Forecasting Model for Air Taxi, Commercial Airline, and Automobile Demand in the United States

Hojong Baik, Antonio A. Trani, Nicolas Hinze, Howard Swingle, Senanu Ashiabor, and Anand Seshadri

A nationwide model predicts the annual county-to-county person round-trips for air taxi, commercial airline, and automobile at 1-year intervals through 2030. The transportation systems analysis model (TSAM) uses the four-step transportation systems modeling process to calculate trip generation, trip distribution, and mode choice for each county origin–destination pair. Network assignment is formulated for commercial airline and air taxi demand. TSAM classifies trip rates by trip purpose, household income group, and type of metropolitan statistical area from which the round-trip started. A graphical user interface with geographic information systems capability is included in the model. Potential applications of the model are nationwide impact studies of transportation policies and technologies, such as those envisioned with the introduction of extensive air taxi service using very light jets, the next-generation air transportation system, and the introduction of new aerospace technologies.

Demand for transportation is outpacing supply, as evidenced by increasing congestion in the commercial air and automobile networks. FAA, NASA, state departments of transportation, and metropolitan planning organizations are developing transportation policies to address this situation. Study of national-level transportation policy requires multimodal travel demand models. Few such models are available today.

This paper presents a nationwide, multimodal, and intercity transportation demand forecasting model. The transportation systems analysis model (TSAM) is an effort to understand the interrelationships among ground and air transportation systems in the United States. TSAM has potential use in quantifying the impact of fielding a next-generation air transportation system (NextGen) (1), studying future mobility trends in the nation with many secondary airports offering more point-to-point services, and understanding the impact of new aerospace technologies—such as very light jets—operating into the National Airspace System.

The model was developed to predict the demand and effects of the Small Aircraft Transportation System (SATS) (2) with the planned framework as described by Trani et al. (3). This NASA SATS program promoted the use of small aircraft and small airports to supplement existing transportation modes, reduce travel times, and improve mobility to small communities. With increasing computing power, the availability of the American Travel Survey (ATS) (4), and county-level socioeconomic projections from Woods and Poole (5), such analysis could be achieved. The model has evolved to include forecast of commercial and cargo flights. TSAM forecasts annual person round-trips by air taxi, commercial airline, and automobile between all counties in the United States. The demand estimation process differentiates between business and nonbusiness trip purposes and five household income group levels. Additional TSAM functions forecast airborne cargo flows and flights, international traveler enplanements, and international aircraft operations. Airport capacity modeling quantifies the capacity and level of congestion at all public airports in the United States. TSAM displays all the results in tables, plots, and maps by using Geographic Information System technology.

**TSAM Modeling Framework**

TSAM uses the four-step modeling process for transportation systems analysis: trip generation, trip distribution, mode choice, and network assignment. The trip generation model calculates the number of produced and attracted person round-trips for each county in the United States. In the context of TSAM, a trip is an intercity journey with a one-way travel route distance greater than 100 statute miles, excluding commuter travel. The trip distribution model distributes the produced person round-trips to all other counties. The mode choice model assigns a transportation mode to each traveler on the basis of the door-to-door travel time and travel cost of each mode of transportation. The network assignment model loads the commercial airline and air taxi demand onto the National Airspace System. Basic outputs of TSAM include total annual person round-trips between all counties in the United States by trip purpose (business and non-business), by mode (air taxi, commercial airline, and automobile), and by five household income groups. Outputs of TSAM serve as inputs to other models, such as the airspace concept evaluation system (ACES) (6), for further analysis. The modeling framework used in TSAM is shown in Figure 1.

**Trip Generation Model**

The trip generation model in TSAM forecasts the number of produced and attracted person round-trips for the 3,091 counties in the United States as defined by Woods and Poole (5). The trip generation forecast...
is based on the ATS (4) trip rates and Woods and Poole’s socio-economic projections to the year 2030 (5). Trip rates from the ATS are a function of socioeconomic variables and region types (4). Future trips are generated by combining these trip rates with Woods and Poole’s future projections of socioeconomic variables (5). The ATS contains detailed information of 540,000 intercity person round-trips made in the United States in 1995. To date, ATS is one of the best data sets known to be publicly available for intercity transportation studies. The 2001 National Household Travel Survey is also available (7), but it contains only 22,000 records.

