An Advanced State-of-the-Practice Hybrid Travel Demand Model for the North Carolina Research Triangle Region

Vincent L. Bernardin, Jr., Ph.D.
Caliper Corporation
1172 Beacon St., Suite 300, Newton, MA 02461
Tel: 617-527-4700; E-mail: vince@caliper.com
ORCiD: 0000-0002-5894-0789

Kyle Ward
Caliper Corporation
1172 Beacon St., Suite 300, Newton, MA 02461
Tel: 617-527-4700; E-mail: kyle@caliper.com

Leta Huntsinger, Ph.D., P.E.
Institute for Transportation Research and Education, North Carolina State University
Research IV, 909 Capability Dr, Raleigh, NC 27606
Tel: 919-515-8899; E-mail: leta_huntsinger@ncsu.edu

Ramachandran Balakrishna, Ph.D.
Caliper Corporation
1172 Beacon St., Suite 300, Newton, MA 02461
Tel: 617-527-4700; E-mail: rama@caliper.com

Srinivasan Sundaram
Caliper Corporation
1172 Beacon St., Suite 300, Newton, MA 02461
Tel: 617-527-4700; E-mail: srini@caliper.com

Bernardin, Ward, Huntsinger, Balakrishna, and Sundaram

ABSTRACT
Since their emergence over a decade ago, hybrid travel demand models have become a popular alternative to both traditional four-step trip-based models and advanced activity-based models. A new generation of the Triangle Regional Model, or TRMG2, was recently developed for the Research Triangle region of North Carolina and serves as a contemporary example of a hybrid travel demand model. This paper provides an overview of the structure of the model and highlights innovative and distinguishing characteristics. The novel innovations include the use of machine learning methods for disaggregate, person-level trip production modeling and the use of nested logit models for destination choice. The TRMG2 also uses the linkage of non-home-based trips to home-based trips by location and mode as has become standard practice in recent hybrid models in place of the stop location and sequence choice models in earlier hybrids. The paper concludes with a summary of several sensitivity analyses performed on the model which demonstrate that the TRMG2 provides nuanced and realistic behavioral responses to various policy interventions.

Keywords: hybrid, Triangle, TRMG2, machine learning, nested destination choice, non-home-based

INTRODUCTION

Many Metropolitan Planning Organizations (MPOs) across the country have moved or are moving towards activity based models. This shift is driven in large part by advanced models' ability to better capture travel choice, and to provide a greater array of performance measures. On the other hand, many MPOs have been reluctant to make such a shift given the cost, complexity, and technical requirements of activity based models. The choice of whether to move away from an advanced trip based model to an activity based model was one recently considered by the Research Triangle region of North Carolina. With a greater focus on more diverse transportation projects and policies, a desire for better performance measures, and the demand being placed on the exiting trip based model by the consulting community, it was clear that a new paradigm was needed. What was not clear was whether that new paradigm would be an activity based model. While the needs of the region would seem to make an activity based model a clear choice, many were reluctant to take such a leap citing the aforementioned challenges. For the Research Triangle region, the choice became one that is not fully trip based, not fully activity based, but rather a hybrid approach. In a sense, a path that allows the region to evolve the Triangle Regional Model (TRM) over time while taking advantage of more immediate improvements. This stepped approach resulted in a model that is simple and easy to use, without sacrificing technical excellence and best practice. It is a model that was built with the end user in mind.

OVERVIEW OF THE TRIANGLE REGION

The Research Triangle region is unique in many ways. Rather than being defined by a strong central urban core, the region is multinucleated with several strong cities with varied and sometimes competing economies. Transportation planning in the region is led by two MPOs: Durham-Chapel Hill-Carrboro MPO to the west, and the Capital Area MPO to the east. At the heart of the region is the Raleigh-Durham International Airport, and Research Triangle Park (RTP). The region's name stems from the confluence of three tier one research universities, North Carolina State University, Duke University, and the University of North Carolina at Chapel Hill; together these universities form something of a Triangle, with RTP somewhat at the center, see Figure 1. (The economy of the region is largely driven by technology and biotech industries. The region is also home to North Carolina state government and has a very affluent and highly educated population.

Like many regions across the country, the region's demographics are changing. Recent Census and travel survey data show a population that is aging, but also a senior cohort that is active and traveling more. Household size is getting smaller with more single person and 2-person households with no kids. Surprising, given the auto dominance of the region, is an increase in the number of zero car households. And finally, the region is seeing emerging development patterns and increased interest in higher density mixed use development.
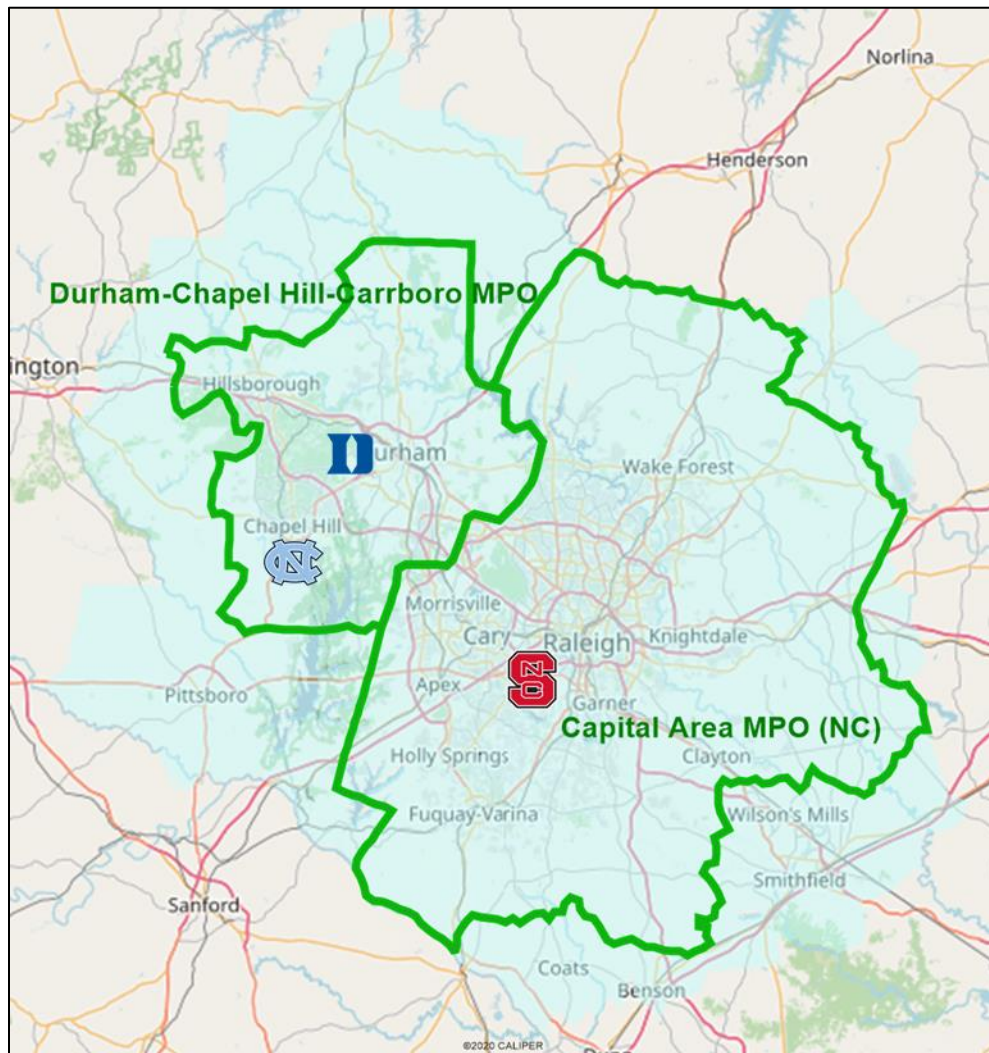
FIGURE 1 The Research Triangle Region of North Carolina

BACKGROUND

Hybrid travel demand models emerged over a decade ago as an attempt to balance the benefits and costs of advanced travel modeling (*1,2,3*). Since then, hybrid models have also been developed as an incremental step on the way to development of a full activity-based model (*4*). Hybrid models have now been developed for numerous metropolitan areas including Jerusalem; Vancouver, BC; Charleston, SC; Hampton Roads, VA; Knoxville, TN; Indianapolis, IN; Ann Arbor, MI; and several smaller metropolitan areas in Indiana and Virginia as well as the Raleigh-Durham-Chapel Hill Triangle region presented here. The Las Vegas MPO has also recently begun development of a hybrid model. The hybrid approach has also been particularly popular for statewide models being used in those for North Carolina, Tennessee, Illinois, Michigan, Iowa, Nebraska, and New Mexico. The spread of hybrid models has been aided by publications by the Federal Highway Administration's Travel Model Improvement Program (TMIP) which have showcased hybrid methods (*5,6*).

