicated. How well do the concurrent GPS devices perform in tracking vehicle mobility? To what extent can the accuracy information provided by the manufactures be trusted? As argued by Shoval (2008), the device can function as an effective and reliable tool for data collection only if it does not affect the nature, quality or authenticity of the data collected.

Following this, the assessment of the reliability of GPS tracking needs to be scrutinized. In this paper, we examine how well GPS tracking data matches the travelled route for a bike, a car, and a bus for which the route, the speed, and the altitude are pre-set within the experiment. In the experiment, we vary the type of vehicle, speed, altitude, sampling frequency, and filtering level.

Section 2 provides a review of research using GPS tracking data with a focus on studies in which the reliability of such data is examined. Section 3 presents the experimental design and the data collection process. Section 4 gives the experimental results. Section 5 ends the paper with a concluding discussion of the findings.

### 2. Literature review

We have conducted a thorough search for literature relevant to the use of GPS based traffic data. It goes without saying that the use of GPS has penetrated into various transportation applications, such as mobility pattern recognition, vehicle navigation, fleet management, route tracking, and schedule information systems. As a consequence, there is a vast body of studies reporting on applications of GPS with a brief discussion about the reliability of the data. In Table 1 we list most such studies under *Applications*. The discussions contained in these studies do not add any new knowledge and we therefore turn to studies with reliability as the primary concern.

Obviously the quality of the hardware and the surroundings where the GPS is being used may affect the reliability of the device. There are some studies that have looked into these aspects (*Hardware* in Table 1). The starting point is typically that a GPS device requires a clear sight
with at least four satellites to determine spatial positions. The accuracy may be enhanced by advanced hardware chipsets, dual-frequency receivers, and carrier-phase measurements supported by augmentation systems (e.g. SBASs, WAAS, EGNOS and MSAS). Under optimal conditions it is possible to have a real-time positional accuracy within decimeters. The required receivers are however too expensive for the use in, for instance, commercial in-car navigation systems. Moreover, optimal conditions do not only call for sophisticated GPS devices, sensors, vehicles, and map information, but also puts requirements on trajectory dynamics and surrounding environment (Skog and Handel, 2009). In urban environments, buildings may partly block satellite signals, forcing the GPS device to work with a poor geometric constellation of satellites, thereby reducing the accuracy of the positional estimates. Multipath propagation of the radio signal due to reflection in surrounding objects may further lead to decreased positional accuracy without notification by the GPS device, thereby reducing the integrity of the navigation solution.

Table 1: Relevant empirical studies on the reliability of GPS-based data for transportation

<table>
<thead>
<tr>
<th>Area of research</th>
<th>Research topics</th>
<th>Typical method</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Hardware</strong></td>
<td><strong>Correction methods</strong></td>
<td>GPS device with advanced technology settings and ideal conditions (open area, many available satellites, and augmentation systems) generates higher accuracy.</td>
</tr>
<tr>
<td>[11], [19], [24], [26-27], [33], [40], [48], [51-52]</td>
<td>How does the configuration of the hardware affect the precision?</td>
<td>Map-matching, Differential GPS, Dead reckoning</td>
<td>The inaccurate information acquired from GPS devices can be rectified.</td>
</tr>
<tr>
<td></td>
<td><strong>Empirical assessments</strong></td>
<td><strong>Applications</strong></td>
<td>The positional accuracy varies from a few centimetres to hundred meters. The error in speed is 1% or much more.</td>
</tr>
<tr>
<td>[4-5], [7], [9], [17], [20-21], [28] [31], [33], [62], [65], [70], [77], [81-82], [87-89], [94]</td>
<td>How well are the objects positioned? Do the recorded speeds coincide with the speedometer?</td>
<td>Travel data collection, Vehicle navigation, Fleet management, Route tracking, Mobility pattern recognition</td>
<td>Stand-alone GPS devices are helpful and useful in transportation analysis, but vulnerable and need to be combined with auxiliary information for accuracy and integrity</td>
</tr>
<tr>
<td></td>
<td><strong>Typical method</strong></td>
<td><strong>Main findings</strong></td>
<td><strong>Note:</strong> The number refers to the reference in the reference list in the end of the paper.</td>
</tr>
<tr>
<td></td>
<td><strong>Deductive reasoning</strong></td>
<td>GPS device with advanced technology settings and ideal conditions (open area, many available satellites, and augmentation systems) generates higher accuracy.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Laboratory studies</strong></td>
<td>The inaccurate information acquired from GPS devices can be rectified.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Observational data for a single device</strong></td>
<td>The positional accuracy varies from a few centimetres to hundred meters. The error in speed is 1% or much more.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Theoretical calculation</strong></td>
<td>Stand-alone GPS devices are helpful and useful in transportation analysis, but vulnerable and need to be combined with auxiliary information for accuracy and integrity</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Simulation</strong></td>
<td>Stand-alone GPS devices are helpful and useful in transportation analysis, but vulnerable and need to be combined with auxiliary information for accuracy and integrity</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Observational data tests</strong></td>
<td>Stand-alone GPS devices are helpful and useful in transportation analysis, but vulnerable and need to be combined with auxiliary information for accuracy and integrity</td>
<td></td>
</tr>
</tbody>
</table>

