

## CHAPTER 68

# European Air- and Rail-Transport

Dominik Grether

This chapter discusses simulation of air- and rail-transport technology and passengers using MATSim. There is no great difference in overall travel times between middle-range rail and air transportation. Airports and railway stations are affected by capacity and opening time constraints. For passengers and goods, geospatial location is an important property. Both modes, but especially air transport, are faced with difficult capacity restrictions at certain departure times.

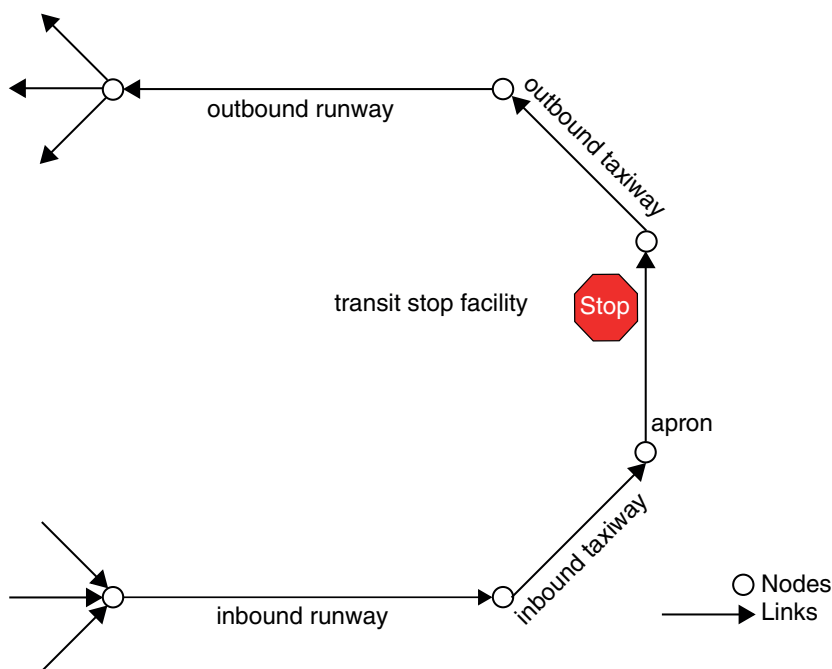
This chapter discusses how MATSim can be applied to capture these constraints and how interaction between passenger demand and constraints on technology supply can be modeled. The public transit model of MATSim (Chapter 16) is applied. Airports and aircraft are microscopically modeled the same way as bus stops and buses. Passengers are represented microscopically as multi-agent demand for air transportation. Their choices of transport mode, routes, and departure time are restricted by the air transport technology simulation model's capacity. The modeling of rail transport is based on teleportation. With appropriate data, the modeling approach for air transport could also be applied to rail transport (Quick, 2012).

The modeling of technology and demand is sketched in Section 68.1. On the basis of simulation results for a pure air transport model, rail transport is added and effects of mode choice are presented (Section 68.2). Section 68.3 then interprets simulation results and highlights some modeling aspects requiring further study. The choice set generation and plans removal algorithm of MATSim is discussed in detail; that is also the subject of Section 97.4. Modeling, results, and studies of this chapter present the highlights of Grether (2014, Chapter 6, pp. 119), in more detail.

---

### How to cite this book chapter:

Grether, D. 2016. European Air- and Rail-Transport. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) *The Multi-Agent Transport Simulation MATSim*, Pp. 419–428. London: Ubiquity Press. DOI: <http://dx.doi.org/10.5334/baw.68>. License: CC-BY 4.0



**Figure 68.1:** Layout of airports in the air transport network: In- and outbound runways are modeled by separate links connected by taxiways and a link representing the apron. There the transit stop facility is attached.

