Traffic assignment for strategic urban transport model systems

Gunnar Flötteröd, KTH Royal Institute of Technology

Stefan Flügel, Institute of Transport Economics (TØI)

1 Introduction

Strategic transport models predict travel demand and network conditions at the time scale of years to decades. The specification of scenarios of interest includes socio-demographic projections, changes in land use, parameters having an impact on travel demand (such as economic development or fuel costs) and transport network developments (such as road constructions or public transport schedules).

The fundamental building blocks of a strategic transport model are (i) a behavioural model of travel demand, (ii) a physical model of network flows (as from now on also called (network) supply model), and (iii) a specification of how demand and supply interact. The combination of different travel demand and network supply models can yield a variety of model system configurations. This text focuses on the role of the road network assignment package within such a system, attempting to account for its effect on the functioning of the strategic model system as a whole.

The remainder of this document is organized as follows. Section 0 provides an overview of the structure of strategic transport model systems. Section 3 then focuses in greater detail on ways in which a network assignment package is coupled system such as system, followed by Section 4, which discusses the implications of this for concrete application contexts. Section 5 summarizes the presentation.

This paper is based on the recent report: Flügel S., Flötteröd G., Kwong C.K. and Steinsland C. (2014) *"Evaluation of methods for calculating traffic assignment and travel times in congested urban areas with strategic transport models"* TØI-Report 1358/2014. It was commissioned by the Norwegian Public Roads Administration (SVV). While the report has some focus on discussing current Norwegian transport models, this paper is concentrating on the conceptual discussing of transport model systems in general. The evaluation of methods (section 4.3.) is unavoidable influenced by the authors own experience and opinions.

2 Typology of transport model systems

Models can, quite generally and independently of the transport context, be classified in terms of how they represent time, their resolution, and how they deal with uncertainty.

- **Static** models do not incorporate the notion of time, whereas **dynamic** models explicitly account for the notion of time and the temporal interdependence of the processes included in the model system.
- **Microscopic** models maintain the integrity of all entities throughout the modelling process, whereas **macroscopic** models are defined in terms of aggregate quantities.

- **Deterministic** model exhibit for given input parameters no randomness in their outputs, whereas **stochastic** models define a whole distribution of outputs given a fixed set of input parameters.

The combination of model components with different properties requires ensuring that the output produced by one component is compatible with the input data requirements of another component. While simplifications are in general rather straightforward to implement (removing the time dimension by averaging over the time window of interest, aggregating the outputs of a microscopic model into aggregate quantities that can be processed by a macroscopic model, filtering out the randomness produced by a stochastic model), the opposite process of enriching “simple” data can be quite challenging, often requiring additional modelling efforts or precluding certain model component combinations.

2.1 Travel demand models

A minimalistic instance of a congestion-sensitive strategic travel demand model predicts how many individuals, possibly stratified by socio-economic groups, wish to travel on average by what mode (for instance car, public transport, other) between the zones of a study region, given inter-zonal travel impedances such as costs or travel times. This minimalistic model hence represents the decision to travel at all plus destination and mode choice, and its output is a set of origin/destination (OD) matrices, one per demand segment and mode.

Travel demand models are subsequently classified by the way in which they explain the occurrence of travel, the way in which they represent time and temporal dependencies, and their representation of traveller heterogeneity. The presentation focuses on the modelling of individual persons.

2.1.1 Explanation of the occurrence of travel

A first classification of travel demand models is based on their assumptions about the mechanisms leading to the occurrence of travel.

The traditional way to model the generation and spatial distribution of travel is based on **zonal attraction**. This is the case in the classical gravity models where the amount of trips between two zones is based on the “size” (a function of area, population, facilities etc.) of each zone and the “distance” between the zones (the actual distance or a function of inter-zonal travel impedance). This approach is most often structured in two stages: Its first step predicts trip productions and attractions as functions of zonal properties like number of households or work places present in the zone. Its second step then distributes trip productions to different zones based on the attraction of the destination-zone and the generalized cost of travel needed to get from one zone to another. The vast majority of zonal-attraction-based models are macroscopic and deterministic (Ortuzar and Willumsen, 2011).

The more recent **activity-based travel demand models** (ABDMs) are based on the behavioural assumption that travel is undertaken in order to perform activities in different locations. Examples: work at the office, shop at the mall, socialize at a café. These models hence predict activities and associated locations. Travel demand then results from activities being located outside of the individual’s dwelling. Most activity-based models are microscopic and stochastic (Bowman and Ben-Akiva, 1996).

2.1.2 Representation of time and temporal dependencies

A second classification of travel demand models is based on their representation of time and temporal dependencies in the travel demand. For this purpose, it is instructive to first distinguish trip-based, tour-based, and all-day travel demand models.

Trip-based models represent travel demand in terms of trips, which are annotated with an origin, a destination, a mode, and possibly additional trip attributes. **Tour-based** models represent travel demand in terms of trip sequences starting and ending at the same location. **All-day** models represent travel demand in terms of a complete trip sequence for an entire day.

While trips may or may not be annotated with starting times, it is difficult to imagine tours or all-day travel plans, both of which consist of trip sequences, without a time dimension. This is reflected by the fact that most **static travel demand** models are trip-based and most **dynamic travel demand** models are tour- or all-day travel plan based. Static demand models consequently apply to a stationary analysis period, typically the morning or evening rush hour, for which they predict the rate at which individuals travel from each origin to each destination by each considered mode. While it is possible to model travel demand independently per time slice, it is rather awkward to establish a temporal interdependency between trips in different time slices. ABDMs, on the other hand, are typically dynamic and tour- or all-day travel plan based. They hence predict trip sequences that have a temporal structure. The ability to link trips also ensures logical consistency between trips, for instance in that starting a second trip is only possible after the first trip is completed or in that driving to the shopping mall by car also requires using the same mode on the return trip.

