A COMPREHENSIVE TAXI ASSESSMENT INDEX USING
FLOATING CAR DATA

Dao-Zheng Zhang
Center for UAV Applications and ITS Research
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai, 200240, China
Phone: (86-21) 3420 6674
Email: zdzwd123@sjtu.edu.cn

Zhong-Ren Peng*, Ph.D
Professor
Center for UAV Applications and ITS Research
Shanghai Jiaotong University
800 Dongchuan Road, Shanghai, P.R. China
And
Department of Urban and Regional Planning
University of Florida
P. O. Box 115706
Gainesville, FL 32611-5706
Email: zpeng@dcp.ufl.edu

Daniel (Jian) Sun, Ph.D
Associate Professor
Center for UAV Applications and ITS Research
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai, 200240, China
Email: danielsun@sjtu.edu.cn

Word count: 5,242+ 6 Figures + 4 Tables = 7,742

Paper submitted for possible Presentation at the 93rd Annual Meeting of
Transportation Research Board
August 1, 2013

* Corresponding author
ABSTRACT

With the expansion of urban area and development of taxi system, problems such as low operation efficiency, high taxi-idling rate, and long passenger waiting-time arise. Although various studies have been conducted, only limited overview of the factors towards urban taxi system has been provided. Consequently, a comprehensive evaluation of taxi system is essential for the urban planner to analyze the current situation and take effective measures. This paper, by using Floating Car Data (FCD), proposes a Comprehensive Taxi Assessment Index (CTAI) to quantify the quality of existing urban taxi system with the assistance of Geographic Information System (GIS) technology. The proposed index system extracts and classifies key factors, reflecting the taxi system from the perspectives of operation efficiency, customer and taxi-driver satisfaction. The system contributes to improving the organization and operation of urban taxi system. Based on the data obtained from the city of Shenzhen, Guangdong province, China, for both weekday and weekends (Dec., 2011), the proposed CTAI was illustrated using the Principal Component Analysis (PCA) with ArcGIS 10.0 platform. Results indicate the system provides a good multi-dimensional view to delve into the existing urban taxi operation, thus to point out the most sensitive indices towards the entire system, which consequently provides guidelines for future improvement and management of urban taxi system.

Key Words: Taxi Evaluation System, Principle Component Analysis (PCA), Floating Car Data (FCD), Geographic Information System (GIS).
1. INTRODUCTION

The urban taxi system is an important component of local transit system due to its relative fast speed and convenience. However, with the expansion of cities and development of the urban taxi system, problems arise such as low operation efficiency, high taxi-idling rate, and long passenger waiting time. The imbalance of supply and demand in urban taxi system stirs up concerns of both city planners and urban citizens. Previous studies focus on increasing the operation efficiency and reducing idling rates (1-3). Other research tends to assess the performance of urban taxi system from the perspective of passengers (15, 16) and taxi-drivers (7, 8). However, few studies present an overall assessment framework to evaluate the performance of the entire taxi system, including the system operational efficiency, the passenger satisfaction and even the benefits of taxi drivers. This research aims to establish a comprehensive assessment framework from the perspectives of system operators, passengers and taxi drivers. A quantitative analysis model is provided to validate the assessment framework by using floating car data (FCD). More specifically, the three sub-objectives are to:

1. Determine the indices which could reflect different aspects of the urban taxi system;
2. Develop a comprehensive framework to model the goodness of the urban taxi system from the perspectives of system operators, passengers and taxi-drivers;
3. Validate the proposed model using the FCD data and propose an overall index for the performance evaluation.

The remainder of this paper is structured as follows: First, the literature related to various aspects of the urban taxi system is summarized in Section 2. Then, the framework of CTAI is presented and key indices are defined in the following Section 3. Next, in Section 4, model implementation, sensitivity analysis and data comparison are described. Finally, conclusion and recommendations for future work are presented in Section 5.

2. LITERATURE REVIEW

Large amounts of taxi-related studies have been conducted, including topics of overall operation management, best path choice of taxi drivers, taxi availability probability to passengers, measurement to improve passenger service, and so on. Focus of the past researches could be categorized from perspective of system operator, taxi driver and passenger.