Source: Grether et al. (2013)

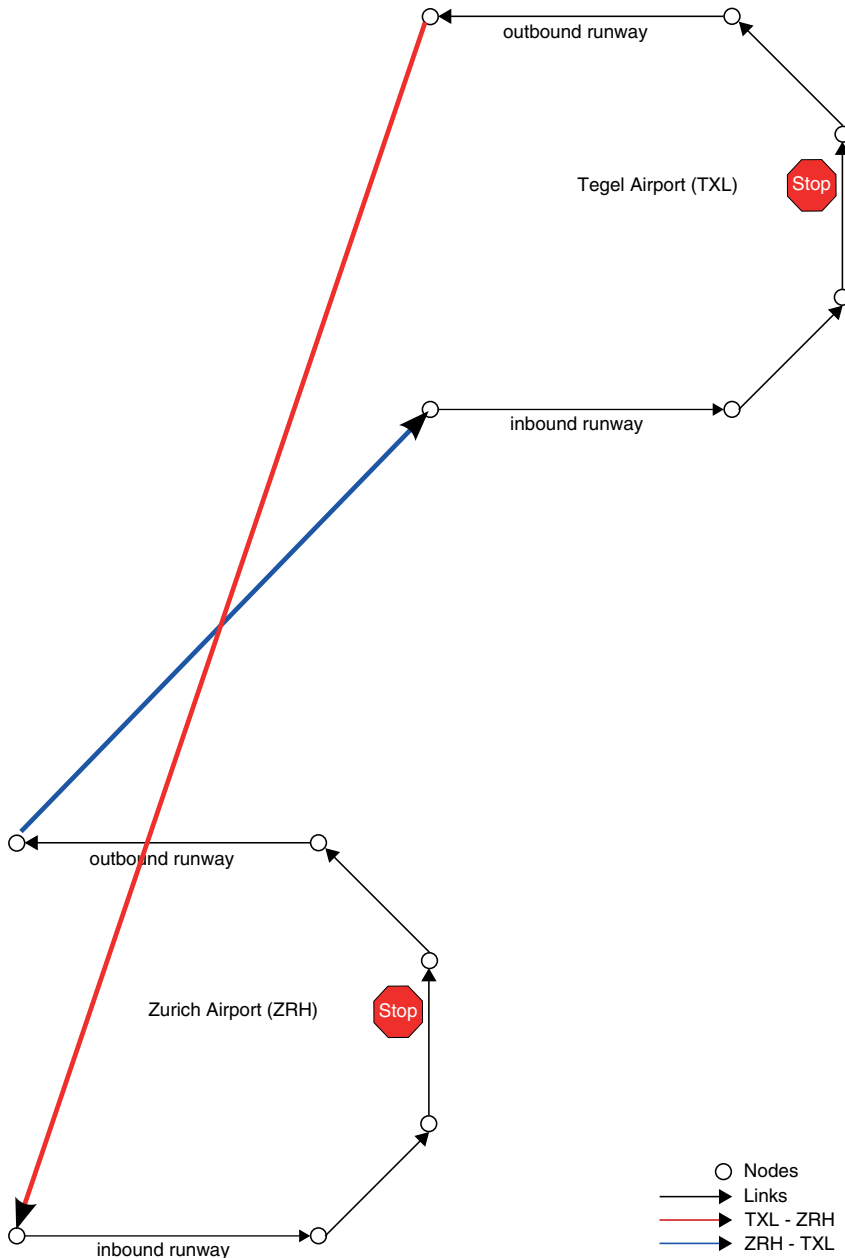
## 68.1 Air Transport Scenario

### 68.1.1 Modeling & Simulation of Air Transport Technology

The air traffic technology model uses data provided by OAG Aviation.<sup>1</sup> Relevant data for schedule and network generation is taken from the September 2009 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.

The air network modeling aims at a simulation with MATSim. The network consists of airports, each showing an identical layout and point-to-point connections in between. 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, i.e., the terminal, where flight movements originate and terminate (Figure 68.1). Each airport pair is directly connected by airway links, one for each flight and direction of travel (Figure 68.2). Maximum speed on any of these links is calculated based on distance and flight duration provided by OAG. Times for taxi, take-off, and landing are also taken into account, i.e., flight duration is reduced by the time needed from push-back to airborne before the maximum speed for an airway link is calculated. Each flight has an individual link that could be interpreted as route, each possessing individual characteristics. Figure 68.3 shows parts of the network for European air traffic.