Hybrid models can be distinguished from activity-based models by the fact that they include at least some aggregate components, while they can be distinguished from traditional trip-based models in that they make some connections between trips, offer improved consistency with tours, and generally

have at least some disaggregate components (*6*). Not all advanced trip-based models are hybrids. Time-of-day or destination choice modeling or more refined trip purposes or modes do not fundamentally alter the four-step framework. Hybrid models include fundamental changes to the model structure and resulting behavior through connecting home-based and non-home-based trips and ensuring basic consistency of the model with tours. Hybrid models often make use of both trip-based and tour-based segmentation as in the new generation Triangle Regional Model, or TRMG2. Hybrid models typically begin with population synthesis followed by some disaggregate model components. Sometimes these components are the same as those used in activity-based models, but other times, as in the TRMG2, these components are different in that they do not make use of Monte Carlo simulation but rather use expected values. Midway through the model stream hybrid models shift to aggregate components. Usually, but not always, destination choice is modeled in aggregate. Mode choices can be modeled in the disaggregate before destination choice (*7*) or in aggregate after it. The TRMG2 illustrates both options with disaggregate non-motorized mode choice prior to destination choice and aggregate motorized mode choice after it.

OVERVIEW OF THE TRMG2

Several data collection efforts together provided a firm foundation for the development of the new TRM. The partner agencies began conducting a recurring regional household travel survey in 2016 with additional sample collected in 2018. The transit agencies in the region had last conducted on-board surveys in 2014 and 2015, and the partner agencies also conducted a parking behavior survey in 2016 which was instrumental in developing the new parking models. A survey of NCSU students from 2012 and a 2010 commercial vehicle travel survey were also used. External travel patterns were taken from the North Carolina Statewide Travel Model.

The overall structure of the TRMG2 is presented in Figure 2. Strictly speaking, this structure is only for the resident passenger models. The TRMG2 also includes auxiliary demand models for trucks and commercial vehicles, university students, the airport, and external trips. The university student models use a similar, but slightly simplified aggregate hybrid framework with a single mode choice. The other auxiliary models are simple trip-based models.
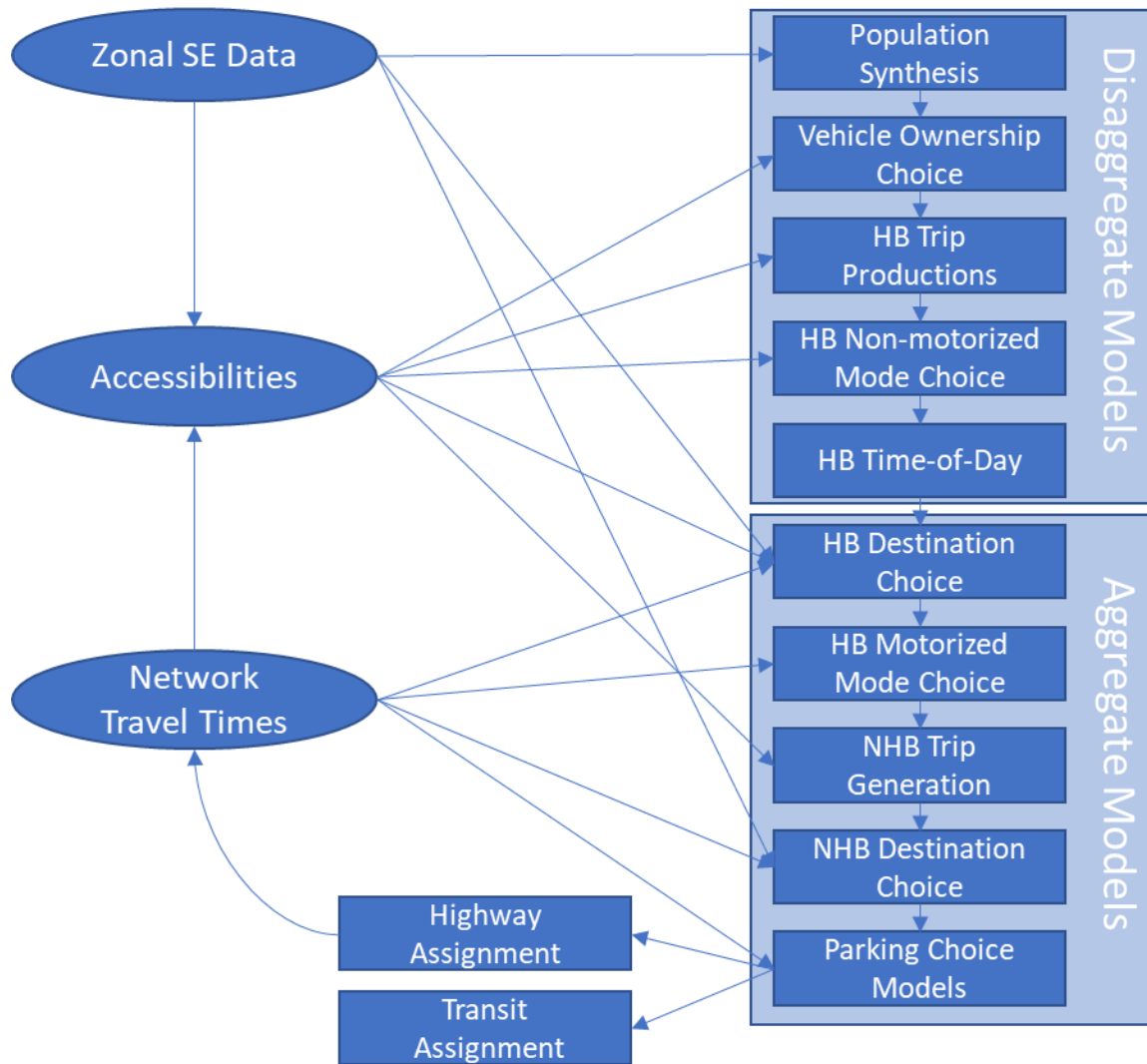
FIGURE 2 The TRMG2's Hybrid Structure

Component Variables

Zonal Socioeconomic Data

The TRMG2 contains 2,965 internal traffic analysis zones and 97 external stations. The demographic variables include the number of households, median household income, household population, percent of the population working, percent of the population under 18, percent of the population age 65 and above, and the number of university students by school. Employment is provided in five categories as well as the percent of employees earning more than $40,000 per year.

Networks and Travel Times

A new network was developed for the TRMG2 including all local streets. The base networks include 234 transit routes and 128,015 links modeling 16,477 miles of roadway. While the model's traffic assignment still uses centroid connectors and ignores local streets and roads, the inclusion of minor streets was instrumental for both the non-motorized and transit components of the model. The non-motorized models are sensitive to the density and connectivity of local streets with increased walking and

biking with denser and connected networks. The transit network uses the local streets as an elegant walk access alternative to a multitude of centroid connectors. The inclusion of the local streets also eliminates the common problem of transit routes traveling over roads not included in the traffic assignment network.

Accessibilities

      At the beginning of the model's execution, the zonal socioeconomic data and network travel times are used together to calculate several accessibility variables. Accessibilities are calculated for roadway, transit, and non-motorized modes, to various attractions and with different distance decay functions. As shown in Figure 2, the accessibilities are used in many of the model components to capture sensitivity to behavioral responses to urban form, area type, and proximity of attractions nearby.

**Disaggregate Demand Models**

Population Synthesis

      The TRMG2 makes use of the population synthesis native in the TransCAD software in which the model is implemented. This iterative proportional updating (IPU) algorithm is fast and generates a synthetic population of over 1.8 million people for the TRMG2 in roughly two minutes. This speed allows the population synthesis to be included as an automated part of the model run rather than a separate pre-process.

      The synthetic population is controlled at the household level for household size, number of workers and income and at the person level based on age groups (children under 18, adults aged 18 to 64, and seniors aged 65 and up). Figure 3 illustrates the ability of the iterative proportional updating to match person level controls where iterative proportional fitting alone without IPU fails.
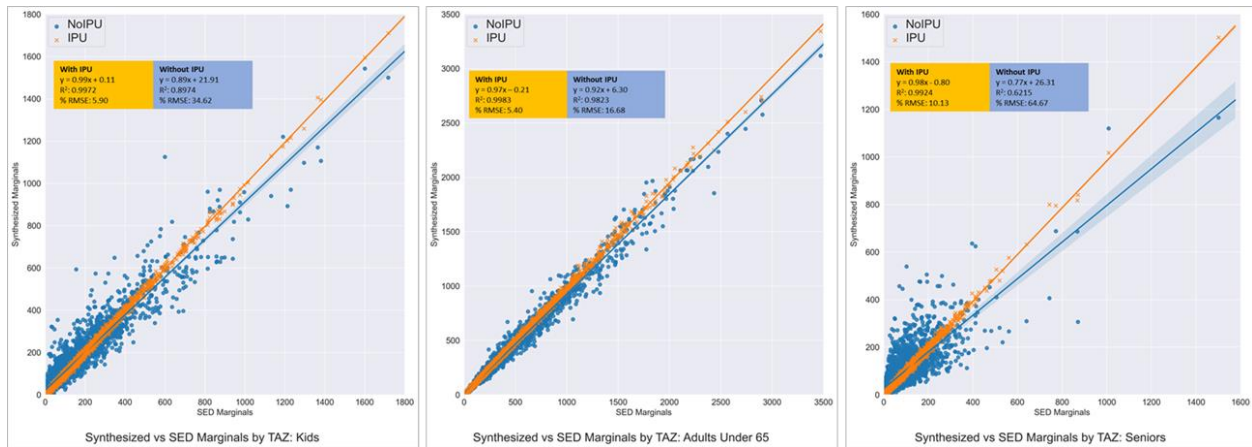
FIGURE 3 The Effect of IPU on Matching Person Level Controls

