CALLING FOR VALIDATION: DEMONSTRATING THE USE OF MOBILE PHONE DATA TO VALIDATE INTEGRATED LAND USE TRANSPORTATION MODEL S

Shan Jiang
Department of Urban Studies and Planning
Massachusetts Institute of Technology
E55-19E, 77 Massachusetts Avenue, Cambridge, MA 02139
857-654-5066
shanjiang@mit.edu

Laura B. Viña-Arias
Department of Civil and Environmental Engineering
77 Massachusetts Avenue, Cambridge, MA 02139
787-374-4913
laurav@mit.edu

Christopher Zegras
Department of Urban Studies and Planning
Massachusetts Institute of Technology
Room 10-403, 77 Massachusetts Avenue, Cambridge, MA 02139
617-452-2433
czegras@mit.edu

Joseph Ferreira, Jr.
Department of Urban Studies and Planning
Massachusetts Institute of Technology
MIT Room 9-532, 77 Massachusetts Avenue, Cambridge, MA 02139
617-253-7410
jf@mit.edu

Marta C. González
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
MIT Room 1-153, 77 Massachusetts Avenue, Cambridge, MA 02139
617-715-4140
martag@mit.edu

Word count: 5,115
Number of tables: 2
Number of figures: 7
Total number of words: 5,115+2,250=7,365

Submission Date: August 1, 2011.

1 Corresponding Author
ABSTRACT

In recent years, massive amounts of data generated by mobile phone activity have been increasingly used to help understand human mobility patterns, trajectories, and travel behavior. Mobile phones are becoming devices to track movement with no cost added to their usage and are greatly distributed worldwide. These capabilities provide a new source of information that can be used to improve model estimation and validation for traditional metropolitan accessibility systems. In this paper, we use mobile phone data provided by a local service carrier to validate travel demand and travel times generated from an integrated land use and transportation (LUT) model calibrated for the Lisbon Metropolitan Area, Portugal. From the mobile phone data, a morning home-to-work origin-destination (OD) matrix was estimated based on the two most frequently visited mobile phone towers. The results show the potential for using this new type of technology to improve the use of complex integrated land use and transportation models.

Keywords: mobile phone data, land use transportation models, validation, travel demand, OD matrices
INTRODUCTION

Paradoxically, understanding the dynamics of the modern metropolis has grown both increasingly possible and increasingly complex in the information age (1, 2). On the one hand, the interaction of advanced information and communication technologies (ICTs) with “traditional” elements of the metropolitan accessibility system – i.e., land use and transportation – poses new questions that can open doors to new applications in the field of transportation research. For example, it is still not clear how such modern technologies might substitute for traditional mobility, enhance certain modes, and/or induce fundamentally new behaviors that radically change human-space interactions in urban space (e.g. see (3, 4)). On the other hand, the digital traces that such technologies produce hold great promise in generating new data sources to improve model estimation and validation (5, 6, 7). In the cases of mobile phones, while increased computational power allows for more complex but tractable models advance, validations by placing data on top of urban infrastructures are of paramount importance (8).

In this paper we demonstrate the use of a particular source of ICT data, mobile phones, treated in such a way that they help us validate an integrated land use transportation (LUT) model calibrated for the Lisbon, Portugal, metropolitan area (hereafter LMA). An anonymous mobile phone service carrier provided all the mobile phone activity for a one-month period. Each time a call was made, the location of the nearest tower routing the communication was recorded. We used these cellular phone towers to generate zones of analysis that were consistent with existing statistical and administrative boundaries (i.e. census blocks and civil parishes, i.e. BGRIs and “freguesias”). The two most frequent towers or anchor points for each user during a specific time span were identified as the residential and employment locations in order to determine a person based journey-to-work seed Origin-Destination (OD) matrix. This seed matrix was later scaled to account for the total employed population in the metropolitan area, based on population and employment data from the census, administrative survey data and travel survey data, to generate morning peak hour OD matrices. Lastly, traffic assignment to the roadway network was conducted based on these scaled ODs.

This information is then used to compare travel demand and travel times from two models. The first validation is for an evolved travel demand model based on the traditional Four-Step model developed in TransCAD, which we use in combination with a land development and household/firm location choice model, UrbanSim (9), an integrated simulation system calibrated for the LMA using a range of data sources from different years. The second is a mobile phone travel flow estimation model.

