VALIDATION OF A TAXI-FCD SYSTEM

BY GPS -TESTDRIVES

Elmar Brockfeld
Scientist
German Aerospace Center
DLR-Institute of Transport Systems, Rutherfordstr. 2, 12489 Berlin, Germany
phone: +49 30 67055 231, fax: +49 30 67055 291, elmar.brockfeld@dlr.de
Alexander Sohr
Scientist
German Aerospace Center
DLR-Institute of Transport Systems, Rutherfordstr. 2, 12489 Berlin, Germany
phone: +49 30 67055 458, fax: +49 30 67055 291, alexander.sohr@dlr.de
Rüdiger Ebendt
Scientist
German Aerospace Center
DLR-Institute of Transport Systems, Rutherfordstr. 2, 12489 Berlin, Germany
phone: +49 30 67055 278, fax: +49 30 67055 291, ruediger.ebendt @dlr.de

ABSTRACT

In contrast to traditional traffic sensors like induction loops, GPS-position data from Floating Car Data fleets (FCD) are an excellent technology to determine travel times which are needed for traffic information and route guidance systems. This contribution is about the validation of a taxi-FCD system in Berlin with more than 4000 taxis, from which a traffic state map and travel times are generated continuously. The validation is done by analyzing the data measured on five weekdays by test drives with GPS-equipped vehicles on three selected arterial routes. The travel times which are calculated from the taxi-FCD system are in this contribution – in contrast to earlier publications – only based on the most current taxi data delivered and thus do not take "historical" information into account. These data are primarily validated via the travel times of the test vehicles, which can be measured exactly with the GPS-information. This way, occurring differences between travel times of the taxi-FCD system and test drives can be analyzed accurately. As a result, the travel time estimates of the taxi-FCD system are validated and stressed for the three routes as a whole as well as for parts of the route.

INTRODUCTION

Traffic monitoring using Floating Car Data fleets (FCD) has become an important issue of transport research in the recent years. There are three reasons for this: First, popular mobility services require such data which are not available with common data collection technologies like inductive loops, infrared sensors or video images. Especially the boom of navigation devices drives an increasing demand for true and up-to-date travel time information not observable with conventional measuring methods. Second, a lot of communities suffer from heavy traffic on the one hand and very short resources for additional infrastructure investments on the other hand. And third, more and more fleet operators use positioning techniques to manage their vehicles producing some kind of Floating Car Data as a side effect with nearly zero additional costs.

The FCD technology nowadays becomes more and more popular. In contrast to usual approaches, vehicles floating with the traffic stream are used as sensors to give information about traffic states. In recent years, the DLR-Institute of Transportation Systems (DLR-TS) has developed algorithms and technologies to exploit these data. Especially taxi-fleets have been used in several applications as probe vehicles, realizing implementations in various cities in Europe and Asia (see (1), (2), (3) for some examples).

The quality of this sensor is - by now - not extensively analyzed. The big challenge is that in contrast to usual traffic sensors only data of a small amount of vehicles can be used to represent traffic states of a complete traffic stream. Thiessenhusen et al did some first analysis of travel times on urban roads on the basis of some hundred probe vehicles in some cities (see (4)). Jang et al presented analyses concerning the optimum number of probe vehicles to get reliable travel times in Seoul (see (5)). Reinthaler et al analyzed travel times estimation based on FCD especially for hazardous good transports with a field test in Vienna (6) and did some analyses on average path speeds in Düsseldorf (7). Brockfeld et al (8) performed an analysis of travel time information generated from a taxi-FCD system in Hamburg by GPS test drives for some selected routes. The main flaw is that there the test drives – and thus the validation – were done for travel times only between 9 A.M. and 6 P.M, which especially does not cover the important morning peak hours. Furthermore, only single test vehicles drove the defined routes.

In this contribution the taxi-FCD system of DLR-TS - which continuously generates a traffic state map on the basis of more than 4000 data-delivering taxis and monitors travel times on important arterial routes in Berlin - is validated in a special way taking only the most-recent taxi-data for the comparison to the test drives. This means there is no assumption about typical traffic states based on “historical” data of the past days or past equal weekdays. The set-up of the measurement campaign and the validation methodology are described. The results in form of daily variation curves of travel times are validated with the travel times obtained from the FCD system for complete arterial routes. In addition, a detailed analysis of a short segment is presented where the trajectories obtained from the test vehicles are directly compared to single trajectories from taxis which drove that complete segment at the same time.

