Presented at: World Conference of Air Transport Research Society (ATRS),
Bergamo, Italy, June 2013. Also VSP WP 13-23, see www.vsp.tu-berlin.de/publications .

Agent-Based Modelling of Air Transport Demand

Dominik Grether*, Kai Nagel
Transport Systems Planning and Transport Telematics
TU Berlin
Salzufer 17-19
D-10587 Berlin
+49-30-314 23308
+49-30-314 26269
grether@vsp.tu-berlin.de
* Presenter and corresponding author

May 31, 2013

Abstract

Constraints such as opening hours or passenger capacities influence travel options that can be offered by an airport and by the connecting airlines. If infrastructure, policy or technological measures modify transport options, then the benefits do not only depend on the technology, but also on possibly heterogeneous user preferences such as desired arrival times or on the availability of alternative travel modes.

This paper proposes an agent-based, iterative assignment procedure to model European air traffic and German passenger demand on a microscopic level, capturing individual passenger preferences. Air transport technology is simulated microscopically, i.e. each aircraft is represented as single unit with attached attributes such as departure time, flight duration or seat availability. Trip-chaining and delay propagation can be added. Microsimulation is used to verify and assess passengers' choices of travel alternatives, where those choices improve over iterations until an agent-based stochastic user equilibrium is reached. This requires fast simulation models, thus, similar to other approaches in air traffic modelling a queue model is used. In contrast to those approaches, the queue model in this work is solved algorithmically. Overall, the approach is suited to analyze, forecast and evaluate the consequences of mid-distance transport measures.

Keywords: Transportation Systems Modelling, Multi-Agent Simulation, Air Transport Demand, Long Distance Travel

1 Motivation

In Italy, recently a private company started providing 2.5 h non-stop train rides between Milano and Rome.[1] From Paris, nearly all major French cities can be reached by high-speed train in 2–4 h trips.[2] For the journey Berlin–Frankfurt in Germany, a 4h non-stop rail connection is provided.[3] Many airlines provide flights between all these destinations that take between 1 and 2 h. When comparing travel times, the additional access time to the airport or railway railway station needs to be included. Overall travel times are often not that different between middle range rail on the one side, and air transportation on the other.

Following recent forecasts, in 2030 13 major EU airports will operate at full capacity at least eight hours a day [Commission, 2011]. Legal opening hour constraints limit operations to a certain time frame. Yet even increasing opening hours for airports may not resolve capacity bottlenecks since it may not be possible to move enough demand away from the peak hours.

In contrast, railway stations are normally not as much exposed to restrictions of opening hours due to noise protection as airports are. Also, in comparison with airports, railway stations mostly feature a more central geospatial location in urban areas. Slightly longer travel times can be compensated for by shorter access times and longer opening hours. Passenger demand and technology supply for middle distance railway or air transportation may interact and are time dependent over a day or even a longer period.

In order to provide more capacity, railway or air transport networks may be target of planned extensions. New infrastructure is often accompanied by new emissions of noise and pollutants and is thus subject to lenghty planning, negotiation, and high private and public costs [Bubalo and Daduna, 2012]. However, improvements on infrastructure may improve quality of journeys or offer even new possibilities of transportation. Identification and appraisal of these disadvantages and benefits is one of the key subjects in infrastructure planning.

Many commercial simulation tools for air traffic are available, e.g. SIMMOD[4], CAST[5], AirTOp[6], RAMSrams plus[7] or Total Airspace and Airport Modeler (TAAM)[8]. All of them provide high level of detail modelling of airports and airspace; some of them use multi-agent architectures for different actors of the scene, e.g. for airport controllers, air traffic manangement, etc. Also in research, simulation toolkits of a high level of detail are available, [e.g. see Bilimoria et al., 2000, Sweet et al., 2002, Alam et al., 2008]. All of them aim at detailed simulations of air traffic in order to improve air traffic management concepts. Neither commercial nor scientific simulation frameworks support agent-based modelling of individual passengers on all stages of a flight.

[1] `http://www.italotreno.it` last access 19.12.2012
[2] `www.tgv-europe.com/en/`, last access 11.09.2012
[3] `www.bahn.com`, last access 11.09.2012
[4] `www.airporttools.com`, last access 22.10.2012
[5] `www.airport-consultants.com`, last access 22.10.2012
[6] `www.airtopsoft.com`, last access 22.10.2012
[7] `www.ramsplus.com`, last access 22.10.2012
[8] `www.jeppesen.com/taam`, last access 22.10.2012

Queueing theory and queueing models are widely used to model the technology of air transportation systems. For example, Pyrgiotis et al. [2011] use queueing theory to model the propagation of delay through the network. Effects of new airspace management technologies are studied by Nikoleris and Hansen [2012]. The models for traffic flow on roads are usually more complex, e.g. "cell transmission models" [Daganzo, 1994, 1995], which model traffic based on discretized partial differential equations, "car following models" [Wiedemann, 1974, Gipps, 1981], which model traffic by following each car individually, typically using discretized time but continuous space, "cellular automata models" [Nagel and Schreckenberg, 1992], which are similar to car following models but also discretize space. Yet, all these models are computationally rather expensive. For that reason, also queue models are in use, which are computationally much faster [Gawron, 1998, Cetin et al., 2003, Cetin, 2005, Cremer and Landenfeld, 1998].

This paper uses a queue model. The model provides several parameters for an explicitly modelled segment of transport systems: The maximum flow that can pass a segment, the maximum amount of vehicles on the segment and a maximum velocity per segment or vehicle. Several segments can be connected via nodes, building a transport network, on which individual vehicles can be simulated. Segments are modelled as FIFO (first-in first-out) queues, nodes can be interpreted as servers. Thus, the modelling of the transport network is quite similar to queueing theory approaches in air transport [e.g. Pyrgiotis et al., 2011]. However, the proposed model is not solved analytically but by simulation. While analytical solveable models may conserve computational resources, a computational fast simulation model enables an agent-based modelling of every individual throughout the complete simulation lifecycle in complex scenarios.

The agent-based, dynamic modelling used in this paper enables a highly detailed view on important aspects of air transport systems. Aircraft are modelled as individual vehicles possessing attributes such as speed or available seats. Passengers are represented as individual virtual persons that want to perform activities at certain locations for a specific durations. Each person plans journeys between geospatial locations of her or his activities. The performance of individual planning is assessed by the joint simulation of aircraft and passengers. In an iterative procedure virtual persons learn the constraints of the transport system until an agent-based stochastic user equilibrium [Nagel and Flötteröd, 2012] is reached. As result one get data capturing the trace of every single agent that, aggregated appropriately, can be helpful to assess and improve the overall performance of medium distance travel options.

The paper is organized as follows. The next section introduces the simulation model and underlying theory. Then, Sec. 3 reviews shortly how air transport technology is represented. Sec. 4 describes how passengers are attached to the model and presents results for simulations covering trips within Germany. The paper ends with a discussion and conclusion.