Nationwide trip production rates are derived from the 1995 ATS for business and nonbusiness trip purposes; for five household income groups (in 2000 dollars): <$30,000, $30,000 to $60,000, $60,000 to $100,000, $100,000 to $150,000, and >$150,000; and for three region types: consolidated metropolitan statistical area (CSA), metropolitan statistical area (MSA), and non-MSA. Table 1 displays the resulting nationwide trip rates. A state-based adjustment factor is used to correct the total produced person round-trips from each state to the ATS published totals (4). Future trips are estimated by multiplying the projected number of households in each income group by the corresponding statewide trip production rate for each trip purpose, income group, and region type.

Business and nonbusiness purpose trip attraction rates are calculated by dividing the ATS statewide attracted person round-trips by the Woods and Poole state total employment and service employment data for business and nonbusiness purposes, respectively. The trip generation model calculates the county trip attraction by multiplying the statewide trip attraction rate tables and the county total employment and service employment data for the business and nonbusiness trip purposes, respectively.

![TSAM modeling framework](image-url)
TSAM displays the output of the trip generation model as tables and maps. The tables show the produced and attracted person round-trips by trip purpose and household income group for all the counties. The maps display the demand by each trip purpose and household income group. Figure 2 shows the produced business person round-trips in 2015 for all income groups.

### TABLE 1 Annual Intercity Round Trips per Household per Year

<table>
<thead>
<tr>
<th>Household Income Group</th>
<th>&lt;$30K</th>
<th>$30K–$60K</th>
<th>$60K–$100K</th>
<th>$100K–$150K</th>
<th>&gt;$150K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSA</td>
<td>0.73</td>
<td>1.30</td>
<td>2.45</td>
<td>4.20</td>
<td>5.89</td>
</tr>
<tr>
<td>MSA</td>
<td>0.82</td>
<td>1.76</td>
<td>2.97</td>
<td>5.28</td>
<td>7.27</td>
</tr>
<tr>
<td>Non-MSA</td>
<td>1.30</td>
<td>2.75</td>
<td>4.23</td>
<td>5.67</td>
<td>8.11</td>
</tr>
<tr>
<td>Nonbusiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSA</td>
<td>4.29</td>
<td>6.63</td>
<td>9.00</td>
<td>10.61</td>
<td>13.95</td>
</tr>
<tr>
<td>MSA</td>
<td>4.62</td>
<td>7.45</td>
<td>9.72</td>
<td>11.92</td>
<td>13.00</td>
</tr>
<tr>
<td>Non-MSA</td>
<td>6.40</td>
<td>10.48</td>
<td>12.74</td>
<td>14.32</td>
<td>16.84</td>
</tr>
</tbody>
</table>

**FIGURE 2** Business person round-trip production for all U.S. counties in 2015.

Trip Distribution Model

The trip distribution model distributes the trip generation of produced person round-trips to each of the 3,091 counties in the United States. This process is done for both business and nonbusiness trip purposes and the five household income groups by using the standard
The trip distribution model forecasts the county-to-county demand to the year 2030.

\[ T_{ij} = \frac{A_{i} F_{ij} K_{ij}}{\sum_{j} A_{i} F_{ij} K_{ij}} \]  

where

\( T_{ij} \) = trips from county \( i \) to county \( j \),
\( P_{i} \) = trips produced from county \( i \),
\( A_{j} \) = trips attracted to county \( j \),
\( F_{ij} \) = friction factor defined as an inverse function of travel distance, and
\( K_{ij} \) = socioeconomic adjustment factor for interchange \( ij \).