THE TAXI-FCD SYSTEM BERLIN

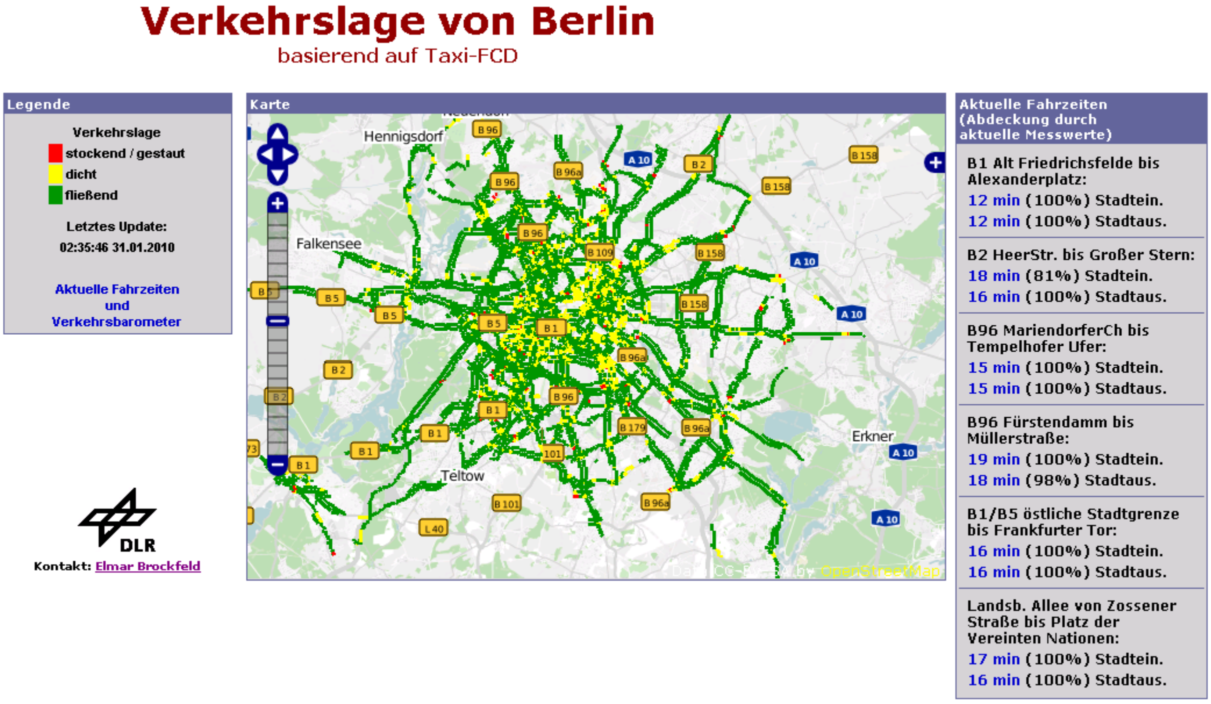

Figure 1. Traffic state map and provision of travel times by monitoring arterials.

The taxi-FCD system for Berlin is running for several years now and has the same architecture as exemplarily described a bit more in detail in (8). Taxis equipped with GPS and GPRS deliver their current position in frequencies of about 30 to 60 seconds to the taxi headquarters. These data then are forwarded to the taxi-FCD system of DLR-TS which matches them on a digital road network, generates a traffic state map with states on each road section and provides travel time information on selected routes (see Figure 1 for an exemplary visualization).

VALIDATION OF THE TAXI-FCD SYSTEM BERLIN

DESIGN OF THE MEASUREMENT CAMPAIGN

For the validation of the taxi-FCD system a measurement campaign with duration of five days was conducted from 17th to 21st of May 2010 as this is a most usual month to conduct representative traffic surveys. Two vehicles were equipped with GPS-sensors which saved their current positions every second. The vehicles drove on three selected arterials in both directions as sketched in the time schedule in Table 1. An emphasis is on the relatively long time period in which the vehicles drove to cover especially the typical peak hours in the morning and the afternoon.