2.1.3 Representation of traveler heterogeneity

A third classification of travel demand models is based on their ability to account for heterogeneity in the traveller population (in particular their socio-demographics) when predicting travel demand. It turns out that this classification is strongly related to the choice of a microscopic vs. a macroscopic modelling approach.

Travellers are different in many ways – for instance age, gender, income, car ownership, marital status, and ethnicity – and many of these dimensions may play a role when it comes to strategic planning that aims at, for instance, cost-benefit analysis or an equity assessment (Lemp et al., 2007). The combinatorial complexity of adequately representing these dimensions has led to essentially two approaches.

The traditional approach is to perform a problem-specific **stratification into a limited number of population segments**. This inevitably comes with a certain aggregation bias, and it also requires performing the stratification before any analysis can be performed. This representation is typically chosen in conjunction with (static and macroscopic) four-step models, where trip production and attraction are then computed separately per population segment. It is, however, also possible to apply activity-based demand models per population stratum (Ortuzar and Willumsen, 2011; Ben-Akiva and Lerman, 1985).

A more recent approach is **population synthesis**. It samples synthetic individuals based on a mechanism that ensures that the resulting synthetic population is statistically consistent with the real population. Statistical consistency means here that all summary statistics available from the real traveller population are reproduced in the synthetic population; it does not mean that there is a one-to-one mapping between real and synthetic individuals. Synthetic populations are preferably used in conjunction with ABDMs, which are then used to predict activity, location, and trip sequences for each individual (Müller and Axhausen, 2010; Farooq et al., 2013).

2.2 Traffic assignment models

The vast majority of traffic assignment packages take (time-dependent) OD matrices as inputs, equilibrate in one way or another route choice, and calculate (time-dependent) link flows, link travel times, and inter-zonal impedances (examples are Traffic Suite (PTV)[1], Cube (Citilabs)[2], Emme/Dynameq (Inro)[3], Aimsun (TSS)[4] or TransModeler (Caliper)[5]). This review hence also concentrates on route choice being the sole dimension of travel behaviour. The two constituent building blocks of a network assignment package are a route choice model and a network flow model.

2.2.1 Route choice models

A route choice model requires at least an origin, a destination, and a set of route alternatives to choose from. The choice of a route follows more or less behavioural principles and is guided by the routes' properties, foremost cost in the form of (congestion-dependent) travel time (Prato, 2009).

Stochastic route choice models are generally considered to be more realistic than their **deterministic** counterparts, but they also are more difficult to specify and calibrate. Two problems stand out.

- The specification of the random component in the utility function depends in a nontrivial way on the overlap of routes. While the underlying modelling principles are well developed, the specification of both realistic and operational models for large networks remains a challenge (Frejinger, 2007; Ben-Akiva et al., 2004).
- How to define the choice set is not clear, even though this can have a strong effect on the resulting route choice. The underlying main difficulty is the combinatorial number of possible routes that render a complete enumeration impossible (Flötteröd and Bierlaire, 2013; Frejinger and Bierlaire, 2010).

The differentiation between **static and dynamic route choice models** is rather straightforward, in that dynamic models evaluate route costs (and hence utilities) at particular points in time. One typically distinguishes between **reactive travel times** (deriving travel times from the instantaneous network conditions at the time of starting a trip) and **predictive travel times** (taking into account that the experienced network conditions depend on when a vehicle has reached a particular location in the network) (Sloboden et al., 2012).

Microscopic route choice models select an individual route per vehicle (and hence are able to account for vehicle and/or driver characteristics in the route choice). **Macroscopic route choice models** distribute vehicle flows deterministically across the routes – deterministic macroscopic route choice models concentrate the flows on routes of minimum cost, whereas stochastic macroscopic route choice models *deterministically* ensure that the share of vehicle flow on a route equals the probability of choosing that route (Watling and Hazelton, 2003).

---

[1] http://vision-traffic.ptvgroup.com/en-us/products/

[2] http://citilabs.com/software/products/cube

[3] http://www.inro.ca/en/products/

[4] http://www.aimsun.com/wp/

[5] http://www.caliper.com/transmodeler/

2.2.2 Network flow models

A minimalistic network flow model specification consists of (i) a topology where links are connected through nodes, (ii) a definition of flow entry and exit points, and (iii) some definition of how links perform under congestion.

Static network flow models assume that route flows propagate instantaneously through the network. The vast majority of these are based on volume-delay functions that compute link travel times from link flows. These models, although still widely applied, are inadequate to model congested conditions: they predict flows beyond capacity, they do not capture the spatial propagation of queues, and they implicitly assume that delay is experienced inside a bottleneck, not upstream of it. In practice, static network flow models are usually macroscopic and deterministic (Ortuzar and Willumsen, 2011).

Dynamic network flow models come in a variety of guises. Macroscopic instances are typically deterministic and often based on the Kinematic Wave Model (Lighthill and Witham, 1955; Richards, 1956); simpler macroscopic flow models lack realism, whereas more complex models quickly become too complex to be applied for network modelling. Microscopic instances move individual vehicles through detailed network geometries according to driving rules of car-following, lane changing, and gap acceptance. “Mesoscopic” models trade precision for speed and often move groups of vehicles according to aggregate speed/density relationships but maintain the identifiability of individual vehicle units. Both microscopic and mesoscopic network flow models are usually stochastic, with the mesoscopic mechanisms often “averaging away” some of the stochasticity that is explicitly captured in microscopic driving rules (Barcelo, 2010; Hoogendoorn and Bovy, 2001).

2.2.3 Network assignment models

The main classification criterion here is time.