      The TRMG2's population synthesis also includes curves for producing ordinal level controls (e.g., one-person households, two-person households, etc.) from average zonal characteristics (e.g., average household size). These curves, illustrated in Figure 4, allow the partner agencies to only forecast zonal average characteristics for future year scenarios rather than discrete distributions for each zone.
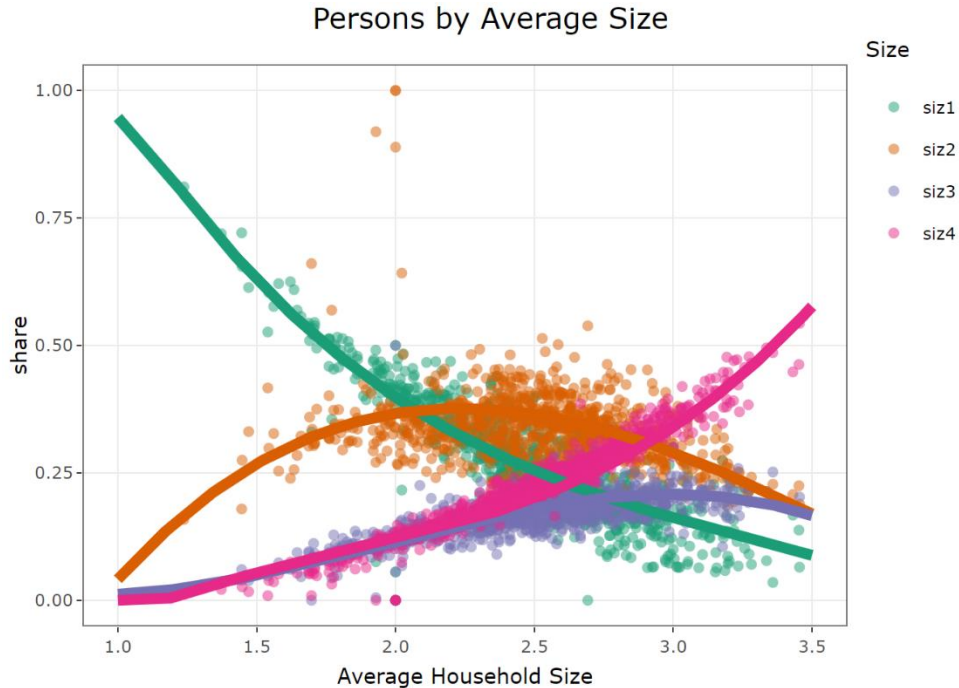
FIGURE 4 Ordinal Number of Persons per Household by Average Zonal Household Size

Vehicle Ownership Choice

       Vehicle ownership is modeled in the TRMG2 as a long-term choice made by each household. The multinomial logit model is used to predict the choice of zero, one, two, three, and four or more vehicles for each household. The model is sensitive to household income and composition as well as accessibility by walk, transit, and driving. Thus, improvements to any mode can in the long run impact the number of vehicles in the region and in that way indirectly as well as directly the resulting mode choices.

Home-based Trip Productions

       The TRMG2 generates the expected daily number of trips for each individual person in the region for each of 14 types of trips. The machine learning methodology for these models is highlighted in more detail in a subsequent section as one of the novel innovations of the TRMG2. At the highest level, trips were segmented based on whether they were on a work tour (36.2%) or nonwork tour (63.8%) since these tour types have very distinct characteristics in terms of destinations, modes, and time-of-day. Trips on each tour type were then divided into home-based trips (HB) (trips with one end at home) and non-home-based trips (NHB). The trips were further divided into the full 14 trip types through an exploratory analysis of the household survey data, distinguishing trips with different lengths, mode shares, and time of day distributions. Home-based trips on work tours were divided into work (W-HB-W), escort to school (W-HB-EK12), and other (W-HB-O). Non-home-based trips on work tours were divided into school escort (W-NH-EK12), work related (W-NH-WR), and other (W-NH-O). Home-based trips on nonwork tours were divided into school (N-HB-K12), other long duration discretionary activities (N-HB-ODL), other short duration discretionary activities (N-HB-ODS), other maintenance, shopping and eating (N-HB-OME), and other medical trips (N-HB-OMED). Non-home-based trips on nonwork tours were divided into school (N-NH-K12), other maintenance/eat out (N-NH-OME), and other (N-NH-O).

Figure 5 shows the contribution of each of the 14 trip types to total person-miles-of-travel for the region based on the survey. Figure 6 and Figure 7 illustrate the distinct modal and temporal distributions of the trip types.
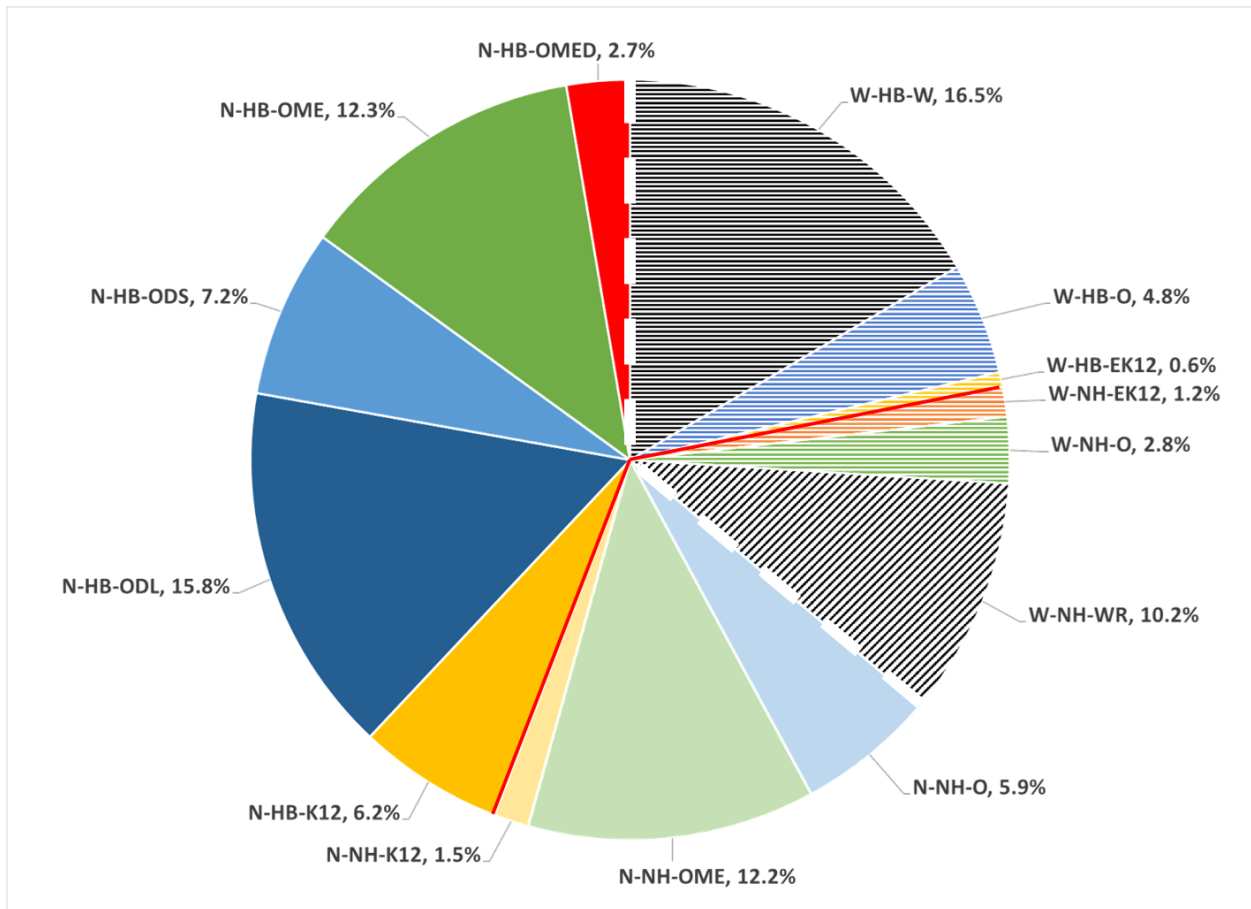
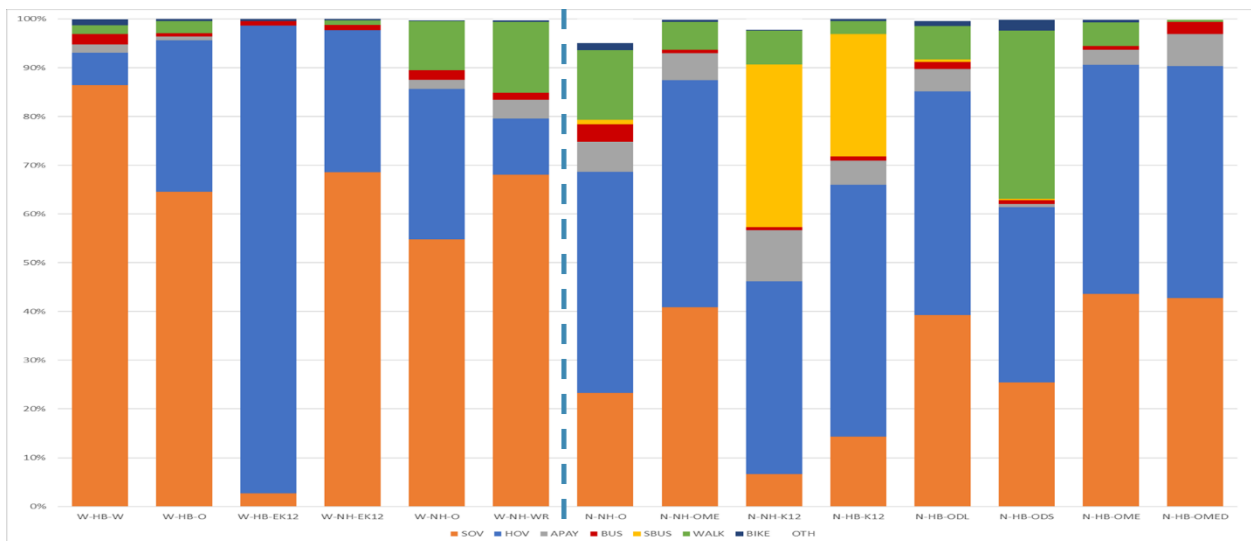**FIGURE 5 Person Miles of Travel by Trip Type**