2 Multi-Agent Transport Simulation

The simulation approach used in this paper is based on the software tool MATSim[9]. The next paragraphs provide an overview of the simulation approach and highlight the most important details used in this work. For more detailed information on technical aspects, please see Raney and Nagel [2006] or Balmer et al. [2005]. For a detailed discussion of methodology, see, e.g., Nagel and Flötteröd [2012]. Regarding economic concepts used in the simulation approach, see, e.g. Nagel et al. [2008], Kickhöfer et al. [2011].

2.1 Simulation Overview

In MATSim, each traveler of the real system is modeled as an individual virtual person. The approach consists of an iterative loop that has the following important steps:

1. **Plans generation**: All virtual persons independently generate daily *plans* that encode, among other things, their desired activities during a typical day as well as the transportation mode. Virtual persons typically have more than one plan ("plan database").
2. **Mobility Simulation**: All selected plans are simultaneously executed in a simulation of the physical system (often called "network loading" or "traffic flow simulation").
3. **Scoring**: All executed plans are scored by an *utility function* which can be personalized for every individual.
4. **Learning**: At the beginning of every iteration, some virtual persons obtain new plans by modifying copies of existing plans. This is done by several *modules* that correspond to the choice dimensions available, e.g. time choice, route choice, and mode choice. In this paper, time and route choice will be used. Virtual persons choose between their plans according to a Random Utility Model (RUM).

The repetition of the iteration cycle coupled with the plan database enables the virtual persons to improve (learn) their plans over many iterations. This is why it is also called **learning mechanism** which is described in more detail by Balmer et al. [2005]. The iteration cycle continues until the system has reached a relaxed state. At this point, there is no quantitative measure of when the system is "relaxed"; we just allow the cycle to continue until the outcome is stable.

In the steady state, the model is equivalent to the standard multinomial logit model

$$p_j = \frac{e^{\mu \cdot V_j}}{\sum_i e^{\mu \cdot V_i}} \,, \tag{1}$$

where p_j is the probability for plan j to be selected and μ is a sensitivity parameter, set to 2 for the simulations in this paper. In consequence, V corresponds to the systematic component

[9] Multi-Agent Transport Simulation, see `www.matsim.org`.

of utility in Random Utility Models (RUM) e.g. [Ben-Akiva and Lerman, 1985, Train, 2003], where utility is defined as $U = V + \varepsilon$. In RUM, the ε is called random component of utility. In the steady state and assuming a Gumbel distribution for ε, the choice model used in this paper is thus equivalent to the standard multinomial logit model.

2.2 Scoring

In order to measure the quality of a plan after execution and to compare plans, it is necessary to assign a quantitative score to the performance of each plan. For this purpose the utility function of the virtual persons is used. The total utility of a plan is computed as the sum of individual contributions:

$$V_{total} = \sum_{i=1}^{n} V_{perf,i} + \sum_{j=1}^{n} V_{tr,j} \ , \tag{2}$$

where V_{total} is the total utility for a given plan; n is the number of activities, which equals the number of trips (the first and the last activity are counted as one); $V_{perf,i}$ is the (positive) utility earned for performing activity i; and $V_{tr,j}$ is the (usually negative) utility earned for travelling during trip j. For calculation of $V_{perf,i}$ a logarithmic form is used

$$V_{perf,i}(t_{perf,i}) = \beta_{perf} \cdot t_{*,i} \cdot \ln\left(\frac{t_{perf,i}}{t_{0,i}}\right) \tag{3}$$

where t_{perf} is the actual performed duration of the activity, t_* is the "typical" duration of an activity, and β_{perf} is the marginal utility of an activity at its typical duration. β_{perf} is the same for all activities, since in equilibrium all activities at their typical duration need to have the same marginal utility. In this paper a β_{perf} of 6 utils is used. The (dis)utility of traveling is linear in travel time, i.e. $V_{tr,j}(t_{tr,j}) = \beta_{tr} \cdot t_{tr}$. In this work, β_{tr} is set to -6 for all virtual persons.

Further details on the default MATSim utility function can be found in Charypar and Nagel [2005] while one of the most recent discussions of this utility based approach is in Kickhöfer et al. [2011].

2.3 Mobility Simulation

The mobility simulation consists of a model of the physical environment, several agent-representations and a model for traffic flow. The physical environment comprises at least a model of the transportation network. Agent-representations exist for virtual persons, public transit vehicle drivers, traffic lights, etc. The traffic flow model is a queue model, that moves vehicles through the transportation network. Queue models for traffic flow disregard most of the details of vehicle movements on a road. Traffic networks are modelled as graphs. Each vertex models a crossing. Vertexes are connected by *links*, a directed edge that describes a road segment. Each link of a road network is described by the following attributes: maximum flow

(capacity) c_{flow}, length l, and the amount of vehicles that fit on the link $c_{storage}$ if cars stand bumper to bumper.

Vehicles entering a link have to stay on that link at least as long as they would travel at their desired velocity or as fast as the speed limit on the link permits. During this time no computation needs to be done, the vehicles are stored in a priority queue. Afterwards the vehicle is placed into another FIFO-queue. In each simulated timestep $floor(c_{flow})$ vehicles may leave the FIFO-queue plus one additional vehicle when the accumulation over the last timesteps of fractional part of c_{flow} is equal or greater than 1. If there is space available on the downstream link, i.e. the number of vehicles is less than $c_{storage}$, a vehicle is moved to the downstream link. This makes the model capable to model spill-back.

2.4 Modelling of Public Transit

The public transit module of MATSim aims at modelling microscopic public transit simulation with a focus on several types of ground transportation, e.g. buses, streetcars or para transit [Rieser, 2010]. In a transit schedule transit stop facilities, *lines* and routes are specified. Passengers can access and leave vehicles at transit stop facilities. Each transit line contains one or more transit routes. Transit routes specify the order in which stops are lined up to a route and the departure time of a vehicle at the beginning of the route. Furthermore each route specifies which links in the network are used to connect stop facilites.

Characteristics of transit vehicles are specified using the default configuration of the MATSim framework[10]. Several *vehicle type*s can be defined that contain information as length, width, passenger capacity, maximum velocity and energy consumption. How fast passengers can access and leave a vehicle is also specified via the vehicle type. In addition to the different vehicle types a set of particular vehicles can be defined. Each vehicle has exactly one type assigned and inherits all its attributes. The individual vehicles are inserted into the traffic flow simulation and moved by the queue model along their routes.

3 Modelling and Simulation of Air Transport Technology

This section focusses on the technology side of air transport networks. It is reviewed how airports and aircraft can be modelled microscopically using the simulation framework presented in Sec. 2. For a more detailled desciption, see [Grether et al., 2013].