Combining a **static route choice model with a static network flow model** yields the typical static assignment package implementing the 4th stage of the four-step model. This class of models is typically macroscopic and deterministic. The interpretation of the solution algorithm is typically in terms of a mathematical program that aims at satisfying some equilibrium condition (Sheffi, 1985).

Combining a **dynamic route choice model with a dynamic network flow model** leads to what is commonly labelled as “dynamic traffic assignment” (DTA). Typical configurations are (i) macroscopic route choice and macroscopic network flows, (ii) microscopic route choice and meso- or microscopic network flows. The fully macroscopic approach is most often again tackled from the mathematical programming side, whereas the meso-/microscopic approach is more frequently taken from the heuristic perspective of explicitly emulating demand/supply interactions (although efforts exist to also introduce mathematical rigor into this process) (Barcelo, 2010; Nagel and Flötteröd, 2012; Peeta and Ziliaskopoulos, 2001).

Intermediate approaches exist that most often simplify the network flow dynamics in that they specify a number of in the first instance independent static assignment problems per time slice but then again introduce simplified dynamic coupling mechanisms that mimic the carry-over of vehicle queues from one time slice to the next (Bliemer et al 2013).

3 Coupling of travel demand and network assignment models

The traditional way to structure strategic transport models is that of a **four-step approach** (Ortuzar and Willumsen, 2011). At the first step, trip generation (or trip frequency), the total amount of daily trips in the future year of interest, is estimated. These trips are distributed between origin and destination zones by a trip distribution model. In the third step, the resulting OD matrix is further subdivided between transport modes by means of a mode split model. Finally, the mode-specific OD matrices, representing the travel demand, are input in the **network assignment package** that models the allocations of vehicles to road sections (network links). The assignment package includes a route choice model with travel times as a central determinant. Travel times are traditionally calculated based on characteristics of the network (the physical environment as the length of links) and increases in travel times (delays) due to increases in traffic volumes.

That route choice and network flow propagation are modelled jointly in a traffic assignment package is not self-evident. Consider Figure 1. In terms of the traditional four-step model, what is labelled as "higher level" behaviour in this figure corresponds to stages one (trip generation) and two (trip distribution). In terms of disaggregated activity-based demand model (see section 3.2.1), "higher level" behaviour corresponds to the choice of activities and corresponding locations. Mode choice corresponds to stage three of the four-step model, whereas departure time choice is often not explicitly modelled in this approach. Both choice dimensions are typically incorporated in activity-based demand model systems.

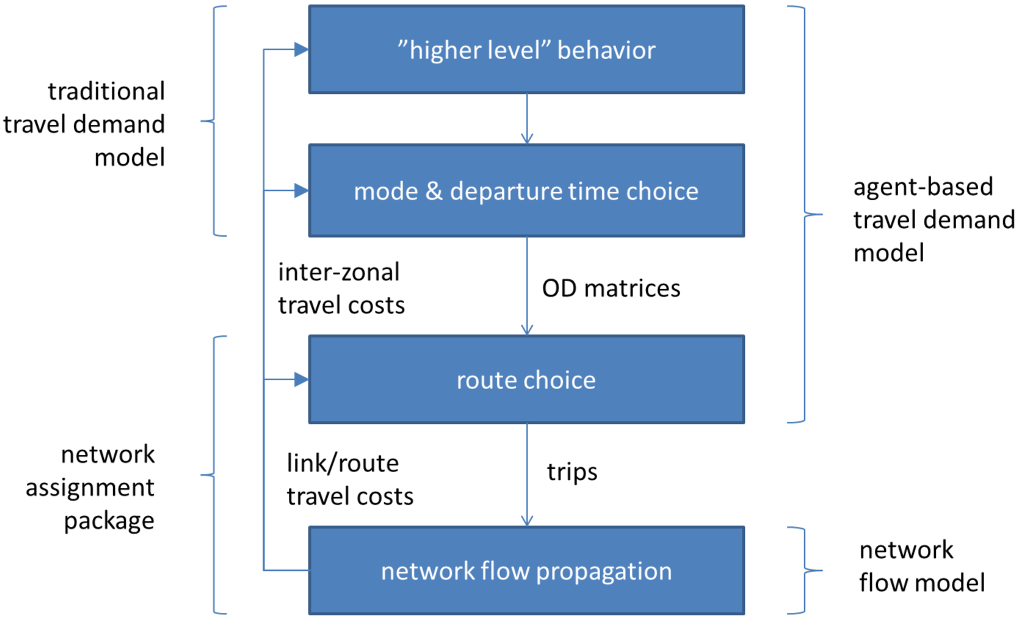

Figure 1: Strategic transport model system components (adopted from Berglund et al. (2014))

Route choice is an aspect of travel behaviour, such that its inclusion in a travel demand model appears more plausible from the behavioural modelling perspective. Despite of its advantages, this approach is not yet widespread but present in what is called the **agent-based** modelling approach (Nagel and Flötteröd, 2012). Travel behaviour that is more detailed than the route choice (car following, lane changing, and the like) is typically not included in the travel demand model but represented in the network flow model, both in traditional and agent-based models. The justification of this simplification is that given that a traveller is able to follow a route, it does from a strategic planning perspective not matter what detailed driving actions were necessary to perform that trip.

The interactions between a traditional travel demand model and a network assignment package is most frequently in terms of OD matrices (output of the demand model, input to the network assignment) and travel time matrices (output of the network assignment, input to the travel demand model). More disaggregate data structures are possible in the traditional approach, and they are characteristic for the agent based approach. In any case, the solution of the models systems is found by bringing the demand and supply side in an equilibrium or a relaxing state. This is usually achieved by iterating between the different components of the model system. The integration of all model components is essential for strategic transport models, which require the demand side to be endogenous and sensitive to conditions and changes of the network.