FIGURE 6 Mode Shares by Trip Type

Home-based Nonmotorized Mode Choice

The choice to walk or bike for a trip is made at the level of an individual's trips. A binary logit model for each HB trip type predicts the expected number of the total trips which are motorized and nonmotorized. The models are sensitive to vehicle ownership, age, presence of children in the household, income, employment status, walkability of the zone and accessibility of attractions by walking.

Home-based Time-of-Day

An individual's home-based trips are apportioned to four time-of-day periods (AM: 7:00 AM – 9:00 AM, MD: 9:00 AM – 3:30 PM, PM: 3:30 PM – 6:15 PM, NT: 6:15 PM – 7:00 AM) based on fixed factors from the survey for each trip type. The factors for all trip types from the survey are shown in Figure 77. Following the time-of-day factoring, trips are aggregated to zones for subsequent models.
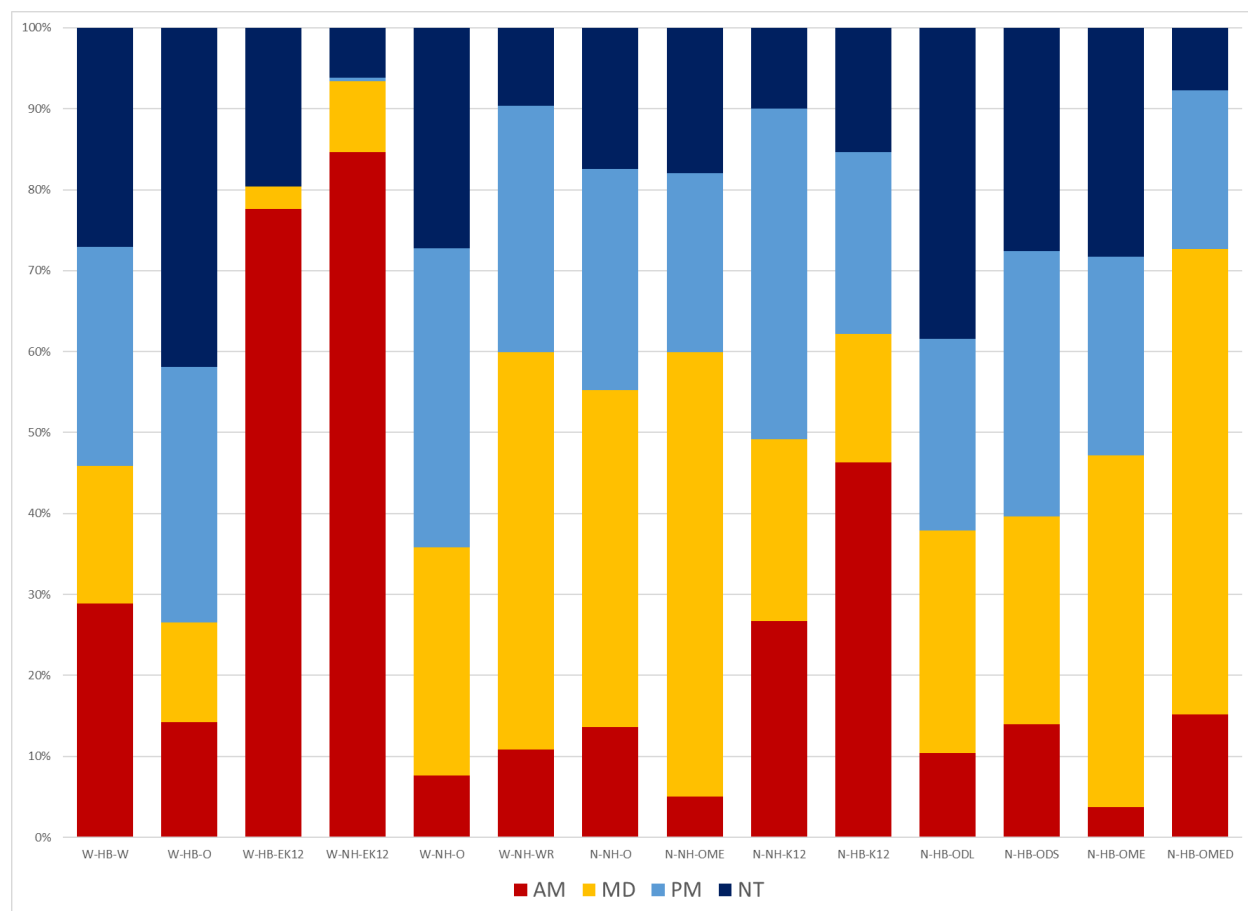

FIGURE 7 Time-of-Day Shares by Trip Type

Home-based Destination Choice

The HB destination choice models are the first aggregate model component in the TRMG2. They are discussed in more depth in subsequent section as they are one of the innovations of the TRMG2 in using nested logit models for destination choice.

Home-based Motorized Mode Choice

The HB motorized mode choice models are nested logit models as are common for this purpose. The nesting structure for the HB work trips on work tours (W-HB-W) is shown in Figure 8.
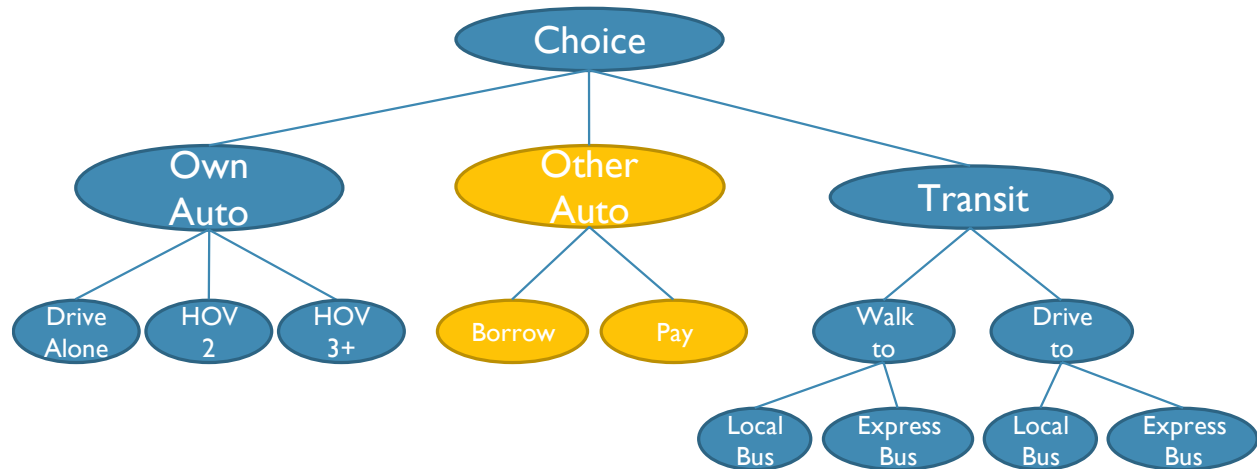
FIGURE 8 Work Trip Mode Choice Nesting Structure

The model departs slightly from traditional structures in that it includes separate nests for automobiles owned by the household and other automobiles not owned by the household. The other automobile nest includes two modes: borrowed car and paid car. Since the rise in popularity of Uber and Lyft, paid car services (also including traditional taxis and rental cars in the TRMG2) have received more attention in travel modeling. The phenomenon of people borrowing cars from friends or neighbors on the other hand has been largely ignored in traveling modeling even though household travel surveys consistently indicate the significance of this mode for travelers who do not own a vehicle. As the survey for the Triangle also clearly confirmed this mode as significant, it was modeled explicitly in the TRMG2's mode choice.