Yang and Wong proposed a network to describe the cruising strategy for both vacant and occupied taxi in a road network, which determined a number of system performance measurements at equilibrium (7). Wong et al. modified the model by considering the relationship between variable demand and multiclass vehicle assignment (2). After that, Yang et al. proposed another equilibrium model to characterize the bilateral searching and meeting between passengers and taxi on road
network (3). The models above contributed to the understanding of the taxi operation system by providing guidelines to preview its efficiency. Moreover, the taxi dispatching system was studied during the last several years to promote system operation efficiency. Lee initially proposed a dispatching system based on real-time taxi demand and traffic condition (4). In his research, a real-time, two-step taxi pooling dispatching system was proposed to improve the passenger satisfaction (5). Chang et al. proposed a data mining process based on hotspots analysis to predict taxi demand considering the factors of time, weather and taxi location (6). Previous researches provided deep insight about the indicators that could influence the performance of taxi operation system, while few discussed the relationship between such indicators.

From the perspective of a taxi-driver, Liu et al. categorized taxi drivers into top driver and ordinary driver by their daily income (7). Yuan et al. detailed the research, considering the fare (occupied) distance per unit of working time. Three groups were divided- top 10% as high-profit drivers, bottom 10% as low-profit drivers and the rest as medium profit drivers (8). The similarity of above research is the use of Global Positioning System (GPS) installed in taxi. The application of GPS in transportation was largely studied and implemented (9-11). FCD data, obtained from GPS, provided another aspect to view the taxi system (12-15). The difference between (7, 8) and this study is that an evaluation system providing here could illustrate a systematic way to view top-drivers’ performance in the whole taxi system.

From the perspective of a passenger, Li et al. proposed the hotspots by counting all the pickup and dropdown events in partitioned grids (16). Yuan et al. improved this method by providing a parking candidate detection algorithm, considering the spatial-temporal feature, point-of-interest (POI) feature, and collaborative feature. The popularity of area was calculated via filtering and clustering the historical taxi data. A probability model was used to minimize passengers waiting time and hail-walking distance (8). The study of a passenger’s decision and preference also include a case study modeled passenger’s preference for a dispatching system (17). Sheu et al. investigated taxi hailing in vehicular hotspot networks and proposed a protocol for taxi hailing, which indicated a desired improvement of taxi operation efficiency (18). These researches tended to analyze customer satisfaction from different aspects individually, while, less discussion covered the inner relationship between these aspects.

In summary, previous studies on the urban taxi system focused on specific aspects, hindering an all-round assessment towards the whole system. To this end, this study delved into the problem by offering a series of indices closely related to the assessment of the urban taxi system, from three different levels, which are operation efficiency, customer and driver satisfaction.

3. RESEARCH FRAMEWORK AND KEY INDICES

In this study, an urban taxi evaluation system is proposed based on multi-dimensional indices. As illustrated in Figure 1, the Comprehensive Taxi Assessment Index (CTAI) could be analyzed from three different perspectives including an operation system,
customers and taxi-drivers.

**FIGURE 1 Comprehensive Taxi Assessment Index Framework.**

The system-level indices focus on the evaluation of the system’s operation efficiency. In other words, the operation efficiency is the reflection of the demand-supply equilibrium of the urban taxi system. Although there are plenty of indicators which could be used to describe the efficiency of the urban taxi system, most of them such as road idling rate, taxi space idling rate and etc., are closely related, reflecting the system operation performance from slightly different angles. Based on FCD data, the system operation efficiency could be quantified and analyzed from the indicators of time idling rate and space idling rate, through which, the demand-supply relationship between taxi and passenger could be assessed.

The customer-level indices aim to reflect the satisfaction of customers (passengers), which could be further separated based on two scenarios. Before getting in the vehicle, a customer’s satisfaction is determined by indicators which could reflect the relationship between their location and potential hotspots (area with higher concentration of taxis). When a customer is driven by the taxi, the satisfaction of the customer could be roughly determined by the comparison between real path and the shortest path. Therefore, the customer’s satisfaction could be generalized by Walking Distance Index, Taxi Availability Index, Driving Distance Ratio, and Driving Time Ratio as explained in the following section.

The driver-level indices are indicators to reflect the satisfaction of individual taxi-driver. To this end, the driver’s daily working profit, working time and occupied working speed are three significant indicators (2), which could perform relative sound reflection about driver’s feeling towards the working condition.