Another strand of the literature presumes erroneous recordings of the GPS device and focuses on methods for correcting the error (Correction methods in Table 1). The Dead Reckoning (DR)
system and map matching algorithms integrated with differential GPS (DGPS) are examples of commonly used hybrid systems for enhancing the positioning of vehicles on land. The DR system can smooth the error of the GPS and provide continuous positioning even in times when the GPS is unavailable (Meng, et al., 2004). The DR produces however an accumulating drift in the error, but this can be corrected by the DGPS. In situations where an underlying network is available, map matching has become a popular solution to remedy the inherent error of the GPS. In essence, map matching is to use a digital map of the road network to impose constraints on the GPS navigation or recordings (Skog and Handel, 2009).

The literature most relevant to our study is the studies attempting to assess the reliability of GPS data by comparing them to known conditions (Empirical assessments in Table 1). These studies aims at evaluating the reliability of GPS device, but are typically not conducted as experiments. They examine one transportation mode, one environment, one aspect of tracked information, and one configuration of the device. The studies are also examining the static accuracy using small samples without controlling for external condition. A notable exception is the recent work of Schipperijn et al. (2014). They tested the dynamic accuracy of a GPS device (Qstarz Q1000XT portable GPS receiver) for the use in public health applications under varying real-world environmental conditions, for four modes of transportation, and at three levels of sampling frequency. They found that not even a half of the positional recordings were within 2.5 meters of the actual position with the proportions varying by travel mode and area.

As claimed by Schipperijn et al. (2014), mobile objects in free-living studies are likely to move dynamically. It is therefore vital to know the dynamic accuracy for various travel modes in changing surroundings. However, Schipperijn et al. (2014) only studied the influence on positional accuracy by changing the sampling frequency of the GPS device neglecting other factors possibly affecting the accuracy. To conclude, the number of studies that have evaluated the reliability of standard GPS devices employing different configurations for tracking various types of vehicles on real road networks is limited.

3. Experimental design and data collection

We want to examine how well GPS tracking data matches an actual route travelled. Vehicles are in focus for this study and we therefore assume them being restricted by an underlying road network. We consider the vehicles bike, car, and bus being the dominating means of private transportations. In the experiment, the vehicles travel on pre-set routes of known geographical position and altitude with speeds decided in advance. While they are travelling their mobility is being tracked by a GPS device.

For the experiments, a standard and integrated GPS device that could be broadly used in different vehicles under various circumstances is preferable. Smart phone with GPS application restricted to cellular network or wireless network is therefore not considered. Other important features in selecting the device are that the device is user friendly, easy to operate and has a durable battery. BT-338 (X) was finally chosen after a survey in the product market, this device is a combination
of a GPS receiver and a data logger. According to the manufacturer, the device should provide a geographical positioning within an error of 5 meters and a measurement error of velocity less than 0.4 km/h. The manufacturer makes no claims about the precision in the measurement of altitude.

Figure 1 illustrates the interface in configuring the device with regard to some of the factors in the experiment. We consider intensive sampling by the device with measurements every one and five seconds as well as sampling every 30 seconds. Note that the latter implies that some of the vehicles will easily travel more than 500 meters between recordings. Such setting implies a coarse assessment of the vehicle’s mobility pattern. Hence, the levels of sampling frequency represent both dense and sparse data. We set the data logging format to track position, time, date, speed, and altitude. The WAAS/EGNOS/MSAS feature is enabled to acquire more precise position as suggested by the manufacturer. We consider both enable and disable data logging when distance is less than the selected radius 20 meters.

