Mapping emissions to individuals
New insights with multi-agent transport simulations

Benjamin KICKHöFER
Transport Systems Planning and Transport Telematics Group
Berlin Institute of Technology
Salzufer 17-19, Sekr. SG 12
10587 Berlin
Germany
Tel: (+49-30) 314 22 542
E-mail: kickhoefer@vsp.tu-berlin.de

Kai NAGEL
Transport Systems Planning and Transport Telematics Group
Berlin Institute of Technology
Salzufer 17-19, Sekr. SG 12
10587 Berlin
Germany
Tel: (+49-30) 314 23 308
E-mail: nagel@vsp.tu-berlin.de

Abstract: It is estimated that metropolitan areas will continue to contribute a large proportion of a country's economic power and will thus attract people from rural areas. By the year 2030, more than one third of the world's population is expected to be living in major cities. However, life quality in urban areas is, among other things, negatively affected by air pollution. Additionally, greenhouse gases due to our life style and mobility behavior will harm globally. This is why the idea of emissions budgets was developed: in order to obtain a sustainable emissions level, individuals should in average not consume primary energy at a rate of more than 2000 Watts.

For this purpose, it is of interest to perform a mapping of emissions back to their producers. This would, in the case of transportation, be on an individual or a household level. In this paper, a tool is presented that addresses this question: For a real-world scenario of the Munich metropolitan area in Germany, a scenario is set up and simulated with the large-scale multi-agent micro-simulation MATSim. The software is capable of simulating complete daily plans of several million individuals and allows emission calculations on a detailed level, e.g. for a single road section or a single vehicle over time of day. A differentiation between emission types (CO_2, NO_x, PM, etc.) can be done, depending on vehicle type (engine type, age and cubic capacity) and traffic state (derived from road category and actual speed). In particular, it is investigated in this paper how differences between urban and suburban life styles, resulting in different mobility behavior, influence individual emission levels. Furthermore, it is shown how disaggregated emission data can be visualized without the need of averaging over predefined zones.

Keywords: Urban Transportation, External Effects, Greenhouse Gas Emissions, Agent Based Modeling

1. INTRODUCTION

By the year 2030, more than one third of the world's population is expected to live in major cities. This raises the question on how to design and organize urban transport systems so that (i) a desired level of service is realized and (ii) the resulting air pollutant emissions do not reduce life quality on a local scale and greenhouse gas emissions not on a global scale. Clearly, especially road transport – a major source of negative external effects – has to become more environmentally friendly. Against this background, the idea of emissions budgets was developed: in order to obtain a sustainable emissions level, individuals should in average not consume primary energy at a rate of more than 2000 Watts. Gaining intuition about the actual level of transport related emissions per person and how this distribution might change with respect to different policy measures seems to be a crucial task. It is therefore of interest to develop a tool that is able to map emissions back to their producers. This would lay the foundations for any attempt to include emission costs into the decision making process of individuals. Also it could provide very detailed data for transport planners and decision makers in order to better communicate policies that aim at reducing emissions.

The goal of this paper is to show that it is possible to perform a mapping of car emissions back to individuals while still being applicable for large-scale scenarios. Therefore, a transport model is needed that provides enough information about the mobility behavior of individuals, and is able to model an entire urban area with several million inhabitants. The multi-agent transport simulation MATSim[1] is capable to deal with large-scale scenarios and is particularly suitable for calculating person-fine emission levels since the traveler's identity is kept throughout the simulation process. In this paper, the software is coupled with the methodology of the 'Handbook on Emission Factors for Road Transport' (HBEFA), developed by INFRAS (2010). For demonstration purposes, it is first analyzed how life style (defined as the choice of home locations in more urban or more suburban areas, respectively) influences people's mobility behavior. In a second step, it is shown how this mobility behavior influences individual emission levels. It is then discussed how the disaggregated results can be visualized in a more suitable way without losing information about the spatial distribution of individual emission levels.

2. METHODOLOGY

This section (i) gives a brief overview of the general simulation approach the software tool MATSim uses and (ii) describes shortly the emission modeling tool that is used in this paper. At this point, only the general idea can be presented. For further information please refer to Raney and Nagel (2006) and the Appendix or to Hülsmann et al. (2011), respectively.

[1] 'Multi-Agent Transport Simulation', see www.matsim.org

2.1 Transport Simulation with MATSim

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

1. **Plans generation**: All agents independently generate daily plans that encode among other things his or her desired activities during a typical day as well as the transport mode for every trip.

