Statewide Transportation Models in the U.S.: A Review of the State of Practice

Rolf Moeckel
Technical University of Munich
Arcisstr. 21, 80333 Munich, Germany
Tel (+49 89.289.22699)
ORCID number: 0000-0002-6874-0393
rolf.moeckel@tum.de

Rick Donnelly
WSP, Inc.
6100 NE Uptown Blvd, Suite 700
Albuquerque, NM 87501, USA
Phone +1 (505) 878-6524
ORCID number: 0000-0002-1883-8528
rick.donnelly@wsp.com

Joanna Ji
PTV UK LTD.
Mercury House
South Bank, London SE1 8UL, UK
Phone: +44 7825 211863
ORCID number: 0000-0002-1747-5163
joanna.ji@ptvgroup.com

7,161 words and 1 table

Abstract
Out of 50 U.S. states, 34 operate a statewide transportation model, and another 4 have a statewide models under development. In 2006, only 20 states had statewide models. To better understand the increasing interest in statewide modeling, a survey of the departments of transportation of all 50 states was conducted. The goal of this survey was to understand the design principles of statewide models, their application in scenario analysis and the challenges faced by state departments of transportation. The research found that it is more common for statewide models than for urban models to include other model domains, such as long-distance travel, freight flows, economic models or environmental impact models. Only two states apply activity-based models, while the majority continues to apply models that follow the four-step approach. A major challenge many states reported is to attract and maintain highly skilled employees. Scenarios that relate to highway extensions are the most common application of statewide models, though the arrival of autonomous vehicles is a significant concern for many agencies that operate statewide models.

Keywords: statewide models; statewide modeling, long-distance travel; person travel demand model; freight model

1 Introduction

Statewide travel forecasting models have become an essential tool for transportation planners. They are used in a variety of planning and programming activities at the state and regional level to evaluate policy and investment options. They also provide information to metropolitan models, and many agencies use their statewide modeling program as the impetus for developing best practices in travel modeling at all levels of geography within the state.

However, many of these advances have occurred only recently. The number of operational statewide models increased from 20 a decade ago [1] to 34 of all 50 states today, with an additional 4 states having a model under development. They are used to formulate plans and policies, evaluate and prioritize projects and programs, and to assess the economic and social impacts of major transportation investments. Typical studies conducted with them include bypasses around congested areas and assessment of improved commuter rail lines on auto dependency before costly projects are constructed.

Closely related are megaregional models that are not defined by the geography of states but rather follow major urban agglomerations. Megaregional modeling is still in its infancy, despite the revised interest in this level of analysis in the U.S. [2, 3]. Except for the Chesapeake Bay Megaregional Model [4], no megaregion in the U.S. has an operational transportation model as of 2017.

Statewide modeling is a diverse field of practice. Model frameworks reach from rather simple aggregate three-step models to quite sophisticated activity-based microscopic modeling approaches. Therefore, an overview of statewide models also describes the range of models used in practice. A survey of all 50 states and an informal survey of leading practitioners were undertaken to gather information about the variety of statewide modeling approaches used in the U.S.

The purpose of this paper is to document the current state of statewide modeling, to identify limitations in current model design and to provide ideas and trends for the future development of such models.

1.1 Rationale for statewide models

Statewide models are used to analyse the impact of policies and trends that are implemented or addressed by state governments, but not captured with urban or national models. This includes several travel markets not commonly included or underrepresented in urban models:

- Foremost, statewide models capture travel occurring in areas outside of major metropolitan areas. Smaller MPOs and agencies outside of them often do not operate their own transportation models, yet there is a need to analyse transportation planning in these areas as well. In practice, state DOTs commonly fulfil this role using statewide models.
- Metropolitan transportation models cannot capture travel between MPOs. For example, the three MPOs in Arizona – the Maricopa Association of Governments (MAG) in Phoenix, Sun Corridor Metropolitan Planning Organization in Pinal County, and Pima Association of Governments (PAG) in Tucson – share borders and have substantial travel

between them. However, their models focus on travel within each area. The Arizona statewide model is used to fully represent travel demand in the corridor between Phoenix and Tucson. Similar examples include travel between the San Diego and Los Angeles areas, or between Baltimore and Washington, D.C.

- Long-distance travel is another focus of statewide models. MPO models commonly represent travel with origins or destinations outside their MPO area as external trips. They enter or leave at external stations, which do not distinguish between final destinations two miles versus two hundred miles away. Both ends of such trips are often explicitly represented in statewide models, due to the size of the study areas for most states. There is no universal definition of long-distance travel. The National Household Travel Survey (NHTS) defines long-distance travel as trips with 50 miles or more [5], the Travel Survey of Residents of Canada (TSRC) uses the thresholds 40 km or more on non-habitual trips or stay away from home over night [6]. Long-distance travel behaviour differs substantially from short-distance travel [7, 8]. The former is infrequent, with different destinations (more concentrated in city centres of major metropolitan areas), a mix of modes different from local travel, different hours during the day (less focused on morning and evening peak hours), and different path choices. Traditionally, long-distance travellers tended to stay on highways and major arterials, for they are not as familiar with local streets. With GPS and ITS technologies, however, this pattern has been weakened.
- Freight is a key market represented in most statewide models. Many urban models capture truck flows as well, but many of them travel to or from places outside of the modeled urban area. Freight moving by rail, water, and air are irrelevant in urban models, apart from where they transfer to trucks.
- Statewide models can provide external traffic volumes for urban models. Sometimes, urban modelers prefer the use traffic counts at external stations and distribute those volumes between internal zones and other external stations using gravity models. Using a statewide model for this purpose, however, offers two advantages. First, the statewide model explicitly distinguishes between origins and destinations that are internal to the MPO region from those that are external, and thereby identify internal-to-external, external-to-internal and through trips, Secondly, a statewide model may be used to forecast traffic under different scenarios. For instance, a major highway project within an MPO region may affect the routing of long-distance trips with origins and destinations outside of the MPO.

On the other hand, statewide models have limitations. While there are no known domain reasons why statewide models cannot be used for certain analyses, insufficient spatial and temporal resolution in statewide models may be an impediment to analysing scenarios that are more commonly addressed with urban models. For example, most statewide models are too coarse to adequately represent non-motorized travel (i.e., walk or bike). The impact of transit-oriented development is another example that requires higher spatial resolution around transit stations than commonly found in statewide models. Likewise, impacts of very small projects, such as a new highway ramp, tend to be overlooked by statewide models for lack of spatial resolution.

1.2 Survey of statewide modeling practices

A detailed survey of statewide modeling practices was distributed to voting members of the AASHTO Standing Committee on Planning. It was designed as an online survey. The survey attempted to balance the desire for obtaining detailed information with the burden of completing it. A pre-test was conducted with the states of Ohio and Oregon. The final version with 27 questions was longer than initially anticipated but necessary for collecting sufficient detail on statewide modeling.

Forty-six states responded to the survey in Spring 2016. For the remaining four states, Greg Giaimo of the Ohio DOT helped answering at least some of the questions by personally contacting colleagues in these four states.

Table 1 shows the response rate.

Table 1. Survey summary

Response category	Number of states	Percent
Invited to participate in survey	50	100%
Responded to online survey	46	92%
Responded after personal contact through Ohio DOT	4	8%
Operate a statewide model	34	68%
Develop statewide model (not yet operational)	4	8%
Neither develop nor operate a statewide model	12	24%

Statewide modeling is becoming mainstream in the U.S. In 2005, 20 out of 50 U.S. states had implemented operational statewide models [1]. In 2016, 34 states operated statewide models, an increase of 70 percent over 11 years.

The distribution of states with and without models is shown in Figure 1. Not surprisingly, many states without models are predominately rural with little congestion. At the other end of the spectrum, New York State does not operate a statewide model despite being the 7th densest state in the U.S. However, New York's density is heavily concentrated in the New York City Metropolitan Area, which operates its own (urban) transportation model.

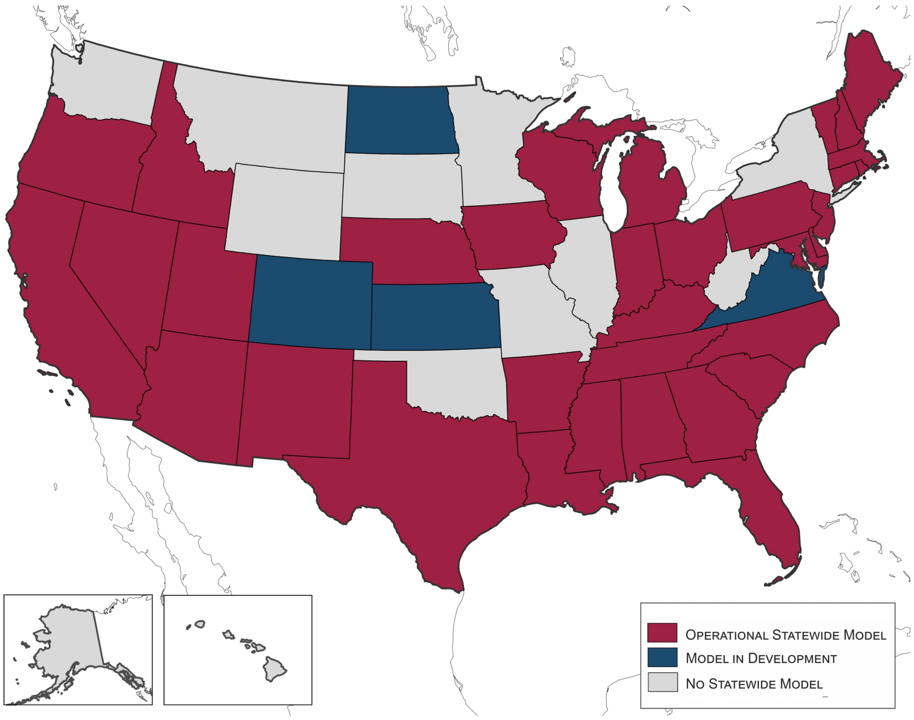

Figure 1. States with operational statewide models

The survey also revealed that many states have developed add-on models that support the transportation model. Fifteen states operate separate long-distance travel models, 26 states have a long-distance freight model, nine states model environmental impacts, and two states formally model land use changes.