Figure 1: Interface of setting configurations for the GPS device BT-338(X)

Table 2 illustrates the factors and corresponding levels in the experimental design. We are in possession of 15 identical GPS devices with a unique identifying number. They are randomly assigned to one of three groups of equal size for which the sampling interval is set to 1, 5, and 30 seconds respectively. In each group two randomly selected devices have the data logging disabled if distance is less than the radius of 20 meters.

On the bike, all the 15 devices are carried by the rider in a backpack. Moreover, the devices are in the backpack in the back seat of the car while the backpack is kept in the front seat of the bus.

---

The data collection of the bike and the car is undertaken in Borlänge in Sweden. The data collection of the bus is undertaken along the bus line 151 between Borlänge and its neighboring city Falun.

Table 2: Experimental design of collecting GPS tracking data

<table>
<thead>
<tr>
<th>Sample Interval</th>
<th>1s</th>
<th>5s</th>
<th>30s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device No.</td>
<td>3</td>
<td>29</td>
<td>37</td>
</tr>
<tr>
<td>Bicycle</td>
<td>36</td>
<td>42</td>
<td>72</td>
</tr>
<tr>
<td>Distance Restriction</td>
<td>Distance radius 0m</td>
<td>Distance radius20m</td>
<td>Distance radius0m</td>
</tr>
<tr>
<td>15km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>15km/h</td>
<td>20km/h</td>
<td>30km/h</td>
</tr>
<tr>
<td>40km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bus</td>
<td>80-100km/h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It was difficult to fix the velocity of the bus in advance as would be preferable. The velocity varied along the scheduled route due to the traffic and the behavior of the drivers. For this reason, only a segment of the route, where the velocity varied smoothly between 80 km/h and 100 km/h, was used for GPS tracking. Meanwhile the bus trip was filmed. The bike followed a strict setting of velocities ranging from 15-50 km/h in six levels. For the car velocities of 15-70 km/h were considered. Travel diaries were used to note unexpected changes in route, velocity, and emergent situation. The bike was ridden by the same rider and the driver of the car was the same throughout the experiment.

Data for the bike was collected at noon in order to reduce the risk of deviation from the protocol caused by other people on the route. Likewise, data collection for the car was undertaken between 3 and 4 in the afternoon to avoid peaks in the traffic. The data collection for the bus was conducted after 6 in the afternoon thereby minimizing the variation in velocity due to people waiting at bus stop. The data collection took part on a cloudy summer day with an air temperature of about 22 degrees and almost no wind.
An accurate speedometer of the vehicles is essential for the experiment. To ensure this we first considered the speedometer of the bike. The speedometer works by counting the wheel revolutions per time unit adjusted by the circumference of the tire. Crucial for the accuracy is the measurement of the circumference. The tires were inflated immediately prior to the experiment and the circumference was measured by two different tape measurers. Thereafter we calibrated the car speedometer by riding the bike and driving the car side by side and recording the speeds simultaneously. We checked the relationship between the recordings from the bike speedometer and the car speedometer by means of linear regression: \( \text{Car} = 1.0385 \times \text{Bike} \). The relationship is strong with a correlation of 0.998. The speedometer of the car was adjusted accordingly in the experiment.

![Figure 2: (a) The bike route; (b) The car route; (c) The bus route](image)

The routes for the experiment were chosen having the need for maintaining a constant velocity in mind. In the choice of routes, we tried to avoid places where the GPS signal was likely to be
disturbed. This means that the routes do not pass high buildings, strong magnetic fields or are in valleys. As for the car, we also needed to consider the speed limits of the roads while a bike may be ridden at any speed on a bike path.

Figure 2(a) depicts the route for the bike with arrows indicating the riding direction. The route is about 2 kilometers and it is a paved bike path. The route was used consecutively for each velocity at a time. For instance, at the velocity of 20 km/h the route took 6 minutes meaning that there could be 360, 72, and 12 recordings per GPS device for the three levels of sampling frequency. The variation in altitude of the route is only a few meters.