2. **Traffic flow simulation**: All selected plans are simultaneously executed in the simulation of the physical system.

3. **Scoring**: All executed plans are scored by a utility function which encodes in this paper the perception of time for the available transport modes.

4. **Learning**: Some agents obtain new plans for the next iteration by modifying copies of existing plans. The modification is done by several modules that correspond to the available choice dimensions. In this paper, agents only adapt their routes (see Sec. 0). The choice between plans is performed with respect to a Random Utility Model (RUM).

The repetition of the iteration cycle coupled with the agent database enables the agents to improve their plans over many iterations. This is why it is also called **learning mechanism** (see Appendix). 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.

2.2 Emission Modeling

The emission modeling tool has been developed by Hülsmann et al., (2011) based on earlier work by Hatzopoulou and Miller (2009). It is composed of two main steps: first, the deduction of kinematic characteristics from MATSim simulations and, second, the generation of emission factors. For the first step, the tool needs to process MATSim output from the traffic flow simulation, in order to deduce kinematic information per agent and road section (= link). This is done as follows: whenever an agent enters or leaves a link, a timestamp is created; thereby, it is possible to calculate the free flow travel time and the travel time in a loaded network. MATSim keeps information about drivers and their vehicles throughout the simulation process. It is therefore possible to access demographic or vehicle-related information at any time.

In the second step, emission factors per air pollutant are identified. Generally, they can vary over vehicle type and traffic state. The former may contain information about engine type, cubic capacity, or age of the vehicle. The latter is derived from road categories (see Sec. 3.1) and actual speed. Emission factors are based on the methodology of the ‘Handbook on Emission Factors for Road Transport’ (HBEFA), developed by INFRAS (2010). The handbook provides emission factors for warm emissions depending on four traffic states: free flow, heavy, saturated, and stop&go traffic. Furthermore, emission factors for cold starts are available which depend, amongst others, of the vehicle’s holding time. For warm emissions, the tool calculates two emission factors for every vehicle and link: first, an emission factor for the part of the link where free flow occurs and, second, an emission factor for the part of the link where stop&go is assumed. These parts are calculated by comparing the minimal travel time to the actual travel time. The resulting emission factors are then assigned to every agent and link whenever an agent leaves a link. For cold start emissions, the main input are activity durations: the longer a vehicle has not been moved (e.g. parked) and therefore cooled down, the higher the cold start emissions.

For this paper, warm emissions vary among road categories and actual speed; cold emissions vary among activity durations. Vehicle types are the same for all individuals, using average vehicles from the HBEFA vehicle fleet (INFRAS, 2010). As of now, public transit is assumed to run emission free. These limitations are addressed in Sec. 5.

3. SCENARIO: MUNICH, GERMANY

The methodology described in Sec. 2 has already been applied in a test scenario for a section of a major ring road in Munich, Germany (Hülsmann et al., 2011). In this paper, it is now applied to the large-scale scenario of the Munich metropolitan area with about two million individuals. Therefore, the scenario has to be set up based on network and survey data. This process is described in Sec. 3.1, followed by a specification of the simulation procedure in Sec. 0 and a validation in Sec. 3.3 where it is discussed to what extent the simulation reproduces reality.

3.1 Setting up the Scenario

Network (supply side)

Network data was provided by the municipality of Munich (RSB, 2005). The data matches the format of the aggregated static transport planning tool VISUM[2]. It represents the road network of the federal state Bavaria, being more detailed in and around the city of Munich and less detailed further away. It consists of 92’259 nodes and 222’502 connecting edges (= links). Most road attributes, such as free speed, capacity, number of lanes, etc. are defined by the road type. Only geographical position and length are attributes of each single link. These data are converted to MATSim format by taking length, free speed, capacity, number of lanes, and road type from VISUM data. VISUM road capacities are meant for 24-hour origin-destination matrices. Since the network is almost empty during night hours, peak hour capacity is set to VISUM capacity divided by 16 (not 24). This results in an hourly capacity of about 2000 vehicles per lane on an urban motorway. In order to speed up computation, some road categories corresponding to small local roads are removed from the network. Furthermore, nodes with only one ingoing and one outgoing link are removed. The two resulting links are then merged, bringing the size of the network down to 17’888 nodes and 41’942 links. When merging, the two link lengths are summed up; free speed is calculated based on the minimal time needed for passing the original links; capacity is set to the minimum of the two links; the number of lanes is calculated based on the number of vehicles that fit on the original

---

[2] ‘Verkehr In Städten Umlegung‘ developed by PTV AG (see www.ptv.de)

links; and the road type – important input for the emission calculation – is set to the one of the outgoing link.