The results show the potential for using mobile phone data to improve complex integrated urban land use and transportation (LUT) models and demonstrate a viable approach for doing so. Our work represents the first steps of integrating urban simulations combined with increasingly available “real-time” data arising from the ad-hoc sensor networks that modern ICT systems provide in our cities. We believe these data sources will fulfill the promise of practical, scalable, and useful integrated LUT modeling for improved urban system management.

LITERATURE REVIEW

In the past few years, new communication technologies epitomized by the mobile phone have emerged and made it possible to collect massive amounts of data at individual levels with high levels of accuracy in time and space (10). This has enabled researchers to study daily human
mobility patterns (6), approximate travel trajectories of individual users (11, 12, 13), study
spatiotemporal patterns in cities (14, 15, 16), estimate residential and employment location based
on anchor points (17), generate OD matrices (18, 19) and estimate other traffic-related
parameters such as travel times and travel speeds (20, 21).

These new capabilities have been undertaken in several pilot projects and simulation
studies, although many issues (including reliability and sample size, privacy, socio-economic
differences of mobile phone owners, and conversion of mobile phone counts to vehicle counts)
concerning the application of this technology are still being addressed by the research
community. Among the first attempts in investigating mobile phones as vehicle probes is
CAPITAL, a 1994 project by the University of Maryland (11). In 1999 the Transportation
Research Laboratory (TRL) investigated the feasibility of using billing data to obtain OD
information using only a subset of all monitored phones in the area of Kent, UK (22).

Projects focused on obtaining reliable travel times and travel speeds from the mobile
phone network include: (i) STRIP (System for Traffic Information and Positioning) in Lyon,
France (20); (ii) Finnra (Finish Road Administration) in 2002 in Finland (11, 20); (iii) a
collaboration between the Institute of Transportation Research of Germany Aerospace Center
and Vodafone Munich, Germany in 2003 (19, 20); (iv) MTS (Mobile Traffic System) in Noord-
Brabant, the Netherlands in 2004 (11); (v) Bar-Gera’s case study of Tel-Aviv, Israel in 2005 (23);
(vi) a collaboration between the operator Proximus and ITIS Holdings in 2006 in Vlaanderen,
Belgium (11); (vii) TrafficOnLine in Germany in 2006 (24); and (viii) Bangkok, Thailand in
2007 (11). Initiatives in the United States include: Hampton AirSage system in the road region of
Hampton, Virginia in 2005 (25); collaboration between ITIS Holding and Delcan Corporation in
Maryland in 2005 (11); and Mobile Millenium Project in the San Francisco Bay Area in 2008.
The majority of these projects have validated their estimations with inductive loop detectors, file
data on floating car, license plate recognition technology, and GPS equipped probe vehicles.

Other projects aiming at urban analysis and the spatial-temporal dynamics of a city
include: (i) the Mobile Landscape case studies in Milan, Italy in 2004 (15) and in Graz, Austria
in 2006 (26) at MIT SENSEable City Laboratory; (ii) Real Time Rome Project, a collaboration
between Telecom Italia and MIT’s SENSEable City Laboratory in 2006 (27); (iii) and the
collaboration of Current City Consortium (MIT’s SENSEable City Laboratory and Salzburg
University) and the Dutch Ministry of Transportation in Amsterdam, the Netherlands in 2007
(26). Two other projects, conducted by Cáceres et al. in Huelva and Sevilla, Spain and Sohn et al.
in Korea between 2007 and 2008, are the only two projects found where simulation
evironments were used for OD matrix estimation. The Spanish project converts simulated
mobile phone data to an OD matrix and the latter uses a simulation environment to validate the
estimated matrix (18, 19).

From the aforementioned literature, we see that few studies have focused on employing
mobile phone data to validate integrated land use and transportation models. In this study, we try
to bridge the gap between the two so as to exploit the massive and rich information embedded in
mobile phone data to fulfill the need for validating land use and transportation models.