**Operational Efficiency Index**

At the system level, taxi-idling rates are key indicators to reflect the overall system performance of the operation of urban taxi system. In general, the idling rates should be in a certain range for a supply-demand balanced taxi system. High taxi-idling rate indicates surplus of urban taxi vehicles and redundant cruising vehicles. On the other hand, low taxi-idling rate reveals that relative high vehicles occupancy, which could lead to a larger amount of passengers waiting for taxis on the road. From this perspective, the idling rates, including time-idling rate and space-idling rate, are
significant indicator to the efficiency of taxi system.

The indices of time idling rate and space idling rate could be defined as follows.

Time idling vectors: \( r_i = (R_i^t, R_i^t, ..., R_i^t) \);

Space idling vectors: \( r_i = (R_i^s, R_i^s, ..., R_i^s) \),

\( R_i^t \) and \( R_i^s \) indicate the time idling rate and space idling rate for a specific vehicle during a period of time.

\[
R_i^t = \frac{1}{T_i} \sum_{j=1}^{N_i} T_{ij} \quad (i=1,2,\ldots,n) \quad , \quad R_i^s \in [0,1],
\]

\[
R_i^s = \frac{1}{S_i} \sum_{j=1}^{N_i} S_{ij} \quad (i=1,2,\ldots,n) \quad , \quad R_i^s \in [0,1]
\]

where,

\( N = \) the total number of passenger-taken times;

\( T_{ij} = \) the time in \( j^{th} \) taxi-taken interval for \( i^{th} \) vehicle;

\( S_{ij} = \) the length in \( j^{th} \) taxi-taken interval for \( i^{th} \) vehicle.

From the definition above, the rate concerning both time idling and space idling of the whole system in a specific period is:

**Time Idling Rate (TIR)**: \( \sum_{i=1}^{n} R_i^t \)

**Space Idling Rate (SIR)**: \( \sum_{i=1}^{n} R_i^s \)

Practical operations and management experience of urban taxi system indicates that the taxi-idling rate should be rationally kept within the range of 30%-40% (19). As a result, the Time Idling Rate Index (TIRI) and Space Idling Rate Index (SIRI) were calculated and are showed in Figure 2.

**FIGURE 2** Index calculation of idling rate indices.

<table>
<thead>
<tr>
<th>Score of TIRI or SIRI</th>
<th>10</th>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
</table>
| Space or Time Idling Rate | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100

**Customer Satisfaction Index**

The satisfaction of a customer could be assessed from two periods: customer waiting
period and in-vehicle period. Indicators were proposed to reflect the customer satisfaction under the assumption that customers in the same local area (with radius less than 1km) are generally with similar travel characteristics (8).

The ascertained procedure of customer-level index is shown in Figure 3:

1. Data processing. The original FCD data contains invalid points produced by equipment error and signal transmission missing, so it is essential to filter the data before usage.

2. Hotspot determination. Pick-up points in each fishnet grid (1km*1km) are counted and ranked in decreasing order numerically. The areas with high popularity (hotspots), indicating high taxi-demanding areas, are chosen as the target areas. After then, the customer satisfaction in such areas could be quantified.

3. Indices obtaining. Indicators covering both customer waiting period and in-vehicle period are chosen and calculated in CTAI scope.

**FIGURE 3** Ascertained Procedure of customer-level index.

*Walking Distance Index (WDI)*
Previous research indicated that passenger preferred to walk less than 400m when waiting for taxi (20). Due to the difficulty in obtaining actual pick-up location based on FCD data, collected data points where the taxi status changes from “available” to “occupied” could be roughly regarded as the real pick-up points. The radius of hotspot is set at 50m and the buffer 400m in order to count all the pick-up points. Detailed illustrations on hotspot and buffer area are provided in Figure 4.

FIGURE 4 Illustration for hotspot and buffer area.