[10] `http://matsim.org/files/dtd/vehicleDefinitions_v1.0.xsd`

3.1 Data Sources

Geospatial location of airports are retrieved by two data sources. OpenNav[11] is an online database of aeronautical navigation information featuring airport coordinates that may be retrieved with a web query based on the IATA airport code. Coordinates for those airports not available on openNav are prompted in the same manner from the Great Circle Mapper[12], which also includes a searchable database of airports. Worldwide, a total of 3670 airport coordinates for IATA codes is retrieved from these data sources.

The air traffic technology model takes advantage of data provided by OAG Aviation[13]. An OAG snapshot of worldwide direct flights in September 2009 is available for schedule generation. All flights with IATA airport codes, flight times, flight numbers and designators, aircraft types, available seats and distance between airport are gathered from the database and processed. Codeshares, multi-stop flights, buses and trains with flight numbers and cargo flights are filtered out of the schedule during the generation process.

3.2 Flight schedule

Relevent data for schedule and network generation is excerpted from the OAG data using all flights departing on a Tuesday, taking each specific flight number into account only once. This may not always result in complete flight cycles, e.g. when the outbound and inbound flight operate on different days of the week. Compared to using all flights of an entire week, the network may be incomplete, as certain destinations are only served on specific days.

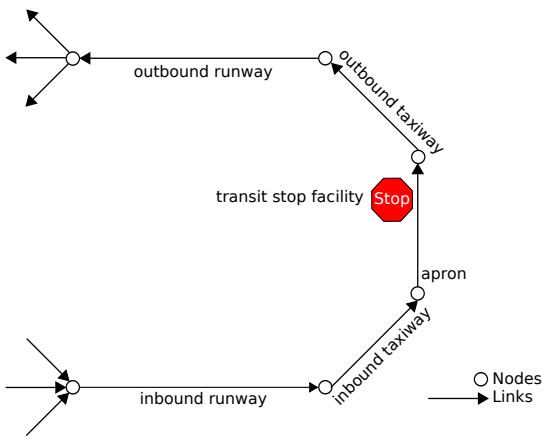
For matters of consistency all local times are converted into Universal Time Coordinated (UTC). This ensures aircraft taking off and landing at the scheduled times throughout all time zones and also enables the model to reflect incoming and outgoing waves at hub airports worldwide at the appropriate times.

The scenario used in this paper contains all Europe to worldwide, non-stop flights for that data is available completely. Airports for which no coordinates or UTC data were available were removed for the present study. For this scenario 16 airports are missing in our database while for the majority of 600 airports data is complete. For each airline that offers a connection between two available airports a flight (*transit line*) is generated. On 11252 O-D pairs 18716 flights are operated, while 575 flights had to be removed because of data inconsistencies.

3.3 Network and Airport Layout

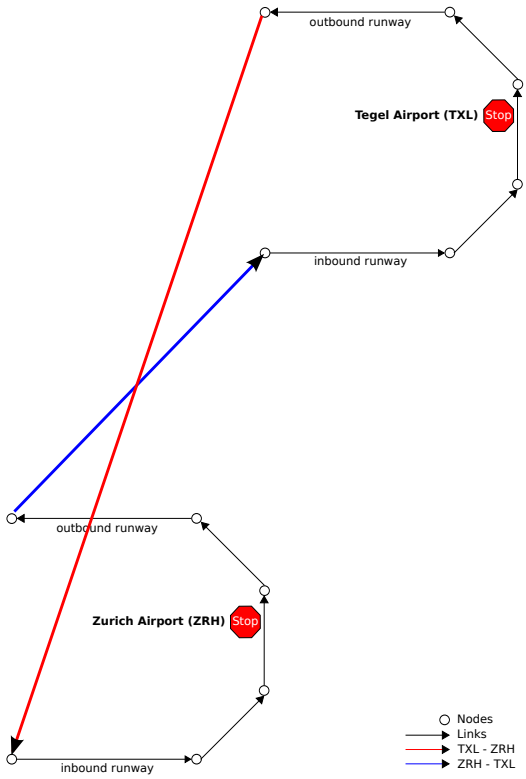
The air network consists of airports, each showing an identical layout, and point-to-point connections in between. Each runway may possess a restriction of flow capacity that is set

[11] `www.opennav.com`, last access 09.08.2012
[12] `www.gcmap.com`, last access 09.08.2012
[13] `www.oagaviation.com`, last access 08.08.2012



Linktype	Count	Length [m]	Speed [km/h]
Runway	2	1500	$length \cdot c_{flow}$
Taxiway	2	500	20
Apron	1	500	20

(a) Airport layout and characteristics



(b) Network layout

Figure 1: Air Network

to a unrealistic high value in this paper to keep influence of capacity restrictions exogeneous, see Grether et al. [2013] for a detailed study. But, not more than one aircraft can be simultaneously on a runway. Every runway is solely used either for inbound or outbound flights with taxiways connecting the runways to the apron. The latter accommodates a transit stop where flight movements originate or terminate (see Fig. 1a) and passengers may enter or leave aircraft.

Each airport pair is directly connected by airway links, one for each flight and direction of travel (see Fig. 1b). Thus, mutual interferences of aircrafts en-route are not included in the model. The maximum speed on any of these links is calculated based on the distance and flight duration provided by OAG. Times for taxi, take-off and landing are also taken into account, i.e. the flight duration is reduced by the time needed from push-back to airborne before the maximum speed for an airway link is calculated.

Fig. 2 shows the network for European air traffic.

