What's your MoCho? Real-time Mode Choice Prediction Using Discrete Choice Models and a HCI Platform

Ronan Doorley*

MIT Media Lab

Ariel Noyman†
noyman@mit.edu
MIT Media Lab

Yasushi Sakai
MIT Media Lab

Kent Larson
MIT Media Lab

ABSTRACT

The impact of city-planning on mobility habits of urban dwellers has been proven crucial to well functioning cities. Nevertheless, the correlation between discrete urban interventions and metropolitan scale mobility mode-choices (MC) is challenging to predict and communicate. This paper presents the design and deployment of 'MoCho', a real-time MC modelling, prediction and collaboration platform. MoCho aims to predict and simulate MC of individuals in a metro region in response to real-time urban design iterations. The prediction models consider individual characteristics and attributes of available alternatives and are calibrated using survey data. To explore MoCho MC predictions, users interact with CityScope, a compu-tangible user-interface which triggers new MC predictions and their impacts based on interactive design of land-use, density or spatial proximity. Finally, a distributed computational system delivers real-time predictions onto a web-based user-interface. In 2018, a MoCho instance has been developed and deployed to simulate MC for the Boston metro area, focusing on a 14 acres development site in Kendall Sq. Cambridge, MA. The choice model was well fitted and the parameters showed significant associations with a range of explanatory variables including travel times, residential and employment densities and personal attributes like age, gender, education-level and home-ownership. Such a combination of an intuitive TUI and well-calibrated prediction models can allow experts and non-experts alike to participate in an evidence-based urban design process. Code for MoCho MC model and front-end is available here: https://github.com/CityScope/CS_choiceModels

KEYWORDS

Machine Learning, Mode Choices, Transportation, Urban Modeling, Prediction, Tangible, Interface, CityScope

*Both authors contributed equally to this research.
†Both authors contributed equally to this research.

UrbComp '19, August 5, 2019, Anchorage, AK

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference format:
Ronan Doorley, Ariel Noyman, Yasushi Sakai, and Kent Larson. 2019. What's your MoCho? Real-time Mode Choice Prediction Using Discrete Choice Models and a HCI Platform. In *Proceedings of UrbComp '19: 8th SIGKDD International Workshop on Urban Computing, Anchorage, AK, August 5, 2019 (UrbComp '19),* 8 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The individual choices urban dwellers make about their mobility and transportation behavior have profound impacts on their own lives, as well as society as a whole. Motorized transportation leads to negative external impacts such as carbon emissions and air pollution, whereas active mode choices (MC) such as walking and cycling improve the physical and mental health of travelers [8, 22]. Urban planning can influence these mobility choices and their societal impacts by organizing spatial land-uses and infrastructure in such a way as to encourage short trips using active modes [14, 27].

Statistical learning methods have been used successfully both in research and in practice to predict the mode choices of travellers in response to urban interventions such as changes in land use or road infrastructure [5, 7, 12, 32]. Additionally, emission and epidemiological models may be used to predict the health and environmental impacts of changes in transportation behavior [19]. Recently, some researchers have developed end-to-end models which can directly predict the health and environmental impacts resulting from infrastructure or policy changes [21, 28] and models which can optimise network infrastructure for these impacts [9].

Making predictions through statistical models generally involves the use of either statistical computing platforms like R [15] or Stata [10] or specialised transportation modelling software such as SimMobility [1] or Python and GIS plugins. Such platforms may carry steep learning curves, require specialised skills and feature limited capabilities for real-time interaction [25]. This challenges both experts and non-experts to evaluate a manifold of 'what-if' scenarios or to rapidly iterate on designs alternatives. In recent years, Tangible User Interfaces (TUI) have been developed to facilitate a more collaborative process of urban design, augmented by computation and data analytics. Examples of such TUIs are the Augmented Urban Planning Workbench [17] and The Clay Table [18] which were designed to facilitate a collaborative urban design process. Since 2013, the City Science group at the MIT Media Lab has been developing CityScope, a collaborative urban platform which combines a TUI with real-time predictive analytics and visualisation of urban dynamics. Previous CityScope instances have included capabilities for modelling energy performance, accessibility of housing, employment and amenities among other metrics [2, 23].

This paper describes the development of MoCho, a CityScope instance which aims to allow decision-makers, planners and community members alike to experiment with different land-uses and spatial organisations in a city district and to understand in real-time how those designs would influence mobility behavior, MC and other societal impacts. An application of this tool to the 'Volpe' development project in Kendall Sq., Cambridge, MA is also reported.