Non-home-based Trip Generation

Non-home-based trips are generated in the TRMG2 based on the destinations of HB trips by mode. Because this is a distinguishing feature of the TRMG2 from earlier hybrid models, these models are highlighted in detail in a subsequent section.

Non-home-based Destination Choice

The spatial distribution of NHB trips is accomplished in the TRMG2 through the use of doubly constrained multinomial logit destination choice models. Since there were not enough observations to support the estimation of separate models for each combination of trip type and mode, the trips were grouped into NHB auto trips on work tours, NHB auto trips on nonwork tours, NHB transit trips, and NHB walk/bike trips.

Parking Choice Models

Parking is limited and priced in the downtown and campus areas of the Triangle region, and some of these areas have free transit between them and remote parking. Therefore, the new TRMG2 incorporates parking choice models. These models, based on the 2016 parking survey, predict the zone where the driver will park and whether they will walk or ride a free shuttle based on their destination and the number and price of parking spaces in the zones. Separate models were specified for downtown and campus areas. As shown in Figure 9, both models were combined mode and destination choice models with mode (walk or shuttle) nested over parking zones. By including the root logsum of these choice models in the utility of the auto modes, travelers' mode and destination choices are also made sensitive to parking considerations.
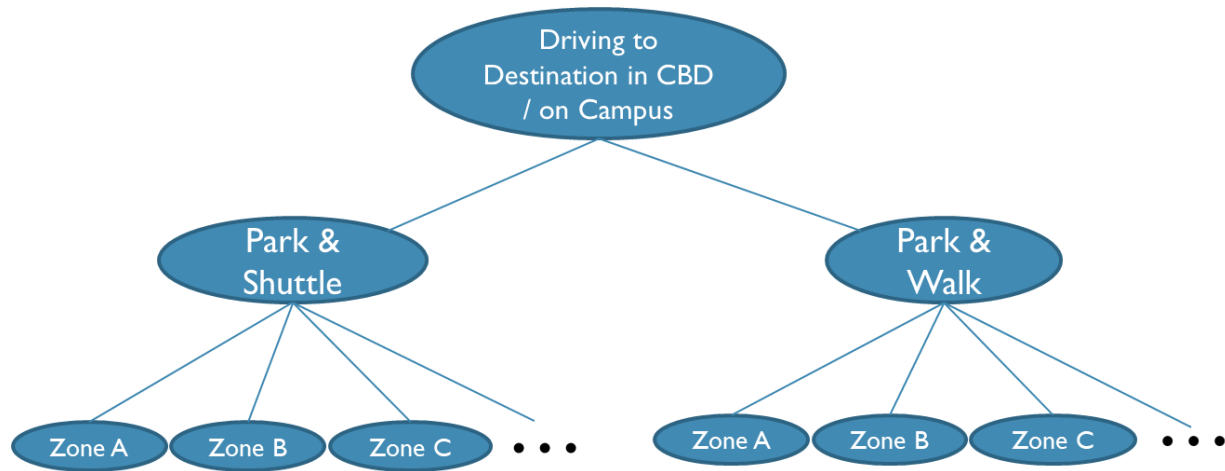
FIGURE 9 Parking Choice Models

TRIP PRODUCTION MODELS

While the TRMG2 shares many characteristics with other hybrid models, it also introduced two new innovations. The first of which are the HB trip production models which make use of machine learning (ML) methods. The HB trip production models are person-level models, applied to individual people in the synthetic population. These models can and do still make use of household characteristics but can also use person characteristics like age as explanatory variables. Although a form of decision trees was ultimately selected for the models, all the traditional forms of trip generation model were also tested. Initially, the tested models included cross-classified trip rates, generalized linear regression models, logit models, and XGBoost (*8*), a form of decision tree model similar in some ways to random forests.

School trips (N-HB-K12) will be taken as an example, although the results were similar across the various trip types. A cross-classification model was estimated to serve as a frame of reference. Average school trip rates (K12) were stratified by the person's age and the total number of children in their household. As is common with cross-classification models, attempting to add more dimensions (like income groups), exhausted the limited samples in the survey and led to non-significant and counter-intuitive estimates. After testing several forms of generalized linear models (GLM), a zero-inflated negative binomial model gave the best results. The model was able to incorporate more explanatory variables, but its goodness-of-fit was inferior to the cross-classification model. GLMs have sometimes been used in hybrid models in order to provide sensitivity to more variables than cross-classification supports, but it has only more recently been recognized that this comes at a cost of being less able to fit the non-linearities in trip rates. Ordinal logistic regression or ordered logit models were also estimated as these had been used in the prior version of the TRM. However, these models performed the poorest of all those tested. The reason seems to be the large number of persons making zero trips. Various attempts to address this, such as excluding adults without children from the model, did not result in any meaningful improvements. Table 1 shows the goodness of fit metrics for the various model forms attempted.

TABLE 1 Goodness-of-Fit of Various Trip Generation Models (School Trips)

Model Type	(Pseudo) r^2
Ordered Logit	0.03
GLM (zero-inflated negative binomial)	0.22
Cross-Classification	0.33
XGBoost	0.60
Rationalized Decision Tree (preferred)	0.53

Decision trees can be simple, but with bagging, boosting, and other methods, they can also become very complex and powerful. The XGBoost model (extreme gradient boosting) leverages all these advanced features and as anticipated, provided the best goodness-of-fit to the data despite taking all measures possible to avoid overfitting including using a holdout sample separate from the training dataset and using cross-validation with the training dataset in developing the model. A grid search of hyper-parameters including the learning rate and boosting weights was also used to achieve the most robust and predictive results. However, despite the predictive power of the XGBoost model, it was not ultimately preferred for two reasons. First, despite the ability to measure feature importance it is not possible to understand how exactly the model uses the various explanatory variables to make predictions. In this sense the model can be considered a proverbial black box. Second, while utmost care was taken not to over-fit the model, the risk of over-fitting remains, particularly with a model that cannot easily be audited with human intuition.