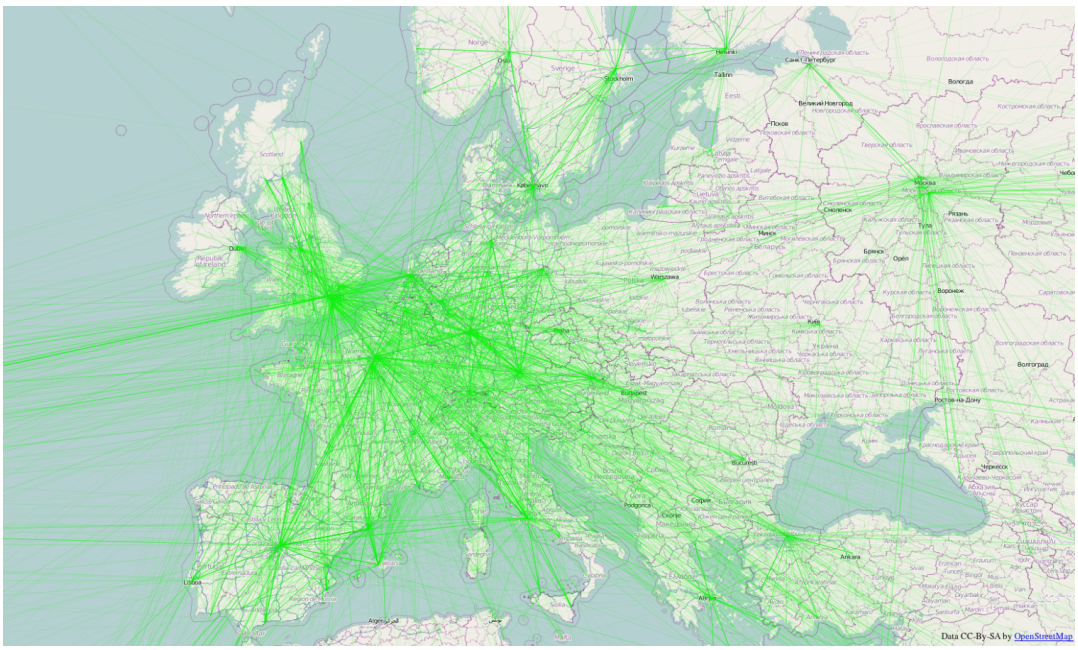


Figure 2: European air network with country boarders in the background (country boarders © `openstreetmap.org`)

3.4 Aircraft

Vehicles are created on the basis of OAG data to represent individual aircraft in the simulation. IATA aircraft codes, operating airlines and seating capacities are reflected in the respective aircraft representation for every flight. Information about boarding times, i.e. passenger flow per door over time, is not available but could be set for each aircraft type. One aircraft per flight is generated, thus delays resulting from a delayed incoming aircraft are not modelled. Accordingly, no aircraft rotations and vehicle trip chains are implemented for the time being. The maximum velocity of aircraft is set to twofold sonic speed, since speed limitations are set for each airway link of the network.

3.5 Simulation Results

The flight schedule, network and aircraft data serves as input for the mobility simulation. The multi-agent approach is, in general, particularly suitable to to model delays (primary and reactive). In this study all functions to simulate delay are switched off thus all flights are exactly on time. Fig. 3 shows the simulated number of aircraft departures and arrivals over time of day. Clearly, one can observe the morning departure peak between 05:00 am and 07:00 am UTC.

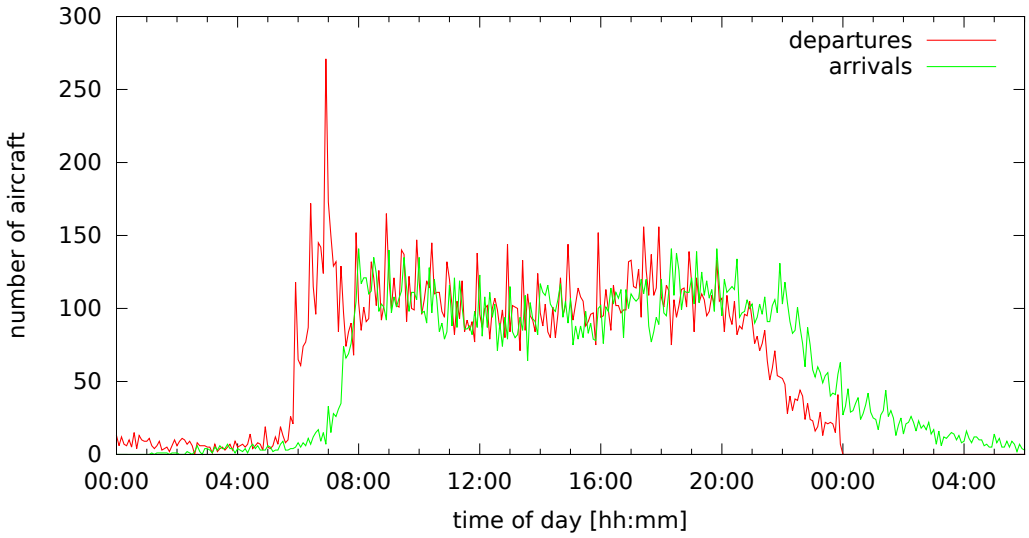


Figure 3: Arrivals and departures over time of day.

4 Modelling and Simulation of Passenger Demand for Air Transport

With the results from Sec. 3 an air transport technology model is available. This section shows how a passenger demand for air transport can be modelled with the multi-agent approach.

4.1 Data sources

There are many different ways in which passenger demand for transport systems can be generated e.g. [Balmer, 2007]. One option is to start with origin-destination (O-D) flows between geographical regions. In a European context, possible data-sources include OAG Aviation[14] and eurostat[15]. They provide data about passengers; O-D flows, however, are not provided. Data-sources geographically limited to Germany as "Der Flughafenverband"[16] or ITP/BVU [2005] do not come with O-D data, neither. The latter may have O-D relationships available in an upcoming version. The German Institute of Air Transport and Airport Research (DLR) provides monthly statistics containing O-D flows,[17] but the pdf format provided is not

[14] `www.oagaviation.com`, last access 08.08.2012
[15] `ec.europa.eu/eurostat`, last access 10.09.2012
[16] `www.adv.aero`, last access 10.09.2012
[17] `http://www.dlr.de/fw/en/desktopdefault.aspx/tabid-2961/9753_read-19683/`, last access 10.09.2012

suited for machine reading, and data is only available up to 09/2010. DESTATIS[18] provides O-D data by airport for German air traffic in a machine readable format. Data is available for whole years or a specific month. DESTATIS data is thus used in the following to create an agent based air transport demand for Germany.

4.2 Passenger Demand

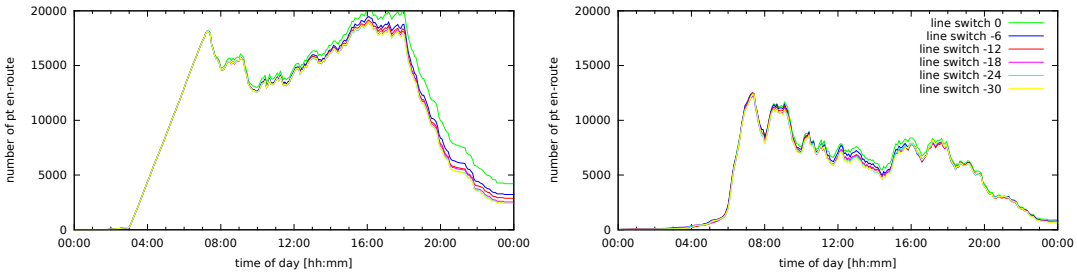
The passenger demand is based on the DESTATIS data for 09-2011 in order to be consistent with Sec. 3. DESTATIS provides data in two different representations (data sets 2.2.1 and 2.2.2). The number of O-D trips between airports is captured in two different ways. For all pairs of airports, the number of direct trips between the airports is given in the data set 2.2.1. Furthermore, the second data set 2.2.2 contains O-D pairs that do not include the first transfer, but provide the second, and possibly final destination. E.g. one person flying from Hamburg (HAM) via Frankfurt (FRA) to Munich (MUC) is contained in the data as one O-D pair: HAM $\rightarrow$ MUC. If a flight starts at Paris (CDG) going via FRA via MUC to HAM it is not clearly stated how the flight is represented in the data. It might be counted as CDG $\rightarrow$ HAM or FRA $\rightarrow$ HAM O-D relation. It is, however, unlikely that passengers will have two *transfer* stops within Germany. Thus an origin or destination abroad may not be the original or final destination, but at least all passenger movements that touch Germany along their itinery are probably included in an unequivocal way.