Walking Distance Index (WDI) indicates the distance between pick-up point and hotspot area.

\[
WDI = \begin{cases} 
10 & r \in [0, r_0] \\
(\frac{r}{R-r_0} + \frac{R}{R-r_0}) \times 10 & r \in (r_0, R] \\
0 & Otherwise
\end{cases}
\]

where, \( r_0 \) = radius of hotspot
\( R \) = radius of buffer area
\( r \) = distance between pick-up point and the hotspot center

Taxi Availability Index (TAI)

When a customer is waiting for the ride, the probability of taxi coming is a key indicator directly influencing the satisfaction of taxi system. A relative high probability means that customer would spend less time on waiting. Under the assumption that no real-time information are provided taxi-driver and no preferred routes and waiting spots for taxi-driver, the driving path chosen is random and aimless. The taxi-coming probability satisfies the Poisson distribution:

\[
P_k = \frac{e^{-\lambda} \lambda^k}{k!}
\]

where, \( P_k \) = the probability of \( k \) taxis available during interval \( x \),
\( \lambda \) = the average taxi coming rate,
In this case, \( \lambda = 1/\frac{l_i}{N_i V_i} \)

Where, \( l_i \) = the length of road segment \( i \),
\( N_i \) = number of taxi on this road,
\( V_i \) = average speed on this road.

The probability of taxi coming

\[ \Pr_{\text{taxis}} = 1 - \sum_{k=0}^{\infty} \frac{(N_i V_i T)^k}{k!} \exp(-\frac{N_i V_i T}{l_i})(1 - R_i^T) \] (1)

Where, \( T \) = time interval,
\( R_i^T \) = time idling rate.

Denote the taxi number density \( n_i = \frac{N_i}{l_i} \), using Taylor’s expansion, the equation (1) could be simplified as follows.

\[ \Pr_{\text{taxis}} = 1 - \exp(-\frac{N_i V_i T}{l_i} - R_i^T) \] (2)

From equation (2), the probability of taxi availability could be represented with the formula with variables of number of the taxis on road, road length, average speed, and time idling rate.

Hence, the Taxi Availability Index could be defined as \( TAI = 10 \times \Pr_{\text{taxis}} \).

Driving Distance Ratio (DDR) and Driving Time Ratio (DTR)

Driving Distance Ratio (DDR) is the ratio between real path mileage and shortest path mileage, while Driving Time Ratio (DTR) reflects the ratio between real path time and shortest path time. Usually, the taxi driver tends to choose the relative fast path instead of the shortest path (7), as shown in Figure 5. However, this strategy might not be a desirable choice under a systematic consideration, which compromising the customer satisfaction for a single trip. In practice, DTR is higher than DDR since unexpected condition such as congestion will delay the actual time.
Figure 5 Taxi’s real path and shortest path.

The definition of DTR and DDR are provided as follows:

\[ DTR = \frac{TR_i - TR_{\text{min}}(N)}{TR_{\text{max}}(N) - TR_{\text{min}}(N)} \times 10 \]

\[ DDR = \frac{DR_i - DR_{\text{min}}(N)}{DR_{\text{max}}(N) - DR_{\text{min}}(N)} \times 10 \]

where, \( TR_i \) = ratio of real trip time and shortest path time,
\( DR_i \) = ratio of real trip length and shortest path length,
\( TR_{\text{min}}(N), \ TR_{\text{max}}(N) \) are the maximal and minimal value of \( TR_i \) of all trips with pick-up point in the hotspot,
\( DR_{\text{max}}(N), \ DR_{\text{min}}(N) \) are the maximal and minimal value of \( DR_i \) of all the trips in hotspot area.

Driver’s Satisfaction Index

The satisfaction of an individual taxi driver is another aspect to assess the quality of existing taxi operation system. In this level, indicators which could reflect the taxi drivers’ concern are proposed. Hence, the Driver’s Satisfaction Index (DSI) considers individual driver’s profits and working condition, reflected by the indices of Profit Index (PI), Working Time Index (WTI) and working Occupied Speed Index (OSI).

Working profit is closely related to individual driver’s satisfaction. Based on the condition that different cities have their unique rules to charge the fee per passenger trip, a driver’s profit differs on the case by case basis. Taking the charging rule in China as an example, the fare is separated into two parts- starting fee (minimum charge) and fees per mile. Starting fees differ between day and night, so the driver’s daily working profit could be calculated separately by revenue of day (Rd) and revenue of the night (Rn).

\[ \text{Daily Working Profit} = R_d + R_n \]
where, 
\( S_i \) = taxi traveling distance when occupied,
\( D \) = starting fee when the traveling distance is less than 3 km,
\( b \) = fees per mile over 3.0km,
\( D_n \) = starting fee at night (23:00-05:00),
\( b_n \) = fees per mile over 3 km at night, usually \( b_n > b \),
\( c \) = gasoline price per mile.