2 Survey of existing practice

The following seven sections cover specific model design principles revealed in the survey.

2.1 Person travel demand modeling

The traditional four-step modeling paradigm is applied in 30 states to model person travel demand, as shown in Figure 2. This category also includes states that split traffic into various periods of time, making it a five-step model. Many states apply this four-step concept to short-distance travel only, using a different modeling approach for long-distance travel. Three states are in the process of developing a four-step transportation model, and another three have an activity-based model under development. Five states have an operational activity-based model.

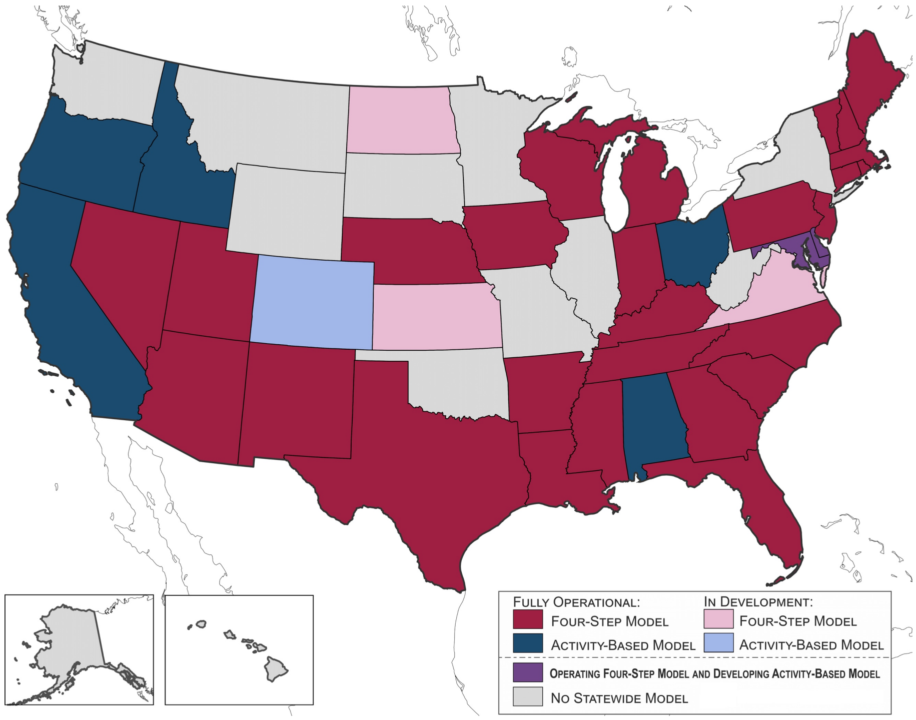

Figure 2. Status of statewide modeling in the U.S.

Hawaii as an island state does not operate a statewide model. Separate models are implemented for the islands of Kauai, Maui, Oahu and Hawaii, covering over 99 percent of the state's population. Other states without statewide models tend to be clustered towards the northern part of the country in regions with lower population densities and less severe levels of congestion, which might have affected the decision against developing a statewide model.

The frequency of trip and tour generation methods is shown in Figure 3. As expected, using trip rates based on cross-classification is the most common approach. In the early years of travel demand modeling, multiple regression was the dominant approach for generating trips. In multiple regression models, trip rates are treated as continuous rather than discrete variables, which may lead to unrealistically high (or sometimes even negative) number of trips in zones with unusual household type compositions. Therefore, the 1970s marked a shift away from aggregate zonal level regression analysis to more disaggregate household cross-classification procedures [9]. However, trip rates based on multiple regression have the advantage of allowing the analyst to consider multiple independent variables, and may work well if generated trips are reviewed carefully for inconsistencies.

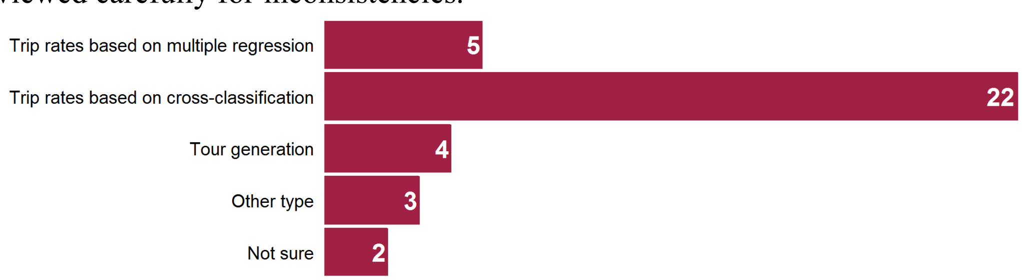

Figure 3. Frequency of trip and tour generation methods for person travel (multiple answers allowed)

Trip distribution is modeled with a gravity model in 22 states (see Figure 4). This model type is easies to implement and calibrate. However, if trips with longer distances (such as trips over 50 miles) are included, gravity models perform poorly because the tail of longer trip lengths cannot be calibrated well in a gravity model. This is less of a concern if a separate long-distance model is used (true for seven out of 22 states that implemented gravity models for short-distance travel).

The logit-based destination choice model, on the other hand, evaluates different destinations against each other, taking into account the attraction in every other zone, the distance to every other zone, and possibly other factors that may affect destination choice, such as rivers or language barriers. Eleven states apply logit-based destination choice models (Figure 4). Even though logit-based destination choice models are capable of handling long-distance trips, eight out of eleven states that use logit-based destination choice models have implemented a separate long-distance travel model. Such a combination commonly ensures the largest model sensitivities for the trip distribution step.

Figure 4. Frequency of trip distribution models for person travel (multiple answers allowed)

About every other statewide model explicitly accounts for mode selection (Figure 5). 13 states responded that they only generate auto trips, obviating the need for a mode choice model. Five states apply static (fixed) modal shares, making it 18 states (or 53 percent) that do not model mode choice. Of the 16 states that do model mode choice, the majority uses nested logit models, with only two states using simple multinomial formulations.

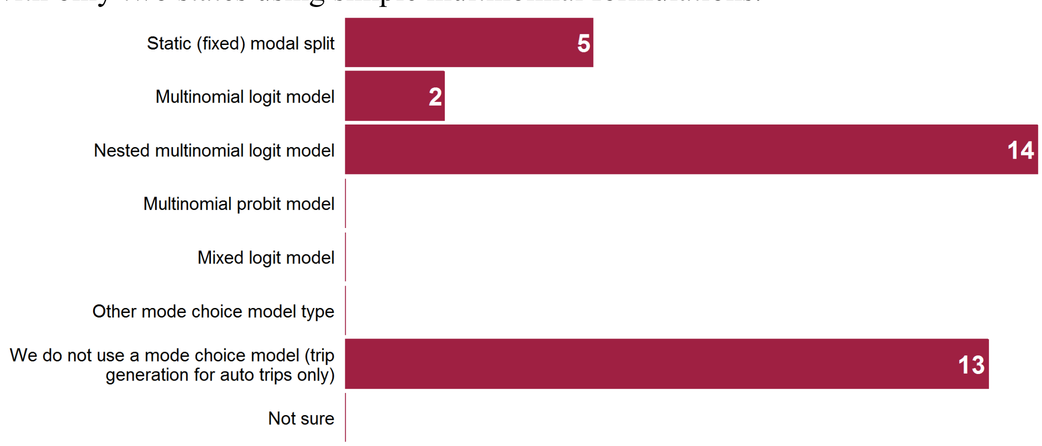

Figure 5. Frequency of mode choice models for person travel

The modes of transport represented in statewide models for person travel are shown in Figure 6. Thirteen states distinguish auto occupancy. Eight models account for non-motorized travel. Often, the resolution in statewide models is too coarse to reasonably account for this mode. However, almost a quarter of all statewide models account for these walking and biking given the rising interest in non-motorized modes as an alternative to auto travel and for population health analyses.

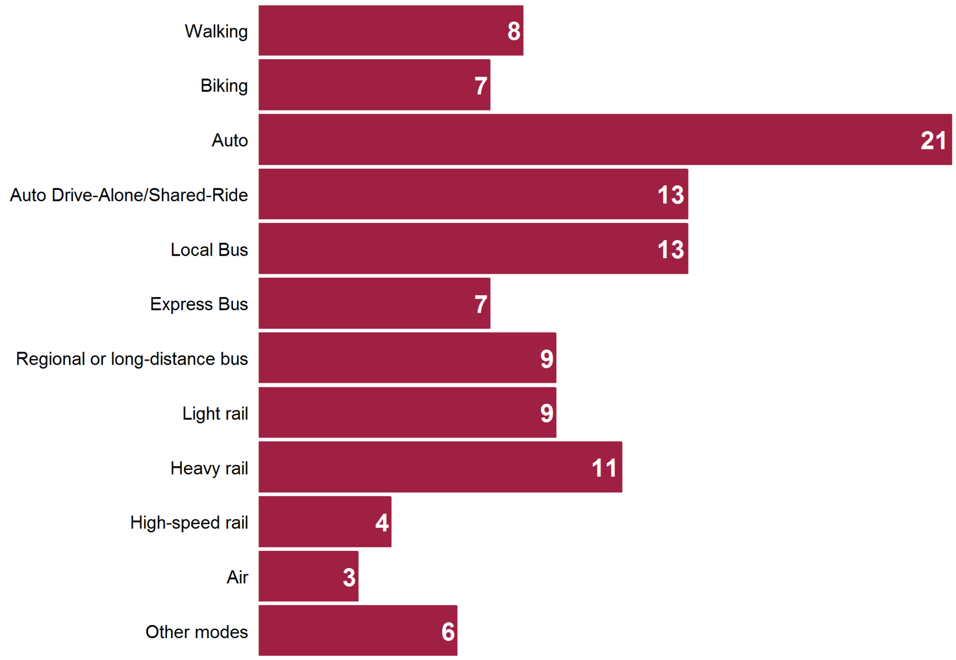

Figure 6. Modes represented in mode choice models for person travel (multiple answers allowed)

Local bus is the most frequently modeled transit mode, followed by heavy rail (which commonly includes commuter rail), regional or long-distance buses and light rail. The time of day in which travel occurs is represented in 12 out of 34 states, or 35 percent (Figure 7). Predominately rural states with very low levels of congestion may omit this step, as travel time will not differ significantly by time of day.