The second data set (2.2.2) is used to create the virtual persons for the passenger demand. For each O-D pair the number of trips is scaled from monthly to daily values by a division by 30. for each O-D pair and trip a virtual person is created. The resulting synthetic population contains 65251 virtual persons, 1304 trips from the original data are neglected as origin and destination are equal. Each virtual person performs two activities, one at the origin and the other at the destination airport. Both activities are of same type, thus time spend performing both activities is accumulated before it is evaluated by the utility function according to Eq. (3). A "typical duration" (t_*) of 21 h is set for this activity type. In between the two activities a flight leg is scheduled, connecting origin and destination. As is common, the demand does not specify if a direct flight from O to D is chosen or the virtual person is on a route containing one or more transfers. The time virtual persons arrive at the origin airport and start waiting for a connection is drawn randomly from a uniform distribution in 04:00 to 18:00, UTC. This reflects estimated typical opening hours of airports in Europe.

4.3 Simulation Setup

The synthetic population is used as input for the simulation. As scenario the European flight model with no delays and no effective runway capacity restrictions from Sec. 3 is used. The assignment of concrete flights to the desired O-D connection, i.e. the passenger routing, is done

[18] `destatis.de`, Fachserie 8 Reihe 6, last access 10.09.2012



(a) After 0th iteration (b) After 600th iteration

Figure 4: Travelers en-route, i.e. waiting for a flight or travelling by plane, as a function of the time-of-day

by the default public transit routing module of MATSim [Rieser, 2010]. The routing basically looks for a least cost path in terms of travel time. The network used for routing is constructed from the information contained in the transit schedule. In order to penalize transfers, the routing assumes an additional cost of $c_{lineswitch}$ for each transfer. The same parameter is also considered by the scoring function, i.e. a (dis-)utility of $-c_{lineswitch}$ is added to the score of the agent for each transfer. The simulation is run several times using different values of the $c_{lineswitch}$ parameter, i.e. $c_{lineswitch} \in \{0, -6, -12, -18, -24, -30\}$[utils/transfer].

The simulation is run for 600 iterations. In each iteration, 10 % of the virtual persons may shift their departure time randomly within a 2 h interval. The amount of shift is drawn from a uniform distribution. Another 10 % may seek a different route, i.e. a different connection between origin and destination. Each passenger chooses out of a set of 5 plans using a multinomial logit model, see Nagel et al. [2008] for details. The outcome is stable after 500 iterations, thus departure time choice and routing are switched off. For another 100 iterations only the logit model is used by the passengers to select a plan. Empirically, fixing the choice set for the last 100 iterations reduces the noise of learning and eases analysis and intepretation of results. All other parameters used for simulation are the "default" values of the MATSim framework. For a detailed discussion, see, e.g., Nagel et al. [2008], Nagel and Flötteröd [2012].

One iteration takes around 10 min. on an Intel Xeon Processor (2.67 GHz) using one core for the execution of mobility simulation and two cores for the replanning modules.

4.4 Results

First, in order to show the effects of routing, the result after the zeroth simulated iteration is presented. Each virtual person gets a connection assigned based on a generalized cost routing for the connection and the preset departure time. Fig 4a shows the number of travellers en-route, i.e. waiting for a flight or travelling by plane, as a function of the time-of-day. Some passengers are still waiting for a flight at midnight. As only one day of operation is simulated, these

$c_{lineswitch}$	σ^2	σ	mean rel error	stuck
$od_{transfer} - od_{direct}$	6715	82	1.56	-
-0	5248	72	0.55	1534
-6	5586	75	0.53	1469
-12	5713	76	0.65	1448
-18	5777	76	0.63	1480
-24	5785	76	0.62	1458
-30	5810	76	0.61	1456

Table 1: Simulation results for different values of $c_{lineswitch}$

passengers are stuck and will not reach their destination. The number of stuck passengers is decreasing with the increasing disutility of line switch.

The output after 600 iterations is depicted in Fig. 4b. The shape of all curves is different from the shape of the 0th iteration. One can identify two morning and two evening peaks. Some passengers still get stuck at the end of the day, but fewer than in the 0th iteration. In addition, the differences between the curves for the $c_{lineswitch}$ parameter are diminishing.

In order to study the influence of the $c_{lineswitch}$ parameter, the simulation results are compared with the input data. Recall that the synthetic population is generated based on O-D pairs that may contain transfers ($od_{transfers}$), while other data directly counts the number of passengers on actual direct flights (od_{direct}). The latter is used to evaluate the accuracy of the model. For comparison, the number of passengers on direct flights is thus calculated for each O-D pair (sim_{direct}) from the simulation results.

Based on these datasets, the mean square error σ^2 is computed as

$$\sigma^2 = \frac{\sum_{i \in OD} (sim_{direct}(i) - od_{direct}(i))^2}{|OD|} .$$

The (unsigned) mean relative error for each O-D relation is calculated as

$$\text{mean rel error} = \frac{\sum_{i \in OD} |(sim_{direct}(i) - od_{direct}(i))| / od_{direct}(i)}{|OD|}.$$

Tab. 1 shows the results for these calculations. The first line contains the comparison of the two input data sets from DESTATIS, i.e. in the above formulas sim_{direct} is replaced by $od_{transfers}$. This serves as reference as it would assume that *all* demand is served by direct flights. All simulation runs explain the data better than that reference. The values for all simulation runs are then quite similar.

The last column of tab. 1 shows the number of passengers stuck at the end of day. Values for all parameter settings are around 1500 passengers, i.e. around 2 % of the 65'251 simulated passengers.

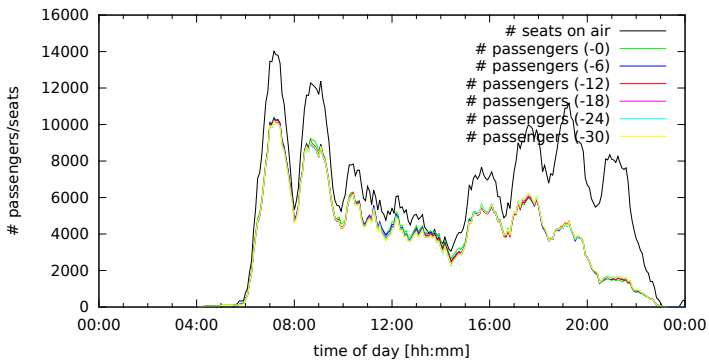


Figure 5: Passengers and available seats over time within Germany

4.5 Passengers Stuck

Overall, the results show that a microscopic, agent-based simulation of passenger demand for air transport is feasible. Most passengers are able to learn the constraints of air transport technology and arrive at their desired destination.