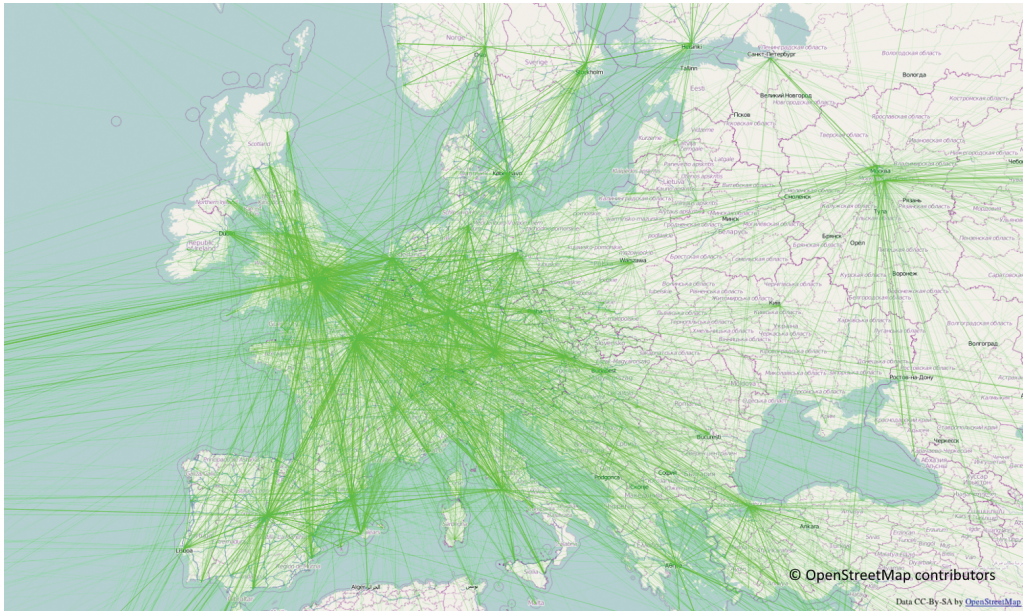
<sup>1</sup> <http://www.oagaviation.com>, last access 08.08.2012



**Figure 68.2:** Layout of airways in the air transport network: each airport pair is directly connected by two airway links, one for each flight and direction.

Source: Grether et al. (2013)

Flight schedules are taken from the OAG data and translated to a MATSim transit schedule containing information about each line, route, and departure. For each airline offering a connection between two airports, a transit line is generated. A transit route, which represents the route on the air traffic network, is created for each flight offered by this airline. Mutual interferences of aircrafts en-route are not included in the studies presented in this chapter.



**Figure 68.3:** European air network with country borders in the background (country borders © <http://www.openstreetmap.org>).  
 Source: Grether et al. (2013)

To represent individual aircraft in the simulation, transit vehicles are created on the basis of OAG data. 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 modeled. Accordingly, no aircraft rotations and vehicle trip chains are implemented at this time. The maximum velocity of each aircraft is set to twice sonic speed, since speed limitations are set for each network airway link.

### 68.1.2 Passenger Demand

As soon as the technology side of air transport is modeled, passenger demand simulation can begin. The passenger demand for trips in Germany created and used for the results of this section is based on O-D data of DESTATIS.<sup>2</sup> For each O-D pair and trip a virtual person is created. 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 spent performing both activities is accumulated before it is evaluated by the utility function according to Section 3.2. A typical duration,  $t_{typ,q}$ , of 21 hours is set for this activity type. The time virtual persons arrive at the origin airport and start waiting for a connection is drawn randomly from a uniform distribution in 4 am to 6 pm, UTC. This reflects estimated typical opening hours of European airports. No other time constraints are set, thus the only incentive for virtual persons is to reduce overall travel time and maximize time spent at the activity. A flight leg is scheduled between the two activities, connecting origin and destination. As

<sup>2</sup> Deutsches Statistisches Bundesamt, <http://www.destatis.de>, Fachserie 8 Reihe 6, last access 10.09.2012

usual, 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 synthetic population contains 51 832 virtual persons, 1 550 trips from the original data are neglected as origin and destination are equal.