2 DATA AND MODELLING

MoCho makes use of publicly available datasets and APIs which cover all major urban settlements in the USA. This ensures that the methodology presented here can be replicated for the vast majority of American cities. The main resources used are listed in Table 1. These data sources are used to generate a synthetic population of people and to calibrate models which could predict their mobility behaviors in response to changes such as new residential or commercial development. The modelling can be described in four steps: (1) population synthesis, (2) home and job location choices, (3) transportation mode choice and (4) impact assessment. Each of these steps are outlined below.

2.1 Population Synthesis

Mobility behaviors may be analyzed using two main approaches: (1) an aggregate approach which divides the area into zones and predicts aggregates inter-zonal flows, or (2) using a disaggregate approach which recognizes that urban mobility patterns are the result of many decisions made by individuals. The disaggregate approaches directly explain why an individual makes a choice given their circumstances, and therefore they are better able to predict how those choices may change in different circumstances [20]. Due to privacy and data availability constraints, it is generally not practical to make predictions with respect to real population. Instead, it is common practice to use population synthesis techniques in order to produce a "synthetic population" and to make predictions with respect to these individuals. The methods take individual/household-level demographic profiles and zonal aggregate demographic data and allocate the individual records to zones in order to create the synthetic population. Some common techniques include Iterative Proportional Fitting, [4, 11], convex optimisation [31] and Bayesian methods [29]. The PUMS survey data used in MoCho include the home Public Use Microdata Area (PUMA) and place-of-work PUMA for each respondent where each PUMA corresponds to a set of census tracts. In order to model commuting trips at a tract-to-tract level, the population synthesis process needs to allocate each individual to a home and work census tract pair. A simple Bayesian method is utilized for this purpose, using the origin-destination flows from the CTPP data as well as aggregate demographic data from the ACS. PUMS individuals with attributes A, home PUMA H and work PUMA POW are assigned to an origin-destination (O-D) pair w_{ij} according to the probability calculated with equation 1.

$$P(w_{ij}|A) \propto \delta_{ij} \prod_{a \in A} P(a|w_{ij})P(w_{ij}) \quad where \quad \delta_{ij} = \begin{cases} 1, & \text{if } H \subset i \text{ and } POW \subset j \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The prior probabilities $P(a|w_{ij})$ and $P(w_{ij})$ may be obtained from the ACS aggregate demographic data and the CTPP O-D data.

2.2 Home and Employment Location Choices

The model needed to be able to simulate the changes in home and work locations of the population in response to edits made on the CityScope (CS) platform in respect to land-use, density and spatial organisation. It was assumed that when a residential or employment unit appears in a census tract, new residents/workers appear. The number of new people is determined by the density of the unit and some demographic attributes (such as income or job sector) may be determined by the unit type. It is assumed that the conditional probability distribution of the new residents' demographic attributes and work locations is similar to that of the existing residents of that same tract, conditional on the known attributes. Therefore, the new resident can be simulated by cloning a randomly sampled person with the same home location tract and attributes from the baseline synthetic population. A similar process is used to assign home locations to new workers. This simple location choice can be expected be accurate for small changes to the land-uses and densities in a district but for more substantial changes, a more sophisticated model may be required. In future iterations of MoCho, location choices will be predicted using a discrete choice model, similar to the mode choice model described in the next section.

2.3 Mode Choices

The choice of transportation mode for each synthetic individual's commute needs to be predicted in response to each simulated intervention. The MC are modelled using a logit-based discrete choice model. This class of models has been used extensively by researchers and practitioners in modelling of decisions including home location, work location and mode of transportation. An advantage of discrete choice models over many other classification models is that their estimated parameters have economic interpretations. This means that the final model results can be easily understood by practitioners. Discrete choice models assume that individual decision makers select the alternative from their available options which maximises their utility. The utility is composed of a systematic component and a stochastic component. The systematic portion of the utility is an additive function of attributes of the decision maker, attributes of the alternative and interactions between both. This stochastic component is needed because in reality, two people with the same measured attributes may take different decisions when faced with similar alternatives. This component is typically assumed to be Gumbel distributed due to computational advantages and this leads to the logit formulation. These models

Resource	Description
Public Use Microdata Sample (PUMS)	Individual person and household level survey data for the USA
American Community Survey (ACS)	Aggregated demographic data for administrative zones in the USA
OpenStreetMap (OSM)	An open-source editable map of the world [13]
Open Source Routing Machine (OSRM)	An API which provides routing information using OSM data [16]
Open Trip Planner (OTP)	A server which computes multi-modal transport itineraries [30]
Census Transportation Planning Products (CTPP)	A special tabulation of the ACS data for commuting characteristics

Table 1: Data resources used for US cities in the MoCho framework

can be defined by the following three expressions [20].

$$
\begin{aligned}
U(X_i, S_t) \geq U(X_j, S_t) \forall j \in C \\
U_{it} = V_{it} + \epsilon_{it} \\
V_{it} = V(S_t) + V(X_i) + V(S_t, X_i)
\end{aligned}
\tag{2}
$$

where C is the choice set, i is the alternative chosen by decision maker t, U_{it} is the true utility of alternative i to decision maker t, X_i are the attributes of option i, S_t are the attributes of person t, V is the systematic utility and ϵ is the stochastic utility.