Some passengers fail to reach their destination; they get "stuck". As only trips within Germany are modelled, which are usually completed within a few hours without any requirement for an overnight stay at an airport, this is considered unrealistic. Fig. 5 shows that this is not a consequence of a general lack of seats: at any time-of-day, there are more seats than demand. There are many reasons why stuck passengers can arise in such a situation. Further analysis of the simulation results leads to the following insights for the $c_{lineswitch} = 0$ scenario:

- 813 passengers are stuck because there is no seat, *and* there is no other flight by the same airline later during the day to which they would be shifted otherwise.
- 721 passengers are stuck at an airport because there is no connection after their departure time between that airport and their destination airport.

In order to study the influence of departure time on available connections, several simulations are run that set the departure time of each passenger being stuck to 04:00 UTC, i.e. before the first aircraft is departing. Simulation results produced similar findings as presented above.

Thus, it is worth looking more closely at the relation between agents being stuck and the capacity of seats offered for each O-D pair. For each O-D pair one can obtain the number of travelers that plan to travel from O to D, i.e. the demand. Furthermore, the number of seats offered on that O-D pair can be retrieved from simulation input data. Fig. 6 plots the number of travelers that are stuck on their planned O-D connection over the difference between seats offered and demand. In order to improve visibility Fig. 6 is cut values where available seats increase demand by more than 800 - the number of stucked persons is always 0. Apparently, passengers get more likely stuck the more the requested demand is equal or greater than overall capacity.

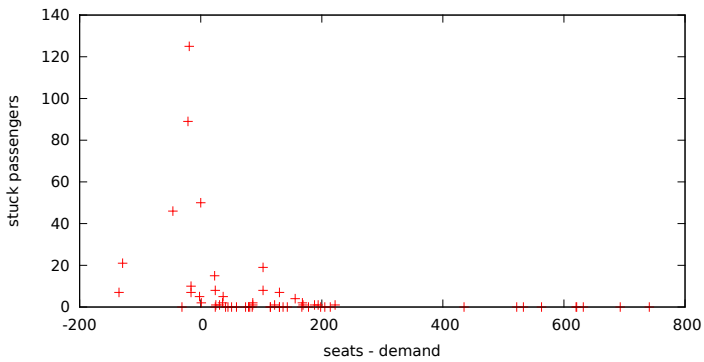


Figure 6: Correlation between the available seats, the demand for seats and the number of passengers being stuck

4.6 Adding an Alternative Mode

To gain further insights, in the following a slightly different simulation setup is used. The additional cost for each transfer is fixed to $c_{lineswitch} = 0$ and has no influence on the model. Instead, a second option for mode choice is added. Each virtual person can now choose between the micro-simulated air transport options and an alternative mode. The alternative mode has no capacity restrictions. Furthermore, passengers that travel with the alternative mode can start directly at their desired departure time. The travel time t is computed by the microsimulation with an estimation of the beeline distance between the O-D pair d and a velocity v, i.e. $t = d/v$. This velocity is varied in several simulation runs, i.e. $v \in \{100, 150, 200, 250, 300\}[km/h]$. If the alternative mode is chosen, the (dis-)utilities for travelling in the scoring are calculated accordingly.

Each person in the synthetic population obtains a second plan that uses the alternative mode. With this population the simulation is again run for 600 iterations. Like in the previous simulations 10 % of the virtual persons may shift their departure times while another 10 % seek a different route between origin and destination in the air transport network. Additionally, further 10 % of virtual persons may change mode, i.e. they can switch between the air traffic mode or the alternative mode. After 500 iterations all choice modules are switched off, thus for the last 100 iterations the logit model is used by by passengers to select one of their plans.

From the output of the 600th iteration the same numbers as for the previous simulation runs are calculated and depicted in Tab. 2. If the speed of the alternative mode is 100 or 150 km/h the mean square error is quite similar to the previous results while the mean relative error is even less. The number of stuck passengers however is remarkable reduced from approx. 1500 to 185 or even 69. Alternative mode speeds higher than 150 km/h further reduce the number of stuck passengers while the relative error is quite similar. In contrast, the mean square error is increasing the higher the speed for the alternative mode is set.

Effects of the speed increase on the modal split are shown in Tab. 3. While for a $v = 100$ km/h

$v[km/h]$	σ^2	σ	mean rel error	stuck
$od_{transfer} - od_{direct}$	6715	82	1.56	-
100	5380	73	0.40	185
150	5605	75	0.39	69
200	6334	80	0.38	33
250	7580	87	0.43	29
300	11239	106	0.37	9

Table 2: Simulation results for different values of v

the alternative mode is used by 3.59 % of the passengers only, a mode alternative with a speed of 300 km/h would attract 17.09 % of travelers.

$v[km/h]$	# air mode	# alt. mode	# stuck	# total	air mode [%]	alt. mode [%]	stuck [%]
100	62726	2340	185	65251	96.13	03.59	00.28
150	61379	3803	69	65251	94.07	05.83	00.11
200	59491	5727	33	65251	91.17	08.78	00.05
250	57248	7974	29	65251	87.74	12.22	00.04
300	54089	11153	9	65251	82.89	17.09	00.01

Table 3: Modal split for different train speeds

Temporal effects between the two modes are also illustrated looking at speeds of 100 km/h and 300 km/h for the alternative mode. Fig. 7 shows the passengers over time for both simulation runs per mode. One can observe that passengers using air transport follow the time distribution of the offered capacity. In contrast, users of the alternative mode are rather equally spread over time of day. This is plausible considering the setup of simulation: Passengers have no time constraints that force them to arrive at a certain time at their destination. Departure times are equally distributed between 04:00 and 18:00, UTC and then randomly distributed over the iterations. As the alternative mode is always available there is no constraint within the model that ties passengers to any departure time.

5 Discussion

Overall, the results show that a microscopic, agent-based simulation of passenger demand for air transport is feasible. Most passengers are able to learn the constraints of air transport technology and arrive at their desired destination.

Some passengers get stuck at the end of the day. But, the number of stuck passengers depends on the setup of the model.