INTEGRATED LAND USE TRANSPORTATION (LUT) MODEL FOR LMA

As the focus of this paper is to demonstrate the validation of a Land Use and Transportation
Model by exploiting mobile phone data using Lisbon as an example, not to discuss the details of
the specific LUT model, in this section, we will only briefly discuss the framework and results of
the LMA LUT model as shown in FIGURE 1.
FIGURE 1 The LMA LUT (UrbanSim—TransCAD) model structure.
An Integrated UrbanSim—Travel Demand Model

The integrated land use and transportation model for the LMA was implemented jointly in UrbanSim (a software-based simulation system for supporting planning and analysis of urban development (9)) and TransCAD (a GIS-based travel demand modeling software) under the “Strategic Options for Integrating Transportation Innovations and Urban Revitalization” (SOTUR) project within the Transportation Systems of the MIT-Portugal Program (MPP). This Integrated LUT model was developed to evaluate urban development and transportation solutions for urban revitalization in Portugal. The model was calibrated for 2001 as a base year and simulated the subsequent decade (2001 to 2011) for model testing.

FIGURE 1 demonstrates the different modules of the SOTUR-LUT Model for LMA. The right hand side of the figure depicts modules of the UrbanSim model, which include a real estate price model, real estate development model, household mobility model, firm mobility model, household location choice model, and firm location choice model. These modules were developed and calibrated based on the default UrbanSim model with several variations in some of the sub-modules using various sources of census data, economic activity data, and administrative surveys for the Lisbon Metropolitan Area (28, 29).

The left hand side of FIGURE 1 shows the travel demand model evolved from a traditional Four-Step Model (FSM) developed for LMA (30). The first three steps of the traditional FSM (31) were replaced with newly customized modules, including a car-ownership model, trip generation model, nested mode choice/destination choice model for both home-based-work (HBW) and home-based-other (HBO) trip purposes. The model estimates AM peak hour OD matrices by mode (including driving, car-passenger, transit, and walking) and by hour (7:00 to 8:00 a.m., and 8:00 to 9:00 a.m.). Due to data limitations, the transit network was not integrated with the same road network, so in the last step of the evolved FSM, we only assign car trips onto the road network. Transit performance measures (such as travel time) are modeled separately using an external empirical econometric model.

Between the UrbanSim and Travel Demand Model (the right and left hand side) of FIGURE 1, the measure of accessibility was calculated as a byproduct of the Travel Demand Model for each simulation year, which was then used as an input variable for different modules of the UrbanSim side of the model to determine real estate price and household/firm location choice for the next simulation year. The loop interchange between the right and left hand side of the LUT model continued for future simulated years.

Validation Approach

As mentioned, we are aiming to validate the base year (2001) road network performance estimated from the integrated LMA LUT model by using mobile phone data. In most existing studies in the literature, mobile phone data were more often used to generate ODs. However, as detailed in the following section, the basic spatial analysis units for our LUT model and the mobile phone model are quite different – the LUT uses units generated from administratively defined census zonal boundaries (freguesia, akin to a civil parish), while the mobile phone model uses the Voronoi Lattice of the mobile phone towers. Therefore, the OD pairs generated by LUT model are not directly comparable to those in the mobile phone model. For these reasons, we created the same road network (consisting of both centroids of “freguesia” and towers as nodes
in the network) and used the mobile phone data to validate the LUT model’s road network performance (see FIGURE 2) instead of comparing the OD matrices from the two models.

FIGURE 2 Traffic assignment on the road network estimated from the LMA LUT model (2001).

OD AND TRAFFIC FLOW ESTIMATION USING MOBILE PHONE DATA

The mobile phone data used for this study record the time and tower location of on-going calls for more than 0.3 million users for an entire month in 2009 (LMA has approximately 2.8 million population in 2009). A total of 601 towers were identified in the Lisbon Metropolitan Area. FIGURE 3 illustrates all trips detected by the mobile phone usage in an average day in the Lisbon Metropolitan Area. A time frame was determined in order to identify two most frequently used towers, assuming that they were the residential and employment locations for a mobile phone user. A tower routing a signal from a mobile phone user during evenings was inferred as home location, and during daytime as the employment location. These locations were used to estimate a 601 by 601 seed OD matrix.
FIGURE 3 Average daily trips (290,085) for all trip purposes in LMA estimated by mobile phone data.

Spatial Analysis Units

In the LMA LUT model, population and employment data are aggregated from the census block (BGRI) level (the LMA contains 32,762 BGRIs), to the freguesia level. We use 216 freguesias as the spatial units of analysis in that model. As for the mobile phone data analysis, Voronoi lattice is often used to derive service area, however, because the Voronoi areas usually do not match with administrative boundaries, we dissolved the census blocks (BGRIs) into 601 zones based on the 601 Voronoi Lattice of mobile phone towers. This enables us to estimate population and employment data that fall in the tower service areas (see FIGURE 4).
FIGURE 4 LMA’s spatial analysis units: census blocks (BGRIs), mobile phone tower service area, tower-BGRIs, and parishes (*freguesia*). [Note: In our methodology, mobile phone tower service-areas (gray lines) initially estimated are dissolved with the BGRIs (yellow blocks) to generate our zones of analysis (red lines)].