Figure 2(b) depicts the route for the car with arrows showing the directions. The route is segmented by color representing the attained velocity. The route was travelled several times to ensure sufficiently many recordings per cell in the experimental design. The range in altitude is 40 meters. Maintaining a constant velocity with a car in an ordinary traffic situation is of course difficult. The circles in figure 2(b) represent segments identified in advance as impossible to maintain the speed due to intersections and speed bumps. Afterwards the experimental recordings, pertaining to segments where the intended velocity was not met according to the travel diary, were removed. Figure 2(c) depicts the bus route. This route has a variation in altitude with a range of 37 meters.

All the GPS devices were turned on before initiating the data collection. The reason was that there is acquisition time for the device to start recording. The original GPS tracking data were kept into DataLogger files. The files may be loaded from the device to a computer by using the software GlobalSat Data Logger PC Utility. We retrieved the data directly after the experiment was completed. The device number 4 was malfunctioning and did not record any data. The other 14 devices worked well and we obtained in total 25,901 recordings of the car, 9,224 recordings of the bike, and 8,688 recordings of the bus.

As a final remark we note that there is a trade-off between sampling interval and battery lifespan (Ryan et al., 2004). We checked whether the duration of the battery of the device differed for various settings of the sampling interval. The check was conducted by randomly selecting 6 of the GPS devices and letting 3 of them with intervals 1, 5, and 30 seconds and letting the other 3 of them with intervals 1, 5, and 30 seconds and data tracking within 20 meters distance radius disabled. It turned out that the duration of the battery was unrelated to these two factors.

4. Experimental results

We begin by examining the positional reliability, followed by examining the reliability of velocity and end with a check on the measurement of altitude obtained from the GPS device.

4.1 Geographical positioning

The geographical positions of the mobile object are necessary to identify the objects trajectory. In the experiment the trajectory of the vehicles is known by the road network and its digital
The location and the trajectory of a car are restricted by the road network\(^2\) (Skog and Handel, 2009). As a statistics to assess the reliability of the geographical positioning obtained from the GPS device we measure the concordance of the recordings and the road network. Ideally the positional recordings should be on the underlying road network\(^3\).

Figure 3 shows by an example some of the positional recordings on the road network. The green circles indicate the recordings that match the road network. The yellow circles indicate recordings on the edge of the road network, by us regarded as matching the road network well enough. The red squares indicate inaccurate recordings off the road network. In this example, 8 of the 42 recordings failed in giving an accurate position of the car. The width of the road is 14-20 meters meaning that an error of 5 meters is tolerated even if one considers that the car was not driven in the middle of the road.

![Figure 3: Example of positional recordings and the road network](image)

A bike-path in NVDB is represented by a line, not a polygon, although its width is 3.5 meters according to the department of motor vehicles in Sweden. In assessing the positional recordings of the bike to the underlying road network we allowed for a tolerance distance of 5 meters.

---


\(^3\) The road network is provided by the National Road Data Base (NVDB) and is operated by the Swedish Transport Agency. NVDB classifies the road network into national roads, local roads and streets. The national roads are owned by the national public authorities. The local roads and streets are built and owned by municipalities or companies or private persons. The positional accuracy of the road segments used in this study is within 0.2 meter.
Table 3 gives the proportion of positional recordings that match the road network. Considering that the manufacturer of the GPS device claims that the error in positioning is at the most 5 meters, it is to be expected that almost all recordings should match the road network. This is generally not the case. 75% to 90% of positional recordings for the bus are accurate. The positioning of the car was more reliable with about 90% of the recordings being accurate. As for the bike, the recordings frequently fail to identify its travel on the network.

As an overall finding drawing on Table 3, there is no clear pattern emerging from the factors considered in the experiment. Possibly the longest sampling interval tends to lead to better positioning, the device generally gives higher accuracy in positioning for the car but tends to have large variation on bike. However, we have noted a serial correlation of the recordings implying that an inaccurate recording is likely to be followed by another if the time interval is short.