The objective of the following presentation is to elaborate on if and how different types of travel demand models and network assignment packages can be integrated into one strategic transport model system. This depends on the properties of each component, the relevant dimensions of which have been laid out before. Strongly related to these model properties are the properties of the data moved back and forth between the model components. The size and complexity of the resulting model system also raises computational concerns, which may enforce otherwise undesirable simplifications.

Section 3.1 presents the relevant data structures through which a travel demand model and a network assignment package may interact, drawing from Nagel and Flötteröd (2012). The following essential dimensions of their integration into one strategic transport model system are subsequently discussed: representation of time, resolution, and representation of uncertainty (Section 4.4), building on Flötteröd and Bierlaire (2012).

3.1 Relevant data structures

3.1.1 Representations of travel demand

Travel demand is the output of the travel demand model and the input of the network assignment package. One can distinguish two typical data structures.

Classically, travel demand is represented through **OD matrices**. These may be time-dependent, typically in the form of separate OD matrices per time slice. The pure trip count information contained in each of these matrices (per demand segment and/or time slice) can be annotated with additional demand parameters, such as socio-demographic summary statistics per OD matrix.

A more recent approach is to replace the aggregate OD matrix information by disaggregate **trip lists**, which contain one entry for each “one” in the OD matrix. Given that OD matrices for large study regions contain many almost-zero entries,[6] this approach can have computational advantages. More importantly, it offers a higher resolution when representing time (which can be made a real-valued attribute of the trip) and the possibility to annotate trips with arbitrary demand information, for

---

[6] Both the number of trips and the number of zones grow roughly linearly with the network size. The number of zone-to-zone relations hence grows quadratically with the network size, and the ratio of number of trips to number of zone-to-zone relations sinks anti-proportionally with the network size.

instance in terms of driver income (relevant at the route choice level) or vehicle type (important for instance in the presence of vehicle-specific tolls or network access constraints).

3.1.2 Representations of network impedances

Network performance (impedance) is the output of the network assignment package and the input of the travel demand model.

Classically, network performance is represented through **inter-zonal impedance matrices** (also called "level-of-service matrices" or "skim matrices"). These may be time-dependent, typically by defining separate matrices per time slice. Matrices may also be distinguished by impedance measure: Apart from travel time, monetary costs such as tolls come to mind. Impedance matrices per travel demand segment are also thinkable, for instance in the presence of vehicle-specific tolls or speed limits. This representation has the same structure as an OD-matrix-based travel demand representation.

Virtually every network assignment package internally also computes network performance measures at the link level (in particular travel times), which can be output directly, leaving the (possible) aggregation step into inter-zonal impedances to the travel demand model. This has advantages in terms of spatial resolution, for instance when it comes to the modelling of intra-zonal travel.

If the network assignment package is capable of assigning individual trips, it also becomes possible to write out the travel times that were actually experienced by individual trip makers. This further enables the annotation of trips with additional information collected at the network level. However, this representation of network performance constitutes an incomplete counterpart to the trip-list based travel demand representation unless all possible paths considered by a trip maker are evaluated, including those alternatives that were never chosen in the assignment package.

3.2 Coupling models with different representations of time

Coupling a static travel demand model to a static network assignment package is (in terms of time representation) straightforward, given that both models refer to the same analysis period. The same holds in principle to the coupling of a dynamic travel demand model to a dynamic network assignment package, as long as consistent time resolutions are maintained. Both OD matrices and trip lists are possible demand data structures in either setting, and both inter-zonal impedance matrices and link travel time lists can be used to represent network performance.

Coupling a static travel demand model to a dynamic network assignment package can make sense if the static travel demand represents a peak hour and the off-peak traffic is rather low. In this configuration, the static travel demand is spread out over the analysis. This allows capturing the spatiotemporal dynamics of congestion build-up and dissipation in the dynamic assignment package. A series of independent static demand predictions per time slice could be used to create a (possibly even all-day) temporal demand profile, allowing also for an all-day analysis within the dynamic network assignment package. The realism of this approach is limited by the independence assumption across time slices in the travel demand model and (hence) does not make full use of the network performance information provided by the dynamic assignment package. Travel demand can again be represented by OD matrices or trip lists, and both inter-zonal impedance matrices and link travel time lists are feasible to describe network performance.

When coupling a dynamic travel demand model to a static traffic assignment package, the time-dependent demand is discretized into time slices and the average demand per time slice is fed into separate static assignment packages. This approach comes with all weaknesses of using static assignment packages, as enumerated in Section 2.2. Static assignment packages typically accept only OD matrices (and no trip lists) as inputs and output network performance measures in terms of inter-zonal impedance matrices and/or link travel times.

3.3 Coupling models with different resolutions

Coupling a macroscopic travel demand model to a macroscopic network assignment package, which have mutually consistent representations of travel demand and network performance measures in terms of real-valued quantities, is rather straightforward. Data exchange is based on OD matrices as demand representations and inter-zonal travel time matrices or link travel times as descriptions of network performance. If the demand model is capable of producing separate OD matrices per demand segment and/or per mode, the network assignment package should be able to handle these different classes. Likewise, the different types of travel costs (time, monetary cost, summary representations in terms of (dis)utility) produced by the macroscopic network assignment package should be compatible to the network performance measures expected by the travel demand model.

Coupling a macroscopic travel demand model to a microscopic network assignment package requires disaggregating the travel demand before feeding it into the network assignment package. For this purpose, each OD matrix produced by the travel demand model is taken as a representation of the expected number of trip makers. Discrete trips, as needed by the meso- or microscopic assignment, are then created in a manner such that their number corresponds on average to what the OD matrix prescribes (e.g. by sampling or rounding). Often, this disaggregation step is performed by the network assignment package, which then accepts real-valued OD matrices as inputs. However, despite of the disaggregate representation of trip-makers at the network level, the resolution at which travel demand information can be attached to individuals is limited by the number of different (demand-class specific) OD matrices produced by the demand model. The representation of network performance measures is, on the other hand, typically unproblematic and based on impedance matrices or link-specific data.