Research shows that the income of a taxi driver satisfies the normal distribution (7). Driver’s daily profit equals to income subtracted by expense of gasoline, which also satisfies the normal distribution. The Profit Index (PI) could be defined:

\[
PI_i = \frac{P_i - P_{\min} (N)}{P_{\max} (N) - P_{\min} (N)} * 10
\]

where \( P_i \) is the daily profit of driver \( i \), \( P_{\min} (N) \) is the minimal profit of the total \( N \) drivers, \( P_{\max} (N) \) is the maximal profit among all drivers.

Working Time Index (WTI) is determined by the daily working hours of each driver. Widely recognized viewpoint is that the suitable working time per day is 8 hours. In China, most taxi-drivers tend to adopt the two-shift rules to maximize the profit, which means that two drivers operate the same vehicle in an entire day. Taking this factor into WTI definition:

\[
WTI = \begin{cases} 
1.25 \times \frac{H}{2} & H \in (0, 16) \\
10 - 1.25 \times \left( \frac{H}{2} - 8 \right) & H \in (16, 24) 
\end{cases}
\]

where \( H \) is the total working hours for each vehicle.

The occupied speed of taxi in work is an indicator to reflect the speed when there are passengers in the vehicle. Experience shows that drivers tend to reach the destination with a relative higher speed to maximize profit (22). Considering the speed limitation of different roads, Occupied Speed Index (OSI) could be defined as follows:

\[
OSI = \sum_k \left( \sum_i \left( \frac{\sum_{j=1}^n \left( \frac{v_i^j + v_i^j + \cdots + v_i^j}{V_i^j} \times 10 \right)}{i} / k \right) / 10 \right)
\]

where, \( k \) = the number of passengers who were taken during an interval,
\( i \) = the sequencing label of road in the \( k^{th} \) passenger taken event,
\( n \) = the sequencing number of record in \( i^{th} \) road
\( V_i^j \) = the speed limitation of Road \( i \),
\( v_{n}^{i} \) = the nth instantaneous speed on travel segment \( i \).

**Principle Component Analysis Approach to the System**

To analyze large amounts of dynamic data, Principal Component Analysis (PCA) is widely used in various fields of research. PCA extracts the component from data structures that explains most of the variance. In this paper, PCA is used to evaluate the CTAI of urban taxi system.

Since the largest eigenvectors contain the most useful information relating to the problem while the remaining eigenvectors mostly contribute to the noise. In the case of urban taxi system assessment, the cumulative percentage of total variance explained could be considered to explore the dimension reduction. The detailed discussion about using PCA method to assess the CTAI could be found in Section 4.

4. IMPLEMENTATION AND ANALYSIS USING GIS

**Data Description**

The Floating Car Data (FCD) is collected by the GPS equipment in each taxi, which contains specific information about each taxi in operation. The information collected includes taxi location in longitude, latitude, timestamp, vehicle identification, operation status (empty or occupied), spot speed, and azimuth.

The CTAI system was implemented with the FCD data from Shenzhen, China on Dec. 2nd, 2011 (weekday) and Dec. 4th, 2011 (weekends), including the information from 3096 taxis with more than 5.73 million records.

**PCA Method Implementation and Sensitivity Analysis**

Based on the established CTAI, PCA could be implemented with FCD data from the levels of system, customer and driver respectively. The statistical results are shown as in Table 1.

| TABLE 1 Original Statistics for Various Indices Calculation (Dec. 4th, 2011) |
|-----------------------------------------------|---------------|------------------|-----------------|-----------------|------------------|
| Description                        | Index | Min  | Max   | Mean | Std. Dev | Mean Score |
|-----------------------------------------------|---------------|------------------|-----------------|-----------------|------------------|
| Time Idling Rate                      | TIRI            | 0.29 | 0.46  | 0.35 | 0.036   | 10              |
| Space Idling Rate                     | SIRI            | 0.27 | 0.67  | 0.49 | 0.095   | 8.54           |
| Distance to Hotspot(m)                | WDI             | 137  | 281   | 233  | 31.3    | 5.54           |
| Taxi Availability Probability        | TAI             | 0.60 | 0.99  | 0.86 | 0.084   | 8.61           |
| Driving Time Ratio                   | DDR             | 1.1  | 5.5   | 3.1  | 0.808   | 5.51           |
| Driving Distance Ratio               | DTR             | 1.19 | 1.59  | 1.37 | 0.081   | 5.47           |
| Daily Profit (RMB)                   | PI              | 192  | 1285  | 752.8 | 201.04 | 5.13           |
| Occupied Speed (km/h)                | OSI             | 7.91 | 43.6  | 25.39 | 5.489   | 4.89           |
| Daily Working Time (h)               | WTI             | 14.1 | 24.0  | 23.68 | 1.396   | 5.19           |