The trip distribution model is calibrated at the state level by using the ATS (4). A two-stage process is used to calculate the distribution of the trips at the county level. First, the distribution is done at the state level by using calibrated \( F_{ij} \) and \( K_{ij} \) factors that have been validated with the ATS state-to-state round-trip tables. Then, the trip distribution model calculates the trips between counties by using a Fratar model recursively while constraining the state distribution. The output of the trip distribution model is a set of county-to-county trip tables by trip purpose and for five household income groups. TSAM displays the output of the trip distribution model as tables and maps. The tables and maps display the distributed person round-trips from one county to all other counties by trip purpose and household income group. Figure 3 shows the expected distribution of produced person round-trips from Fairfax County, Virginia, to the United States in 2015.

**Mode Choice Model**

The mode choice model uses the trip distribution and a nested logit model to calculate the air taxi, commercial airline, and automobile county-to-county demand (9). The probability of ridership for each mode is dependent on its utility function \( U_{ij}^{m} \). The utility for each mode is calculated by using its door-to-door travel time and travel cost.

**FIGURE 3** Business person round-trips from Fairfax County, Virginia, to all other counties in 2015.
Model Structure

The nested logit model structure has two tiers so that mode choice (in the first tier) and air route choice (in the second tier) can be considered in a single model. Equation 2 gives the formulation of the nested logit model.

\[
P^m_{ij} = \frac{e^{\alpha_m U^m_{ij}}}{\sum_{m} e^{\alpha_m U^m_{ij}}} \tag{2}
\]

where

\[
P^m_{ij} = \text{probability for a mode } m \text{ to be selected for a given trip between origin county } i \text{ and destination county } j,
\]

\[
sp_m = \text{mode-specific scaling coefficient to be calibrated, and}
\]

\[
U^m_{ij} = \text{utility function for a trip between counties } i \text{ and } j \text{ by mode } m.
\]

The utility function is defined in Equation 3 with two variables: travel cost and travel time. The travel time and cost coefficients (\(\alpha_{\text{time}}\) and \(\alpha_{\text{cost}}\)) vary depending on the trip purpose and the traveler’s household income. The coefficients are calibrated by using the ATS by considering the market share for each mode as a function of distance. Since the ATS lacks travel time and cost information, travel time and cost were developed synthetically to calibrate the nested logit model coefficients for each trip purpose and household income group. Business travelers are more sensitive to travel time than to travel cost. Nonbusiness travelers are more sensitive to travel cost than to travel time. Low-income households are more sensitive to travel cost than are higher-income households.

\[
U^m_{ij} = \alpha_{\text{time}} \cdot \text{time}^m_{ij} + \alpha_{\text{cost}} \cdot \text{cost}^m_{ij} \tag{3}
\]

where

\[
U^m_{ij} = \text{utility value of a mode } m \text{ between counties } i \text{ and } j,
\]

\[
\text{time}^m_{ij} = \text{travel time spent for a trip between counties } i \text{ and } j \text{ by mode } m,
\]

\[
\text{cost}^m_{ij} = \text{travel cost spent for a trip between counties } i \text{ and } j \text{ by mode } m,
\]

The rail mode was investigated but has not been calibrated in the mode choice, because the ATS does not have sufficient data to calibrate the rail mode adequately. The rail network, travel cost, and

<table>
<thead>
<tr>
<th>Table 2 Mode Choice Model Coefficients, by Income Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>Travel time</td>
</tr>
<tr>
<td>&lt;$30K</td>
</tr>
<tr>
<td>$30K–$60K</td>
</tr>
<tr>
<td>$60K–$100K</td>
</tr>
<tr>
<td>$100K–$150K</td>
</tr>
<tr>
<td>$150K</td>
</tr>
<tr>
<td>Travel cost</td>
</tr>
<tr>
<td>&lt;$30K</td>
</tr>
<tr>
<td>$30K–$60K</td>
</tr>
<tr>
<td>$60K–$100K</td>
</tr>
<tr>
<td>$100K–$150K</td>
</tr>
<tr>
<td>$150K</td>
</tr>
<tr>
<td>Scale parameter</td>
</tr>
</tbody>
</table>

\(\alpha_{\text{time}}, \alpha_{\text{cost}}\) = calibration coefficients (see Table 2).

Table 2 displays the mode choice model coefficients, and Figure 4 shows the model structure.