Population and Employment Validation

One limitation of our mobile phone data source is that only one service carrier provided us with data, far from a complete representation for the whole region. According to the 2001 Census for the LMA, the total number of employed residents was 1,293,920 and the total number of jobs was around 1,304,194 (estimated from 2009 data). Our 2009 mobile phone data recorded a total of 320,531 home-end users, and 307,965 employment-end phone users. By analyzing the population (employment) density and mobile phone users’ residential and employment density at the 601 tower zone level, we can observe similar spatial distribution pattern for the employed residents (IR_EP) and the mobile phone residential users (H_user), especially in the LMA downtown area and near the 25 de Abril Bridge in the South of the region. However, the mobile phone service carrier used in this study seems to have a smaller market share of mobile phone users resided in the northwest and the center of the region compared to the Census residential density (see FIGURE 5 a-b). A good representation of the job locations is obtained—with most users routing their signal to towers in the downtown area and in the southern part of the region (sees FIGURE 5 c-d).
FIGURE 5 (a) Employed residents per square kilometer based on 2001 Census. (b) Mobile phone users residing at the tower zone per square kilometer based on 2009 mobile phone records. (c) Employment per square kilometer based on 2001 job estimation. (d) Mobile phone users working at the tower zone per square kilometer based on 2009 mobile phone records.

Scaled OD Matrix: Iterative Proportional Fitting

An OD matrix extracted from mobile phones is generated directly by linking mobile users’ detected “home” and “work” towers, which only accounts for a portion of the total LMA population. In order to compare with the LMA LUT model in the same time periods, two scaled OD matrices were estimated based on the 2001 Census and Employment data using iterative proportional fitting (IPF) method.

IPF, originally presented by Deming and Stephan in 1940 (32) and refined by Fienberg in 1970 (33), was created to adjust sample frequency tables to match known marginal distributions. In 1996, Beckman et al. developed the conventional approach used to estimate joint distribution for household attributes (33). It was first created as a method for estimating cell probabilities, \( p_{ij} \), in an \( n \times r \times c \) table based on observations, where the marginal totals are known and fixed.

Deming and Stephan (32) proposed this method as a way of deriving estimates which minimized
subject to the marginal totals, assuming the number of observations in the (i,j) cell \((n_{ij})\) is more than zero. The procedure runs as follows:

1. There are \(n_{ij}\) observations in the \((i,j)\) cell, where

\[
\sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} = n
\]

and the initial value is taken as

\[
p_{ij}^0 = \frac{n_{ij}}{n} \quad \forall \ i, j
\]

2. At the \((2m)\)th step \((m \geq 1)\)

\[
p_{ij}^{(2m-1)} = p_{ij}^{(2m-2)} \frac{p_i}{p_i^{(2m-2)}} \quad \forall \ i, j
\]

3. At the \((2m+1)\)th step

\[
p_{ij}^{(2m)} = p_{ij}^{(2m-1)} \frac{p_j}{p_j^{(2m-1)}} \quad \forall \ i, j
\]

4. The iteration is continued until two successive sets of values for the cell probabilities agree sufficiently well.

In our study, the seed OD matrix served as the initial values \((p_{ij}^0)\) for the iterations rather than the cell probability. As introduced above, the marginal totals need to be known and fixed. Since these marginal totals – the total employed residents and the total employment in the area – were obtained from different data sources their totals are not equal. Therefore, the total number of jobs in 2001 per tower zone was scaled down to match the total employed residents from the Census in 2001, ensuring that the summation of marginal totals for the IPF were equal.

After having the scaled OD matrix, we applied two more modifications to these marginal totals to obtain morning peak OD matrices by hour and by mode: the morning peak coefficients from two regression models and the mode share model, which will be introduced shortly. This procedure resulted in a total of 251,644 car trips and 309,660 car trips for the 7:00-8:00 AM and 8:00-9:00 AM period respectively.