Population (demand side)

In order to obtain a realistic time-dependent travel demand, several data sources are converted into the MATSim population format. The level of detail for the conversion from disaggregated stated preference and aggregated statistics into individual daily plans depends on the area. Three different populations are created, each corresponding to one of the three different data sources:

- Munich municipality (based on Follmer et al., 2004):
  The synthetic population of Munich is created on the base of very detailed survey data provided by the municipality of Munich (RSB, 2005), named “Mobility in Germany” (MiD 2002). In the area of the municipality of Munich, 3612 households (with 7206 individuals) were interviewed. The data consists of different data sets like household data, person specific data and trip data. A detailed description of survey methods and data structure can be found in Follmer et al., 2004. Individuals were asked to report their activities during a complete day including activity locations, activity start and end times as well as the transport mode for the intervening trips. Due to privacy protection, not the exact coordinates of activity locations is available, but only the corresponding traffic analysis zones (see Figure 2 later in this paper). For the generation of the synthetic MATSim population, individual activity locations are distributed randomly within these zones. Furthermore, all incomplete data sets are removed, e.g. when the location or the starting times of one activity is missing in the survey. The transport modes train, bus and ship are treated as public transit trips, motorbikes and mopeds are treated as car trips. The transport modes ride (= in car as passenger), bike and other (= unknown) are kept for the initial MATSim population. Overall, the data cleaning results in 3957 individuals, the representative sample for demand generation. Finally, these agents are “cloned” while holding activity transport analysis zones constant but finding new random locations within these zones for every clone. This process is performed until the population reaches the real-world size of 1.4 million inhabitants. Thus, the synthetic population living inside the Munich municipality boundaries consists for this study of 1’424’520 individuals.

- Commuter Traffic (based on Böhme and Eigenmüller, 2006):
  Unfortunately, the detailed data for the municipality of Munich does neither contain information about commuters living outside of Munich and working in Munich nor about people living in Munich and working outside of Munich. BAA (2004) delivers data for workers that are subject to the social insurance contribution with the base year 2004. Origin and destination zones are classified corresponding to the European “Nomenclature of Statistical Territorial Units” (NUTS)[3], level 3. Thus, the origin-destination flows between Munich and all other municipalities in Germany are available. However, neither departure times nor transportation mode are provided.

[3] See http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction, last access 18.02.2011

Also, the total number of commuters tends to be underestimated since public servants and education trips are not included in this statistic. Therefore, every origin-destination relation is increased by the factor 1.29 (Guth et al., 2010). Car trips are assumed to be 67% of the total commuter trips and public transit to 33% (MVV, 2007). Departure times were set so that people arrive at their working place according to a normal distribution with *N* (8 am, 2 hours) when routed on an empty network. Work end times are set to nine hours after the arrival at the working place. This results overall in 510’150 commuters from which 306’160 people have their working place in Munich. All these MATSim agents perform a daily plan that encodes two trips: from their home location to work and back. Due to this simplification, they are the first contribution to “background traffic”, as it will be addressed from here on.

- Commercial Traffic (based on ITP/BVU, 2005):
  The second contribution to “background traffic” is given by commercial traffic with the base year 2004. On behalf of the German Ministry of Transport, ITP/BVU (2005) published the origin-destination commodity flows throughout Germany differentiated by mode and ten groups of commodities. Origin and destination zones inside Germany are classified corresponding to NUTS 2 and outside Germany to NUTS 3 level. The number of trucks (> 3.5 tons) between two zones or within a zone is calculated based on the commodity flow in tons and the average loading of trucks.[4] The starting and ending points of the trips are – due to the lack of more detailed data – randomly distributed inside the origin and destination zone, respectively. The resulting MATSim agents obtain therefore a plan that only consists of two activities with one intervening trip. Departure times are set so that the number of “en-route vehicles” in the simulation matches a standard daily trend for freight vehicles.[5] For this scenario, only those trips are considered that are at least once during the day in Bavaria. This results in 158’860 agents with one single commercial traffic trip.