Day	Time period	Vehicle-nr	Route
Monday, 17.05.2010	5 A.M. – 9 P.M.	Veh 1	2 - Frankfurter Allee
	5 A.M. – 9 P.M.	Veh 2	1 - Landsberger Allee
Tuesday, 18.05.2010	5 A.M. – 9 P.M.	Veh 1	2 - Frankfurter Allee
	5 A.M. – 9 P.M.	Veh 2	1 - Landsberger Allee
Wednesday, 19.05.2010	5 A.M. – 9 P.M.	Veh 1, Veh 2	3 - Tempelhofer Damm
Thursday, 20.05.2010	5 A.M. – 9 P.M.	Veh 1	2 - Frankfurter Allee
	5 A.M. – 9 P.M.	Veh 2	1 - Landsberger Allee
Friday, 21.05.2010	5 A.M. – 9 P.M.	Veh 1	2 - Frankfurter Allee
	5 A.M. – 9 P.M.	Veh 2	1 - Landsberger Allee

Table 1. Total time schedule of the two GPS-equipped vehicles during the five days.

Figure 2. The three routes on which the test drives took place.

The routes themselves are depicted in figure 2. They are typical arterial roads with morning and afternoon rush hour and of relatively long length.

RESULTS OF THE MEASUREMENT CAMPAIGN

Table 2 shows the number of traces obtained from the test vehicles for each route, separated by the direction of the routes (second number: -1 towards center, -2 away from center). The numbers for the direction- pairs “to center” / “from center” are not equal because in some cases the logging of the GPS-signal failed, especially at the beginning or end of a trace. In these cases a really measured travel time could not be obtained.

Route	Length [km]	Days in May 2010	Vehicle-nr	# Traces
1-1 Landsberger Allee to center	11.927	17., 18., 20., 21.	Veh 2	67
1-2 Landsberger Allee from center	11.916	17., 18., 20., 21.	Veh 2	63
2-1 Frankfurter Allee to center	14.153	17., 18., 20., 21.	Veh 1	57
2-2 Frankfurter Allee from center	14.160	17., 18., 20., 21.	Veh 1	56
3-1 Tempelhofer Damm to center	11.857	19.	Veh 1, Veh 2	34
3-2 Tempelhofer Damm from center	11.859	19.	Veh 1, Veh 2	32

Table 2. Final route definitions and obtained number of test vehicle traces on the routes.

As the test vehicles delivered GPS trajectories in frequency of one second in the normal case the exact travel time was measured exactly at that points where the definition of the routes start and end. The travel times which are permanently calculated from the taxi-FCD system and related to the exact route definitions, too, are validated via the travel times measured exactly by the two GPS-equipped test vehicles. In contrast to the “normal” methodology the taxi-FCD system at DLR-TS delivers travel times, only the data of exactly the same day at exactly this time of day are taken for these analyses. Thus, time intervals with missing data are not filled by “historical” data (of the days before or some same weekdays before). If for a daytime interval of 15 minutes taxi data are only available for parts of the route, the rest is linearly extrapolated. This means for example if a route is only covered by 80% of the total length with current data, the resulting travel time is calculated for this part and multiplied by 100/80 to get an estimated Taxi-FCD travel time for the whole route.

The following Figures 3 to 5 depict exemplarily the results of one particular day on each of the three routes. Shown are the results delivered by the Taxi-FCD-system with data only of the according day in comparison to the travel times of the test vehicles. In general, typical daily variations can be seen in the taxi data as well as for the test vehicles.

Figure 3 for Route 1 shows that for lower travel times of about 14 to 22 minutes the results are quite similar. But during the typical peak hours from 6-9 A.M. in direction to the center the test vehicle took some minutes more than the FCD-system tells. During the afternoon peak from 15-18 the travel times match relatively good except for an outlyer at 35 minutes.