**OD Matrices by Time-of-Day and by Mode**

Mobile phone data are very helpful in revealing mobile phone users’ whereabouts; however, they are not informative in terms of detecting users’ activities or transportation mode associated with their locations. We need to have external data sources to complete this information.
Trip Estimation by Time-of-Day (Morning Peak)

As discussed earlier, we estimated mobile phone users’ “home” and “work” towers based on their mobile phone usage pattern. The main challenge to face in the presented methodology was that we could not estimate when the users started their journey-to-work trips everyday based on their mobile phone usage alone. As mobile phone data are, after all, not as spatially and temporally detailed as GPS information, and mobile phone users may not (indeed, if they are driving, probably should not) be using their phone during their journey-to-work trips. This represented spatial and temporal limitations that had to be overcome.

In order to determine journey-to-work morning peak OD matrices by hour, we needed additional data to estimate the total OD flow for the time periods that we are interested in. We used the 1994 Lisbon household travel survey which was the most recent survey available to us to estimate the relationship between the number of morning journey-to-work trips and the employed residents at each zone (freguesia). We estimated a regression model [Equation 1] for two time periods (7:00-8:00 am and 8:00-9:00 am):

\[ JTW_{\text{Trips}}_{z,t} = \beta \times EMP_z + \epsilon \]  

[Equation 1]

where \( JTW_{\text{Trips}}_{z,t} \) is the journey-to-work trips originated from zone \( z \) during time period \( t \), and \( EMP_z \) is the total employed residents in zone \( z \). The regression results are shown in TABLE 1.

<table>
<thead>
<tr>
<th>Time of Day</th>
<th>( \beta )</th>
<th>Standard Error</th>
<th>t</th>
<th>Adjusted R Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00-8:00 AM</td>
<td>0.3330</td>
<td>0.0052</td>
<td>63.47</td>
<td>0.951</td>
</tr>
<tr>
<td>8:00-9:00 AM</td>
<td>0.4090</td>
<td>0.0058</td>
<td>70.80</td>
<td>0.960</td>
</tr>
</tbody>
</table>

Mode Share

Since we cannot detect the transportation mode that each mobile phone user chooses for her daily trips, we adopt the same mode share in the LMA LUT model. But as in the LMA LUT model, the spatial analysis units used are 216 freguesias while in the mobile phone model the spatial units are 601 “tower-BGRI zones,” we needed to convert the 216x216 mode share pairs into 601x601 pairs.

As mentioned, both the freguesia and the “tower-BGRI zone” are dissolved from the basic census block unit (BGRI). We therefore use BGRI as the common ground to link freguesia and “tower-BGRI zone,” which gives us a 1272 unique combination of freguesia and “tower-BGRI zone” IDs. By joining the 216x216 unique freguesia-to-freguesia mode share pairs to the 1272x1272 freguesia-to-BGRI to freguesia-to-BGRI pairs, we obtain a long list of 1,617,984 mode share records. We then group by unique “tower-BGRI zone” to “tower-BGRI zone” ID combination, and calculate the average of the mode share for the 601x601 pairs of mode share to generate a “tower-BGRI zone” level mode share matrix.

In these operations, we have made the assumption that the mode share of each freguesia to freguesia pair is evenly distributed within the same freguesia. By multiplying the mode share matrix at the “tower-BGRI zone” level with the OD matrices generated for the two hours in morning peak (7:00 to 8:00 am and 8:00 to 9:00 am), we are able to get the OD matrices by cars for each peak hour. TABLE 2 summarizes the statistics of the key input and output variables for the mobile phone OD generation model.
TABLE 2 Summary of the Key Inputs and Outputs of the Mobile Phone Model

<table>
<thead>
<tr>
<th>Input</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of residents (2001)</td>
<td>2,682,687</td>
</tr>
<tr>
<td>Total number of employed residents (2001)</td>
<td>1,293,902</td>
</tr>
<tr>
<td>Estimated number of Employment (2001)</td>
<td>1,304,194</td>
</tr>
<tr>
<td>Estimated number of mobile phone users at home-end (2009)</td>
<td>320,531</td>
</tr>
<tr>
<td>Estimated number of mobile phone users at works-end (2009)</td>
<td>307,965</td>
</tr>
<tr>
<td>Initial number of trips based on mobile phone activity</td>
<td>290,085</td>
</tr>
<tr>
<td>Total number of all-day trips</td>
<td>1,290,904</td>
</tr>
<tr>
<td>Average car driving rate (2001)</td>
<td>0.62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of morning trips (7-8 AM)</td>
<td>429,860</td>
</tr>
<tr>
<td>Total number of morning trips (8-9 AM)</td>
<td>527,918</td>
</tr>
<tr>
<td>Total number of morning trips by vehicle (7-8 AM)</td>
<td>251,644</td>
</tr>
<tr>
<td>Total number of morning trips by vehicle (8-9 AM)</td>
<td>309,660</td>
</tr>
</tbody>
</table>