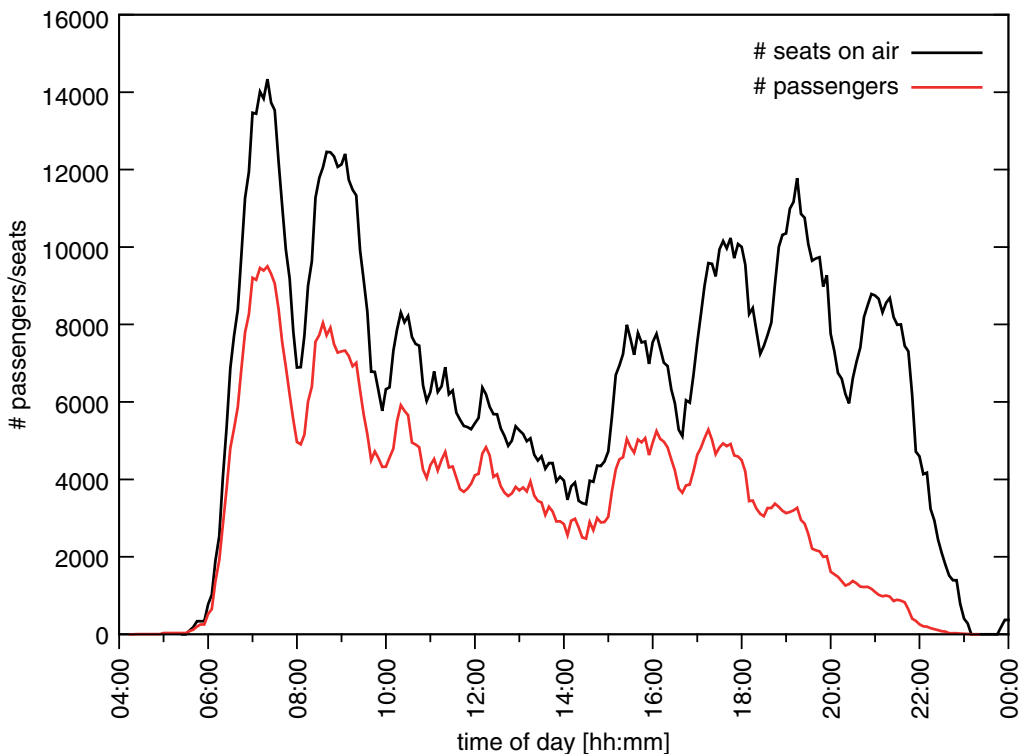
## 68.2 Simulation Results

### 68.2.1 Air Transport

As a scenario for air transport technology, a coverage model from Europe to world wide destinations is used; with the synthetic population, it serves as input for the simulation. The assignment of flights to the desired O-D connection, i.e., the passenger routing, is calculated by MATSim's default public transit routing module.

Each simulation is run for 600 iterations. In each iteration, 10 % of the virtual passengers may shift their departure time randomly within a 2 hour interval. Another 10 % may seek a new route, i.e., a connection between origin and destination. Each passenger chooses from a set of 5 plans using an MNL. The outcome is stable after 500 iterations, then departure time choice and routing are switched off. For another 100 iterations only the MNL is used by passengers to select a plan.

Results are then taken from the output of the 600th iteration. Filtered by flights in Germany, Figure 68.4 depicts passengers in aircraft (red) and seats (black) over time of day and reveals passengers' tendency to depart early.



**Figure 68.4:** Passengers in aircraft and available seats over time in Germany: At any time, there are more seats than passengers. Air transport-only scenario based on O-D data for Germany, iteration 600.

Some passengers fail to reach their destination and are “stranded”. This is unrealistic; only trips within Germany are modeled. These are usually completed within a few hours, with no requirement for an overnight airport stay. 320 passengers are stranded at the end of the day. Getting stranded is not a result of insufficient seating; at any time of day, there are more seats than demand. There are many reasons why passengers could be stranded in such a situation. Further analysis of the  $C_{lineswitch} = 0$  scenario simulation results indicates:

- 92 passengers are stranded because there is no seat and no other flight on the same airline later that day, to which they could be shifted.
- 228 passengers are stuck at an airport because there is no connection after their departure time between that airport and their destination airport.

Behavioral aspects: neither departing early, nor getting stranded, are explicitly modeled.