There are no large-scale travel surveys that link household income and employment data to air taxi mode choice observations and their associated characteristics of travel time and travel cost. To overcome this shortcoming in available data, the new mode of air taxi was calibrated by assuming the mode choice coefficients (travel time and travel cost) for commercial air also apply to travel by air taxi. Unquantifiable air taxi mode characteristics such as comfort and perceived safety are assumed to be similar to that for commercial airline. Limited stated-preference surveys at three different locations confirmed these assumptions (10). A detailed explanation of the mode choice calibration and validation with the ATS was explained by Ashiabor et al. (9).
Automobile Travel Cost and Time Calculations

The driving distances and driving times are calculated between the population centroids of the counties by using Microsoft MapPoint 2004 (12). The county population centroids are derived by using a weighted average of the census tract population. The mode choice model obtains the automobile travel cost by multiplying the driving distance by the cost per mile of driving and adding the appropriate lodging cost in case of an overnight trip. Travel times are modeled by adding the driving time and the appropriate time for overnight lodging.

Commercial Airline Travel Cost and Time Calculations

Commercial airline travel costs are based on the Bureau of Transportation Statistics’ airline origin and destination survey (DB1B) (13). The DB1B is a 10% sample of all airline ticket sales from reporting carriers. The model uses an average business and coach class fare between an airport origin–destination pair. The business class fare is calculated by averaging together the DB1B first class, business class, and unrestricted coach fares. Coach class fares are calculated by averaging the restricted coach class fares. The averaging excludes coach fares less than $50 and business fares less than $100 to eliminate the frequent flyer and upgrade fares. Fares higher than $5,000 are also excluded. Fares for certain airlines are corrected because they were not reported correctly. Southwest Airlines, for example, reports all their fares to be unrestricted and restricted first class.

These fares are reclassified to unrestricted and restricted coach class fares in this analysis. The model calculates the door-to-door commercial airline travel cost by adding the access cost to the origin airport, the average fare between origin and destination airports, and the egress cost from the destination airport. Lodging cost is also added in case the round-trip is not possible in 1 day.

The commercial airline travel time is calculated by using the Official Airline Guide (OAG) (14). The OAG is the official commercial airline schedule. The door-to-door travel time is composed of the access time to the origin airport, the processing time at the origin airport, the schedule delay, the average travel time between the origin and destination airports, the processing time at the destination airport, and the egress time from the destination airport. Additional time is added in case of an overnight trip. The model calculates the travel time between all the commercial airports by averaging the travel time of all the feasible routes between two airports. Three candidate airports are selected for each origin and destination county: the closest airport by driving distance, the cheapest airport by average fare, and the airport that are open for public use. The FAA’s landing facility database provides the needed airport characteristics (15).

Air Taxi Travel Cost and Time Calculations

The air taxi door-to-door travel cost and travel time components are similar to commercial airlines. However, the flying time between the origin and destination airports is calculated by using the air taxi aircraft performance and the great circle distance adjusted by typical detour factors observed in the system. The flying cost is the cost per seat mile multiplied by the great circle distance. The origin and destination airports are selected by assigning the closest useable airport to each county by using driving time as the selection criteria. An airport is useable depending on its runway characteristics and the air taxi vehicle chosen. For example, if a very light jet is used, then only airports that have at least an effective 3,000-ft paved runway and that are open for public use are considered. The FAA’s landing facility database provides the needed airport characteristics (15).

Model Outputs Description

The mode choice model outputs the person round-trip demand for each transportation mode by trip purpose and household income group for all county pairs between the 3,091 counties in the United States. For commercial airline and air taxi, the model also records the person round-trip demand at each airport. Additionally, the model reports the total revenue and the total travel time per mode.

These statistics are useful for measuring the effects of new policies and technologies on the transportation system. Travel time and revenue savings and losses can be quantified. Finally, the model calculates the market share for each mode of transportation between all the county pairs. Figure 5 shows the forecast number of commercial airline round-trips taken by business travelers from Washington Dulles International in 2015. Figure 6 displays the expected distribution of automobile and commercial airline round-trips by household income groups over distance in 2015.