Coupling a disaggregate travel demand model to a macroscopic network assignment package requires to aggregate the travel demand before feeding it into the network assignment package. The ability of macroscopic assignment packages to handle heterogeneous demand representations (in terms of many class-specific OD matrices) has computational limitations. This aggregation hence comes typically with information loss, in that no matter how rich the original output of the disaggregate travel demand model may be, it is aggregated rather coarsely. The representation of network performance measures is at the basic level of travel times and costs again rather unproblematic and based on impedance matrices or link-specific data. However, the resolution of subgroup-specific network performance measures that can be provided to the travel demand model is limited to the granularity at which the travel demand can be distinguished at the network level, i.e. depending on the number of different OD matrices used.

Coupling a disaggregate travel demand model to a meso- or microscopic network assignment package does, ideally, not require any intermediate aggregation of the travel demand: Every single trip of an individual in the travel demand model can be processed individually by the network assignment package. Socio-demographic information about the traveller is hence available at the network level. In the route choice model, this means that heterogeneous values of time and other person-specific attributes such as trip purpose or income can be accounted for. In the network flow model, the trip-maker's vehicle type is uniquely identified, meaning that vehicle-specific tolls or restrictions are experienced exactly by those individuals owning the respective vehicles. Further, information about the network experience of a traveller can, at least in principle, be made available to the travel demand model at the level of individual travellers.

A fully disaggregate transport model system hence appears to be able of avoiding any aggregation bias. This capability, however, can only be exploited if the representations of travel demand and network performance also maintain this level of resolution. The use of aggregate OD matrices to represent travel demand, although possible, is therefore not adequate. Such a configuration would require an (unnecessary) aggregation step of the disaggregate travel demand into one or several OD matrices, which then would (unnecessarily) have to be disaggregated again in the network assignment package. Trip lists, on the other hand, are an adequate travel demand data structure. Things are somewhat different when it comes to the representation of network performance, which defines the attributes of the behavioural alternatives in the travel demand model. Since person-specific information is available anyway in the (fully disaggregated) travel demand model, it is for many purposes sufficient to output only non-person specific information (in particular link travel times and costs).[7]

3.4 Coupling models with different representations of uncertainty

If both the travel demand model and the network assignment package are deterministic, then the model system resulting from their combination is also deterministic.

As soon as one component in the model system is stochastic, the entire model system is stochastic. If, for instance, the travel demand model is stochastic but the network assignment package is deterministic, then the output of the network assignment package is only deterministic given a particular realization of the travel demand model's output – overall the network assignment package makes stochastic predictions due to the stochasticity of its inputs. A symmetric statement holds for the case of a deterministic travel demand model and a stochastic network assignment package.

Despite of the increasing acceptance of the fact that (adequately designed) stochastic model systems come with the advantage of representing modelling uncertainty and hence allow at least in principle to account for this uncertainty when using model systems for strategic planning, truly stochastic transport model systems have not yet entered general practice. A possible cause for this is that the presence of stochasticity in such model systems adds yet another level of complexity that may be seen more as a hindrance in the analysis than an added value. A more practical reason may be that either the demand model or the network assignment package expects input values representing average conditions: Discrete choice models typically expect the attributes of the alternatives (in particular travel times or inter-zonal impedances) to be expected values. OD-matrix based network assignment

---

[7] These observations make a strong case for moving the route choice model out of the assignment package and into the travel demand model, cf. Section 2: This information would naturally make all person-specific information available to the route choice model and would only leave the representation of the "physical environment" of the travellers to the network assignment package, which accordingly would be reduced to a representation of network flows.

packages interpret these matrices as average trip levels per time slice. Identifying and adequately characterizing the effect of allowing for randomness in such parameters is more a research question than a matter of model application. Stochasticity in the demand model or the network assignment package may therefore be removed (typically by filtering or switching off random number generators) in an attempt to obtain an overall almost deterministic model system, the predictions of which may then be taken as average values.

Table 1: Coupling models with different time representations and resolutions. summarizes this discussion.

Table 1: Coupling models with different time representations and resolutions.

		Network assignment package		
		Static macro	**Dynamic macro**	**Dynamic meso/micro**
Travel demand model	**Static macro**	Adequate.	Adequate for studying peak-period congestion dynamics, but dynamic network performance data cannot be utilized by demand model.	Adequate for studying peak-period congestion dynamics, but dynamic network performance data cannot be utilized by demand model.
	Dynamic macro	Suffers from simplistic representation of congestion.	Adequate.	Adequate.
	Dynamic micro	Suffers from simplistic representation of congestion. May suffer from coarse representation of traveler heterogeneity in the network.	May suffer from coarse representation of traveler heterogeneity in the network.	Adequate if demand is represented through trip lists.

4 Scope and feasibility of different network assignment packages

The adequacy of a network assignment package is only defined with respect to a particular objective. This text hence first lays out possible strategic application contexts of network assignment packages in Section 4.1. Section 4.2 discusses model properties that apply equally in all application contexts. Section 4.3 finally synthesizes the applicability of concrete network assignment package types in concrete application fields.

4.1 Possible application contexts of traffic assignment packages

Two important and interrelated application contexts at the network level need to be mentioned.

A first question relates to the (expected) level of congestion in the network. This is particularly important if one is interested in **studying the effect of congestion-mitigating measures** (road capacity expansions, the introduction of road pricing rings around the city center, the redirection of substantial traffic streams for instance from through-city-traffic onto bypasses). If congestion can be expected to be low then there is little added value in accounting for it in the model system. This in turn renders its detailed representation in the network assignment package unimportant. Given that without congestion there is only limited physical coupling between the network conditions of different time slices, a static network assignment package may be fully sufficient. If congestion cannot be expected to be negligible, a network assignment package that captures spatiotemporal congestion dynamics is needed. Virtually all packages that come with this capability are also dynamic.