The parameters of the logit model must be calibrated with individual-level stated-choice or revealed-choice data. For MoCho, individual observations from the PUMS data could be used for the calibration. The PUMS survey data contains 12 different options for MC but for the purpose of this work, this list is simplified to 4 major classes: car, bicycle, walk and public transportation modes. The explanatory variables considered for inclusion in the model include person attributes from the PUMS data (such as: age, income, gender, education level and employment type), attributes of the home and workplace census tracts from the ACS data, and the estimated travel times and costs for each mode and each trip. The travel times for each census tract pair were estimated by querying the OSRM API (for walking, cycling and driving times) and Open Trip Planner (for public transit travel times). The PUMS data and ACS data contain hundreds of variables and so some exploratory analysis and feature engineering needs to be done to create a list of candidate features prior to model fitting. Once the features have been selected, the coefficients of each features can be estimated by maximum likelihood estimation, in this case using the python library 'pylogit'.

2.4 Impact Assessment

The framework currently includes impact calculations for carbon emissions and the effects of physical activity on mortality rates. The carbon emission calculations are based on simple per-km estimates for each mode. More accurate estimates could be obtained by using data specific to the vehicle fleet of the study region and detailed modelling of emissions rates. However, in order to make the methodology easily replicable, the aggregated method is preferred. The physical activity impacts are based on the approach of the World Health Organisation's Health Economic Assessment Tool (HEAT). This method calculates the expected change in mortality due to a given amount of walking and/or cycling in a population using Relative Risk estimates from meta-analysis of multiple epidemiological studies. All impact estimates are normalised by the normal of people in the study region so that changes in population do not not bias the results.

$$
\begin{aligned}
RR = RRref \times (V/V_{ref}) \\
\Delta D = N \times (1 - RR) \times MR_{base}
\end{aligned}
\tag{3}
$$

where RR is the predicted Relative Risk of mortality in the population due to the walking or cycling, $RRref$ is the Relative Risk of mortality due to walking/walking in the reference study, V is the volume of walking or cycling undertaken, V_{ref} is the volume of walking or cycling undertaken in the reference study, ΔD is the expected number of avoided deaths, N is the population size and MR_{base} is the baseline mortality rate in the population.

3 SYSTEM ARCHITECTURE

This section describes the system which allows the models described in section 2 to support raid experimentation in real-life design processes. The system design, software and hardware, user-interface, interaction and experience are described below. The hardware of the MoCho system consists of an MIT CityScope instance, a TUI dedicated to rapid urban prototyping and real-time feedback. The software architecture of MoCho, illustrated in Figure 1 consists of four parts: (1) Human Computer Interaction (HCI) front-end as well as three services each handling (2) Tangible User Interface (TUI) data management, (3) computation of MC model and (4) a spatial Geo-Server. The rest of the section describes each of the system's different components, the data-flow and networking between each part.

3.1 MIT CityScope

Since 2013, MIT City Science Group researchers are developing CityScope (CS): a human-centered, urban modeling, simulation and decision-making platform. CS sits in the intersection of urban-planning, Human-Computer Interaction and social sciences with a goal to support an evidence-based discourse around the nature of the built environments. Through a series of lab experiments and real-world deployments, CS has been successful in providing insights, predictions and consensus in various real-world urban questions [3, 23, 24]. For the purpose of MoCho, a CS instance was redesigned and constructed in an active demo area situated at the MIT Media Lab in Cambridge, MA.

3.2 CS User Interaction

A common CS instance features a few key components: A tangible scaled urban model (city, neighborhood or at street scale), a computational acquisition unit and a feedback module. The scaled urban model of a CS instance is usually built over a translucent table-top