Figure 7. Time of day representation in statewide models

The number of time-of-day periods distinguished by individual models is shown in Figure 8. Most models deal with four time periods, usually defined as AM Peak, Midday, PM Peak and Night. In general, a more fine-grained resolution of time is desirable. This will enable a model to represent better the time-dependent effects of congestion, which some travellers will attempt to avoid by traveling before or after those periods. However, more time intervals increase the computational burden and the need to model departure time shifts.

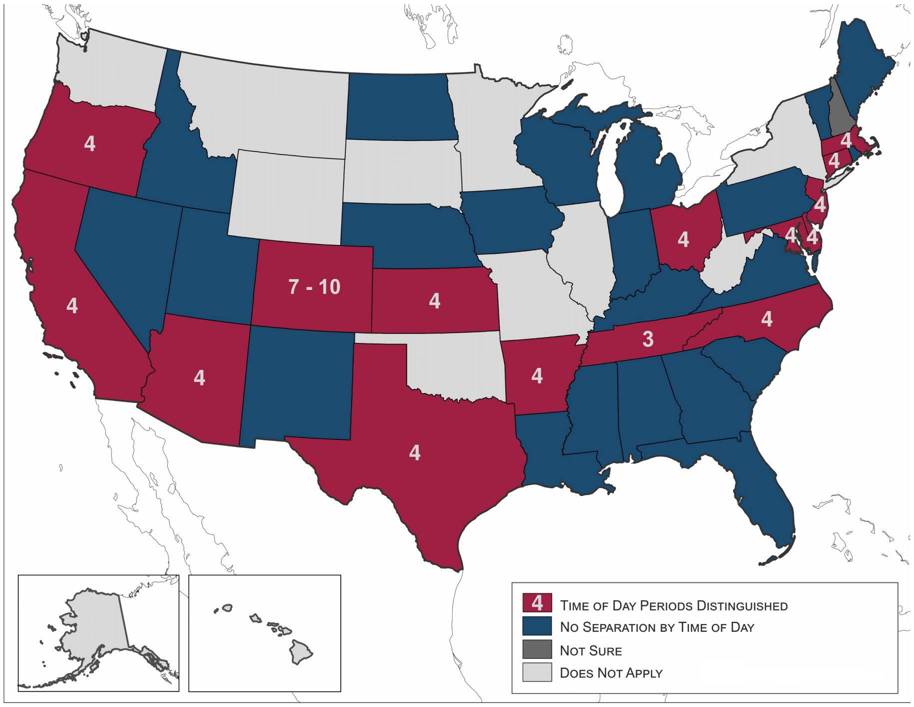

Figure 8. Number of time periods distinguished by individual statewide models

Four out of five statewide models use the traditional static user equilibrium algorithm for the assignment of highway travel (Figure 9). Three mostly rural states apply the all-or-nothing assignment, where congested travel times do not differ much from free-flow travel times.

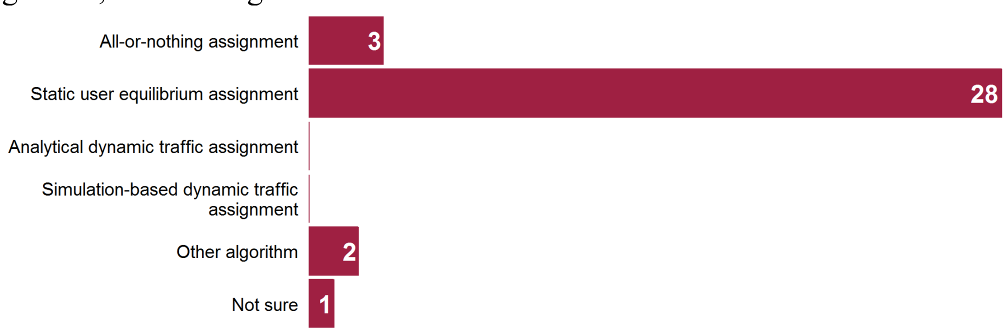

Figure 9. Frequency of assignment algorithms in statewide models

Some models provide feedback from the assignment back to previous steps of the model. For example, under congested conditions some travellers may choose other destinations or other modes. By feeding back travel times to previous steps, an equilibrium between different submodules and congested travel times may be reached. The concept is shown for traditional four-step models in Figure 10.

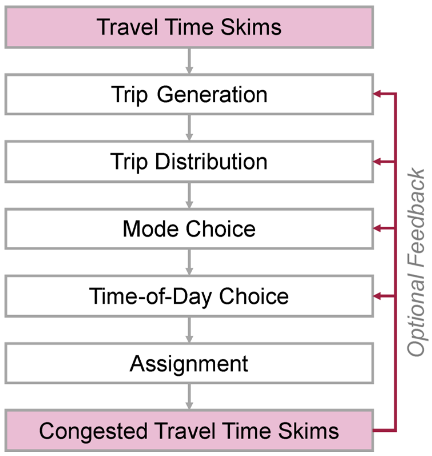

Figure 10. Trip-based model feedback process

Figure 11 shows how many statewide models apply feedback. Congested travel times are fed back into the trip distribution step in 20 out of 34 models (59 percent). As the mode choice model is run after the trip distribution model, congested travel times may also affect mode choice in those models. Six models feed congested travel times back to trip generation.

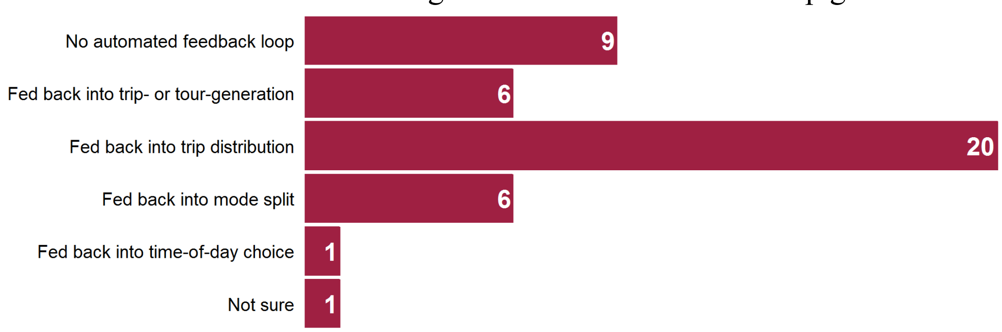

Figure 11. Frequency of feedback of congested travel times (multiple answers allowed)

2.2 Person long-distance travel

15 states (or 44 percent) have implemented explicit long-distance person travel demand models (Figure 12). Arizona and North Carolina exclude long-distance commute trips (which are handled by the short-distance model because they are unlike most other habitual long-distance trips). For Alabama, long-distance trips are those that either cross the state boundary or travel across more than one MPO boundary, and for Colorado, trips that cross the state boundary are handled separately. Iowa is the only state that defines long-distance trips by travel time, namely greater than 60 minutes.

Figure 12. Frequency of explicit long-distance models for person travel

Long-distance models tend to be implemented in larger states by area (Figure 13). Some states, such as Georgia, capture long-distance travel with a separate trip purpose in the short-distance travel model.

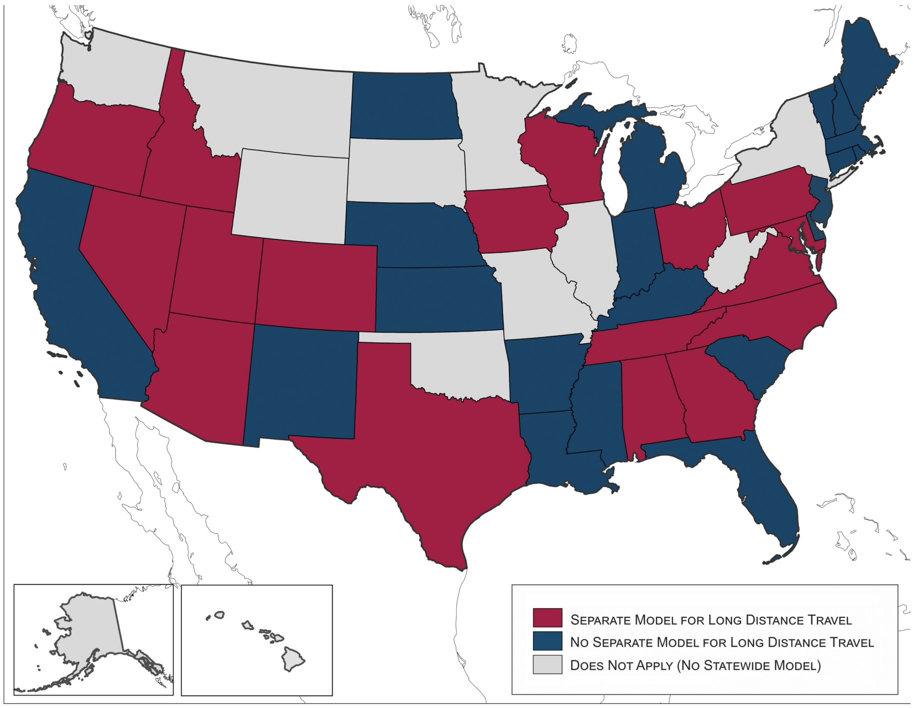

Figure 13. States that operate separate long-distance models for person travel

A wide variety of sources are used for trip generation of long-distance trips (Figure 14). Iowa is currently the only state that uses FHWA's national long-distance person model [10] and long-distance trip rates provided in NCHRP Report 735 [11]. Arizona, Maryland, and North Carolina use the long-distance model NELDT [12], which is based on trip frequencies reported in the long-distance element of the 2001 NHTS [5].