Ultimately, a simpler decision tree approach was taken following the principle of XAI or explainable artificial intelligence (*9*) that predictive methods should be intelligible to their users. It was possible to estimate more explainable decision trees without sacrificing much explanatory power. In the tree illustrated in Figure 10, each node lists the average trip rate as well as the number of observations that node represents. This conveys the overall average trip rate (top of the tree) and how it changes as you segment the surveyed population. The tree was created with the assistance of ML and preserves the ability to explain the model structure. Machine learning in this case uses simple ANOVA to determine which questions best divide the population into distinct groups. Branches that represent spurious divisions can be manually removed, the depth of branching controlled, and rules set to maintain a minimum number of samples in each group. These features allow a human auditor to be much more comfortable with the resulting model compared to full ML that may feel opaque. This approach also makes over-fitting easy to identify and avoid.
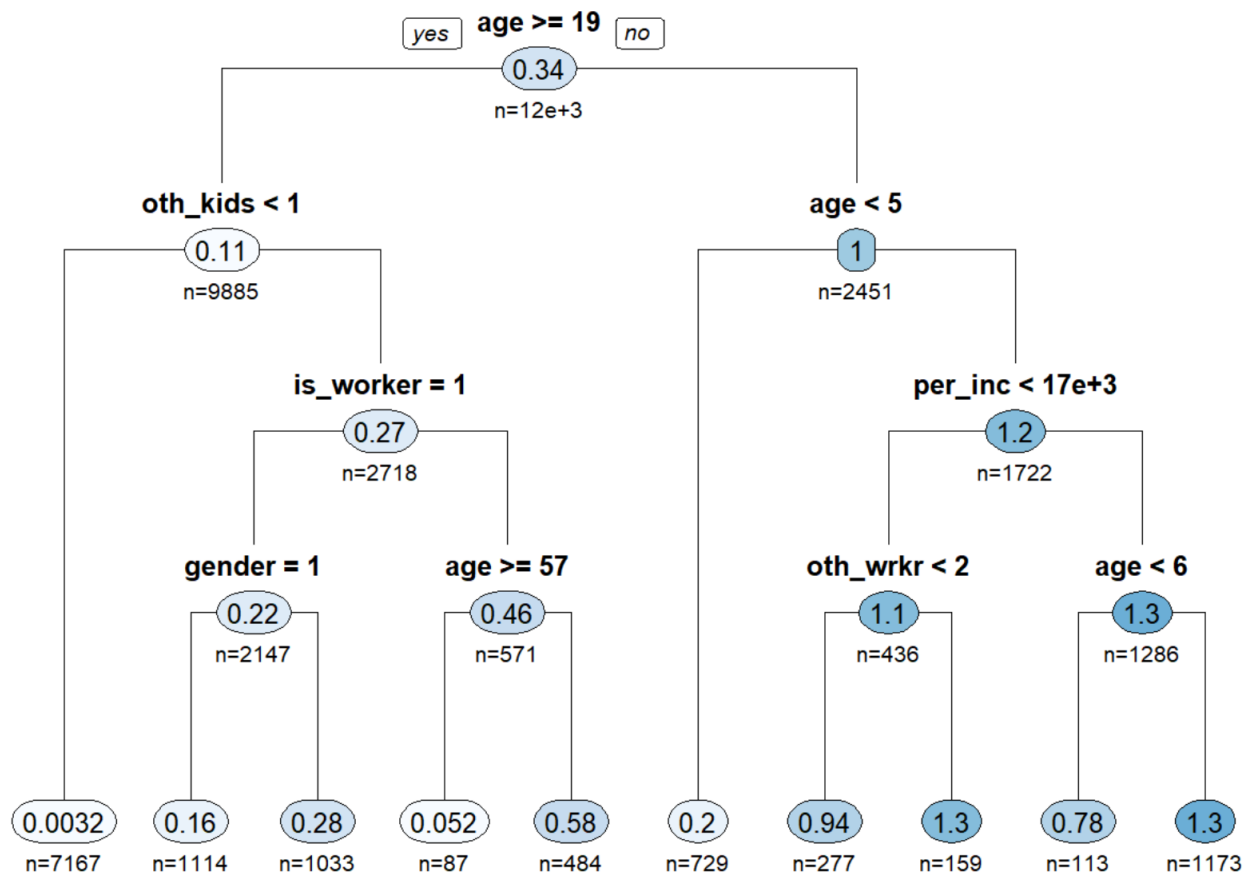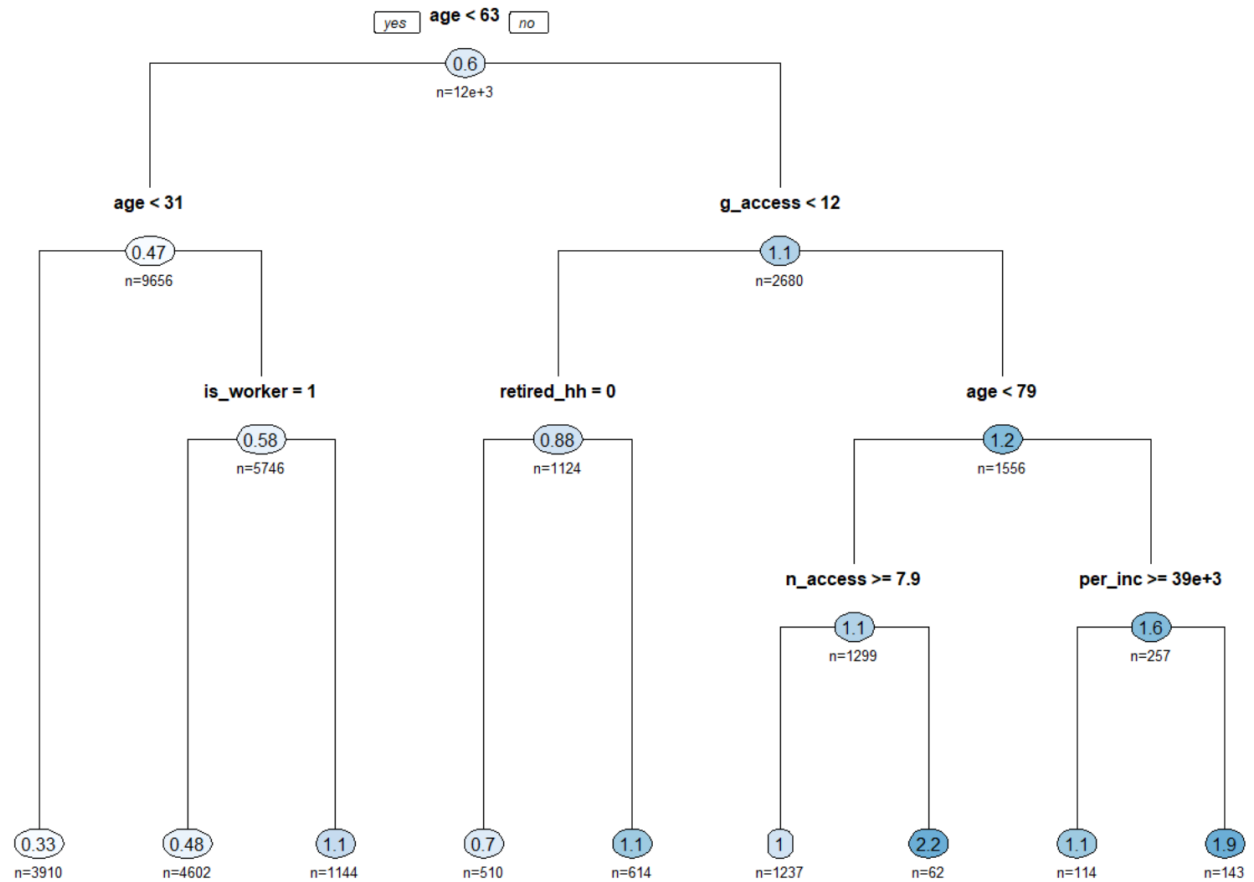
FIGURE 10 Simple, Rationalized Decision Tree for School Trips

The result is a logical structure that is easy to explain and not overfit. For example, the first question asked is if the person is over 18 (an adult). The left and right halves of the tree then describe the behavior of adults and children separately, which is intuitive for school trips. The left side of the tree describes the behavior of adults, and if an adult has no children in the household (oth_kids < 1), they effectively make no school trips. If they do have children, then work status, age, and gender help explain differences in trip making. The right half of the tree describes the behavior of children. Children under 5 make few school trips on average (usually only those with older siblings) while school-age kids make many more. Per-capita income and the number of workers in the household further influence their behavior.

Although the ability to interpret the model comes at a cost in goodness-of-fit, the cost is modest, and the method still significantly outperforms all traditional methods. The decision tree is able to capture the highly non-linear aspects of trip making like cross-classification, while also incorporating more explanatory variables, maintaining appropriate sample sizes for rate estimation, and partitioning the data in a more intelligent way.

The tree for maintenance/eat-out trips (N-HB-OME, including shopping, eating, personal business, etc.) provides another example. As can be seen in Figure 11 in addition to age and work status, this model recognizes differences based on the accessibilities where people live. The model also recognizes that the factors influencing shopping trips for non-seniors (under 63) are relatively simple: work status and age. At the same time, a larger number of factors (including accessibility) influence seniors over 63. The role of accessibility here is particularly interesting. With poor general accessibility

14

(g_access), best associated with rural areas, seniors make fewer home-based shopping trips. On the other hand, the trip rate falls again for seniors in downtowns (high nearby accessibility, n_access) reflecting their ability to chain multiple trips together in a single outing, making more NHB and fewer HB trips.

FIGURE 11 Rationalized Decision Tree for Maintenance/Eat-Out Trips

NESTED DESTINATION CHOICE MODELS

The second innovation introduced in the TRMG2 is the use of nested logit models for HB destination choice with the choice of destination zone conditional on a choice of destination district. Nested logit models have long and widely been used for mode choice (*10,11*). It has also been common to consider destination and mode choices together as a nested logit model (*12*). The possibility of using nested logit models to address unobserved homogeneity between alternative destinations related to spatial autocorrelation has long been acknowledged but was initially generally rejected either as too difficult computationally or because of the potentially arbitrary nature of spatial nesting structures (*14*, *15*). However, with the advent of more computing power after the turn of the millennium, there have been at least two successful demonstrations of the use of the nested logit model for destination choice itself (*16*, *17*). While these papers demonstrate the value of the approach, both were foreign academic applications, and this is believed to be the first fully-realized application of the approach in practice in the United States.

The Triangle region being highly multinucleated with several distinct communities was thought to be an ideal case for testing spatial nesting in destination choice. Twelve districts,

shown in Figure 12, were developed in consultation with the local agencies to delineate distinct communities or subareas of the region.

FIGURE 12 District Scheme for Spatial Nesting

The use of these districts as nests in a hierarchical destination choice process helps overcome the loss of spatial information such as adjacency in treating destination choice as a discrete choice among competing alternatives. Alternative means of capturing spatial information that is otherwise lost in discrete choice approaches have generally involved the use of one or more accessibility variables (*13*,*18*,*19*) (but see *20* for a more complex approach). For the TRMG2, rather than relying exclusively on nesting or accessibilities, nested logit destination choice models with accessibility terms were estimated, combining these approaches for the first time. As both the nesting and accessibility approaches are limited in different ways in their ability to capture spatial autocorrelation, it was hypothesized that a model employing both would be more successful than a model relying on either approach alone. The significance of both accessibility variables and nesting coefficients in the estimated models confirms this hypothesis.

The estimation results for the work (W_HB_W) trip model are presented in
Table 2 as an example.