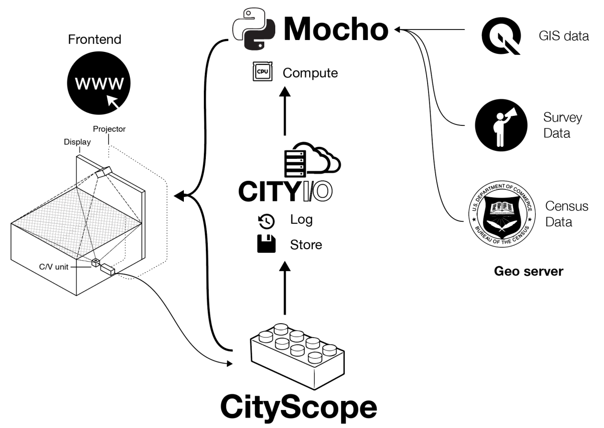

Figure 1: software architecture
The CS table data is sent to CityIO from the tangible interface made with LEGO. CityIO stores that data to have it available to the MoCho computation module. MoCho combines this data with other geo-data from the Geo server and performs computation. The Front-end visualizes the table state and mode choices. Each arrow indicates a HTTP response.

and includes a set of pre-tagged LEGO bricks acting as intractable building-blocks or massing elements. These allow users to rapidly modify land-use (ranging from roads, parks, amenities, residential, office buildings or other), density and other urban properties by manipulating the tiles. The computational acquisition unit has sensors and cameras for real-time scanning of the scene. Each user interaction is detected and recorded via a computer-vision scanning system and is then transmitted to cityIO, a distributed cloud service. In this case, an openCV-based Python tool was designed to recognize the tile-tags and sliders interaction. CityIO role is to act as a mediator between the different computational modules and aggregate their results for displaying on the CS instance. Lastly, a feedback module contains display screens, projectors and other representation tools to communicate the analysis outcomes.

3.3 cityIO API

CityIO is a cloud service holding the CS TUI interaction data and is responsible for distributing these interactions to local or remote urban-analysis modules. In the case of this project, the MoCho service (described in section 3.4) listens to CityIO to attain the CS table state, and combines this with data from the Geo Server to make predictions. As with most web based systems, the communication between the components was HTTP based.

The flow of data within this system has four steps: First, the CS interface reads the tags and sends it to CityIO. Next, CityIO receives the CS data and exposes it as a HTTP API making it available to MoCho and other modules. Then, in combination with the data in the Geo-Server, MoCho predicts and exposes the result by it's own HTTP endpoint. Finally, the front-end collects data from the different APIs and visualizes the overall result on the CS instance. Additionally, cityIO can hash and cache prior analysis results to avoid redundant computation leading to faster response time, crucial for user-interaction. This micro-service architecture allows for an extendable CS framework, in which multiple demanding computation modules (such as noise, energy or mobility) can cross-interact without affecting performance and experience of end-user.

Figure 2: CityScope MoCho
User interaction with CS TUI: Groups of up to 4 users can freely interact with the CS TUI and amend the urban design setup on the CS table-top. Interaction results are computed and visualized on both vertical (metro-scale) and horizontal (parcel-scale) planes. A TUI slider (bottom left corner) allows density and building-height iteration for the land-use type poised into the slider slot.

3.4 MoCho API

The MoCho API is the component responsible for predicting the mobility choices of each simulated individual in response to the user interactions. This module listens to the cityIO API for changes to the state of the TUI interaction data. When a change is detected, the module creates new synthetic individuals corresponding to the new residential and commercial buildings and assigns their home and work locations as described in section 2. The residential and employment densities of the tracts containing the new buildings are also updated. The modes of transportation for each commuting trip are then predicted using the calibrated MC model. Finally, the carbon emissions and mortality impacts of the change in mobility patterns are calculated. The results are composed in JSON format and exposed as HTTP end-points.

The user interactions can lead to changes in the mobility choice and impact predictions in three ways. Firstly, when residential and commercial units are added to denser, more central parts of the city, the newly spawned individuals will be more likely to choose workplaces with shorter commutes. Secondly, when new units are added to parts of the city where the current population of residents/workers has personal characteristics which tend to favor cycling (for example), the newly spawned people will tend to also have those characteristics. Lastly, the addition of new buildings affects the attributes of the census tract, such as overall residential and employment density and these attributes can affect the mode choice probabilities of all people living and working in the census tract.

3.5 Geodata API

Census tracts and other geo-spatial data are required by multiple modules to compute MC as well as to visualized the model's predictions. In order to ensure that the geo-data used by the different modules is always consistent, a data-service module is dissociated from the computation API and generalized as a service. For this purpose, an additional HTTP end-point was designed to service GeoJson data for the study area, in the case study of section 4, geo-data for the Boston metro area and for a selected interactive region in Cambridge, MA were served. This API can be easily scaled to serve MoCho or other modules in different regions, given access and availability of spatial data.

4 CASE STUDY: MOCHO FOR 'VOLPE' DEVELOPMENT

To examine the MC model in a real-world mobility environment, a site under development in Kendall Sq., Cambridge MA was selected. In addition, a CS TUI platform and a cityIO system were constructed to support real-time users' interaction and visualization of the model's prediction. This Section will explore the selected geography, the CS interaction and model's prediction in details.