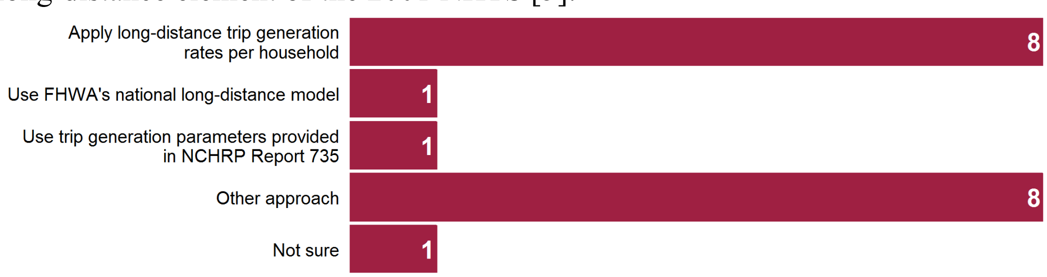

Figure 14. Travel demand generation rates for person long-distance travel (multiple answers allowed)

Most long-distance models use traditional gravity models for trip distribution (Figure 15). Five states use advanced logit-based destination choice models, making the share of one-third of statewide models similar to the pattern found for short-distance travel models.

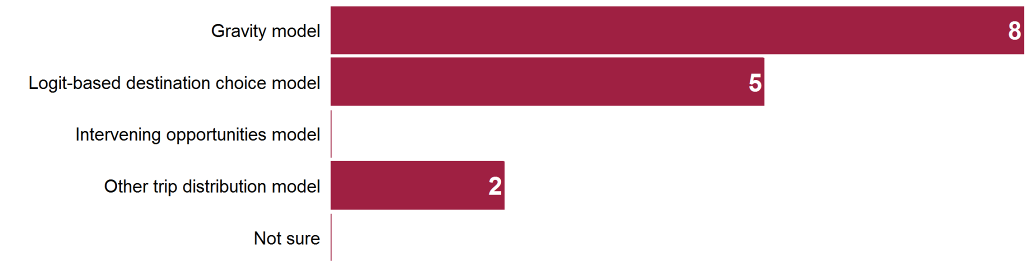

Figure 15. Frequency of trip distribution models for long-distance person travel

About half of all long-distance models apply nested multinomial mode choice models (Figure 16). Wisconsin uses a multinomial model, and Utah applies static mode shares (consistent with their short-distance mode choice model). Alabama, Arizona, Maryland and Nevada generate long-distance trips for autos only.

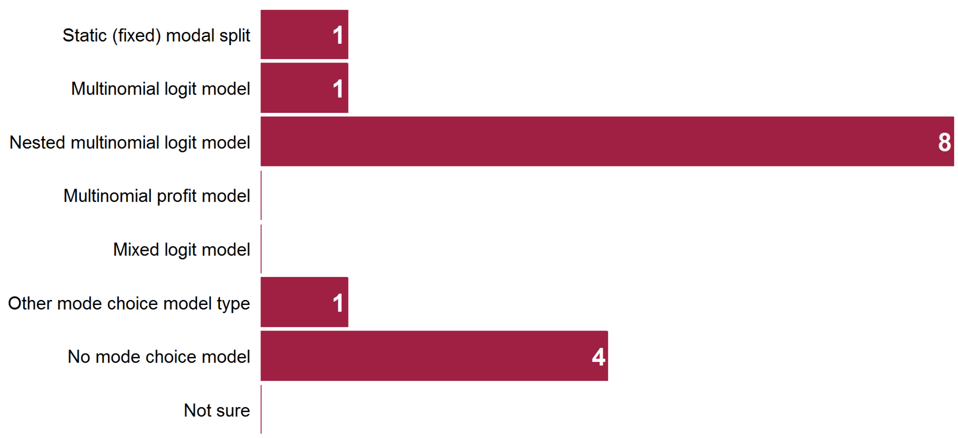

Figure 16. Frequency of mode choice models for long-distance travel

The modes of transportation represented by long-distance travel models are shown in Figure 17. Obviously, non-motorized travel is not modeled for long-distance travel. All states model the auto mode, and five models distinguish drive-alone from shared-ride. Many include bus, rail, and air.

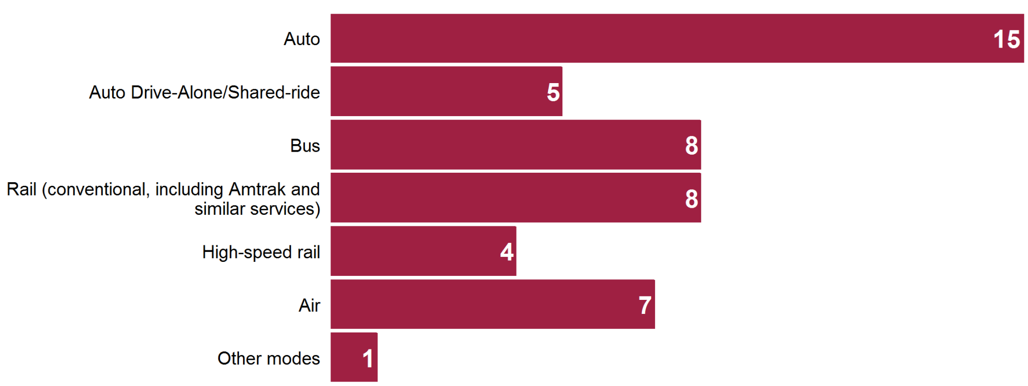

Figure 17. Modes represented in long-distance person mode choice models (multiple answers allowed)

2.3 Freight models

Growth in freight transportation is expected to significantly outpace growth in passenger transportation [13]. Also, easing freight travel has become a mantra for economic development [14]. The ratio between freight-miles travelled and the Gross Domestic Product, also known as the freight-transportation intensity, shows a strong (yet gradually declining) relationship between freight activity and economic growth.

Given their disproportional impact on the transportation system [15], it is not surprising that most statewide models account for freight modeling (Figure 18), particularly in areas with high levels of congestion. As freight tends to make up a higher share of traffic on rural roads, statewide models tend to have a larger share of freight traffic than urban models. Therefore, statewide models tend to pay more attention to freight flows, often distinguishing short- and long-distance freight flows. While short-distance trucks are covered by 21 states (62 percent of all states with statewide models), long-distance trucks are modeled by 26 states (76 percent).

Figure 18. Frequency of freight models in statewide models

Of the 21 states that model short-distance trucks, 19 use trip-based models, and only Ohio and Oregon use tour-based truck models. The limitations of trip-based truck models have been discussed in the literature [16], yet tour-based models are uncommon in statewide models. The heterogeneous travel behaviour of trucks (depending, among other factors, on truck type and commodities carried) and the limited freight data availability (much more so than for auto travel) make it inherently challenging to represent tour-based travel behaviour for trucks at the statewide level.

The spatial distribution of long-distance freight models is shown in Figure 19. Freight modeling appears to be less common in states in the northern parts of the U.S. The Interstate 10 corridor and possibly the I-65 corridor are the only ones that are covered completely by statewide truck models. Several states in the Midwest and New England have not tackled freight flow models yet. Given the especially large volumes of long-distance truck flows on east-west highway corridors, many states could benefit from explicitly modeling them.

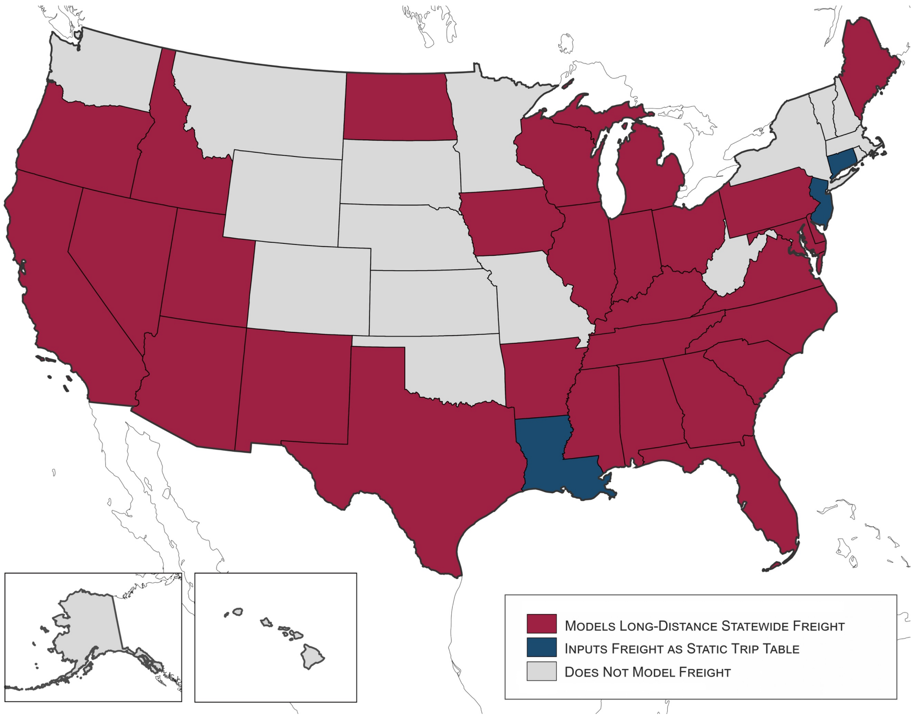

Figure 19. States operating long-distance freight models

Long-distance truck modeling is dominated by commodity flow models (Figure 20). Most of the respondents who reported using commodity flow models in the survey reported that they are based, at least in part, upon origin-destination freight flow data from the Freight Analysis Framework (FAF).