Overall, the synthetic population now consists of 2’093’530 agents. To speed up computations, a 10%-sample is used in the subsequent simulations since other studies indicate that this seems to be an appropriate percentage in order to achieve realistic results (Chen et al., 2008).

[4] Estimations based on personal correspondence with Dr. Gernot Liedke from Karlsruhe Institute of Technology (October, 2010).

[5] Estimations based on personal correspondence with Dr. Gernot Liedke from Karlsruhe Institute of Technology (October, 2010).

3.2 Simulation

In this paper, the mental layer within MATSim which describes the planning of activities and the behavioral learning of agents (see Appendix) is reduced to one choice dimension: the simulation allows individuals to only adapt their routes on the road network. Departure time choice and a change of the transport mode for trips is switched off. This is due to the fact that this paper aims at describing the status quo. Since the routes on the road network are not provided by survey data, the outcome of all route choices are calculated with MATSim. Therefore, the following approach is used:

- For 750 iterations, 20% of the agents perform route adaption (discovering new routes) and 80% of the agents switch between their existing plans.
- Between iteration 751 and 1000 route adaption is switched off; in consequence, agents only switch between existing options.

When choosing routes corresponding to a Random Utility Model (see Appendix), agents that go by car are assumed to only include travel times in their decision. These are provided by the mobility simulation of the previous iteration. For all other modes, route choice is not possible. Travel times for ride are also taken from the mobility simulation. Public transit travel times are approximated by twice the free flow car travel time on an empty road network. For bike, walk and the unknown mode, travel times are approximated by taking the geographical distance times a mode specific speed of 15 km/h, 3 km/h and 50 km/h, respectively. Please note that for this paper, only the car travel times are crucial for the decision making process of the agents.

3.3 Verification

Modal split

While converting the input data from MiD 2002 into the MATSim synthetic population, quite a large number of individuals was omitted due to a lack of coordinates or activity times. Therefore, Table 1 shows differences in the modal split over all legs for the two populations. Note that only the mode share of the synthetic population living inside Munich is analyzed (taken from MiD and commuters living inside Munich). As one can see, the synthetic population overestimates the percentage of walk trips by 1.03% and bike trips by 1.47%, while underestimating the percentage of car trips by 1.35% and of ride trips by 2.16%.

Table 1: Trips per transport mode as percentage of total trips; Comparison between input data (MiD 2002) and the MATSim synthetic population.

	MiD 2002	Synthetic population	Difference
Car	26.00	24.65	- 1.35
Ride	13.00	10.84	- 2.16
Public transit	22.00	22.50	- 0.50
Walk	29.00	30.03	+ 1.03
Bike	10.00	11.47	+ 1.47
Unknown	---	0.52	+ 0.52

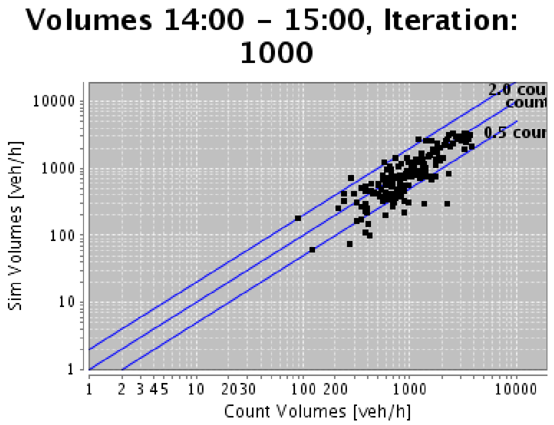
Public transit trips remain almost unchanged and the unknown mode is not discussed further due to the small number of trips. The error seems to be acceptable since no major differences occur. Furthermore, it is estimated that, when opening mode choice in consecutive studies, the modal split will slightly change. Thus, the synthetic MATSim population seems to be a good starting point for analyzing the mobility behavior of the two subgroups with different life styles.