Another question very related to the first one is if one intends to **study the long-term benefits of intelligent transportation systems (ITS)**. Since the introduction of such systems mostly comes with the intention to (also) provide congestion relief, the network assignment package again needs to be able to describe the build-up and dissipation of congestion. Beyond this, the benefits of ITS are strongly dependent on information availability (E.g.: Who receives the real-time congestion information?), technical equipment (Who will then follow a recommended path?), representation of time (Where are travellers in the very moment of an incident?), and individual properties (Who is willing and/or capable to at all react to a congestion warning?). Apart from time and congestion, a detailed representation of vehicle types, vehicle equipment, and drivers may become necessary. Further traffic control measures that fall under the umbrella of ITS are intelligent (traffic responsive) signals, the dynamic allocation of HOV (high-occupancy vehicle) lanes, and variable speed limits. Vehicle-to-vehicle and vehicle-to-infrastructure communication may also need to be accounted for. These measures require a representation of vehicles and infrastructure that goes down to the level of detailed vehicle movements and hence require a disaggregate representation of network flows.

When focusing on travel behaviour, two again interrelated and relevant application contexts require attention.

Travel demand management recognizes that the efficiency of supply-side measures (capacity improvements, ITS) is limited and aims at affecting travel behaviour in a way that leads to an overall improved system performance. This may comprise campaigns to promote using public transport in order to reduce the number of vehicles on the road, the introduction of staggered working hours to spread out peak traffic, and the introduction of road pricing. In any case, a good understanding of the (expected) response of travellers to the demand management measure is essential. This requires an adequate level of behavioural model resolution. Two key aspects to be considered here are (i) the fact that the different trips made by a traveller are interconnected in a rather complex manner and that (ii) the responses of different travellers to the same travel demand measure may be very different. Item (i) calls for a dynamic behavioural model, and item (ii) requires a disaggregate behavioural model. ABDMs meet these requirements and are (hence) considered the most adequate behavioural modelling approach in this application context. As explained in Section 3, the best match of an ABDM in a strategic transport model system is a dynamic and disaggregate network assignment package.

Similar observations hold when it comes to **equity and winner-and-losers analysis**. Focusing on the effect of a measure on (the welfare of) individuals, it is essential to adequately represent the heterogeneous and complex interrelations between socio-demographics and mobility, calling again for a disaggregate behavioural model. Since similarly complex processes guide the perception, valuation, and use of time, the model should also be dynamic. ABDMs are hence again considered as the most adequate option. The use of a synthetic population has in this context an important practical advantage over group-specific OD-matrices: These matrices need to be defined before a model-based investigation is performed, meaning that the a priori design of the demand representation already frames the set of possible (group-specific) answers that can be given. A synthetic population, on the other hand, requires no a priori aggregation, meaning that summary statistics over arbitrary subsets of the demand can be computed after the model has been evaluated. This is particularly important when it comes to equity analysis, where the winners and losers of a particular measure may be identifiable only in hindsight (Flötteröd et al., 2012). Again, an ABDM is best matched by a dynamic and disaggregate network assignment package.

4.2 Further relevant properties

In the following, a number of network assignment package properties are discussed that deserve consideration independently of a concrete application context.

4.2.1 Robustness and accountability

Strategic decisions are by definition long-term. The longer the time horizon until the measure under consideration takes effect, the larger the influence of unforeseeable processes that may affect the transport system. One hence is interested in making robust decisions and in coming to accountable conclusions (Bliemer et al., 2013).

Robustness refers to the management of uncertainty in the model system. This uncertainty can be classified into (i) uncertainty in the (exogenous) boundary conditions (inputs) of the model system and (ii) uncertainty in the (endogenous) processes inside of the model system. Uncertainty in boundary conditions refers to the limited predictability of input variables needed by the model. These may comprise socio-demographics (income development, migration), oil prices and economic growth, and weather-induced network deterioration. Such uncertainty should be accounted for by repeatedly evaluating the model system with different realizations of these boundary conditions. This poses no particular requirements on the transport model system (Ross, 2006). Things are different, however, when it comes to the representation of uncertainty within the processes of the model system. An example of this is travel time variability, which results from complex demand/supply interactions and has a systematic effect on network performance and travel behaviour. In theory, such effects are best captured by a stochastic transport model system. In practice, the understanding of how uncertainty propagates through a complex transport system is still limited, and hence it is difficult to adequately account for. However, even if one is not able to precisely model all sources of uncertainty, one still is interested in a model system that is capable of indicating that its predictive power is exhausted beyond a given time horizon. Arguably, disaggregate simulations are by design the most adequate approach to address strategic transport planning problems in the presence of uncertainty.

Accountability refers to the model being transparent and at least conceptually understandable. This is absolutely necessary, given that a model as complex as a transport system cannot be applied as a "black box" – it is fair to say that an analyst that is capable of interpreting its outputs and identifying inconsistencies comes as an integral part of such a prediction system. The more intuitive the workings of a model, the easier it is for the analyst to make sense of its outputs. A more abstract model requires an equally abstractly trained analyst to adequately deploy it. Microscopic simulations, which attempt to truthfully mimic real-world processes, have the clear advantage that their workings have real counterparts, which supports intuition. On the other hand, the relatively large number of fine-grained processes evaluated in a micro-simulation makes it difficult to intuitively understand its detailed cause-effect relationships. In addition, the inherent stochasticity of micro-simulators requires some statistical training to be adequately managed. The opposite statement is true for macroscopic simulations: their typically mathematical problem formulation is rather difficult to understand, but once it is understood it reveals relatively clearly the underlying cause-effect mechanisms. The often guaranteed solution uniqueness of macroscopic model may be seen as a practical advantage that eases interpretation when comparing scenarios. However, it comes with the danger of ignoring other solutions of the systems that may be equally valid. Stochastic models (correctly) allow for different system solutions.