### 68.2.2 Adding an Alternative Mode

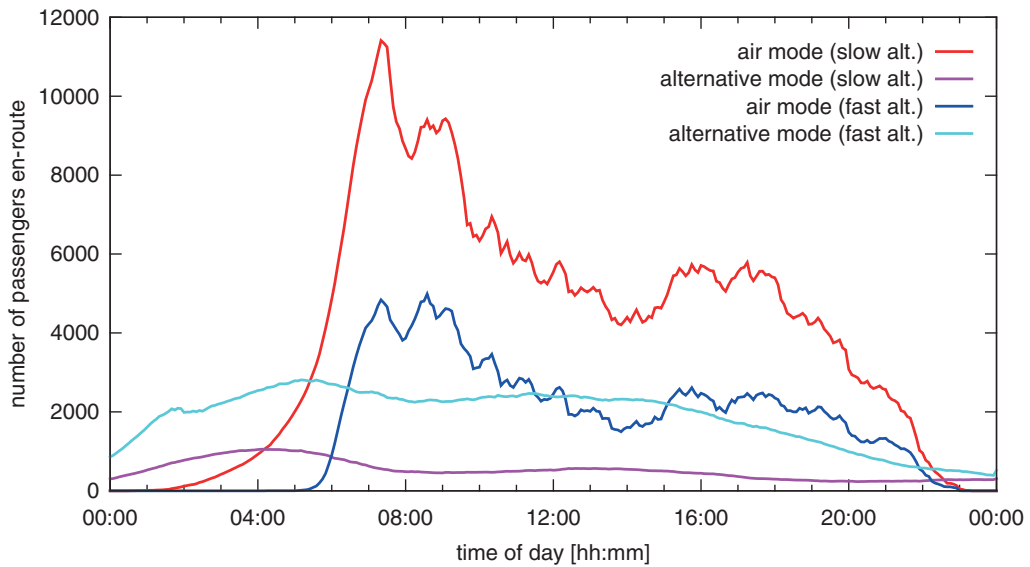
To gain further insights, in the following a slightly different simulation setup is applied. A second option for mode choice is added. Each virtual passenger can now choose between the microsimulated air transport options and an alternative mode. The alternative mode has no capacity restrictions. Passengers traveling with the alternative mode can start directly at their randomly selected departure time. The travel time,  $tt$ , is computed by the microsimulation, with an estimation of the beeline distance between the O-D pair  $d$  and a velocity  $v$ , i.e.,  $tt = d/v$ . 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 traveling are calculated accordingly in the scoring.

With this population, the simulation is again run for 600 iterations. As in the previous simulations 10 % of the virtual passengers 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 and the alternative mode. After 500 iterations all choice modules are switched off; thus, for the last 100 iterations, passengers use the logit model to select a plan.

Simulation results for the 600th iteration show that the increasing speed of the alternative mode affects the modal split. While for a  $v = 100 km/h$  the alternative mode is chosen by 1.2 % of the passengers, a mode alternative with a speed of 300 kilometers per hour attracts 15.69 % of travelers. The number of stranded passengers for the alternative mode with  $v = 100$  kilometers per hour is substantially reduced, from approximately 320 to 67. Higher speeds of the alternative mode further reduce the number of stranded passengers. Slow speeds of the alternative mode imply dominance of the air transport mode. If there is a seat on a flight, travelers receive a higher score than when they use the alternative mode. However, travelers risk getting stranded, which can be hard to analyze and interpret. The implemented algorithm is also an open issue; if the number of plans per traveler exceeds a threshold of 5, the plan with the lowest score is removed from the plan database.

Instead of this deterministic plan removal, a probabilistic algorithm can be implemented: e.g., plans for removal can be selected based on a path size logit model. With this modification, simulation runs are repeated. Figure 68.5 shows the resulting travel patterns over time for alternative modes at speed 100 kilometers per hour and 300 kilometers per hour. Traveler distribution on the alternative mode over time of day is quite homogeneous. The alternative mode speed increase attracts more passengers, as reflected by the modal splits in Table 68.1. At most, one passenger is stranded at the end of day.

Simulation results are compared in more detail with DESTATIS data serving as a base for the virtual population. Synthetic population is generated based on O-D pairs that may contain transfers ( $od_{transfers}$ ), while other DESTATIS data counts the number of passengers on actual direct flights



**Figure 68.5:** Passengers waiting for a flight, traveling by plane, or by alternative mode over time of day. Air transport and alternative mode scenario for Germany, iteration 600. Results with random selector for plan removal.