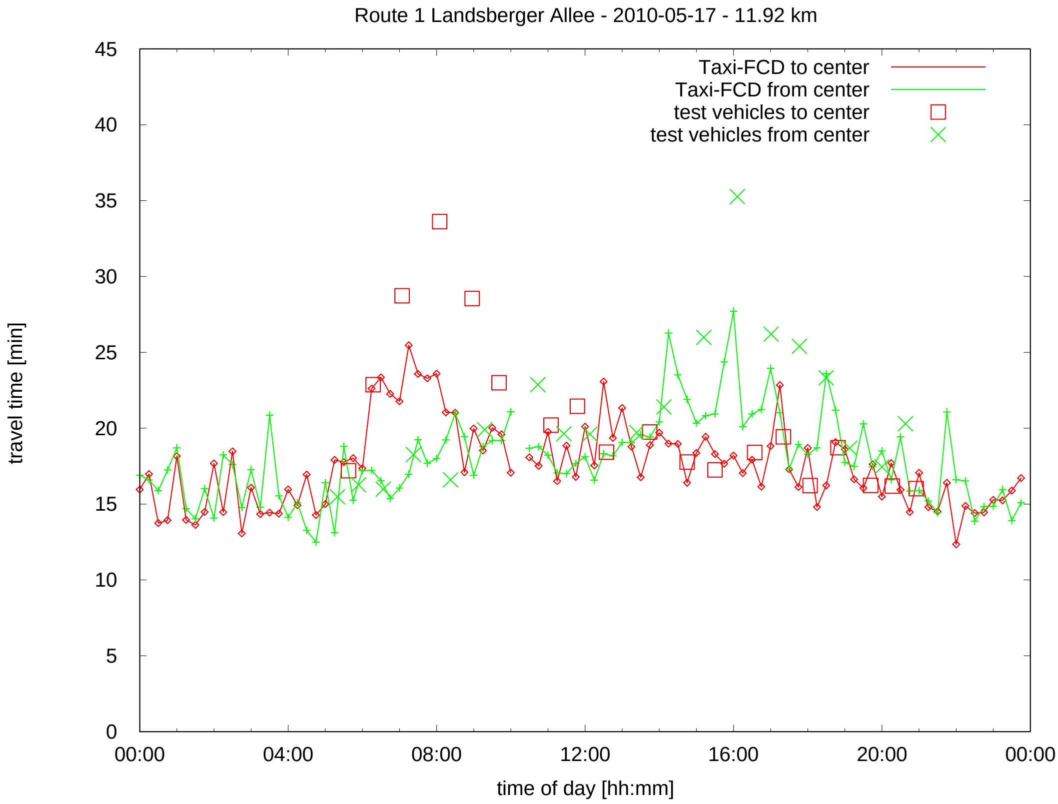

Figure 3. Route 1 - Examples of average daily variation curves of travel times from Taxi-FCD in comparison to those obtained from the test vehicle.

Figure 4 for Route 2 shows a similar picture as seen for Route 1. In general the data match quite well. Especially in the morning peak hour from 6-9 A.M. the test vehicles follow the Taxi-FCD data, but again at about 8 A.M. with about 4 minutes longer travel times. The afternoon peak at this particular day seems to be very short as for Taxi-FCD it is only around 17 o'clock. The test vehicles show a similar picture with two values in the peak area around this time, which seems to be in line.

Figure 5 for Route 3 shows also good results with a similar structure as on the other routes. Remarkably here that for the time of the relatively smooth morning peak obtained for Taxi-FCD the test vehicles took in some cases much longer travel times.

Figure 4. Route 2 - Examples of average daily variation curves of travel times from Taxi-FCD in comparison to those obtained from the test vehicle.