Figure 20. Frequency of long-distance freight modeling methods (multiple answers allowed)

A growing number of states apply mode choice models to freight flows as well (Figure 21). Out of 26 states that model long-distance freight flows, six states (23 percent) apply rule-based freight mode choice models. Such models do not attempt to econometrically estimate mode shares, but rather apply simple rules of modal allocation that can be reviewed and changed. Logit-based freight mode choice models were implemented by Florida, Georgia, Illinois, Ohio, Oregon, Texas, and Virginia. Many of these logit-based models are designed as so-called freight diversion models (i.e., they model the shift from one mode, such as truck, to another mode, such as rail). Starting with the observed mode share and modeling only the potential shift from one mode to another is a powerful way to deal with data limitations in freight modeling while maintaining some freight mode sensitivities to policy scenarios.

Figure 21. Frequency of freight mode choice models (multiple answers allowed)

Of the 11 statewide models that represent freight mode choice, all include truck and rail as modal options. Figure 22 shows the modes modeled.

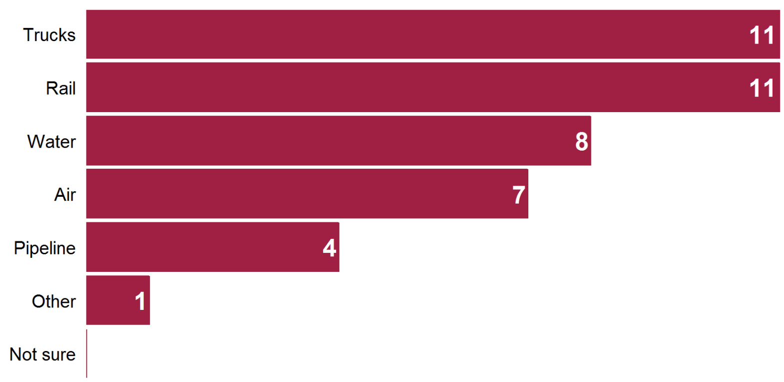

Figure 22. Modes represented in long-distance freight mode choice models (multiple answers allowed)

2.4 Economic models

Traditionally, statewide transportation models worked with static socio-economic input data. Given the large uncertainty of economic forecasts, many states have moved towards integrating their transportation model with economic forecast models (Figure 23). Externally prepared commercial forecasts are the most frequent source of future socio-economic data, closely followed by forecasts developed by other state agencies.

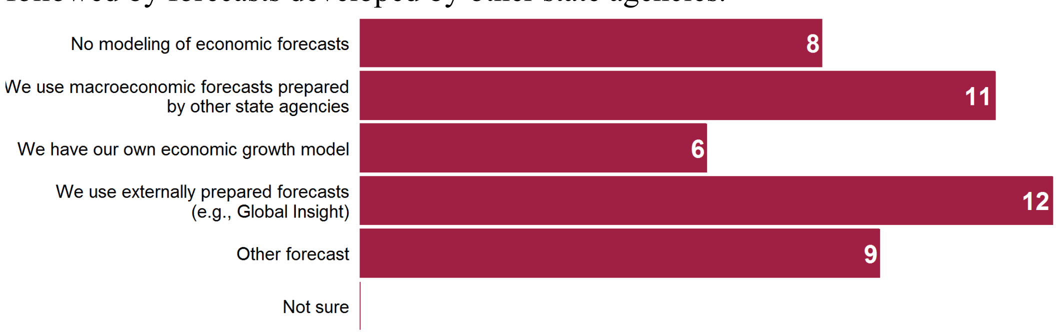

Figure 23. Frequency of economic forecast models (multiple answers allowed)

The distribution of base and future model years are visualized in Figure 24. Base years cluster around 2010, as expected, while future years dominate in 2015, 2020, 2030, 2035 and particularly in 2040. Beyond that, New York models 2044, California 2050 and Nevada 2060. Several models can provide forecasts for any future year within their model time frame. This is achieved by interpolating between five or 10-year model runs. While this approach assumes a somewhat artificial linear growth between two modeled years, interpolation provides additional data for years the model cannot be run.

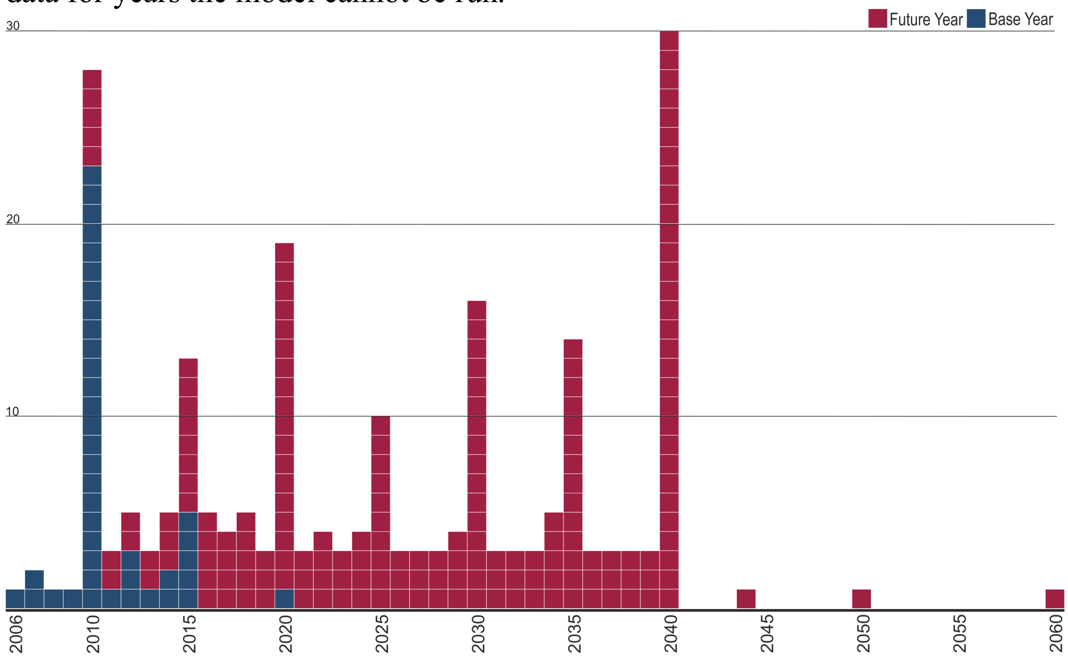

Figure 24. Distribution of base and future years in statewide models

2.5 Land use models

Land use models can be integrated with travel demand models to reflect the interactions between the transportation system and land use development. The integration of land use with transportation models has proven to improve model sensitivities in scenario analyses [17]. For example, if the model is used to test the expansion of a rail line, households may decide to relocate because the rail line may make certain neighbourhoods more attractive. As another example, if congestion increases substantially, urban sprawl might be slowed down. Only two states, Ohio and Oregon, have operational land use models at the statewide level Figure 25. Nevada and Indiana are currently developing land use models. While at least three more states (California, Florida and Maryland) have operational land use models as well, they have not been integrated with the official version of the statewide transportation model.

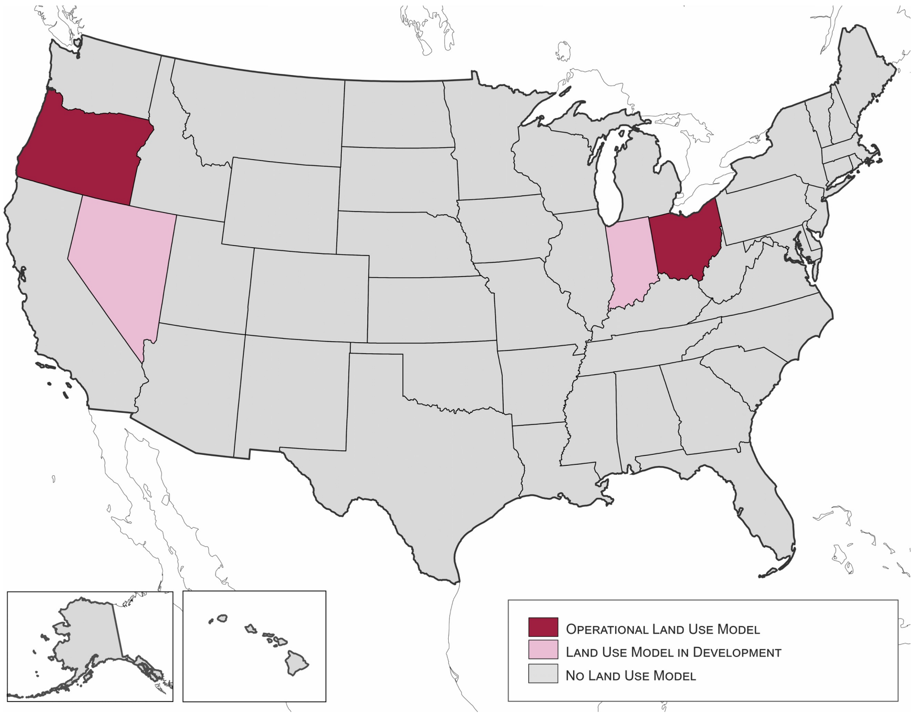

Figure 25. Distribution of states with land use models

2.6 Environmental impact models

Nine states explicitly model the environmental impacts of traffic flows, as shown in Figure 26. All cases reported referred only to air quality.

Figure 26. Frequency of environmental impact modeling within statewide models

The spatial distribution of states with environmental impact models is shown in Figure 27. It is notable that all West Coast states (of the lower 48 states) model environmental impacts. Two Southern states and Michigan also model environmental impacts, and a New England cluster can be seen as well.