(Observation: 50)

Using PCA to assess the CTAI system, choose the principal components with
higher Variance Explained. The result of applying this method is presented in Table 2, from which, it could be seen that when choosing six components, the total variance explained reaches 90.9%.

**TABLE 2** Analysis of variance in PCA

<table>
<thead>
<tr>
<th>Comp.</th>
<th>Value</th>
<th>Variance Explained (%)</th>
<th>Total Variance Explained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.720</td>
<td>30.224</td>
<td>30.224</td>
</tr>
<tr>
<td>2</td>
<td>1.535</td>
<td>17.055</td>
<td>47.279</td>
</tr>
<tr>
<td>3</td>
<td>1.249</td>
<td>13.882</td>
<td>61.161</td>
</tr>
<tr>
<td>4</td>
<td>1.022</td>
<td>11.358</td>
<td>72.519</td>
</tr>
<tr>
<td>5</td>
<td>0.936</td>
<td>10.397</td>
<td>82.916</td>
</tr>
<tr>
<td>6</td>
<td>0.722</td>
<td>8.018</td>
<td>90.933</td>
</tr>
<tr>
<td>7</td>
<td>0.656</td>
<td>7.292</td>
<td>98.226</td>
</tr>
<tr>
<td>8</td>
<td>0.145</td>
<td>1.613</td>
<td>99.838</td>
</tr>
<tr>
<td>9</td>
<td>0.015</td>
<td>0.162</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Table 2 shows that the top four principal components occupy over 70% of total variances in data and the first component counts for 30% of total variance. Calculating the initial factor loading matrix of the first four principal components, the result could be shown in Equation (3), it could be seen that indices, such as SIRI, WDI and OSI, own a relative higher weight than others when consisting the most principal component. This observation indicates that the above indices also play a significant role in determining the final evaluation of CTAI system.

\[
V = \begin{bmatrix}
    \text{SIRI}^T \\
    \text{TIRI} \\
    \text{WDI} \\
    \text{TAI} \\
    \text{DDR} \\
    \text{DTR} \\
    \text{PI} \\
    \text{OSI} \\
    \text{WTI}
\end{bmatrix} = \begin{bmatrix}
    .914 & -.217 & .160 & .155 \\
    -.203 & .597 & -.370 & .364 \\
    .907 & -.204 & .206 & .091 \\
    .285 & -.401 & -.494 & .370 \\
    -.008 & .241 & .427 & .777 \\
    -.213 & -.233 & .655 & -.132 \\
    .616 & .316 & -.302 & -.241 \\
    .715 & .568 & .106 & -.187 \\
    .047 & -.628 & -.294 & .083
\end{bmatrix}
\]

(3)

The CTAI model could be calculated as follows:

\[
\text{CTAI} = 0.24\text{SIRI} + 0.06\text{TIRI} + 0.24\text{WDI} - 0.03\text{TAI} + 0.23\text{DDR} - 0.01\text{DTR} + 0.12\text{PI} + 0.28\text{OSI} - 0.14\text{WTI}
\]

(4)

To perform a quantificational analysis of index change in CTAI system, sensitivity analysis was conducted by varying specific index to observe the changing of final output. In sensitivity analysis, the means indices are changed from -20% to 20%. Though the value adjustment, the resulting CTAI could be obtained and the difference from the origin could be calculated. The results are presented in Table 3.
From the table, it could be shown that raising SIRI for 20% could contribute to the largest value improvement (6.52%) of CTAI. As expected, higher SIRI, WDI and OSI will result in a higher value of the whole system. In system-level, measurement on improving the SIRI could be a sensible choice; in customer-level, walking distance index is important to customer satisfaction; in driver-level, increasing the driving speed when passenger in vehicle by 20% could lead to 4.45% increase of CTAI system.