Comparison to Counting Stations

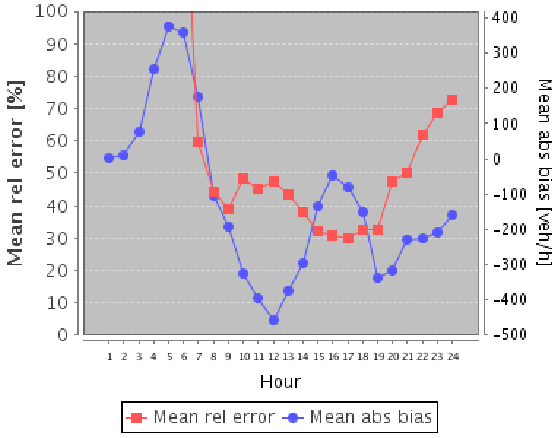
Before analyzing emission levels for two different life styles, the realism of the executed plans in the simulation is verified. The interaction of individuals on the physical representation of the road network is simulated over 1000 iterations as described in Sec. 2.1 and Sec. 0. After reaching a stable outcome, some kind of measurement must exist to determine the quality of the simulation output. For the Munich region, data from 166 traffic counting stations is available and aggregated for every hour over time of day. The best quality of this data is available for Thursday, January 10th 2008. It is used to compare the traffic volumes from the simulation to real-world values. Different statistical values can be calculated, like mean relative error or mean absolute bias. Figure 1 shows two examples of standard reports that MATSim automatically generates. The mean relative error for every sensor and every hour is calculated as:

$$MRE = \left| \frac{Q_{sim} - Q_{real}}{Q_{real}} \right| \tag{1}$$

where Q_{sim} indicates the simulated and Q_{real} the real-world vehicle flow over the corresponding counting station in the corresponding hour. Averages for a given hour are obtained by averaging over all sensors. In the example shown in Figure 1b, the simulation deviates strongly from the reality during the night hours, i.e. from midnight until 7am. However, during daytime, i.e. from 7am until late evening, the hourly mean relative error is between 30% and 50% with better values in the afternoon.



(a) Comparison for one hour (2 pm to 3 pm)



(b) Hourly analysis over time of day

Figure 1: Realism of the simulation results. 166 traffic counting stations provide real-world traffic counts for the Munich municipality area.

In order to reach this accuracy, some adjustments were done, e.g. varying the parameters of the normal distribution that describe work arrival time peak and variance for commuters (Sec. 3.1). For now, since this is a newly set up scenario, the quality of the simulations seems to be adequate. However, by further optimizing travel demand and network information, better values for the mean relative error can be obtained as Chen et al. (2008) or Flötteröd et al. (2011) showed for a scenario of Zurich, Switzerland.

4. ANALYSIS

In order to answer the questions how urban and suburban life styles influence mobility behavior and how this, in turn, affects emission levels, it has to be defined how these "life styles" are used in this paper. As one can depict from Figure 2, the different areas of the municipality of Munich are characterized by different population densities. For illustration purposes, the area of Munich is divided in two different density zones, lower density in blue and higher density in red. The choice of living in the more dense center of the city or the less dense suburban regions depends on many factors. Thus, people living in the red area in Figure 2, are assumed to have an urban lifestyle. Whereas people living in the blue areas have a suburban lifestyle.

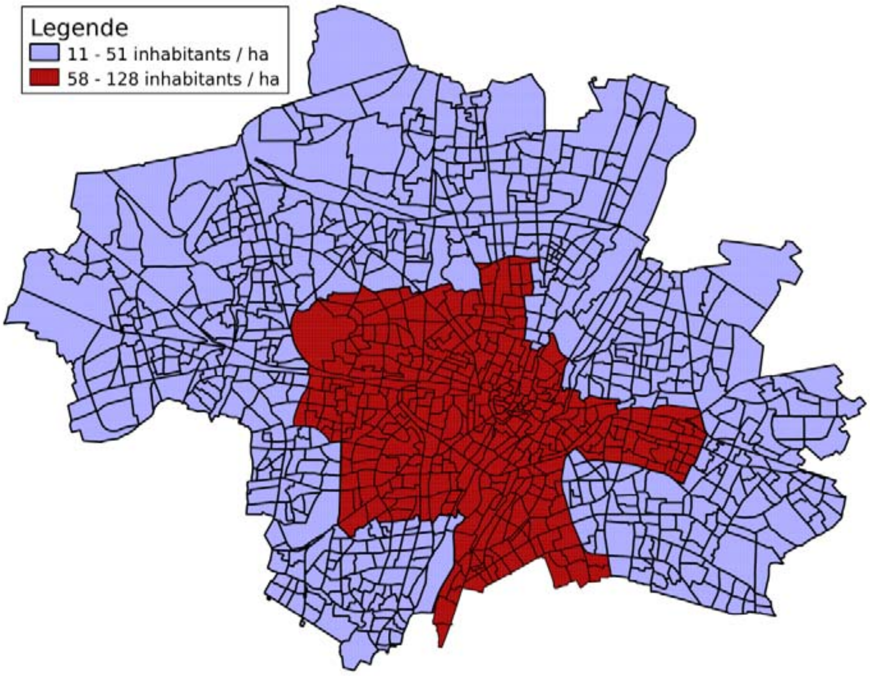


Figure 2: Traffic analysis zones for the municipality of Munich;
population density in inhabitants per hectare: lower (blue area), higher (red area).