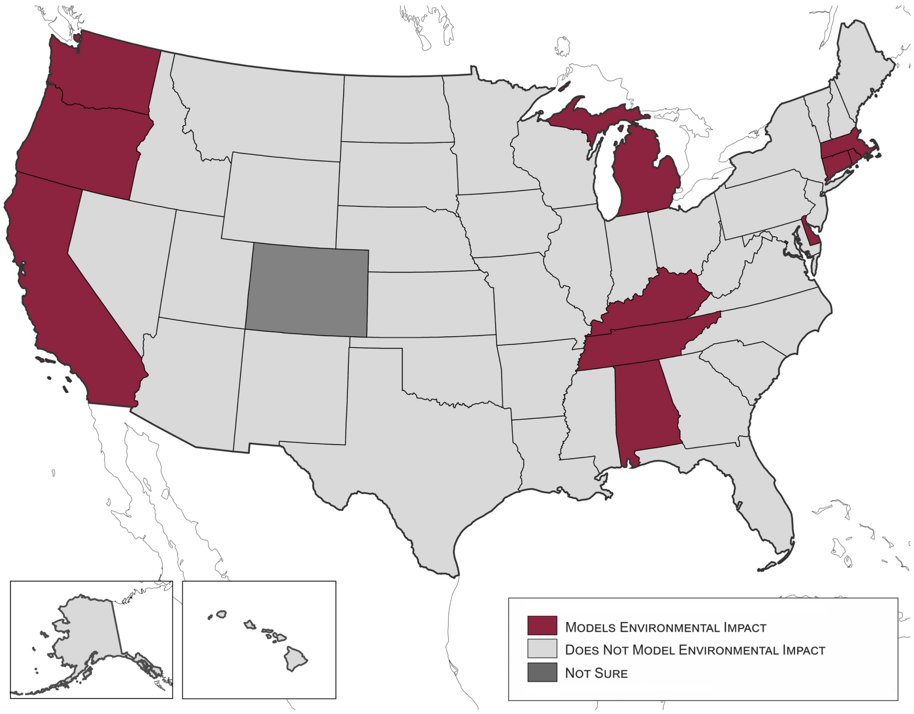

Figure 27. Distribution of states that model environmental impacts

Most states that model environmental impacts use the MOVES model (Figure 28). Oregon uses MOVES in combination with their own greenhouse gas model (GreenStep), and Kentucky uses both MOVES and its predecessor, MOBILE. EMFAC is used in California only, based upon emission rates provided by California EPA's Air Resources Board.

Figure 28. Frequency of environmental impacts models (multiple answers allowed)

The types of emissions covered are listed in Figure 29. CO_2, NO_x and PM are the most common emissions modeled. Oregon and Washington are the only two states that calculate noise emissions, a significant factor that impacts human health and well-being [18].

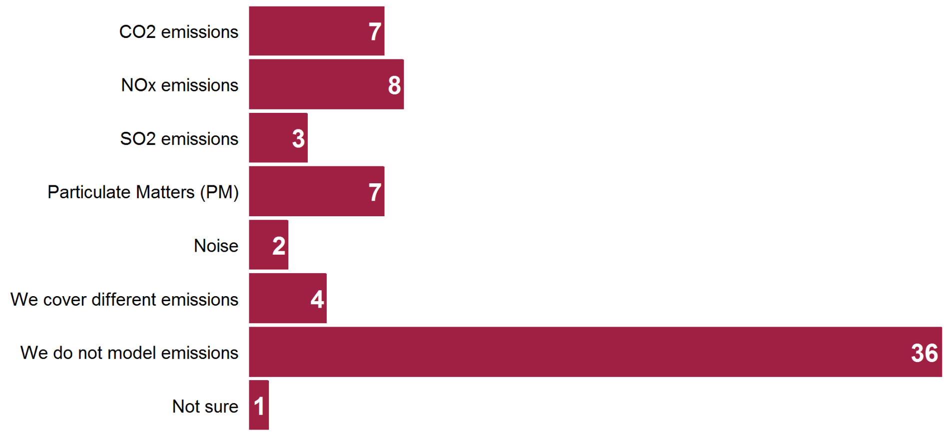

Figure 29. Emissions modeled with statewide models (multiple answers allowed)

2.7 Resources

Agencies that operate statewide models were asked about the resources they have invested in model development and application. The first question asked for the number of full-time equivalent employees (

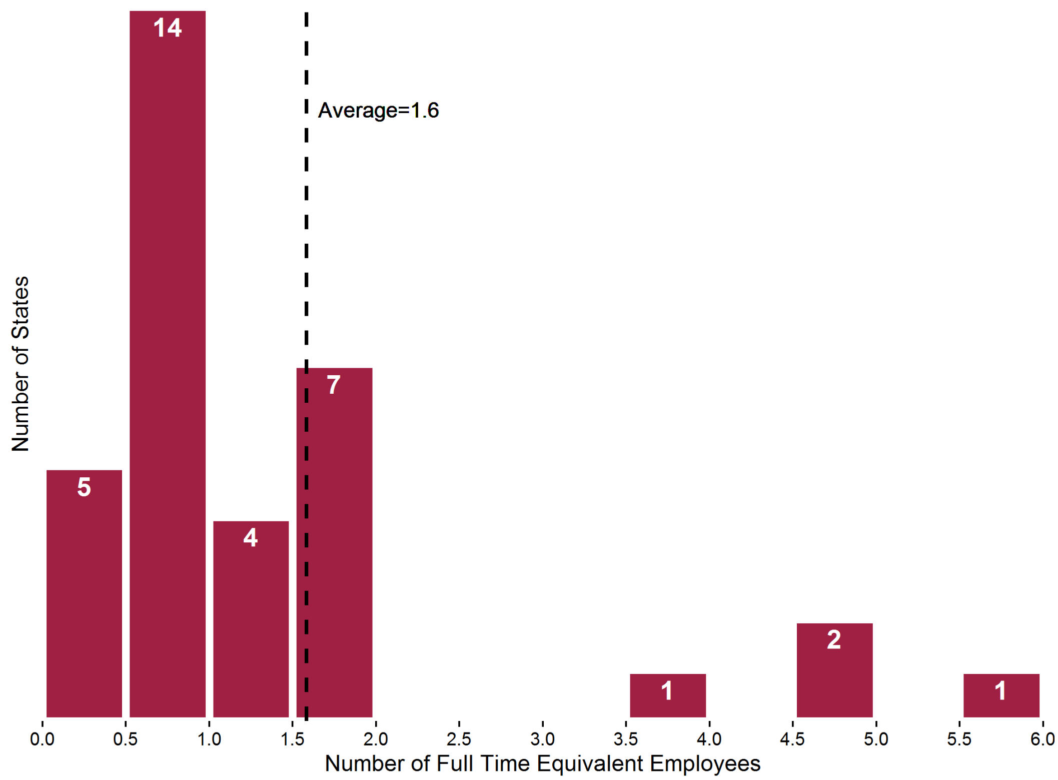

Figure 30).

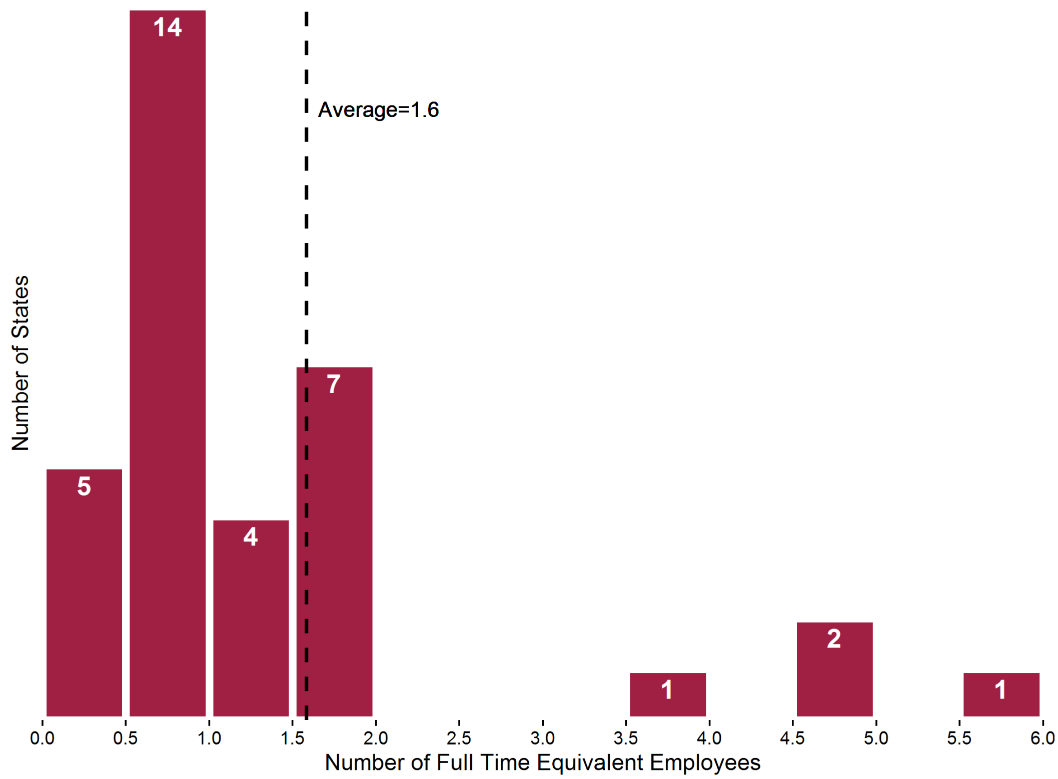

Figure 30: Number of full-time equivalent employees working on statewide modeling

For model development, a relatively large share of resources (71 percent) were allocated to consultants on average, plus another eight percent being allocated to universities (Figure 31). Only 20 percent of the resources were invested in-house or for partner agencies. On the one hand, this means that a lot of expertise in model development is found outside the state agency. On the other hand, it might be considered neither cost efficient nor practical to train staff to build a model, a task faced by the agency maybe every 10 to 20 years.