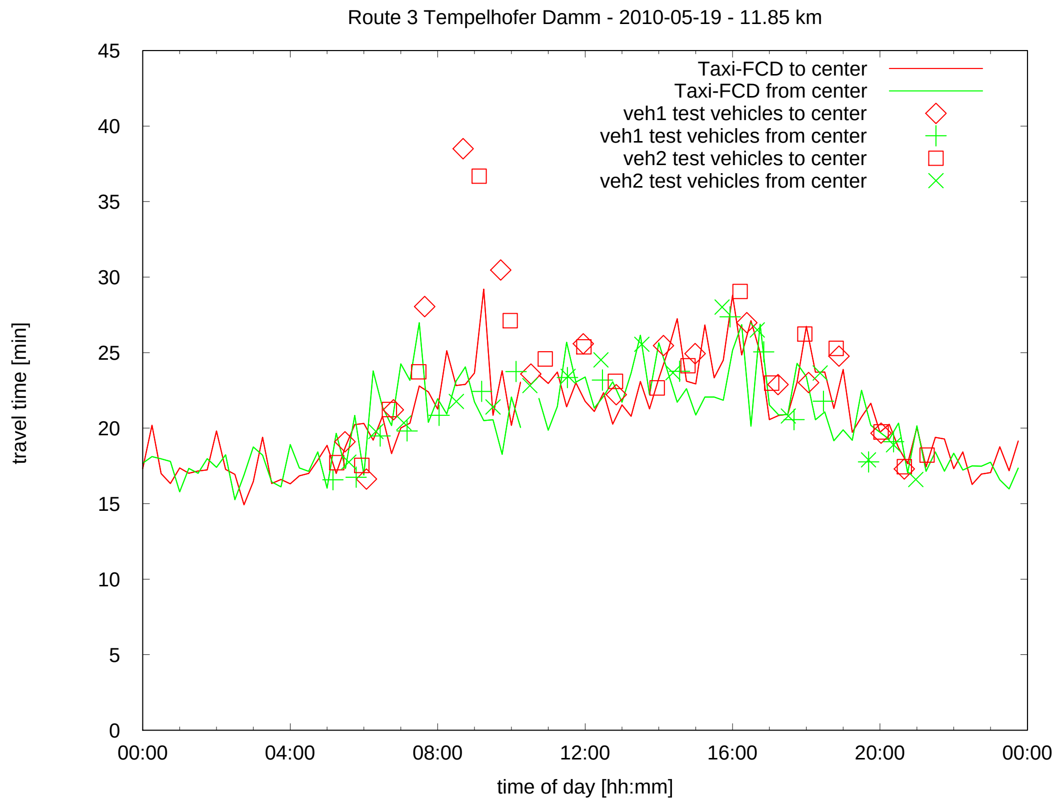

Figure 5. Route 3 - Examples of average daily variation curves of travel times from Taxi-FCD in comparison to those obtained from the test vehicle.

VALIDATION OF TRAVEL TIMES FOR COMPLETE ROUTES

From all obtained results for the three routes and all calculated travel times the symmetric mean absolute percentage error (SMAPE) is calculated defined as:

$$SMAPE = \frac{1}{N}\sum_{1}^{N}\frac{|tt(TEST)_i - tt(FCD)_i|}{(tt(TEST)_i + tt(FCD)_i)/2}$$

N is the number of values to compare, $tt(TEST)_i$ the travel time of a test vehicle and $tt(FCD)_i$ the travel time of the time-corresponding 15-Minute-averaged value obtained from the Taxi-FCD system on the basis of data only from the according day.

Route	SMAPE		
1-1 Landsberger Allee to center	11.19 %	12,73 %	11,04 %
1-2 Landsberger Allee from center	14.38 %		
2-1 Frankfurter Allee to center	7,19 %	8,75 %	
2-2 Frankfurter Allee from center	10.34 %		
3-1 Tempelhofer Damm to center	12,71 %	11,63 %	
3-2 Tempelhofer Damm from center	10,48 %		

Table 3. Error in the comparison of travel times of travel time Final route definitions and obtained number of test vehicle traces on the routes.

The general error obtained is about 11 % as can be seen in Table 3. As Route 2 is a bit longer than the other two the results correspond completely and there is no sign that arterial routes to the city center or away from city center have different results.

VALIDATION OF TRAVEL TIMES FOR ROUTE SEGMENTS

Route 3 (Tempelhofer Damm) was in each driving direction split into six parts which were analyzed for the two driving directions leading to and from city center. In Table 4 only the directions leading to center are given, which are indicated by a “-1” suffix of the Route number. The same principle is for the opposite direction.