Source: Grether (2014)

$v[\text{km/h}]$	# air mode	# alt. mode	# stuck	air mode[%]	alt. mode[%]	stuck[%]
100	49280	2551	1	95.08	04.92	00.00
150	44835	6996	1	86.50	13.50	00.00
200	39929	11902	1	77.04	22.96	00.00
250	34332	17499	1	66.24	33.76	00.00
300	27270	24562	0	52.61	47.39	00.00

**Table 68.1:** Modal split for different speeds  $v$  of the alternative mode. Air transport and alternative mode scenario for Germany, iteration 600. Results with random selector for plan removal.

( $od_{direct}$ ). The latter is used to evaluate model accuracy. For comparison, number of passengers on direct flights is calculated for each O-D pair ( $sim_{direct}$ ) from the simulation results. Based on these data sets, the mean square error and the mean relative error are calculated.<sup>3</sup>

Table 68.2 shows the outcome of these calculations. The first line is the comparison of two input data sets from DESTATIS.<sup>4</sup> 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. Mean square error and variance increase with the speed  $v$  of the alternative mode; logical, as the demand covers only air transport trips.

<sup>3</sup> The mean square error  $\sigma^2$  is computed as  $\sigma^2 = \frac{\sum_{i \in OD} (sim_{direct}(i) - od_{direct}(i))^2}{|OD|}$ , whereby  $|OD|$  denotes the number of O-D pairs,  $sim_{direct}(i)$  the simulated passengers on a direct flight between the O-D pair  $i$ , and  $od_{direct}(i)$  the same, but retrieved from data. With the same values, the (unsigned) mean relative error for each O-D relation is calculated as mean rel error =  $\frac{\sum_{i \in OD} |(sim_{direct}(i) - od_{direct}(i))| / od_{direct}(i)}{|OD|}$ .

<sup>4</sup> In the calculation,  $sim_{direct}$  is replaced by  $od_{transfers}$ .

$v[\text{km/h}]$	$\sigma^2$	$\sigma$	mean rel error	stuck
$od_{transfer} - od_{direct}$	12640	112	1.75	-
100	10367	102	0.35	1
150	13820	118	0.43	1
200	18651	137	0.56	1
250	25291	159	0.68	1
300	36059	190	0.76	0

**Table 68.2:** Error calculations for different speeds  $v$  of the alternative mode. Air transport and alternative mode scenario for Germany, iteration 600. Results with random selector for plan removal.

### 68.3 Interpretation & Discussion

The alternative mode can be defined as a combination of train, bus, or car connection availability. Clearly, the results hinge on the assumption that the alternative mode is always available and not capacity-restricted. All passengers on the alternative mode travel at the same speed, but this assumption is too coarse for the scenario presented. For example, average speed and temporal availability of train connections depends on the O-D pair. In principle, the alternative mode could be refined by including O-D pairs' dependent average speed data. Alternatively, train, bus, and car can be simulated explicitly, featuring capacity restrictions and mutual interactions. Even considering these factors, a homogeneous velocity for the alternative mode seems to be more appropriate for the overall modeling approach illustration. Effects triggered by the alternative mode availability are illustrative. Data for the demand provides O-D pairs for air transport, but not for car, train or bus trips. For more plausible interpretations, further demand data for other modes is required.

All the presented modeling approaches explain passenger routing in more detail than technically possible from the input data. Most passengers use a direct connection, which is very plausible, considering the geospatial demand extent. Flying within Germany is often not worthwhile if the connection includes a transfer; empirically it is faster to travel by train, car, or bus. For further insights, the geospatial extent of the modeled demand could be increased; but this depends on data availability, not on the overall simulation approach.

Passengers are modeled without specific desired departure or arrival times. This study's input data does not contain any information about time distribution. The simulation approach can capture such individual time constraints and the information can be added, without too much effort, with some more data, thus resolving several departure time choice problems.

Stranded passengers are an unwanted product of the simulation. Without the alternative mode, the only available transport mode is a capacity-restricted flight connection provided in discrete, irregular time intervals. The number of stranded passengers is higher than for the simulation runs with the alternative mode. Passengers are more likely to get stuck in O-D pairs, where demand is higher than seat capacity, for extrinsic and intrinsic model reasons.