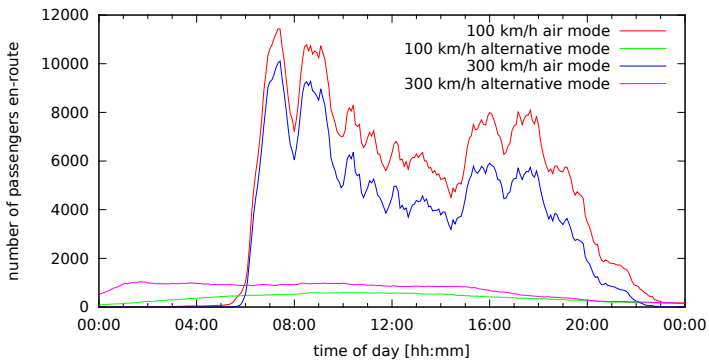


Figure 7: Passengers over time by mode

5.1 Air Transport Only

Without an alternative mode of transport the number of passengers that get stuck is higher than in the case an alternative mode is provided. The only available transport mode is a capacity restricted flight connection that is served in discrete, irregular time intervals. The results show that passengers get more likely stuck on O-D pairs where requested trips reach the number of available seats. This may have model extrinsic and intrinsic reasons.

The choice and quality of available data sources is extrinsic to the modelling approach. Recall that a flight schedule for 09-2009 is used in conjunction with a transport demand for 09-2011. The number of starts of flights within Germany increased slightly between 2009 and 2011 [DLR, 2012, p. 23]. Assuming that the number of available seats is increasing accordingly, the simulation model provides too little capacity. If simulation is setup on consistent historical input data for a specific day, we would request the model to provide a solution in that no passenger get stuck. However, such detailed data is not available.

Furthermore, the problem may be intrinsic to the model. It may, for example, be possible that passengers depart from an origin that only has one –early– connection to a hub per day, and the passengers' departure times are too late to reach that connection, and the random departure time mutation may not be able to find that connection for all passengers. This has been ruled out for the current setup but should be considered in further studies.

Alternatively, it may be the case that passengers have a connection that works in theory, but they are "crowded out" by other passengers who arrive earlier at the gate. They would make it if either of them would take a different route. The current approach would not find such a solution, since passengers do not take into account the costs they impose on others, see Lämmel and Flötteröd [2009] for an approach to take that into account. The real-world solution presumably would be to raise prices on congested seats until one or the other passenger re-routes. The present model does not (yet) include such a process.

An alternative approach to remove some of these shortcomings might be to use a router that generates a larger diversity of routes even for the same departure time. Such a router would be

able to point a passenger to a route where seats are available without by itself knowing about seat availability. That approach would, however, not address the issue that some passengers might need to switch their path in order to allow *others* to obtain a feasible path.

5.2 Alternative Mode

Adding an alternative mode makes the model more plausible in terms of a demand for transport that can be served by a given network. Choice between the alternative mode and the microsimulated air transport mode is consistent with the overall logit assumption, see [Rieser et al., 2009]. Clearly the results hinge at the assumption that the alternative mode is always available and not capacity restricted. The alternative mode can be interpreted as mixture between train, bus or car connection availability. In principle, these alternatives can also be modelled explicitly featuring capacity restrictions and mutual interactions from overall mode choice. We expect that passengers get also stuck when a capacity restricted bus or train connection is used near its capacity. Thus same argumentation applies as for the scenario with air transport only, a better routing or inclusion of prices and costs may improve the model.

In this study the modelling of the alternative mode is rather coarse. All passengers on the alternative mode face the same travel speed. This assumption is too simple for the presented scenario as e.g. average speed of train connections depends on the O-D pair. In principle, the alternative mode could be refined by inclusion of O-D pair dependent average speed data. For illustration of the overall modelling approach, however, a homogeneous velocity for the alternative mode seems to be more appropriate.

Overall, the inclusion of a not capacity restricted alternative mode improves the robustness of the model while staying consistent with existing theory.

5.3 Overall approach

Both modelling approaches can explain the routing in more detail than it can be solely retrieved from the input data. The quantity of reaction, however, seems to be relatively small. Most O-D pairs in the data are served by a direct connection. Considering the geospatial extent of the chosen scenario this is highly plausible. Flying within Germany is mostly not worth it if the connection includes a transfer. Then, it is empirically faster to travel by train, car or bus. Probably, the effects of the model would be better visible if a bigger geospatial extent, e.g. all Europe, would be simulated.

Results show that some departure time structure evolves due to the availability of air transport at certain times of day. Passengers, however, are modelled without explicit desired departure or arrival times. The simulation approach could capture such individual time constraints. Input data for this study, however, contains monthly O-D pairs without any further information about time distribution. We assume that results would improve if data used for this study is refined by additional information describing individual passengers in more detail. This is not limited

to time structures, also more detailed information for activity locations and price sensitivities could be attached to each individual traveler.

Clearly, potential applications of the proposed model depend on type and detail of included information. In general, application in the public sector allows a detailed evaluation of the effects from mid-distance travel policies that includes consideration of mode alternatives. The approach could also be useful for private companies while planning flight-schedules and capacities on distinct connections. The impacts of changes on customers can be assessed on high level of detail.

6 Conclusion

In this paper, an agent-based modelling approach for air transport systems is proposed. An iterative assignment model from urban transport planning is used. Aircraft are represented microscopically featuring attributes as speed, available seats and boarding constraints. The air traffic network as well as flight performance is modelled at a low level of detail as the model is not intended for air traffic management. The computationally affordable simulation technique enables an agent-based modelling of passenger demand and its iterative assignment to flights.

Results are presented for a simulation of German national air transport demand that is assigned on an Europe to worldwide air transport network. The results reveal some potential problems for agent-based air transport assignment models and discuss several solutions. Overall, the model can be used to simulate, forecast and assess changes in air transport systems on a high level of detail.

Models for other modes as rail or car transportation are the subject of current work, following the same approach as the one presented here. The same software and solution procedure is used. In consequence, these models can be integrated into the approach presented in this paper. This might help to get a more detailed picture of middle distance traffic.

Acknowledgements

For the model and simulation approach proposed in this paper, several prototype studies have been conducted [Fürbas, 2011, Böhme, 2011, Treczka, 2012, Krüger, 2012]. These prototypes were then improved and developed further; the paper describes the overall result. Further acknowledgements to all providers of data used in this study. Without accessible data suited for machine processing this work would not have been possible.

References

S. Alam, H.A. Abbass, and M. Barlow. Atoms: Air traffic operations and management simulator. *Intelligent Transportation Systems, IEEE Transactions on*, 9(2):209–225, 06 2008. ISSN 1524-9050. doi: 10.1109/TITS.2008.922877.

M. Balmer. *Travel demand modeling for multi-agent transport simulations: Algorithms and systems*. PhD thesis, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland, 2007.

M. Balmer, B. Raney, and K. Nagel. Adjustment of activity timing and duration in an agent-based traffic flow simulation. In H.J.P. Timmermans, editor, *Progress in activity-based analysis*, pages 91–114. Elsevier, Oxford, UK, 2005.