Route / name	Route part	Description
3-1 / Tempelhofer Damm	3-1_1	Platz der Luftbrücke until end of Hallesches Ufer
3-1 / Tempelhofer Damm	3-1_2	A100 until Platz der Luftbrücke
3-1 / Tempelhofer Damm	3-1_3	Friedrich-Karl-Str until A100
3-1 / Tempelhofer Damm	3-1_4	Alt-Mariendorf until Friedrich-Karl-Str.
3-1 / Tempelhofer Damm	3-1_5	Buckower Chaussee until Alt-Mariendorf
3-1 / Tempelhofer Damm	3-1_6	Start of Gross-Ziethener-Str. until Buckower Chaussee

Table 4. The three routes and their route parts leading to the city center.

While it is very unlikely that a taxi drives one of the three arterial routes completely, this can frequently be observed for the shorter route parts. By selecting all taxis which drove a part of a route completely, it is possible to obtain the travel times for this route part. In contrast to average travel times as calculated by the taxi FCD system, these travel times are the exact travel times experienced by individual taxis. Thus, they can be used to validate the travel times of the data obtained by the test vehicles. This is achieved by analyzing the compliance of the test data with the observed variation of the taxi traffic. In a second step, the measured variation of the taxi traffic on the route parts can also be used to check whether the deviation of the average travel times of the taxi FCD system from the travel times of the test vehicles stays within the range of the variation that is immanent to the traffic.

For every route part as described in Table 4 and for every hour of the day, the taxis which drove a part of the route completely have been selected. The standard deviation of the travel times x_i of the selected N taxis $(1 \le i \le N)$ is defined as:

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

Figure 6 (left) depicts the travel times of the two test vehicles on route part 3-1_2 (leading to center) whereas in Figure 6 (right), the results for route part 3-2_2 (away from center) are shown. In both figures, the travel times of the individual taxis driving on the complete route part together with their standard deviations as aforementioned are included.

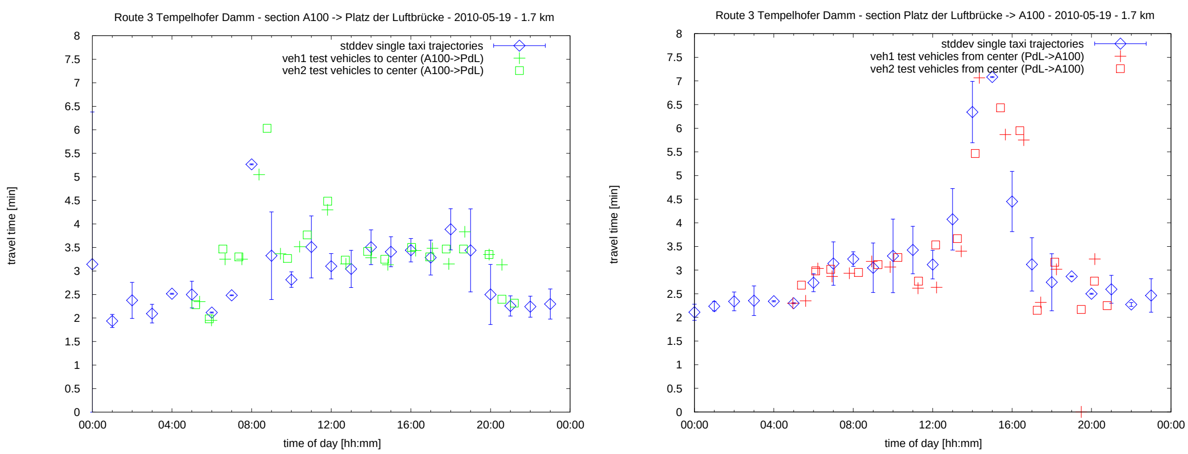

Figure 6. Route 3 – Example for comparison of travel times by the test vehicles with distribution of single taxi-trajectories which drove completely in the same hour of day on a segment of length 1.7 km. (left: direction to center, right: direction from center)

E.g. for Route 3-1_2, around 79% of the measured test data values are within the range of the standard deviations of the taxi travel times. This result shows that the test data complies with the observed taxi traffic on the route part.