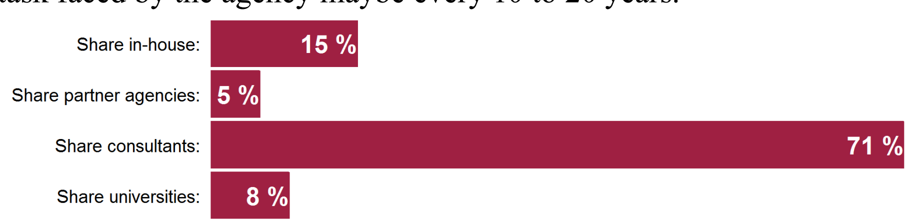

Figure 31: Resource allocation for model development

For model application, the percentages were almost completely reversed (Figure 32). In-house and partner agencies on the average conduct 60 percent of the model application work. Compared to model development, the share for consultants and universities drops in half.

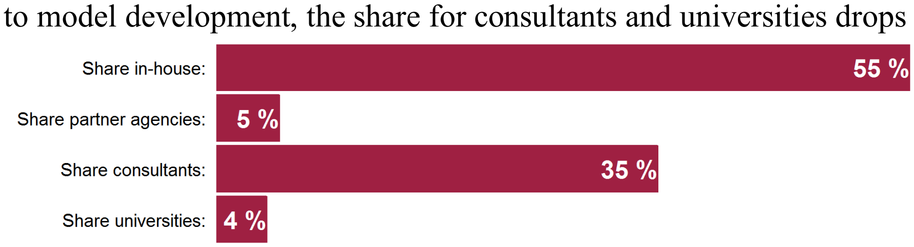

Figure 32: Resource allocation for model application

Finally, the questionnaire asked how much money was invested into the model over the past several years. The estimate does not include costs for staff within the agency, but only expenditures for data purposes, software licenses, and outside help. On the average, agencies spent $700,000 on statewide modeling in the last year. However, last year's expenditures were highly skewed by one agency that reported spending $11 million. The standard deviation for this average is $2 million, almost three times the average. Removing this one outlier reduces last year's average expenditure to $340,000, which appears to represent the average better.

2.8 Scenario analysis

The survey asked what kind of scenarios were evaluated using statewide models. Respondents could select more than one answer, which is why the total of all responses is much larger than the number of states that operate statewide models. A summary of the common scenarios is shown in Figure 33. Light bars indicate scenarios described as most important.

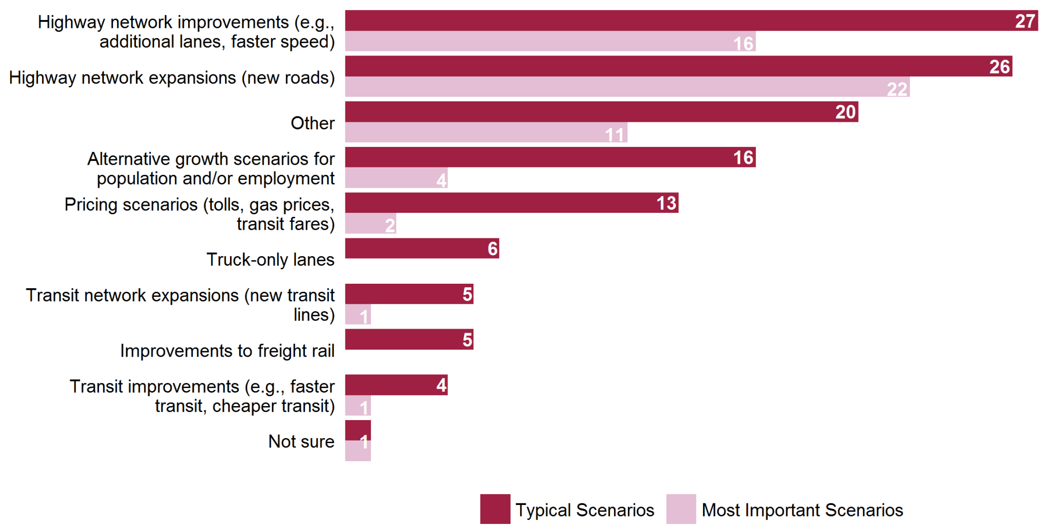

Figure 33 Typical scenarios tested with statewide models (multiple answers allowed)

By far the most common application of statewide models involved highway network improvements (79% of all states operating models). Analysing the impacts of highway expansions was a close second (76%). It is not surprising that highway analysis far exceed the interest in transit expansions (15%), as state and local governments spend six-times as much on highways as they spend on transit [19]. Another important scenario type was testing alternative growth rates (47%) This finding underlines the large uncertainties associated with population and employment forecasts. The interest in testing alternative pricing structures is on the rise, as lack of alternative forms of funding pushes transportation agencies to charge for the use of infrastructure [20].

2.9 Critical review of the survey methodology

Best practices of survey research [21] were applied when conducting the survey on statewide modeling among all U.S. states. The questionnaire was developed and revised based on comments from the review panel overseeing this synthesis report. The online survey tool was reviewed and refined by four different scientists. A pretest was conducted that helped fine-tune contents. Sensitive questions (on staffing and budget) were asked towards the end. The survey request was sent out by TRB with an email explaining the relevance of the study. Contact information for questions was provided, and late respondents were reminded several times by email and telephone. Nevertheless, survey results need to be interpreted with some caution.

Some survey responses were inconsistent. They were reviewed carefully, and corrected if it was obvious what was intended. For example, a few states reported that they do not model environmental impacts, yet did report modeled types of environmental emissions. It could not be determined from the responses whether those states conducted environmental modeling or not. In at least one case it was found that emission estimates from the traffic assignment model were compiled in a manual post-processor rather than using a separate emissions model. Several phone calls and follow-up emails were necessary to disentangle inconsistencies. Future studies should consider setting aside sufficient time and resources to conduct phone interviews instead of online surveys with every state to avoid such inconsistencies.

2.10 Summary findings of the survey

Despite some shortcomings, the survey provided intriguing findings on statewide modeling in the U.S. It is remarkable how statewide modeling has become a standard practice in most states. Given the complexity of the transportation system and the intricacy of policy questions posed by decision makers today, transportation planning agencies cannot continue to rely on intuition and experience alone. Most states make heavy use of statewide models, some of them quite sophisticated, to support decision making in transportation planning.

At the same time, it became obvious that urban models tend to be more advanced than statewide models. When comparing the 34 operational statewide models with the 34 largest urban models, the latter show substantially more complexity and rigor [22]. For example, only five statewide models reported using a tour-based approach, while more than a dozen urban models do so [23]. There are five statewide models still using multiple regression for trip generation, a concept that has mostly disappeared from urban models. While urban models without mode choice models have become rare, six statewide models use static mode shares, and another 12 ignore different modes of transport entirely. Most remarkable is the fact that 20 out of 34 models do not distinguish time of day, but rather generate daily traffic. Reasonable estimations of congestion are very challenging without distinguishing at least between peak and off-peak travel conditions.

However, the fact that statewide models tend to be simpler is not a critique per se. Simpler models may well can answer questions asked in each state. If two models could answer the questions at hand, the simpler model should always be preferred as it limits the risk of model inconsistencies. Moreover, the temporal, spatial, and behavioral resolutions found in many state-of-the-art urban models would be prohibitively costly if extended to cover an entire state.

Efforts to integrate several statewide models with analytics from other related domains are particularly worth noting. Many statewide models operate separate models for person long-distance travel, short-distance truck travel, long-distance freight flows, freight mode choice, economic forecasting model, or a land use model. Statewide models tend to be more advanced than the average urban model in terms of these interdisciplinary modeling approaches.

3 Fundamental Challenges

The largest reported challenges included:

- The lack of precision in the model for detailed corridor studies was often cited. Stated differently, the same levels of spatial, temporal, and behavioural resolution found in urban models were desired at the statewide level. However, stretching an urban model to cover an entire state is impractical and too costly to build and maintain in most cases.
- The lack of data on visitors and their travel patterns was often cited as a significant limitation for some project or corridor studies.
- Too much time went into the development of some models, and too little into the user interface and visualization of model results.
- Coding networks, especially for future years, is very labour-intensive and time-consuming. This was particularly an issue when analysts were asked to include all the capacity or operational improvements across a larger state. The use of Open Street Map has turned into an important facilitator for base year networks [24]

The three most common emerging issues identified revolved around the expected benefits of big data, development of multimodal project evaluation or cost-benefit analysis processes, and the likely effect of autonomous and connected vehicles. It is hoped that big data, in the form of passively collected origin-destination patterns and travel times from cellular devices, will fill data gaps at an affordable price point, particularly for freight and commercial vehicle flows. Some viewed this as adding value to traditional revealed and stated preference surveys, while others felt that it opens the door to new modeling approaches.

Some states use a formal economic impact framework, while others are still working on potential approaches. A recent survey of such models by Holian and McLaughlin [25] suggests the use of ad hoc methods until better solutions become available.

Modeling the effects of autonomous and connected vehicles is a hot topic, for which the respondents felt unprepared to address. How to formulate forecasts for a society where such vehicles dominate, and travel choices change in response to them, remain open questions. Issac [26] summarizes the likely policy impacts, as well as describing a range of different futures, depending on how society, laws, markets, and governments respond to this new mode of transportation. Beyond her work, however, there appears to be little to guide planners seeking to understand the likely impacts of autonomous vehicles.

The largest challenges cited included the difficulty in attracting and retaining well-qualified staff, sporadic or uneven funding for statewide models, the difficulty of integrating or reconciling them with urban models, and the burden of obtaining information about regional or national travel affecting their state. All but one of the respondents felt that the staffing issue was one of the biggest challenges associated with statewide modeling. Often, modelers are classified in pay grades too low to attract top-qualified applicants, forcing agencies to rely upon consultants, contract employees, or sometimes partnerships with MPOs, to gain access to the talent required to build and use such models.