The quality of the simulation model's outcome hinges on the data available. For older studies of air transport passenger demand, DESTATIS data for 09-2011 was used, but the air transport technology model was created on an 09-2009 flight schedule. The number of flight starts within Germany increased slightly between 2009 and 2011 (DLR, 2012, p. 23). Assuming that the number of available seats increased accordingly, the simulation model provided too little capacity, at least on certain O-D pairs. As result, the number of passengers not reaching their destination (stranded) was much higher. With the availability of 09-2009 DESTATIS data, the overall quality of results improved. Replacement of the 2011 data with 2009 data reduced the number of stranded passengers significantly, from around 1 500 to 350 travelers.



Data is provided on a monthly basis, while the simulation model time horizon is one day. Number of trips per day is retrieved using the assumption that trips are uniformly distributed over all days of a month. The remaining 350 stranded passengers might be resolved by a more accurate distribution. Otherwise, a longer time horizon could be simulated.<sup>5</sup> This would also include flights not departing on a Tuesday. With these alternations, the issue of stranded travelers might be solved.

The problem of stranded passenger can be model-intrinsic. The algorithm removing plans is apparently critical to avoid stranded passengers. Replacing the deterministic formulation with the probabilistic resolves most of the stranded passenger problem. The applied path size logit modeling approach seems to be feasible, but requires further studies for parametrization and interpretation. In general, this modeling approach allows the generation of more heterogeneous choice sets, see also Section 97.3. With the deterministic plan, removal plans with a high score (but similar structure) dominate all other generated plans. In combination with capacity restrictions, lack of alternatives results in stranded passengers. All other approaches to simulate more heterogeneity—discussed on the following—should consider these effects.

In further studies, departure time choice and cost structures can be refined. If there is only one early connection to a hub per day, some passengers' departure times might be too late to make connections. The random departure time mutation may not be able to find a connection for all passengers. This has been ruled out for the current setup, but should be considered in further studies.

Alternatively, passengers could have a connection that works in theory, but are “crowded out” by other passengers arriving earlier at the gate; these passengers would reach their destination if they would take a different route. The current approach would not find such a solution, since passengers do not consider costs they impose on others; see Lämmel and Flötteröd (2009) for an approach taking that into account. The real-world solution, presumably, would be to raise prices on seats during congestion periods until a passenger re-routes. Currently, all passengers have homogeneous time values. For a more meaningful price modeling, additional heterogeneous passenger attributes can be included. As the present model is based on only O-D data, it does not include such a process. In principle, other data, e.g., Lorenz curves and median incomes, can be merged with the O-D data (Kickhöfer et al., 2011).

An alternative approach to improve heterogeneity is a router generating a greater route diversity for the same departure time. Such a router would be able to direct a passenger to a route where seats are available, without actually knowing about seat availability. That approach would, however, not address the issue that some passengers might need to switch their path to allow others to obtain a feasible path. In Graf (2013), a first prototype of such a router is tested in a different context, with first tests for the flight model revealing only slight improvement. As more diverse routes are dominated by the direct connection, they are removed by the algorithm similar to routes on slow alternative modes. After this general problem is solved, a more diverse routing should be reconsidered.

## 68.4 Conclusion

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.

The technology modeling is similar to the Clarke et al. (2007) approach, although the level of detail is coarser. In the same way as Clarke et al. (2007), further models for, e.g., gates, taxiing, weather

---

<sup>5</sup> Note, that this requires some changes in the source code that may not be resolved by sole customizations of MATSim. Please ask the developers before running MATSim for a longer time horizon.

or airline operations can be added to the approach. As the open source code of MATSim comes with options for extension, more detailed models of the technology side hinge on the availability of data. In contrast, and going beyond Clarke et al. (2007), passengers are captured at all stages of their trip and passengers traveling on alternative transport modes can be simulated. The chapter discusses certain open general issues not specific to air transport systems. Interested users should support the MATSim team in solving these more general questions first, which will aid the model in achieving a more detailed picture of mid-distance travel patterns.

Clearly, potential applications of the proposed model depend on type and detail of information included. In general, application for policy planning allows a more detailed evaluation of mid-distance travel policy effects, including mode alternative consideration. The approach could also be useful for private companies' planning of flight-schedules and capacities to their connections. The impacts of these changes on customers can be assessed in close detail.