M. Ben-Akiva and S. R. Lerman. *Discrete choice analysis*. The MIT Press, Cambridge, MA, 1985.

K. Bilimoria, B. Sridhar, G. Chatterji, K. Sheth, and S. Grabbe. FACET: Future ATM Concepts Evaluation Tool. 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy, 06 2000. URL `http://www.aviationsystemsdivision.arc.nasa.gov/publications/2000/Bilimoria_ATM2000.pdf`.

Tobias Böhme. Analyse und Simulation von Verspätungen im europäischen Linien-Luftverkehr. Bachelor's thesis, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany, October 2011.

Branko Bubalo and Joachim Daduna. Airport capacity and demand calculations by simulation–the case of Berlin-Brandenburg International Airport. *NETNOMICS*, pages 1–21, 2012. ISSN 1385-9587. doi: 10.1007/s11066-011-9065-6.

N. Cetin. *Large-Scale parallel graph-based simulations*. PhD thesis, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland, 2005.

N. Cetin, A. Burri, and K. Nagel. A large-scale agent-based traffic microsimulation based on queue model. In *Proceedings of the Swiss Transport Research Conference (STRC)*, Monte Verita, CH, 2003. See www.strc.ch. Earlier version, with inferior performance values: Transportation Research Board Annual Meeting 2003 paper number 03-4272.

D. Charypar and K. Nagel. Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32(4):369–397, 2005.

European Commission. Airport policy in the European Union – addressing capacity and quality to promote growth, connectivity and sustainable mobility. COM(2011) 823 final, 12 2011. URL `http://ec.europa.eu/transport/air/airports/airports_en.htm`.

M. Cremer and M. Landenfeld. A mesoscopic model for saturated urban road networks. In M. Schreckenberg and D.E. Wolf, editors, *Traffic and Granular Flow '97*, pages 169 –180. Springer, Singapore, 1998.

C.F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research B*, 28B(4):269–287, 1994.

C.F. Daganzo. The cell transmission model, part II: Network traffic. *Transportation Research B*, 29B(2):79–93, 1995.

DLR. Luftverkehrsbericht 2011 – Daten und Kommentierungen des deutschen und weltweiten Luftverkehrs. Technical report, Deutsches Zentrum für Luft- und Raumfahrt e.V. – DLR, 2012.

Sören Fürbas. Modellierung und Simulation des europäischen Linien–Luftverkehrs. Bachelor's thesis, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany, 03 2011.

C. Gawron. *Simulation-based traffic assignment*. PhD thesis, University of Cologne, Cologne, Germany, 1998. URL `www.zaik.uni-koeln.de/AFS/publications/theses.html`.

P. G. Gipps. A behavioural car-following model for computer simulation. *Transportation Research B*, 15:105–111, 1981.

D. Grether, S. Fürbas, and K. Nagel. Agent-based modelling and simulation of air transport technology. In *The 2nd International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications*, 2013.

ITP/BVU. Prognose der deutschlandweiten Verkehrsverflechtungen 2025. Technical report, 2005. URL `http://daten.clearingstelle-verkehr.de/220/`.

B. Kickhöfer, D. Grether, and K. Nagel. Income-contingent user preferences in policy evaluation: application and discussion based on multi-agent transport simulations. *Transportation*, 38:849–870, 2011. ISSN 0049-4488. doi: 10.1007/s11116-011-9357-6.

Philip Krüger. Verspätungen im Flugverkehr: Modellerweiterung für die Simulation des europäischen Linien-Luftverkehrs. Studienarbeit, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany, 07 2012.

G. Lämmel and G. Flötteröd. Towards system optimum: Finding optimal routing strategies in time-tependent networks for large-scale evacuation problems. volume 5803 of *LNCS (LNAI)*, pages 532–539, Berlin Heidelberg, 2009. Springer. doi: 10.1007/978-3-642-04617-9_67.

K. Nagel and G. Flötteröd. Agent-based traffic assignment: Going from trips to behavioural travelers. In R.M. Pendyala and C.R. Bhat, editors, *Travel Behaviour Research in an Evolving World – Selected papers from the 12th international conference on travel behaviour research*, chapter 12, pages 261–294. International Association for Travel Behaviour Research, 2012. ISBN 978-1-105-47378-4. URL `http://www.lulu.com/shop/ram-m-pendyala-and-chandra-r-bhat/travel-behaviour-research-in-an-evolving-world/hardcover/product-18838195.html`.

K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I France*, 2:2221–2229, 1992.

K. Nagel, D. Grether, U. Beuck, Y. Chen, M. Rieser, and K.W. Axhausen. Multi-agent transport simulations and economic evaluation. *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, 228(2+3):173–194, 2008. Special issue on agent based models for economic policy advice, edited by B. LeBaron and P. Winker.

Tasos Nikoleris and Mark Hansen. Queueing models for trajectory-based aircraft operations. *Transportation Science*, 2012. doi: 10.1287/trsc.1120.0411. URL `http://transci.journal.informs.org/content/early/2012/04/05/trsc.1120.0411.abstract`.

Nikolas Pyrgiotis, Kerry M. Malone, and Amedeo Odoni. Modelling Delay Propagation within an Airport Network. *Transportation Research Part C: Emerging Technologies*, (0):–, 2011. ISSN 0968-090X. doi: 10.1016/j.trc.2011.05.017. URL `http://www.sciencedirect.com/science/article/pii/S0968090X11000878`.

B. Raney and K. Nagel. An improved framework for large-scale multi-agent simulations of travel behaviour. In P. Rietveld, B. Jourquin, and K. Westin, editors, *Towards better performing European Transportation Systems*, pages 305–347. Routledge, London, 2006.

M. Rieser. *Adding transit to an agent-based transportation simulation concepts and implementation*. PhD thesis, TU Berlin, 2010. Also VSP WP 10-05, see www.vsp.tu-berlin.de/publications.

M. Rieser, D. Grether, and K. Nagel. Adding mode choice to a multi-agent transport simulation. *Transportation Research Record: Travel Demand Forecasting 2009*, 2132:50–58, 2009. doi: 10.3141/2132-06.

D. Sweet, V. Manikonda, J. Aronson, K. Roth, and M. Blake. Fast-Time Simulation System for Analysis of Advanced Air Transportation Concepts. In *AIAA Modeling and Simulation Technologies Conference and Exhibit*, number AIAA-2002-4593, Monterey, CA, 08 2002.

K. Train. *Discrete choice methods with simulation*. Cambridge University Press, 2003.

Christopher Manuel Treczka. Modellierung und Simulation der Luftverkehrsnachfrage im deutschen Luftraum. Bachelor's thesis, TU Berlin, Institute for Land and Sea Transport Systems, Berlin, Germany, 01 2012.

R. Wiedemann. Simulation des Strassenverkehrsflusses. Schriftenreihe Heft 8, Institute for Transportation Science, University of Karlsruhe, Germany, 1974.