### Table 3: Proportion of positional recordings matching the road network

<table>
<thead>
<tr>
<th>Factors</th>
<th>Vehicle</th>
<th>Bike</th>
<th>Car</th>
<th>Bus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1s</td>
<td>5s</td>
<td>30s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance radius 0m</td>
<td>60.06%</td>
<td>54.90%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance radius 20m</td>
<td>68.24%</td>
<td>54.90%</td>
</tr>
<tr>
<td>5s</td>
<td>Distance radius 0m</td>
<td>54.90%</td>
<td>87.33%</td>
<td>75.29%</td>
</tr>
<tr>
<td></td>
<td>Distance radius 20m</td>
<td>26.69%</td>
<td>93.27%</td>
<td>74.42%</td>
</tr>
<tr>
<td>30s</td>
<td>Distance radius 0m</td>
<td>73.00%</td>
<td>92.15%</td>
<td>80.95%</td>
</tr>
<tr>
<td></td>
<td>Distance radius 20m</td>
<td>91.18%</td>
<td>92.86%</td>
<td>90.00%</td>
</tr>
</tbody>
</table>

The surprising results for the bike prompted us to run a secondary experiment. We speculated that the positional recordings of the bike were interfered by the surrounding environment. Figure 4 depicts the two routes travelled by the bike at a second occasion. One route coincides with the route used in the original experiment while the second route is a part of the car’s route.

In the first experiment, we had numerous inaccurate recordings in the three areas depicted in Figure 4 by a white circle and two triangles. The circled area is nearby power lines to the north. The areas indicated by triangles have trees with a height of 8-10 meters. In the secondary experiments all settings of the GPS devices were kept as in the first experiment, but the bike travelled both routes at a speed of 20 km/h.

Table 4 gives the proportion of accurate recordings on the two routes. Although the proportion of accurate recordings on the original bike route is higher in the second experiment, it is still rather low. Again most inaccurate recordings happened at the three areas previously identified as problematic. The positional recordings on the car’s route were substantially better. This exercise illustrates that the GPS device may generate (infrequent) errors due to the interferences with the surroundings such as trees and built-ups in a non-obvious way (Modsching et al., 2006).
Table 4: Proportion of positional recordings matching the road network for the bike in the secondary experiment

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Original route</th>
<th>On the car’s route</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance radius 0m</td>
<td>73.83%</td>
<td>89.22%</td>
</tr>
<tr>
<td>Distance radius 20m</td>
<td>58.79%</td>
<td>99.50%</td>
</tr>
<tr>
<td>5s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance radius 0m</td>
<td>50.38%</td>
<td>90.06%</td>
</tr>
<tr>
<td>Distance radius 20m</td>
<td>69.29%</td>
<td>88.38%</td>
</tr>
<tr>
<td>30s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance radius 0m</td>
<td>71.49%</td>
<td>98.78%</td>
</tr>
<tr>
<td>Distance radius 20m</td>
<td>80.13%</td>
<td>100%</td>
</tr>
</tbody>
</table>

4.2 Estimating the velocity

It goes without saying that it is more difficult to estimate a changing velocity than a constant velocity. Drivers (and riders) need to adjust their speed in line with the traffic but also at intersections, roundabouts, tortuous locations, and traffic lights (Jia et al., 2012). This is also true in conducting an experiment of this kind. We used the travel diary of the car and the bike to delete recordings where the intended constant velocity was not possible to maintain. As for the bus, the films were used for deleting recordings where the velocity was not constant.
Figure 5 illustrates how the recorded velocity varies around the pre-set constant velocity. The figure shows the recordings from one device in the car where the device was set to record the velocity in intervals of 30 seconds. There is a tendency that the recorded velocity is generally lower than the actual velocity. Recall that the manufacturer claimed that the error in velocity should be within 0.4 km/h. Table 5 further shows the statistics for the recorded velocity as the average, the standard deviation, and the root mean square error (MSE). The velocity is underestimated by about 5% and the standard deviation exceeds by far 0.4 km/h. The relative error in the recorded velocity seems not however to be related to the setting of the GPS device.

![Figure 5: Recorded velocity versus actual velocity as measured by one GPS device for the car](image)

We have conducted analysis of variance (ANOVA) to formally test for the factors. The error between the recorded and actual velocity was the response variable. The error increased with the velocity. There was no significant difference for whether the distance restriction was on or off. The sampling interval was unrelated to the error, except for the recordings of the bike. In this case the longer sampling interval was associated with a (marginal) increase in the error.