**Traffic Assignment**

We then assign the two OD matrices (7:00 to 8:00 am and 8:00 to 9:00 am) to the same network shared by the two models (the LUT Model and the Mobile Phone Model). We use dynamic traffic assignment method offered in TransCAD, using the first hour as a warm up for the road network; the network performance measures (such as link volume and link travel speed) are the hourly averages for the second hour (8:00 to 9:00 am, which is the peak of the AM peak). FIGURE 6 shows the traffic assignment results generated from the mobile phone model.

**FIGURE 6** Traffic assignment on the road network from the mobile phone model (2001).

**MODEL VALIDATION RESULTS**

With the network performances obtained from traffic assignment from both the LMA LUT model and the mobile phone model, we are able to compare the models, and tentatively validate...
the LUT model. FIGURE 7 demonstrates the comparison of link volume and link speed from these two models. From the left panel, we can see that for links with volume less than 2,500 vehicle per hour, these two models vary in a relatively large range; while for link volume greater than 2,500 the total traffic volume from the mobile phone model generally predicts lower volume. From the right panel of FIGURE 7, we can see more consistency in the congested travel speed (and thus travel time) of the two models. However, for local roads (with relatively low speed limits), the mobile phone model predicts lower speed than the LUT model, even though on average the LUT model behaves very closely to the mobile phone model. Based on these initial results, we will fine tune the LUT model, and also compare the network performance for different testing model years.

![FIGURE 7](image)

**FIGURE 7** Link volume and speed comparison between the LUT model and mobile phone model.

**CONCLUSION**

As the latest trend in integrated land use and transportation modeling has been towards increasing sophistication and complexity, external sources for validation become increasingly important – data collection for validation purposes itself can be an expensive endeavor. Meanwhile, the increasingly accessible, massive data flows generated by mobile phones, combined with newly developed data mining and statistical techniques, present a rich validation opportunity. Mobile phone data have been employed in recent studies of human travel and mobility networks; however, they have, apparently, rarely been used for validating integrated land use and transportation models. We demonstrate new methods and approaches for the model validation purposes, which we hope will open the door for the LUT research/practice community to further explore the potential usage of mobile phone data in improving modeling capabilities and practice.

A crucial advantage of the methodology presented here is that we show that mobile phone data can be scaled in a proper way to capture flows in the road network generated by the journey-to-work trips of users. The kind of mobile phone data we used in this work is just a byproduct of billing activities of any mobile phone service carrier worldwide. This represents a
very low cost alternative to the tremendous amount of time and money spent nowadays in collecting traditional data such as traffic counts and travel surveys, while still ensuring the accuracy of integrated LUT models. Furthermore, we believe that researches in countries that do not have traffic counts or travel survey data available but have well-developed mobile phone markets could greatly benefit from this kind of data as a very valuable source of information. In short, we see that in the future, mobile phone activity and other ICT data can be used as a daily detector of travel demand. There are still several steps to undertake in this direction. For example, in order to better calibrate both the LUT and mobile phone models, information from taxi fleets as a third data source can shed lights on improving the accuracy of travel time/travel speed estimations. This can help us clarify the differences in flows predicted by both models. Additionally, we would like to see analyses at smaller scales for road usage to develop specific applications related to land use. Measurements over more than one month to analyze variations in a particular day and time, and further comparison with other real-time sources of information also remains to be done.

ACKNOWLEDGEMENTS

This work was made possible by the generous support of the Government of Portugal through the Portuguese Foundation for International Cooperation in Science, Technology and Higher Education and was undertaken as part of the MIT-Portugal Program. It was also partially supported by the National Academies Keck Futures Initiative through the Nakfi Complexity Grant to MCG. We acknowledge the contributions of Dr. Pu Wang, Weifeng Li, Yi Zhu, and Jae Seung Lee at MIT, USA.

REFERENCES