Network Assignment

Network assignment in TSAM is composed of the commercial airline and the air taxi network assignment models. The commercial airline network assignment model calculates future enplanements at commercial airports by using the commercial airline origin–destination airport person round-trip demand from the mode choice model (16). The air taxi network assignment model generates daily schedules and flight trajectories from the air taxi origin–destination airport person round-trip demand for input into the ACES model (6).

Commercial Airline Network Assignment Model

The commercial airline network assignment model uses demand generated from the mode choice model (16). The model calculates the number of passengers on each commercial airline route by using the origin–destination airport person round-trip demand. A multinomial logit model is used to calculate the probability of each commercial airline route being selected by a traveler on the basis of travel time and fare paid. A synthetic commercial airline network is developed by using the commercial airline schedule published by the OAG (14). This network provides the average travel time on each commercial airline route. The model was calibrated for this study by using the DB1B (13) and the OAG. The DB1B provides the passenger flows and fares for the different routes between origin and destination airports.

The main output of the commercial airline network assignment model is the future number of originating and transfer enplanements at each of the commercial airports in the continental United States. Results are used to predict airline operations and delays caused by capacity constraints at commercial airports. Furthermore, proposed
changes in operations of commercial airports and airlines can be examined.

**Air Taxi Network Assignment Model**

The air taxi network assignment model in TSAM also uses the mode choice model output of annual air taxi demand. From the origin–destination airport person round-trip demand, the model creates daily flight schedules by using a Monte Carlo distribution simulation model. The conversion from annual demand to daily demand is based on seasonal variations of passenger travel behavior. Thus, TSAM also outputs the flight trajectories of all the daily air taxi flights. The flight trajectories of all the daily air taxi flights and their schedule is an input to ACES (6) to simulate future en route and terminal area sector loads and traffic conflicts.

The air taxi flight trajectories are used to calculate the impact of the additional air taxi flights on the air route traffic control centers (ARTCCs) over time. Figure 7 shows the forecast of additional loads incurred in the Atlanta ARTCC during a typical day in 2015 due to very light jet traffic.

**ADDITIONAL TSAM FUNCTIONS**

TSAM includes three additional forecasts: international enplanements and operations, air cargo flows and operations, and airport capacity estimations.

**International Travel Demand Model**

International operations at some airports are a significant percentage of the total operations and contribute to congestion and delays at these airports. The international travel demand model in TSAM estimates the number of attracted and produced international enplanements at 66 international airports in the continental United States (17). A regression model is developed by using the gross domestic product of nine world regions and historical international enplanements from the Bureau of Transportation Statistics T100 international segment database (18). The nine world regions used are Africa, Asia, Canada, the Caribbean and Central America, Europe, Mexico, the Middle East, Oceania, and South America. The model reports total international enplanements and passenger operations at each
of the 66 international airports by each of the nine world regions until 2030.

**Air Cargo Demand Model**

Most air cargo operations are flown on weekday nights. However, these operations could have a significant effect on congestion and delays in the future. The model forecasts the quantity of freight and mail that is transported by air up to 2025 (19). The air cargo flows and operations are calculated between the 893 domestic airports with cargo operations by using a Fratar model.

**Airport Capacity and Delay Model**

The capacity of up to 2,320 airports is calculated (20) by using the airfield capacity model (21) and the FAA airport capacity benchmark report (22). From the aircraft mix by airport, an aggregate value is derived for the maximum capacity for visual flight rules and instrument flight rules. By using the predicted operations for all the commercial, cargo, international, air taxi, and general aviation operations and the capacity for each airport, this model forecasts future delays at airports. A feedback loop will be integrated to incorporate delay information into the mode choice module.