4.1 Volpe Site

In the past two decades, the Kendall Sq. area has undergone massive transformation as a result of emerging biotech and startup industries, fueled by its proximity to MIT and Harvard campuses. Residential density of 3,000 inhabitants per sq/km, limited housing stock and growing land values force much of the area's workforce to commute. Restrictive zoning ordinance and the lack of affordable-housing incentives demote developers from constructing the necessary range of housing options, thus promoting mono-functional developments. The low residential density is matched with scarcity of services, amenities and 3rd-places [6]. A 14-acres site in the heart of Kendall Sq. was selected as a test-bed for the MC model prediction and CS TUI. Known as 'Volpe', the site is a United States General Services Administration (GSA) parcel which is planned to be fully redeveloped by MIT in the next decade. MIT intends to develop housing, commercial and lab space, retail and open space on 10 out of the site's 14 acres. In total, this amounts to 1.7 million sqft of commercial development and nearly 1,400 housing units with building heights ranging from 170 to 500 feet. 40% of this development would be housing, including 300 mid-and-low income units. 65% of ground floors on the site's main streets would contain retail and active street uses and approximately 2.5 acres would feature open space [26].

4.2 CityScope MoCho

A CS table, shown in Figure 2 was designed to encompass the Volpe site and its immediate surroundings. This covers a region of 0.5sqkm at a scale of 1:500 where each 4x4 LEGO-tile represents a 16sqm or 4sqm per each LEGO stud. This abstraction downsamples the fine detail of building-form and urban-design and instead focuses on the design of zoning envelopes and general urban structure. In the CS MoCho platform, six major classes of land uses were defined: green open spaces, streets, high-income housing ('Housing-1'), mid-to-low income housing ('Housing-2'), large companies' development ('Commercial-1'), and startup and co-working spaces development ('Commercial-2'). Each LEGO tile on the CS TUI is classified with one of these land-uses. An empty cell or a non-type due to scanning issues defaults to open-space. Using the CS TUI, users could edit the position of each grid-cell, add or remove them and change their proximity to one another. This allows a degree of discrete urban design that is more fine-grained than classic zoning exercises. As such, allocating different tiles next to the each other (i.e, two tiles of type 'housing-2' and one tile of type 'commercial-1') would be translated as a mixed-use structure with multistory housing and offices. A TUI slider adds an additional dynamic control which allows the density (height) of all cells of the same class (e.g., 'Housing-2') to be altered at the same time. This control allows rapid iteration over different design scenarios for the Volpe parcel, with variations in spatial organization, land-use distribution as well as the density of each land use.

4.3 Real-time Model Output and UI

As users interact with the CS TUI, a record of their interactions is delivered to the different APIs as previously described. These APIs yield analysis results that are projected back into the TUI table-top surface as well as onto a vertically mounted display using two web-browser instances, as shown in Figure 4.

4.3.1 MoCho Vertical Display. This display converges and visualizes the results of both the MC model, the scanned TUI and a Geo-data service. The Geo-data service parses census tract data of the Greater Boston Area into GeoJson polygons and projects them onto a cartographic background. Than, the results of the MoCho API model are rendered as origin-destination (O-D) arcs connecting pairs of different census-tracts, as shown in Figure 3. Each arc represents the sum of trips between the pair of tracts. The arc color represents the prevailing MC chosen by most trips leaving that tract (i.e, green for bikes, purple for cars, etc.). The arc's thickness corresponds to the sum of trips by that mode. On average, nearly 11,000 arcs are reproduced with each user iteration; To avoid illegibility and visual noise, the UI renders only arcs terminating at a given census tract, selected via user's interaction. For each selected census tract, the breakdown of trips by each mode also displayed in numerical format. The UI was built and deployed as a node.js application, using the ReactJS, mapboxGLJS and Deck.GL libraries.

Figure 3: MC predictions
Two outputs of the MC model: (top) shows trips originating at the Volpe site census tract and (bottom) showing trips from a suburban tracts. The arcs display volumes of trips (thicknesses) between each O-D pair for each mode (color). As clearly shown, the suburban tract yields more car trips with greater distances than those originating at the Volpe site. A breakdown of trips and MC appears on the left.