4.2.2 Computational efficiency

It is often argued that an increased level of disaggregation also comes with higher computational requirements. This statement, however, is inadequate as a general observation because it depends strongly on the problem under consideration.

On the travel demand side, an increased level of population heterogeneity calls for an increased number of stratified submodels, with the computational load increasing roughly linearly with the number of strata. A microscopic model, on the other hand, maintains a constant computational requirement because it is based on a synthetic population that may exhibit arbitrary heterogeneity (one may attach as many background variables to an agent as one likes). The same observation can be made on the network supply side. The larger the number of commodity flows (e.g. demand segments, origin/destination relations, vehicle classes) to be tracked on the network, the larger the computational effort in a macroscopic assignment package. Again, a vehicle microsimulation is by design insensitive to an increase of vehicle hetereogeneity; every simulated vehicle is a realization from an arbitrary distribution.

More important than the degree of disaggregation is the model's time resolution: static models are generally faster than dynamic models.

4.2.3 Implementation and calibration, use and maintenance

A strategic transport model is, no matter what type of model one selects, a highly complex system that requires large amounts of data to be set up and calibrated, as well as expert knowledge to be used and continuous improvements to be maintained. This puts concerns about the computational run times of a transport model system somewhat into perspective; it is more adequate to assess the sum of the time it takes to prepare the model for a particular analysis purpose plus its run-time – with the latter then easily becoming by an order of magnitude smaller than the time invested in preparing the computations.

Dynamic models require the calibration of parameters guiding their dynamical mechanisms and call for an understanding of these mechanisms by the analyst. Examples are time choice models on the travel demand side and network models that are based in realistic traffic flow theory. All of this is not required in static models. Similarly, the more disaggregate a model, the more input data is necessary to initialize all model processes, the more parameters need to be calibrated for the different processes, and the more domain knowledge is needed on the analyst's side to make sense out of the interactions of all of these processes. Examples are fine-grained behavioural model systems that explain trip-making through processes as complex as inter-household negotiations about who takes the car (Bhat et al. 2012) or microscopic traffic flow models that mimic the car-following and lane-changing decisions of individual drivers (Toledo, 2008).

In summary, dynamic models require more effort than static models, and disaggregate models require more effort than aggregate models.

4.2.4 Flexibility and extendability

Flexibility and extendability refer to the possibility to update the model system in reaction to future requirements. This is a nontrivial aspect of a present model selection effort, given the difficulty of anticipating the requirements to come. It is important because the "life span" of a strategic transport model system is in the order of decades.

This has a technological and a modelling dimension. Technologically, it is advantageous that the model system is extendable in the sense that one is able to replace components (e.g. a fixed by an adaptive signalling system or a simple by a more sophisticated destination choice model). Things are more difficult on the modelling side, in that certain modelling paradigms are virtually impossible to enrich: Turning a static into a dynamic model system or turning a macroscopic into a microscopic model system is difficult to impossible without the danger of introducing problematic ad-hoc modifications.

In summary, the choice of a static model system essentially excludes future opportunities for dynamic modelling, and the choice of a macroscopic model system puts strong restrictions on analysing increased detail and heterogeneity.

4.3 Evaluation summary and synthesis

The choice of a network assignment package for a strategic transport model system can be based on the following three considerations.

- **The planned application context of the model system.** The requirements of different application contexts, which are elaborated in Section 4.1, are strong indicators of desirable (and of unnecessary) model properties.
- **Properties of the available (or to-be-developed) travel demand model.** Section 3 establishes the compatibilities of different types of travel demand models and network assignment packages.
- **Further model capabilities and practical features.** A number of such criteria is provided in Section 4.2.

Table 2 and table 3 summarize these considerations.

Table 2: Evaluation of network assignment packages for application purposes

	Static macro	**Dynamic macro**	**Dyn. meso/micro**
Congestion mitigation	Inadequate (S)	Adequate	Adequate
ITS	Inadequate (S,A)	Inferior (A)	Adequate*
Travel demand management	Inferior (A)	Acceptable	Adequate
Equity analysis	Inadequate (A)	Inferior (A)	Adequate
Standard Cost-benefit analysis	Adequate if congestion low	Adequate	Adequate**

Reasons (S): Static, (A): Aggregated; *micro-level might be necessary **if distributions of predictions are compared

Table 3: Evaluation of network assignment packages on general model capabilities and practical features

	Static macro	**Dynamic macro**	**Dyn. meso/micro**
Robust and accountable	Yes but potential biased (S)	Sensitive	Stochastic*
Richness in analysis	Limited (S,A)	Moderate (A)	High
Computation times	Fast**	Slow	Slow***

Implementation, calibration, use & maintenance	Simple (S,A)	Moderate (A)	Involved
Flexibility and extendibility	Low	Moderate	High

Reasons (S): Static, (A): Aggregated; *single model runs not robust **slow if number of segments high ***micro-level may be too slow for large scenarios

Dynamic and dissagregate assignment models are most applicable for all considered application purposes considered. The main reason goes back to the explicit modelling of congestion dynamics and the disaggregated nature allowing a great richness in analysis. The evaluation for “travel demand management” and “equity analysis” rests on the assumption that they are coupled with corresponding disaggregated demand models.

Static/macro models have in general the lowest requirements in practical use and are – arguable for historical reasons – still much better known (and much more often applied) among consultants. Dynamic disaggregate assignment models are more demanding with respect to implementation, calibration and usage and require more expert knowledge on part of the users. If the corresponding budget-, time-, data constraints are not too tight and human resource are available, then a dynamic meso-/microsimulation offers both the greatest application range and the most flexible and far-sighted model structure.