**TSAM VALIDATION AND RESULTS**

TSAM is validated by using historical domestic and international enplanement data at U.S. commercial airports reported on the basis of FAA passenger boarding statistics (23). FAA-reported enplanements are compared to domestic enplanements calculated from the commercial airline network assignment model and to international enplanements calculated from the international travel demand model. TSAM estimates historical enplanements well when compared to the reported FAA enplanements (23), as shown in Figure 8, and the T100 domestic airport enplanements (18), as shown in Figure 9. Between 2,300 and 3,100 very light jet aircraft are estimated to be needed to meet air taxi demands in 2025 (24).
TSAM can forecast the possible benefits of deploying NextGen. The analysis carried out with TSAM assumes that airport processing times are reduced by half in 2025 compared to airport processing times in 2006. The analysis also considers a 5% reduction in the air transportation block time to model expected benefits of NextGen as aircraft would be routed more directly between airports (with a robust deployment of ADS-B and advanced air traffic management technology). The analysis assumes no penalty for interstate users. Figure 10 illustrates that with such assumptions, air transportation demand would grow by 11% if NextGen is implemented compared to a do-nothing scenario (25).

NextGen benefits travelers and increases demand for flying distances of less than 800 mi as it offers the ability to shorten door-to-door travel times. Beyond this distance, travel time reductions due to NextGen, although critically important to reducing delays at individual airports, are not significant compared to the total trip time, and thus predicted demand does not change. Also, at distances greater than 800 mi, travelers have no practical alternatives to air travel for high-speed transportation for these trips, and thus demand for these trips is not forecast to change.

CONCLUSION

An integrated TSAM was developed that uses the four-step transportation systems modeling process. The trip generation model calculates the number of produced and attracted person round-trips between all the 3,091 counties in the United States. The trip distribution model distributes the produced person round-trips to all the other counties by using a gravity model. The mode choice model splits the distributed person round-trip demand among air taxi, commercial airline, and automobile by comparing the door-to-door travel time and travel cost for each mode. Finally, the network assignment model converts the annual commercial airline and air taxi person round-trips into daily flights. The resulting output reports annual person round-trips between all counties for each mode, trip purpose, and household income group. The airport demand output contains the annual person round-trips, enplanements, and flights between all the commercial and air taxi airports. TSAM also provides the annual total time and total cost for each mode, which can be used to quantify the effect of policy changes on the transportation system. In addition, TSAM provides the commercial airline and air taxi potential revenue.
FIGURE 8  TSAM 4.7 enplanements versus FAA historical enplanements.

FIGURE 9  TSAM domestic airport enplanements versus T100D airport enplanements in 2004.
The development of TSAM started as part of the SATS program to forecast the demand for small aircraft and small airports. Today, TSAM is one of the models used to predict very-light-jet traffic for the Joint Program and Development Office, which is developing NextGen. TSAM, in conjunction with noise and emission models, can estimate noise and emission effects at airports when new technologies are introduced to the system (26). TSAM can predict air transportation demand changes affected by fuel prices and taxes charged to use the air transportation system (27).

TSAM’s planned role for NASA studies is to be one of the models to forecast the number of flights by commercial airline, cargo airline, air taxi, and general aviation and forecast mobility metrics (25). Air traffic simulation models use this demand prediction to quantify potential delays in the National Airspace System. TSAM can also be used to predict nationwide demand changes resulting from various FAA policies and concepts of operations (28). Another application of TSAM is cost–benefit assessment of FAA technology investments in the National Airspace System. Airport priority investments also could be studied with TSAM.

FUTURE RESEARCH

Future development will add a feedback loop between the network assignment module and the trip distribution module to model the effects of capacity constraints in the system. A schedule generator will be implemented to forecast the commercial airline schedule and network structure. The automobile forecast will be improved by adding a highway network and demand module.

ACKNOWLEDGMENTS

The National Consortium for Air Mobility, NASA, and the Joint Program and Development Office provided financial support to the Air Transportation Systems Laboratory at Virginia Polytechnic Institute and State University to develop the model. The authors thank Stuart Cooke, Jeff Viken, Jeremy Smith, and Sherry Boroner of NASA, Sam Dollyhigh of ATK, and John Callery of Swales Aerospace for their support in developing TSAM and for constructive criticism and suggestions to improve the model. The authors acknowledge C. Joshi, N. Shen, Y. Xu, D. Sohn, K. Murthy, and Y.-T. Chen for their modeling contributions to TSAM.

REFERENCES


The Aviation Economics and Forecasting Committee sponsored publication of this paper.