4 Conclusions

A small group of modelers met in Florida in late 2004 to share experiences about a dozen statewide models (Giaimo & Schiffer 2005). While they did not review every statewide model that existed then, those discussed were the most actively used. Today most states have a statewide model, and several more are in development. A significant amount of applied research has been carried out, a vibrant TRB subcommittee devoted to the topic has emerged, and a great deal of data have become available that reduce the burden of building and using such models. A definitive practice has yet to evolve, but this more reflects the diversity of uses and scale at which such models are applied, rather than lack of scientific consensus or broad experience with them. In a sense, the diversity of these models has been their strength, as it has allowed different approaches to be tested in practice, which in turn has guided the development of new data, methods, and platforms for statewide modeling.

The diversity of approaches described in this paper also poses challenges, particularly for states uncertain about next steps they should take, or the costs and benefits associated with doing so. Data on long-distance travel remain almost non-existent, precluding the ability to model such flows with confidence or understand their impact upon state and local transportation systems. This is especially so with freight and commercial vehicle data. Thus, progress with such models at all levels has lagged that of person travel demand modeling and forecasting. There is a significant gap between simple freight models based upon borrowed or synthetic data and more sophisticated models based upon robust local data. A major investment is required to bridge the gap, even with the greatly increased availability of federal commodity flow data and experiences with advanced freight models elsewhere. It is further remarkable that only 53% of all statewide models include a mode choice model for person travel. On the one hand, statewide models are less concerned about travel behaviour in urban areas where usually urban models (with mode choice) exist. On the other hand, almost all transportation projects and policies have some influence on mode choice, which is not picked up by models with a static mode share.

It is encouraging to note that considerable progress has been made with statewide models despite these limitations. A wide variety of uses were reported, with most used to evaluate travel between or outside of urban areas, long-distance travel, freight, and to better represent visitor and external travel in urban models. Many early statewide models were developed to address one or two issues. They have scaled to address a wider range of them, and emerging methods and data will enhance both the range of issues that can be evaluated with such models, as well as the levels of resolution they can do so at.

How statewide models advance over the next decade will be shaped by decisions made outside of the states. Influential will be decisions about investments in national person and freight travel data and models made by the federal government. A national model of person long-distance travel has been designed for FHWA [27]. A comparable behavioral freight model is still in the exploratory research phase, and further away from development. The availability of either of these models would significantly improve the data available to statewide modelers, and if designed to accommodate such, policy-sensitive models that significantly augment, if not replace altogether, such models that each state currently develops separately. How states, working singly or together, can develop interim solutions that anticipate the capabilities of national models, remains an unexplored but fertile avenue for development.

Many states have increased their reliance on consultants to fill the in-house staffing gap, with some even preferring outsourcing due to shrinking budgets and positions. In the short term, this arguably benefits the agency, for they gain instant access to expertise that would take years to develop themselves, and can quickly match staffing to current requirements. However, such practices are counter-productive in the long term, for much of the expertise and staff bandwidth required to use and maintain the models are lost when the contract ends or lapses. Even the best data and models cannot overcome the loss or lack of staff with the knowledge and experience required to creatively and competently apply them. To the extent that this problem lingers, it represents as big an impediment to further progress with statewide models as the other factors cited above.

None of these issues are insurmountable, and some perhaps easily tackled. It is useful to consider the challenges faced a decade ago, and the progress made since. Miller [28] wrote a particularly cogent critique of intercity travel models. His review focused upon several intercity rail forecasts under review at that time. But the weaknesses he exposed applied equally well to statewide models. His call for improved data, increased model disaggregation, deeper consideration of access and egress modes for intercity travel, and advancements in model structure and specification are all better addressed in most statewide models today, and much more so in those used to evaluate proposed high-speed rail systems. Some models have been rebuilt from scratch, as he suggested, while others have improved significantly in their ability to meet the information needs of their sponsors. With big data poised to fill long-standing data gaps and machine-learning algorithms promising to analyse such large datasets efficiently, statewide models are expected to undergo further fundamental improvements in the years to come.

Acknowledgements

The research was funded by the National Academies of Sciences under the NCHRP program. The research was also supported by the Technische Universität München – Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement n° 291763.

Author contribution statement

The authors confirm contribution to the paper as follows: study conception and design: R. Moeckel and R. Donnelly; data collection: J. Ji; analysis, interpretation of results: R. Moeckel, R. Donnelly, J. Ji; draft manuscript preparation: R. Moeckel, R. Donnelly. All authors reviewed the results and approved the final version of the manuscript.

6 References

1. Horowitz, A., *Statewide Travel Forecasting Models. A Synthesis of Highway Practice*, in *National Cooperative Highway Research Program*, T.R. Board, Editor. 2006: Washington, D.C. p. 125.
2. Amekudzi, A., L. Thomas-Mobley, and C. Ross, *Transportation Planning and Infrastructure Delivery in Major Cities and Megacities.* Transportation Research Record: Journal of the Transportation Research Board, 2007. **1997**: p. 17-23.
3. Florida, R., T. Gulden, and C. Mellander, *The rise of the mega-region.* Cambridge Journal of Regions, Economy and Society, 2008. **1**: p. 459-476.
4. Moeckel, R., et al., *Modeling complex Megaregion systems: Horizontal and Vertical Integration for a Megaregion Model.* International Journal of Transportation, 2015. **3**(1): p. 69-90.
5. Federal Highway Administration, *2001 National Household Travel Survey. User's Guide. January 2004 (Version 3)*. 2004.

6. Statistics Canada, *Microdata User Guide. Travel Survey of Residents of Canada*. 2011.
7. Erhardt, G.D., et al. *The Ohio Long Distance Travel Model*. in *Transportation Research Board*. 2007. Washington, DC.
8. Moeckel, R., R. Fussell, and R. Donnelly, *Mode choice modeling for long-distance travel.* Transportation Letters, 2015. **7**(1): p. 35-46.
9. Ortúzar, J.d.D. and L.G. Willumsen, *Modelling Transport*. 4th ed. 2011, Chichester: John Wiley & Sons. 375.
10. Outwater, M.L., et al., *A tour-based national model system to forecast long-distance passenger travel in the United States*, in *Annual Meeting of the Transportation Research Board*. 2014: Washington, D.C.
11. Schiffer, R.G., *Long-Distance and Rural Travel Transferable Parameters for Statewide Travel Forecasting Models*, in *National Cooperative Highway Research Program*, T.R. Board, Editor. 2012: Washington, D.C. p. 144.
12. Moeckel, R. and R. Donnelly. *Nationwide Estimate of Long-Distance Travel (NELDT). Generating External Trips for Local Travel Demand Models*. in *Annual Meeting of the Transportation Research Forum*. 2011. Long Beach, CA.
13. Chow, J.Y.J., C.H. Yang, and A.C. Regan, *State-of-the art of freight forecast modeling: lessons learned and the road ahead.* Transportation, 2010. **37**(6): p. 1011-1030.
14. McKinnon, A., *Life without trucks: The impact of a temporary disruption of road freight transport on a national economy.* Journal of Business Logistics, 2006. **27**(2): p. 227-250.
15. Small, K.A., C. Winston, and C.A. Evans, *Road Work. A New Highway Pricing and Investment Policy*. 1989, Washington, D.C.: The Brookings Institution. 128.
16. Holguín-Veras, J., et al., *Urban Freight Tour Models: State of the Art and Practice*, in *Freight Transport Modelling*, M. Ben-Akiva, H. Meersman, and E. van de Voorde, Editors. 2013, Emerald Group: Bingley, UK. p. 335-351.
17. Conder, S. and K. Lawton, *Alternative Futures for Integrated Transportation and Land-Use Models Contrasted with Trend-Delphi Models.* Transportation Research Record: Journal of the Transportation Research Board, 2002. **1805**: p. 99-107.
18. De Coensel, B., et al., *The influence of traffic flow dynamics on urban soundscapes.* Applied Acoustics, 2005. **66**(2): p. 175-194.
19. USDOT, *New Department of Transportation Report on Highway, Transit Conditions Points to Need for More Investment.* . 2014, United States Department of Transportation.
20. Perez, B.G., T. Batac, and P. Vovsha, *Decision-Making Framework*, in *Assessing Highway Tolling and Pricing Options and Impacts*. 2012, Transportation Research Board: Washington, D.C.
21. Babbie, E., *The Practice of Social Research*. 13 ed. 2013, Belmont, Calif.: Wadsworth Cengage Learning. 584.
22. Donnelly, R., et al., *Advanced Practices in Travel Forecasting*, in *National Cooperative Highway Research Program*, Transportation Research Board, Editor. 2010: Washington, D.C. p. 90.
23. Bowman, J.L., *Historical development of activity based model theory and practice.* Traffic Engineering and Control, 2009. **50**(2): p. 59-62.
24. Sall, E., et al. *Making Open Transportation Data Useful and Accessible: Recommendations for Good Practices in Open Data Standards Management*. in *96th Annual Meeting of the Transportation Research Board*. 2017. Washington, D.C.

25. Holian, M. and R. McLaughlin, *Benefit-cost analysis for transportation planning and public policy: towards multimodal demand modeling*. 2016, California Department of Transportation.
26. Issac, L., *Driving towards driverless: a guide for government agencies.* . 2016, WSP - Parsons Brinckerhoff.
27. Bradley, M., M.L. Outwater, and N. Ferdous, *Implementation of a Practical Model System to Predict Long-Distance Travel for the Entire U.S. Population.* Transportation Research Record: Journal of the Transportation Research Board, 2016. **2563**: p. 9-17.
28. Miller, E.J., *The Trouble with Intercity Travel Demand Models.* Transportation Research Record: Journal of the Transportation Research Board, 2004. **1895**: p. 94-101.