We also checked for a relationship between the error in velocity and the geographical error as discussed in section 4.1. We did so by labelling all positional recordings on the road network as accurate and all those off the road network as inaccurate. Thereafter we repeated the ANOVA including the factor Accurate in the model. It was strongly significant suggesting a greater underestimation of the velocity if the positional recording was inaccurate.

Table 5: Statistics of recorded velocity for bike, car and bus
4.3 Altitudes

The GPS device is presumably able to record the altitude of the vehicle as it travels. However, the manufacturer is not specific about the precision in the recorded altitude. We expect the precision of altitude to be poorer than the geographical position considering for instance the requirement for connection to additional satellites for estimating altitude.

In order to check the precision in the recorded altitude, we first acquired the geo-information of altitude in Borlänge from the national altitude database (NNH)\(^4\). We applied spatial join in ArcGIS 10.1 to join the attribute table of the actual altitude layer to the attribute table of the recorded altitude layer. Each position of the vehicle where a recorded altitude occurred is related to the nearest point in the actual altitude layer. The maximum distance between the position of the recording and the actual altitude layer is 21 meters. This is an inconsequential approximation as the road network covered in the experiment does not contain any steep up- and down-hills. Another (trivial) approximation is the fact that the devices were carried by the rider in a backpack, in the back seat of the car, and in the front seat of the bus. Hence, the altitude of the devices was 1-2 meters above the level of the road network.

The error in recorded altitude with respect to the actual altitude is large. Most of the time the error was within the range of -50 meters and 50 meters, but frequently the error exceeded 100 meters. Considering for instance that the bike path travelled in the experiment was essentially flat such a magnitude in error is enormous.

\(^4\) The altitudes data is provided by Sweden’s Mapping, Cadastral and Land Registration Authority (www.lantmateriet.se). The altitude model is made by laser scanning and has an average altitude error of 0.1 meter and 0.4 meter in the plane.
Moen et al. (1996) discussed the concepts of 2-D and 3-D fix and argued that a 3-D fix should offer a greater precision in estimating the altitude. The GPS device used in the experiment generates a 3-D fix. All the same, the results are not impressive.

5. Concluding Discussion

This paper focuses on a method for evaluating the reliability of portable, standard GPS devices in tracking vehicles. The experiment was conducted by equipping a GPS tracking device BT-338(X) on vehicles being car, bike, and bus and then track the geographical position, velocity, and altitude of the vehicles in the road network. Preprocessing and cleaning of the data was necessary and auxiliary information needed.

The GPS tracking data identified the actual positions of the vehicles fairly successfully. The surroundings of the experiment had no obviously interfering attributes like high built-ups, forests, magnetic fields, and so on. The partially poor identification of the bike’s positions by trees and in the vicinity of magnetic fields shows however that the positional error of the GPS is highly vulnerable to the surroundings (see also Modsching et al., 2006). Fortunately, this problem can be rectified by using map-matching algorithms as proposed by Brakatsoulas et al. (2005), Taylor et al. (2006), and Quddus et al. (2007). We believe that the GPS tracking data on position may be useful for routing, mobility pattern recognition, and general navigation of travelers. However, the accuracy is insufficient in cases requiring high geographical precision such as parking, emergency rescue, and the like.

The tracked instantaneous velocities are quite accurate with a tendency of underestimation. The error between recorded velocity and actual velocity is monotonically increasing with the speed. It should however be noted that we did not study the accuracy regarding acceleration and deceleration which are common phenomena in ordinary traffic.

Concerning the recorded altitudes in the tracking data, we found it to be highly inaccurate and we suggested disregarding this parameter in practical use until further investigations.

The reliability seems to be unrelated to the sampling frequency. Of course, intensive positional recordings provide more details regarding the mobility pattern. However, it comes at the expense of more aggressive data rendering communication, storage, data processing, data mining, and data analysis more costly. Balancing between these aspects is necessarily specific to the domain of application.

There is drawback of GPS devices due to a short effective lifespan (Ryan et al., 2004). The data collection part of the experiment in this paper lasted at the most for two hours; the duration of the device was not a concern here as the operational time for the device is about 11 hours after being fully charged and in continuous mode. However, the lifespan may be a costly drawback in full-scale applications.
Finally, this study examined one specific standard GPS device. It would be interesting in the future to conduct further analyses including other types of GPS device by using the experimental method outlined in this study.

References