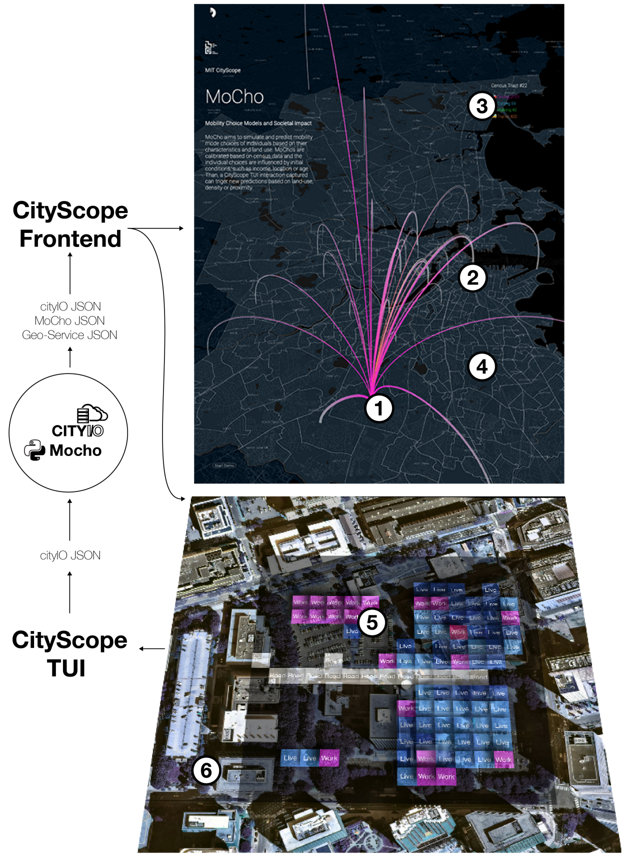

Figure 4: CityScope MoCho Interface Components
Vertical Display: (1) Selected tract showing trips and their MC (2) MC trip ending at the Volpe site (3) Numerical output of tract's MC (4) GeoJson of 500 computed census tracts. Horizontal TUI: (5) Tagged LEGO bricks with projected land-use scheme (6) Immediate context surrounding the Volpe site design space

4.3.2 MoCho TUI Table Top. The horizontal table-top of CS MoCho is used as both the design space as well as a canvas for visualisation. With each user interaction, the canvas updates a schematics land-use diagram. A shadow mapping algorithm adds perceptual depth to the grid tiles so that higher density tiles appear taller than others. A background mapping service contextualizes the design space to the Volpe site. Lastly, animated color dots represent individual trips entering or exiting the site. The dots colors correspond to the arcs on the vertical display (i.e, purple dot is one vehicular trip from one census tract) and animated to move from general direction of that tract to its designated land-use destination. A more advanced version of this UI could assign trips to routes based on Dynamic User Equilibrium or micro-simulation and feed updated travel times back to the mode choice predictions. Together, the visual aids of CS MoCho allow users to easily associate a relatively small-scale urban development with large scale MC impacts and observe the effects of different scenarios in real-time.

4.4 Model Calibration Results

The mode choice model introduced in section 2 must be calibrated for each location context. For the Case Study, data for the Boston metro area were used. As mentioned in section 2.3, some exploratory analysis and feature engineering is required for each case study before model fitting. Through some exploratory analysis of PUMS data for this study region, a number of variables were selected as being likely to affect MC. As well, some features were converted to different formats. For example, the age and income variables were converted from continuous to binary variables by dividing each into three quantiles and using binary variables to indicate records in the lowest and highest groups. The encoding of travel-time also required some experimentation. Time spent in different types of travel activities, such as driving, waiting for a bus or walking, are associated with different perceived costs and therefore should be treated differently in discrete choice models, without being so specific as to cause model under-fitting. In this case study, through some iteration, it was found that using the three variables of walking_time, cycling_time and in_vehicle led to a well fitted model with sensible parameter estimates. The final model had a pseudo-R-squared value of 0.45 which indicates a well fitting model. The full list of parameters estimates is shown in Table 2. In interpreting the parameters, it should be noted that the choice of driving was taken as the reference choice. For example, all of the parameters

Table 2: Multinomial Logit model calibration Results

Features	Cycle	Walk	Transit
alternative_specific_constant	-6.2687	-4.2422	-4.0632
employment_density_home_tract	9.6306	23.0826	9.7915
employment_density_work_tract	4.3132	9.1642	13.4858
residential_density_home_tract	47.2316	75.5508	53.945
residential_density_work_tract	32.9934	-	30.7119
age_youngest	0.6831	0.4743	0.31
age_oldest	-	-	-0.1701
income_lowest	0.3672	0.4887	0.2788
college_degree	0.8974	0.4997	0.1427
grad_degree	0.4789	0.1884	-
female	-0.7425	-	0.0864
renter	0.5955	0.7435	0.7031
non_profit_worker	0.8203	0.376	0.2227
	All		
walking time	-0.0004		
vehicle_time	-0.0002		
cycling time	-0.0005		
cost	-0.1424		

for the density variables are positive, indicating that increases in residential or employment density in one's home or workplace are associated with increased likelihood of cycling, walking or public transit relative to driving. The travel-time parameters show that time spent cycling is perceived as the most costly whereas time spent in vehicles is the least costly. This is in line with previous research and conventional wisdom [20]. Finally, some interesting associations are found between personal characteristics and the likelihood of taking each mode. For example, the model shows that having a college degree, a graduate degree and/or working for a non-profit decreases one's likelihood of driving and in particular, increases one's likelihood of taking active modes. Also, those in the youngest age group and lowest income group are less likely to drive than others.