**TABLE 3 Sensitivity Analysis Results**

<table>
<thead>
<tr>
<th>Value</th>
<th>Change in percent</th>
<th>CTAI Change</th>
<th>Difference</th>
<th>Difference in Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRI</td>
<td>6.68 -20</td>
<td>5.75</td>
<td>-0.40</td>
<td>-6.52</td>
</tr>
<tr>
<td></td>
<td>7.52 -10</td>
<td>5.95</td>
<td>-0.20</td>
<td>-3.26</td>
</tr>
<tr>
<td></td>
<td>8.35 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>9.19 +10</td>
<td>6.35</td>
<td>0.20</td>
<td>3.26</td>
</tr>
<tr>
<td></td>
<td>10.02 +20</td>
<td>6.55</td>
<td>0.40</td>
<td>6.52</td>
</tr>
<tr>
<td>TIRI</td>
<td>7.95 -20</td>
<td>6.03</td>
<td>-0.12</td>
<td>-1.94</td>
</tr>
<tr>
<td></td>
<td>8.95 -10</td>
<td>6.09</td>
<td>-0.06</td>
<td>-0.97</td>
</tr>
<tr>
<td></td>
<td>9.94 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10.93 +10</td>
<td>6.21</td>
<td>0.06</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>11.93 +20</td>
<td>6.27</td>
<td>0.12</td>
<td>1.94</td>
</tr>
<tr>
<td>WDI</td>
<td>4.43 -20</td>
<td>5.88</td>
<td>-0.27</td>
<td>-4.33</td>
</tr>
<tr>
<td></td>
<td>4.99 -10</td>
<td>6.02</td>
<td>-0.13</td>
<td>-2.16</td>
</tr>
<tr>
<td></td>
<td>5.54 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>6.10 +10</td>
<td>6.28</td>
<td>0.13</td>
<td>2.16</td>
</tr>
<tr>
<td></td>
<td>6.65 +20</td>
<td>6.42</td>
<td>0.27</td>
<td>4.33</td>
</tr>
<tr>
<td>TAI</td>
<td>6.89 -20</td>
<td>6.20</td>
<td>0.05</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>7.75 -10</td>
<td>6.18</td>
<td>0.03</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>8.61 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>9.47 +10</td>
<td>6.12</td>
<td>-0.03</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td>10.33 +20</td>
<td>6.10</td>
<td>-0.05</td>
<td>-0.84</td>
</tr>
<tr>
<td>DDR</td>
<td>4.41 -20</td>
<td>5.90</td>
<td>-0.25</td>
<td>-4.12</td>
</tr>
<tr>
<td></td>
<td>4.96 -10</td>
<td>6.02</td>
<td>-0.13</td>
<td>-2.06</td>
</tr>
<tr>
<td></td>
<td>5.51 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>6.06 +10</td>
<td>6.28</td>
<td>0.13</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td>6.61 +20</td>
<td>6.40</td>
<td>0.25</td>
<td>4.12</td>
</tr>
<tr>
<td>DTR</td>
<td>4.38 -20</td>
<td>6.16</td>
<td>0.01</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>4.92 -10</td>
<td>6.15</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>5.47 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>6.02 +10</td>
<td>6.14</td>
<td>-0.01</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>6.57 +20</td>
<td>6.14</td>
<td>-0.01</td>
<td>-0.18</td>
</tr>
<tr>
<td>PI</td>
<td>4.10 -20</td>
<td>6.03</td>
<td>-0.12</td>
<td>-2.00</td>
</tr>
<tr>
<td></td>
<td>4.62 -10</td>
<td>6.09</td>
<td>-0.06</td>
<td>-1.00</td>
</tr>
<tr>
<td></td>
<td>5.13 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>5.64 +10</td>
<td>6.21</td>
<td>0.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>6.15 +20</td>
<td>6.27</td>
<td>0.12</td>
<td>2.00</td>
</tr>
<tr>
<td>OSI</td>
<td>3.91 -20</td>
<td>5.88</td>
<td>-0.27</td>
<td>-4.45</td>
</tr>
<tr>
<td></td>
<td>4.40 -10</td>
<td>6.01</td>
<td>-0.14</td>
<td>-2.23</td>
</tr>
<tr>
<td></td>
<td>4.89 -</td>
<td>6.15</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Meanwhile, the results validate the case in Shenzhen. The space idling rate of the entire taxi system is 49% based on the FCD data, indicating that nearly half driving mileages of the taxi are wasted on cruising for passenger. The current condition in Shenzhen conforms to the result, showing that the supply of taxis is higher than passenger demand. For a government transportation planner, the policy for reducing space idling rate of taxis is the most significant factor, which should be taken into consideration. Besides, from the customer’s point of view, WDI is the key index for improving the CTAI performance. Projects like adding taxi stands in the hotspot area could increase the probability of taxi availability for passengers. In CTAI scope, such projects could lead to an increase of WDI and TAI, which improves the entire system performance.