5 Summary

A strategic transport planning model consists of a travel demand model, a network flow model, and a logic according to which these two model components interact. We classified the functioning of either component according to the properties static/dynamic, aggregate/disaggregate, deterministic/stochastic. The feasibility of combining model components with different properties was subsequently discussed. Based on this general but also comprehensive perspective, are more focused discussion of the applicability of different types of network assignment packages within strategic transport planning models was delivered. Considering both concrete application contexts and the practicalities of model deployment, it was concluded that the superior modeling capabilities of dynamic and disaggregate models over static and macroscopic alternatives outweigh their greater deployment costs.

References

Barcelo, J. (2010). *Fundamentals of Traffic Simulation*, Springer.

Ben-Akiva, M. E., S. R. Lerman (1985). *Discrete Choice Analysis*. The MIT Press, Cambridge, MA.

Ben-Akiva, M., Ramming, M. and Bekhor, S. (2004). Route choice models, *in* M. Schreckenberg and R. Selten (eds), *Human Behaviour and Traffic Networks*, Springer, pp. 23–45.

Berglund, S., Engelson, L., Flötteröd, G. and Jonsson, R. (2014). Further development of Sampers and modeling of urban congestion, *Technical Report IHOP project report*, Trafikverket. http://swopec.hhs.se/ctswps/abs/ctswps2014_010.htm

Bhat, C. R., Paleti, R., Pendyala, R., Goulias, R. G., & Hu, H. (2012). Intrahousehold interaction with MDCEV for SimAGENT. Proceedings annual Transportation Research Board meeting, Washington, DC.

Bliemer M, Raadsen M., de Romph E.,Smits E-S, (2013) Requirements for Traffic Assignment Models for Strategic Transport Planning: A Critical Assessment Australasian Transport Research Forum 2013 Proceedings 2 - 4 October 2013, Brisbane, Australia, Publication website: http://www.patrec.org/atrf.aspx

Bowman, J. L., M. E. Ben-Akiva (1998). Activity based travel demand model systems. P. Marcotte, S. Nguyen, eds. Equilibrium and Advanced Transportation Modelling. Kluwer, Dordrecht, The Netherlands, 27–46.

The Netherlands, 27–46.Farooq, B., Bierlaire, M., Hurtubia, R. and Flötteröd, G. (2013). Simulation based population synthesis, *Transportation Research Part B: Methodological* **58**(C): 243–263.

Flötteröd, G. and Bierlaire, M. (2012). Running urban microsimulations consistently with real-world data, *in* S. Müller Arisona, G. Aschwanden and J. Halatsch (eds), *Digital Urban Modelling and Simulation*, Springer, Heidelberg, Dordrecht, London, New York, pp. 181–199.

Flötteröd, G., Chen, Y. and Nagel, K. (2012). Behavioral calibration and analysis of a large-scale travel microsimulation, *Networks and Spatial Economics* **12**(4): 481–502.

Flötteröd, G. and Bierlaire, M. (2013). Metropolis-Hastings sampling of paths, *Transportation Research Part B* **48**: 53 – 66.

Frejinger, E. and Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models, *Transportation Research Part B* **41**(3): 363–378.

Frejinger, E. and Bierlaire, M. (2010). On path generation algorithms for route choice models, *Choice modelling: the state-of-the-art and the state-of-practice*, Emerald Group Publishing Limited, pp. 307–315.

Hoogendoorn, S. and Bovy, P. (2001). State-of-the-art of vehicular traffic flow modelling, *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering* **215**(4): 283–303.

Lemp, J.D., McWethy, L.B. and K.M. Kockelman (2007). From aggregate methods to microsimulation: assessing the benefits of microscopic activity-based models of travel demand, Transportation Research Record **1994**: 80-88.

Lighthill, M., Witham, J. (1955). On kinematic waves II: a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A* 229 (1178), 317–345.

Müller, K. and Axhausen, K. W. (2010). Population synthesis for microsimulation: State of the art, *10th Swiss Transport Research Conference, Ascona*, ETH Zürich, Institut für Verkehrsplanung.

Nagel, K. and Flötteröd, G. (2012). Agent-based traffic assignment: going from trips to behavioral travelers, *in* R. Pendyala and C. Bhat (eds), *Travel Behaviour Research in an Evolving World*, Emerald Group Publishing, Bingley, United Kingdom, chapter 12, pp. 261–293.

Ortuzar, J. and Willumsen, L. (2011). Modelling Transport, Wiley.

Peeta, S. and Ziliaskopoulos, A. (2001). Foundations of dynamic traffic assignment: the past, the present and the future, Networks and Spatial Economics **1**(3/4): 233–265.

Prato, C. (2009). Route choice modeling: past, present and future research directions, *Journal of Choice Modelling* **2**(1): 65 – 100.

Richards, P. (1956). Shock waves on highways. *Operations Research* 4 (1), 42–51.

Ross, S. (2006). *Simulation*, fourth edn, Elsevier.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall.

Sloboden, J., Lewis J., Alexiandis, V., Chiu, Y.-C. and Nava, E. (2012). Traffic Analysis Toolbox Volume XIV: Guidebook on the Utilization of Dynamic Traffic Assignment in Modeling. US Department of Transportation, Federal Highway Administration report number FHWA-HOP-13-015, website: http://ops.fhwa.dot.gov/publications/fhwahop13015/fhwahop13015.pdf

Toledo T. (2008), Integrated model of driving behavior, VDM Verlag Dr. Müller, Saarbrucken, Germany.

Watling, D. and Hazelton, M. (2003). The dynamics and equilibria of day-to-day assignment models, *Networks and Spatial Economics* **3**(3): 349–370.