5 CONCLUSION

This paper has described the design and deployment of a CityScope instance focused on predicting mobility choices and societal impacts in response to user inputs through a Tangible User Interface. While there already exist tools for predicting MC and for estimating the health and environmental impacts of transport, few efforts have been made to integrate both of these modelling steps in an end-to-end tool. Moreover, using these models for predictions typically requires laborious specification of inputs by professionals with specialised skills. The tool presented in this paper is the first to provide an intuitive user-interface to such models, allowing multiple people with varying levels of expertise to collaboratively experiment with different urban designs scenarios with real-time feedback. The underlying models are rigorously calibrated using publicly available data sources to ensure the credibility of the model predictions. Over the past year, hundreds of users, both professionals and non-experts interacted with the tool at an MIT demo space. Although not designed for external usage, 100 monthly users explored the MoCho endpoint, publicly available as a git repository and website. As observed by Alrashed, et al., CS deployments have proven to streamline complex urban planning and design questions [3]. CS MoCho advanced the CS framework with on-demand MC machine-learning predictions, a distributed computational back-end and an end-to-end online user interface. This allows stakeholders such as public health professionals or mayors – who may not be experts in transportation or statistical modelling – to participate in an evidence-based urban design process.

REFERENCES

[1] Muhammad Adnan, Francisco C Pereira, Carlos Miguel Lima Azevedo, Kakali Basak, Milan Lovric, Sebastián Raveau, Yi Zhu, Joseph Ferreira, Christopher Zegras, and Moshe Ben-Akiva. 2016. SimMobility: A multi-scale integrated agent-based simulation platform. In *95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record*.

[2] Luis Alonso, Yan Ryan Zhang, Arnaud Grignard, Ariel Noyman, Yasushi Sakai, Markus ElKatsha, Ronan Doorley, and Kent Larson. 2018. Cityscope: a data-driven interactive simulation tool for urban design. Use case volpe. In *International Conference on Complex Systems*. Springer, 253–261.

[3] Tarfah Alrashed, Almaha Almalki, Salma Aldawood, Tariq Alhindi, Ira Winder, Ariel Noyman, Anas Alfaris, and Areej Alwabil. 2015. An Observational Study of Usability in Collaborative Tangible Interfaces for Complex Planning Systems. *Procedia Manufacturing* 3 (2015), 1974–1980. https://doi.org/10.1016/j.promfg.2015.07.243

[4] Richard J Beckman, Keith A Baggerly, and Michael D McKay. 1996. Creating synthetic baseline populations. *Transportation Research Part A: Policy and Practice* 30, 6 (1996), 415–429.

[5] Moshe Ben-Akiva and Michel Bierlaire. 1999. Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*. Springer, 5–33.

[6] Michael Blanding and Michael Blanding. 2016. Reimagining Kendall Square. (Jun 2016). https://www.technologyreview.com/s/540206/the-past-and-future-of-kendall-square/

[7] Ralph Buehler and Jennifer Dill. 2016. Bikeway networks: A review of effects on cycling. *Transport Reviews* 36, 1 (2016), 9–27.

[8] Ronan Doorley, Vikram Pakrashi, and Bidisha Ghosh. 2015. Quantifying the health impacts of active travel: assessment of methodologies. *Transport Reviews* 35, 5 (2015), 559–582.

[9] Ronan Doorley, Vikram Pakrashi, W. Y. Szeto, and Bidisha Ghosh. 2019. Designing cycle networks to maximize health, environmental, and travel time impacts: An optimization-based approach. *International Journal of Sustainable Transportation* 0, 0 (2019), 1–14.

[10] Yuanyuan Gu, Arne Risa Hole, and Stephanie Knox. 2013. Fitting the generalized multinomial logit model in Stata. *The Stata Journal* 13, 2 (2013), 382–397.

[11] Jessica Y Guo and Chandra R Bhat. 2007. Population synthesis for microsimulating travel behavior. *Transportation Research Record* 2014, 1 (2007), 92–101.

[12] Khandker M Nurul Habib, Nicholas Day, and Eric J Miller. 2009. An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model. *Transportation Research Part A: Policy and Practice* 43, 7 (2009), 639–653.