Table 2 Nested Logit Destination Choice Model Estimation Results

Category	Parameter	Value	t Stat
ASCs	ASC_Durham	-0.7100	-3.5
	ASC_Other - NW	0.8740	3.1
	ASC_Cary	0.6470	4.5
	ASC_Southern Durham	1.0300	6.9
Size	Size	0.6740	-12.9
Logsums	Transit Logsum (Segment Specific)	0.0601	2.9
Betas	Intrazonal	0.8150	5.6
	Time	-0.0965	-32.2
Accessibilities	Hospital Accessibility (SOV)	0.1950	8.1
	Transit Accessibility	0.0322	2.5
	Walk Accessibililty (Zero-auto HH)	0.3530	2.5
HomeClusters	HomeCluster_Chapel Hill	0.3950	2.5
	HomeCluster_Cary	0.3340	3.1
	HomeCluster_ResearchPark	1.1100	2.1
	HomeCluster_Other - SW	1.2900	3.1
	HomeCluster_Other - SE	0.3280	1.6
	HomeCluster_Other - NE	0.6010	3.5
	HomeCluster_Other - NW	0.5110	2.1
	HomeCluster_Other - S	0.8800	3.7
Nest Coefficients	Chapel Hill	0.8350	-4.6
	Cary	0.7160	-9.0
	Durham	0.9560	-0.9
	Garner	0.7220	-6.5
	Other - NE	0.7320	-7.0
	Other - NW	0.6600	-6.5
	Other - S	0.6860	-6.4
	Other - SE	0.7390	-6.1
	Other - SW	0.7400	-2.9
	Southern Durham	0.6200	-11.8
	Research Park	0.9970	-0.1
	Raleigh	0.8350	-6.7
Attractions	Industrial Employment	0.3012	-1.1
	Office Employment - High Income	1.0000	
	Office Employment - Low Income	1.0000	
	Retail Employment - High Income	0.3012	-4.2
	Retail Employment - Low Income	0.1033	-9.9
	Service Employment - High Income	0.2982	-6.4
	Service Employment - Low Income	0.3535	-9.9

Statistic	Aggregate	PerCase
Log Likelihood at Convergence	-22353.3	-6.36
Log Likelihood at Null Parameters*	-24883.3	-7.08
Rho squared w.r.t Null Parameters*	0.102	
Number of Cases	3,516	

In addition to the standard size, attraction, impedance, and intrazonal variables the model contains the aforementioned accessibility variables and three sets of variables related to the nesting: alternative

specific constants (ASCs) for the nests, "HomeClusters" or intra-nest constants, and the nesting coefficients. As can be seen in

Table 2 for work trips four of the twelve district nests have ASCs significantly different from zero, eight of the twelve have a home cluster effect, and ten of the twelve have a nesting coefficient significantly different from unity indicating that zones within the district have unobserved similarities and therefore compete more directly with each other than with zones in other districts. The t statistics for these variables indicate that the nested model significantly outperforms and strongly rejects the unnested multinomial logit model specification which is standard in practice. The results for other HB trip types were similar with most districts having significant nesting coefficients in most models with the exception of escort to school trips on work tours (W_HB_EK12) for which the lack of significance may likely have been due to a relatively small number of observations (225) in the estimation dataset.

The significance of the nesting structure and superiority of the nested modeling approach are perhaps not surprising given the multinucleated nature of the Triangle region, but the strength with which the nested models reject their multinomial counterparts suggests that nesting should at least be tested even in more traditionally structured regions.

LINKED NON-HOME-BASED MODELS

There are many problems related to non-home-based trips in traditional trip-based models arising from the fact that they are disconnected from the home-based trips with which they comprise complete tours. In order to properly represent non-home-based trips, two spatial distribution or destination/spatial choice models are required to account for both the trip's origin location and destination location. The four-step model architecture is fundamentally flawed because it produces non-home-based trips from only one trip distribution or spatial choice model.

Earlier hybrid models produced linked HB and NHB trips simultaneously from two spatial choice models called stop location choice and stop sequence choice models (*2,3*). The third distinguishing feature of the TRMG2 is its use of a simpler method of linking HB and NHB trips popularized by the FHWA's Travel Model Improvement Program (TMIP) (*21*). In this approach, non-home-based trips are generated separately for each mode based on home-based trip destinations and modes and then distributed. In this framework, the HB destination choice model serves as one spatial choice model and the NHB destination choice model serves as the second necessary to properly produce NHB trips.

This framework ensures the linkage of NHB trips to HB trips on the same tour in both location and mode. For example, NHB single occupant vehicle (SOV) trips on nonwork tours are generated most by SOV HB trips on non-work tours, somewhat less by high occupant vehicle (HOV) HB trips on non-work tours and not at all by non-motorized trips because no car is available after the HB trip. Conversely, NHB walk trips can be generated by HB trips of any mode although they are generated most by paid auto and walk trips. Thus, for example, if a scenario results in more home-based trips to downtown Raleigh, the scenario will also produce more non-home-based trips in and around downtown Raleigh. Also, to the extent that these new home-based trips are auto trips, more of the non-home-based trips will be as well; whereas, to the extent that the new home-based trips are by transit, more of the non-home-based trips will be by walking and a perhaps a few by transit. This is particularly important for Chapel Hill in the Triangle model, which drastically restricts parking availability downtown.

Since the original publication of the method by TMIP, the method has been further enhanced using boosting to incorporate accessibility. In this approach, the number of NHB trips ($\hat{Y}_{p,m}$) of a given mode (m) and NHB trip type (p) is initially estimated as a function of the number of HB destinations ($X_{t,m}$) by mode and HB trip type (t).

$$\hat{Y}_{p,m} = \sum_{t,m} \beta_{t,m} X_{t,m} \qquad (\,1\,)$$

A boosting model can then be estimated using log-log regression on residuals of the log transform of the initial model, where α and γ are the new parameters, A is a measure of accessibility to nearby destinations and $\bar{Y}_{p,m}$ is the observed number of trips.

$$\gamma ln(A) + ln(\alpha) = ln(\bar{Y}_{p,m}) - ln(\hat{Y}_{p,m}) \qquad (2)$$

So that a revised, boosted estimate of the number of NHB trips ($Y_{p,m}$) is given by Equation 3.

$$Y_{p,m} = \alpha A^\gamma \sum_{t,m} \beta_{t,m} X_{t,m} \qquad (3)$$

Figure 13 illustrates the effect of accessibility to nearby destinations on NHB trip rates by mode and type. For SOV and HOV modes, the NHB trip rate is roughly 50% lower in the least accessible (rural) areas; whereas, the NHB trip rate marginally increases up to 50% higher than the regional average for the most accessible areas in the region. For NHB walk trips, walk accessibility has a more modest affect increasing NHB trip rates by 10 to 20% in the most walk accessible areas but asymptotically approaching zero as walk accessibility approaches zero.
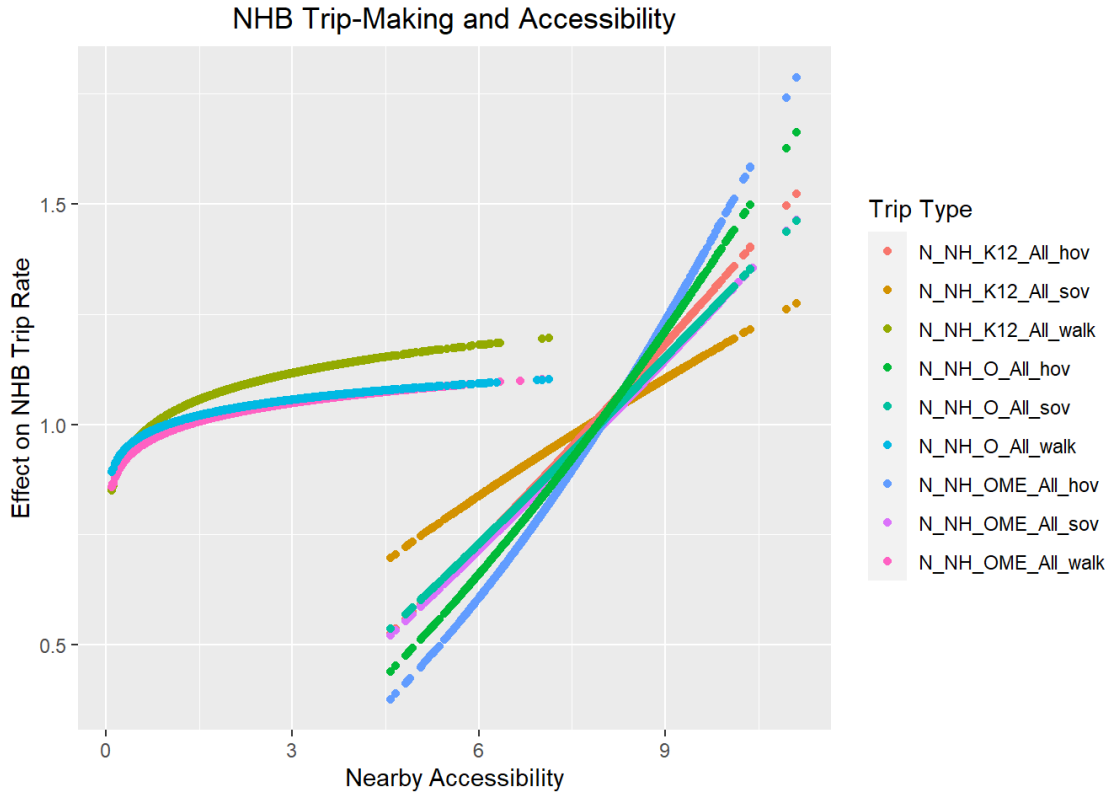

FIGURE 13 Accessibility Boosting of NHB Trip Rates

RESULTS OF SENSITIVITY TESTS

The TRMG2 was subjected to a series of rigorous sensitivity tests to see if it achieved its goal of improved response compared to trip-based models. This paper will present one sensitivity test, the addition of a bus route, in detail, but other sensitivity tests included changes to household income, improvements to pedestrian infrastructure (a pedestrian bridge over a freeway), roadway widening, highway tolling, and parking changes. The TRMG2 met or exceeded expectations in each test.