In a second step, for every route part, the percentage coefficient of variation

$$\mathrm{CV} = \frac{\sigma_x}{\bar{x}} \cdot 100 \quad (\%)$$

has been calculated for every hour of the day, considering the travel times of the taxis selected in that hour. Then a mean CV, averaging over all these single CVs for route parts of the route has been calculated for the date of measurement, the 19th of May 2010. The obtained percentage value of 15.8% is a threshold that should not be exceeded by the (symmetric) mean absolute percentage error of a valid taxi FCD-system, which is the case for this study with a SMAPE of about 11 % for the comparison of taxi with test vehicle travel times.

CONCLUSION

The results show that in general the taxi-FCD system even if only based on most current travel times delivers quite good values during the most time of the day. The variations in travel time are inside the normal range which can be expected. But especially during typical peak periods the test vehicles often showed longer travel times than the taxi-FCD system. It is not clear by now why this is the case. An explanation may be that taxi drivers are generally better informed about the typical traffic states in the city and thus avoid some routes (or parts of it) especially during peak hours. Future research concerning this data set will be to analyze how many taxi data are available for which time of day and especially which parts of the routes are not covered good enough. Additionally, the supplementation of the current taxi data by “historical” data will lead to much better results.

ACKNOWLEDGEMENT

The authors thank the taxi headquarters TAXI Berlin TZB GmbH for the provision of the data for the taxi-FCD system.

REFERENCES

(1) Schäfer, R.-P.; Thiessenhusen, K.-U.; Brockfeld, E.; Wagner, P. (2002): A traffic information system by means of real-time floating-car data. ITS World Congress 2002, 11th-14th Oct 2002, Chicago, USA.

(2) Lorkowski, S.; Brockfeld, E.; Mieth, P.; Passfeld, B.; Thiessenhusen, K.-U.; Schäfer, R.-P. (2003): Erste Mobilitätsdienste auf Basis von "Floating Car Data". In: RWTH Aachen [Hrsg.]: Tagungsband zum 4. Aachener Kolloquium "Mobilität und Stadt", Stadt Region Land, 75, pp. 93 - 100, 4. Aachener Kolloquium "Mobilität und Stadt" AMUS, 31st Jul - 1st Aug 2003, Aachen (Germany), ISBN 3-88354-140-0.

(3) Kühne, R.D.; Schäfer, R.-P.; Mieth, P.; Lorkowski, S.; Bei, X. (2005): Vehicle Probes as Data Collectors for Asian Metropolitan Areas. 4th Asia Pacific Conference, 8th - 10th Nov 2005, Xi'an, China.

(4) Thiessenhusen, K.-U.; Brockfeld, E.; Schäfer, R.-P.; Wagner, P. (2002): Analysis of travel times and routes on urban roads by means of floating-car data. European Transport Conference PTRC, 4th - 9th Sept 2002, Cambridge, UK.

(5) JeongAh Jang; Tae-Ho Yoo; Jeong-Whon You; Seung-Hwan Lee (2006): Determining the optimum number of vehicle probes with accounting the reliability of link travel times. 13th World Congress on Intelligent Transport Systems and Services, 9th - 12th Oct 2006, London, United Kingdom.

(6) Reinthaler, M., Zajicek, J. (2007): Real time route analysis based on floating car technology. In: 18th IASTED International Conference: modelling and simulation - Proceedings, pp. 609 – 612, 2007-05-30 – 2007-06-01, Montreal, Canada.

(7) Reinthaler, M., Nowotony, B., Hildebrandt, R., Weichenmeier, F. (2007): Evaluation of speed estimation by floating car data within the research project DMotion. In: ITS2007 - 14th World Conference on Intelligent Transport Systems - Proceedings on CD-ROM (Paper-ID 2135), 2007-10-09 - 2007-10-13, Beijing Exhibition Center, Beijing, China.

(8) Brockfeld, Elmar, Wagner, Peter und Passfeld, Bert (2007): Validating Travel Times calculated on the basis of taxi floating car data with test drives. In: ITS2007 - 14th World Conference on Intelligent Transport Systems - Proceedings on CD-ROM (Paper-ID 4134), 2007-10-09 - 2007-10-13, Beijing Exhibition Center, Beijing, China.