[13] Mordechai Haklay and Patrick Weber. 2008. Openstreetmap: User-generated street maps. *IEEE Pervasive Computing* 7, 4 (2008), 12–18.

[14] Susan L Handy, Marlon G Boarnet, Reid Ewing, and Richard E Killingsworth. 2002. How the built environment affects physical activity: views from urban planning. *American journal of preventive medicine* 23, 2 (2002), 64–73.

[15] Asad Hasan, Wang Zhiyu, and Alireza S Mahani. 2014. Fast estimation of multinomial logit models: R package mnlogit. *arXiv preprint arXiv:1404.3177* (2014).

[16] Stephan Huber and Christoph Rust. 2016. Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM). *The Stata Journal* 16, 2 (2016), 416–423.

[17] Hiroshi Ishii, Eran Ben-Joseph, John Underkoffler, Luke Yeung, Dan Chak, Zahra Kanji, and Ben Piper. 2002. Augmented urban planning workbench: overlaying drawings, physical models and digital simulation. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 203.

[18] Hiroshi Ishii, Carlo Ratti, Ben Piper, Yao Wang, Assaf Biderman, and Eran Ben-Joseph. 2004. Bringing clay and sand into digital designâĂŤcontinuous tangible user interfaces. *BT technology journal* 22, 4 (2004), 287–299.

[19] Sonja Kahlmeier, Thomas GÃűtschi, Nick Cavill, Alberto Castro Fernandez, Christian Brand, David Rojas Rueda, James Woodcock, Paul Kelly, Christoph Lieb, Pekka Oja, Charlie Foster, Harry Rutter, and Francesca Racioppi. 2017. *Health economic assessment tool (HEAT) for walking and for cycling. Methods and user guide on physical activity, air pollution, injuries and carbon impact assessments.* Technical Report.

[20] Frank S Koppelman and Chandra Bhat. 2006. A self instructing course in mode choice modeling: multinomial and nested logit models. (2006).

[21] Robin Lovelace, Anna Goodman, Rachel Aldred, Nikolai Berkoff, Ali Abbas, and James Woodcock. 2017. The Propensity to Cycle Tool: An open source online system for sustainable transport planning. (2017).

[22] Natalie Mueller, David Rojas-Rueda, Tom Cole-Hunter, Audrey de Nazelle, Evi Dons, Regine Gerike, Thomas Goetschi, Luc Int Panis, Sonja Kahlmeier, and Mark Nieuwenhuijsen. 2015. Health impact assessment of active transportation: a systematic review. *Preventive medicine* 76 (2015), 103–114.

[23] Ariel Noyman. 2015. *POWERSTRUCTURES: The Urban Form of Regulation*. Master's thesis. https://dspace.mit.edu/handle/1721.1/99301

[24] Ariel Noyman, Tobias Holtz, Johannes Kröger, Jörg Rainer Noennig, and Kent Larson. 2017. Finding Places: HCI Platform for Public Participation in Refugees' Accommodation Process. In *Procedia Computer Science*, Vol. 112. https://doi.org/10.1016/j.procs.2017.08.180

[25] Ariel Noyman, Yasushi Sakai, and Kent Larson. 2018. CityScopeAR: Urban Design and Crowdsourced Engagement Platform. *CHI '18 CHI Conference on Human Factors in Computing Systems (to be Published)* (2018).

[26] MIT News Office. 2019. First step on Volpe parcel planned for 2019. (Feb 2019). http://news.mit.edu/2019/john-volpe-transportations-systems-design-0205

[27] Brian E Saelens, James F Sallis, and Lawrence D Frank. 2003. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of behavioral medicine* 25, 2 (2003), 80–91.

[28] Paul Schepers, Elliot Fishman, Rob Beelen, Eva Heinen, Wim Wijnen, and John Parkin. 2015. The mortality impact of bicycle paths and lanes related to physical activity, air pollution exposure and road safety. *Journal of Transport & Health* 2, 4 (2015), 460–473.

[29] Lijun Sun and Alexander Erath. 2015. A Bayesian network approach for population synthesis. *Transportation Research Part C: Emerging Technologies* 61 (2015), 49–62.

[30] TriMet. 2013. Open Trip Planner: An open source multi-modal trip planner. (2013). https://github.com/opentripplanner/OpenTripPlanner

[31] Peter Vovsha, James E Hicks, Binny M Paul, Vladimir Livshits, Petya Maneva, and Kyunghwi Jeon. 2015. *New features of population synthesis*. Technical Report.

[32] Mark Wardman, Miles Tight, and Matthew Page. 2007. Factors influencing the propensity to cycle to work. *Transportation Research Part A: Policy and Practice* 41, 4 (2007), 339–350.