### Comparison of CTAI Performance in Weekdays and Weekends

FCD data from Shenzhen on both Dec. 2\textsuperscript{nd}, 2011 and Dec. 4\textsuperscript{th}, 2011 are used, including weekday (Dec. 2\textsuperscript{nd}) and weekend (Dec. 4\textsuperscript{th}), to analyze the performance and variance of CTAI between different days in a week.

#### TABLE 4 Original Statistics and Score of Various Indices (Dec. 2\textsuperscript{nd} and Dec. 4\textsuperscript{th})

<table>
<thead>
<tr>
<th>Description</th>
<th>Dec. 2\textsuperscript{nd} (weekday)</th>
<th>Dec. 4\textsuperscript{th} (weekend)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev</td>
</tr>
<tr>
<td>Time Idling Rate</td>
<td>0.33</td>
<td>0.032</td>
</tr>
<tr>
<td>Space Idling Rate</td>
<td>0.45</td>
<td>0.080</td>
</tr>
<tr>
<td>Working Distance (m)</td>
<td>233</td>
<td>32.2</td>
</tr>
<tr>
<td>Taxi Availability Prob.</td>
<td>0.87</td>
<td>0.064</td>
</tr>
<tr>
<td>Driving Time Ratio</td>
<td>3.1</td>
<td>0.496</td>
</tr>
<tr>
<td>Driving Distance Ratio</td>
<td>1.35</td>
<td>0.061</td>
</tr>
<tr>
<td>Daily Profit (RMB)</td>
<td>883.9</td>
<td>144.587</td>
</tr>
<tr>
<td>Occupied Speed (km/h)</td>
<td>24.8</td>
<td>3.554</td>
</tr>
<tr>
<td>Daily Working Time (h)</td>
<td>24.0</td>
<td>0.105</td>
</tr>
</tbody>
</table>

(Observation: 50 respectively)

From Table 4, an overview towards the indicators of the urban taxi system could be obtained. At the customer-level, an increasing trend in taxi-idling rate was identified during weekends compared with weekday. Moreover, the high idling rates lead to the drop of daily profit for each taxi-driver, which reduces the CTAI score of the urban taxi system.

In customer-level, it was also found from Figure 6 that the pick-up points in all
the top hotspots drops on the weekend. In addition, the hotspots spread wider since factors similar to weekend-trip could cause new hotspots (the rightest grid in the below graph). From index aspect to review the data above, WDI changes slightly since the distribution of customer demand-occurring point remain relatively same for both weekday and weekends. Based on the information obtained from driver-level that driver’s working occupied speed increases by 2.2%, it is convincing that the value of DDR increases both on average value and variance, which is probably from the reason that drivers tend to give up the shortest path for the relative good road condition on weekends.

![Map of Hotspot Spread]

**FIGURE 6** Hotspot spread in weekday and weekend. (a. Weekday, b. Weekends)

## 5. CONCLUSIONS

This paper proposes an overall evaluation method (CTAI) of the urban taxi system,
which assesses the urban taxi system in levels of system, customer and driver, consisting of nine key indices. The quality of the urban taxi system could be obtained easily and quantitatively from the FCD data collected from urban taxis.

This research extended the traditional research, and significantly further analyzed the correlation between various factors within the urban taxi system. However, as discussed, the SIRI and TIRI were constructed based on operational experiences, which needs to be further validated. In addition, a real time data source input would assist to present a dynamic reflection of the entire system, which is essential to the real time decision-makers.

ACKNOWLEDGE

This research was supported in part by the China National High Technology Research and Development Program (863 program award 2009AA11Z220), and National Natural Science Foundation of China (award 50738004). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors. The authors also appreciate support from Mr. Xiang, Cheng-Cheng about his data processing work.

REFERENCES