The single test that best illustrates the multiple improved sensitivities of the TRMG2 is the addition of a single bus route. In FIGURE 14, the new route is shown in green running next to Lake Lynn in the center of the map.
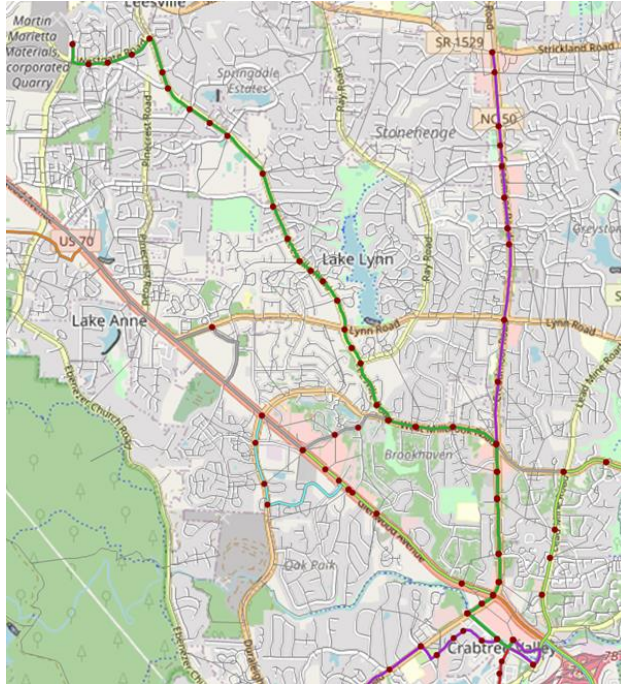
FIGURE 14 Map Showing Addition of a New Bus Route

In a traditional trip model, this addition should lead to a slight change in mode shift, which in turn impacts assignment. In the TRMG2, the model responded with changes in:

- Auto ownership
- Trip productions
- Non-motorized trip making
- Mode choice
- NHB trip making
- Transit/Highway assignments

Auto ownership
The addition of the new bus route caused a slight decrease in auto ownership of 0.6%. This represents the marginal household being able to give up an extra car given new transit access and is captured by the transit accessibility terms in the auto ownership choice model.

HB trip productions
The decrease in auto ownership of 0.6% led to a reduction in home-based trip making by 0.2%. As households give up cars, they make fewer trips and become more strategic about trips they do make employing strategies like trip chaining and carpooling.

Non-motorized
Non-motorized trips increased by 0.12% despite reductions in total trip making and is largely the result of decreased auto ownership.

Mode choice
As expected, transit mode share in the nearby zones increased from 0.7% to 0.9%.

NHB trip production
The improved handling of NHB trips in the TRMG2 meant that the upstream changes to HB trips (in both number and mode) directly impacted NHB trip making. NHB auto trips fell while NHB transit trips increased.

Assignment
Finally, the culmination of upstream impacts resulted in transit trips increasing at the expense of auto trips.

While the impact is small (as expected for a single new bus route), the TRMG2 shifts the neighborhood away from autos and towards multi-modal travel patterns. Each component displayed logical sensitivities and downstream models responded to upstream changes appropriately.

CONCLUSION
This paper presents a new and elegant approach to a hybrid travel demand model. This new framework advances the practice through two notable innovations. First, is the use of a machine learning decision tree structure for the HB trip production models. This approach introduces a disaggregate model that is intuitive yet powerful, allows the use of person and household characteristics, and that outperforms various other trip generation model goodness-of-fit measures. Second is the use of nested logit models for HB destination choice, yielding model results that better reflect the spatial effects of destination choice. In addition to these two innovations, the TRMG2 offers several other advances that will be of interest to other regions. Finally, sensitivity testing shows a model that is appropriately sensitive to changes in ways not found in trip based models. This new framework will be of interest to other MPOs and regional agencies that desire to moving away from traditional trip based models to more advanced models, but without the upfront investment in time and cost of many activity-based models.

ACKNOWLEDGEMENTS

REFERENCES

1. Bernardin, V. An Accessibility-Based Approach to Travel Demand Forecasting: New Alternative to Four-Step and Activity-Based Methods. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.

2. Bernardin, V., Jr. *A Trip-Based Travel Demand Framework Consistent with Tours and Stop Interaction*. Dissertation. Northwestern University, Evanston, Ill., 2008.

3. Bernardin, V., and M. Conger. From Academia to Application: Results from Calibration and Validation of First Hybrid Accessibility-Based Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2010, No. 2176, pp. 50–58. https://doi.org/10.3141/2176-06

4. Vyas, G. Stepping Closer to ABM: Hybrid 4-Step Models. Presented at the 16th TRB National Transportation Planning Applications Conference, Raleigh, NC, 2017.

5. Bernardin, V. and J. Chen. *How-to: Improve Non-Home-Based Trips*. Federal Highway Administration, Travel Model Improvement Program, 2018. https://rosap.ntl.bts.gov/view/dot/55799/dot_55799_DS1.pdf

6. Bernardin, V. *How-to: Think About Model Design for Your Region*. Federal Highway Administration, Travel Model Improvement Program, 2018. https://rosap.ntl.bts.gov/view/dot/55797/dot_55797_DS1.pdf

7. Newman, J., and V. Bernardin. Hierarchical Ordering of Nests in a Joint Mode and Destination Choice Model. *Transportation*, Vol. 37, No. 4, 2010, pp. 677–688. https://doi.org/10.1007/s11116-010-9277-x

8. Chen, T., and C. Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: ACM, 2016, pp. 785–794. http://doi.acm.org/10.1145/2939672.2939785.

9. Arrieta, A., N. Diaz-Rodriguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, Vol. 58, June 2020, pp. 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

10. Ortuzar, J. Nested Logit Models for Mixed-Mode Travel in Urban Corridors. *Transportation Research, Part A: General*, Vol. 17, No. 4, 1983, pp. 283-299. https://doi.org/10.1016/0191-2607(83)90092-4

11. Polydoropoulou, A. and M. Ben-Akiva. Combined Revealed and Stated Preference Nested Logit Access and Mode Choice Model for Multiple Mass Transit Technologies. *Transportation Research Record: Journal of the Transportation Research Board*, 2001, No. 1771, pp. 38–45. https://doi.org/10.3141/1771-05

12. Jonnalagadda, N., J. Freedman, W. Davidson, and J. Hunt. Development of microsimulation activity-based model for San Francisco: destination and mode choice models. *Transportation Research Record: Journal of the Transportation Research Board*, 2001, No. 1777, pp. 25–35. https://doi.org/10.3141/1777-03

13. Fotheringham, A. Modelling hierarchical destination choice. *Environment and Planning A*, 1986, Vol. 18, No. 3, pp. 401-418. https://doi.org/10.1068/a180401

14. Borgers, A. and H. Timmermans. Choice model specification, substitution and spatial structure effects: a Simulation experiment, *Regional Science and Urban Economics,* 1987, Vol. 17, pp. 29-47. https://doi.org/10.1016/0166-0462(87)90067-6

15. Eymann, A., and G. Ronning. Microeconometric models of tourists' destination choice. *Regional Science and Urban Economics*, 1997, Vol. 27, No. 6, pp. 735-761. https://doi.org/10.1016/S0166-0462(97)00006-9

16. Hong, S., J. Kim, H. Jang, and S. Lee. The roles of categorization, affective image and constraints on destination choice: An application of the NMNL model. *Tourism Management*, 2006, Vol. 27, No. 5, pp. 750-761. https://doi.org/10.1016/j.tourman.2005.11.001

17. Hammadou, H., I. Thomas, A. Verhetsel, and F. Witlox, How to incorporate the spatial dimension in destination choice models: The case of Antwerp. *Transportation Planning and Technology*, 2008, Vol. 31, No. 2, pp. 153-181. https://doi.org/10.1080/03081060801948126

18. Bhat, C., A. Govindarajan, and V. Pulugurta. Disaggregate Attraction–End Choice Modeling: Formulation and Empirical Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 1998, No. 1645, pp. 60–68. https://doi.org/10.3141/1645-08

19. Bernardin Jr, V., F. Koppelman, and D. Boyce. Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition. *Transportation Research Record: Journal of the Transportation Research Board*, 2009, No. 2132, pp. 143–151. https://doi.org/10.3141/2132-16

20. Bhat, C., and J. Guo. A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research B*, Vol. 38B, 2004, pp. 147–168. https://doi.org/10.1016/S0191-2615(03)00005-5

21. Bernardin, V. and J. Chen. *How-to: Improve Non-Home-Based Trips*. Federal Highway Administration, Travel Model Improvement Program, 2018. https://rosap.ntl.bts.gov/view/dot/55799