When addressing the first question on how life styles influence mobility behavior, one can take a closer look into the statistics of the suburban and urban population, respectively. As Table 2 shows for the 10%-sample, 23'722 more people live in the suburban areas while travelling in average less often (3.29 compared to 3.52 trips per person and day). The modal split reflects the expectations: An urban life style comes with distinctly less car (- 10.23%) and ride trips (- 2.40%), whereas public transit, bike and especially walk are favored (+ 1.34%, + 1.80% and + 9.53%). To sum up, it can be said that an urban life style goes along with a slightly higher mobility and a higher usage of environmentally friendly transport modes.

Now the question appears how this mobility behavior influences individual emission levels. Clearly, since people from suburban areas use the car more often, the overall emissions are higher as Table 3 shows for all types of emission under consideration. CO_2 as an origin of climate change, Particular Matter (PM) and Nitrogen Oxide (NO_x) as local air pollutants. This is also reflected by a higher emission level per person and day. Moreover, it is interesting to observe that even when leaving out the mode choice effect, emissions per car trip are also higher for suburban life styles. This is due to the fact that people from suburban areas tend to travel longer distances in order reach their activity locations.

Table 2: Comparison of suburban and urban populations in terms of mobility behavior.

	Suburban Population	Urban Population	Difference
No of persons	89'418	65'696	- 23'722
Total no of trips	294'072	231'184	- 62'888
Trips per person and day	3.29	3.52	+ 0.23
Modal split [% of all trips]			
Car	29.15	18.92	- 10.23
Ride	11.90	9.50	- 2.40
Public transit	21.91	23.25	+ 1.34
Walk	25.83	35.36	+ 9.53
Bike	10.67	12.47	+ 1.80
Unknown	0.54	0.50	- 0.04

Table 3: Relationship between life styles and emissions for CO_2, PM and NO_x: total levels, emissions per person and day and per car trip for suburban and urban populations.

	Suburban Population	Urban Population	Difference
CO_2 emissions [kg]	256'053	90'303	- 165'750
- per person and day	2.86	1.37	- 1.49
- per car trip	2.99	2.06	- 0.93
PM emissions [g]	32'001	11'359	- 20'642
- per person and day	0.36	0.17	- 0.19
- per car trip	0.37	0.26	- 0.11
NO_x emissions [g]	729'912	268'018	- 461'894
- per person and day	8.16	4.08	- 4.08
- per car trip	8.51	6.12	- 2.38

The underlying data that is used for the calculation of average emission levels in Table 3 provides much more disaggregated information. Thus, a disaggregated visualization is needed in order to fully capture the advantages of the microscopic approach. Figure 3 is a first attempt into this direction: CO_2 emissions of the 10% inhabitants of Munich used in the simulation are plotted using *Inverse Distance Weighting* (IDW) from ArcGIS[6] v10.0. The emission tool allows the calculation of individual emission levels for every emission type as a result from the person's mobility behavior. Every person's home location is available in space by its coordinate; the individual emission level is attached to that

[6] ArcGIS, developed by ESRI (see www.arcgis.com).

location. ArcGIS then calculates for all points in space (defined by cell size) the *expected* emission level using a search radius and an exponent (power).[7] The result can be seen in Figure 3: low CO_2 emissions are represented by blue areas, high CO_2 emissions by orange/red areas. The figure shows emissions in gram per person and day. It accounts for population density, so it does not show that overall emissions are high where population density is high. The results from the aggregated figures in Table 3 are verified: people with urban lifestyles emit substantially less CO_2 per person and day than those with urban lifestyles. Additionally, one can now see a much more differentiated picture within the two areas. The overall picture indicates that increasing distance to the city center results in higher CO_2 emissions; this effect is more important for lower density areas in the north-west than for higher density areas in the south-east of the city. Please recall that this visualization does not require any averaging over zones or predefined grids.

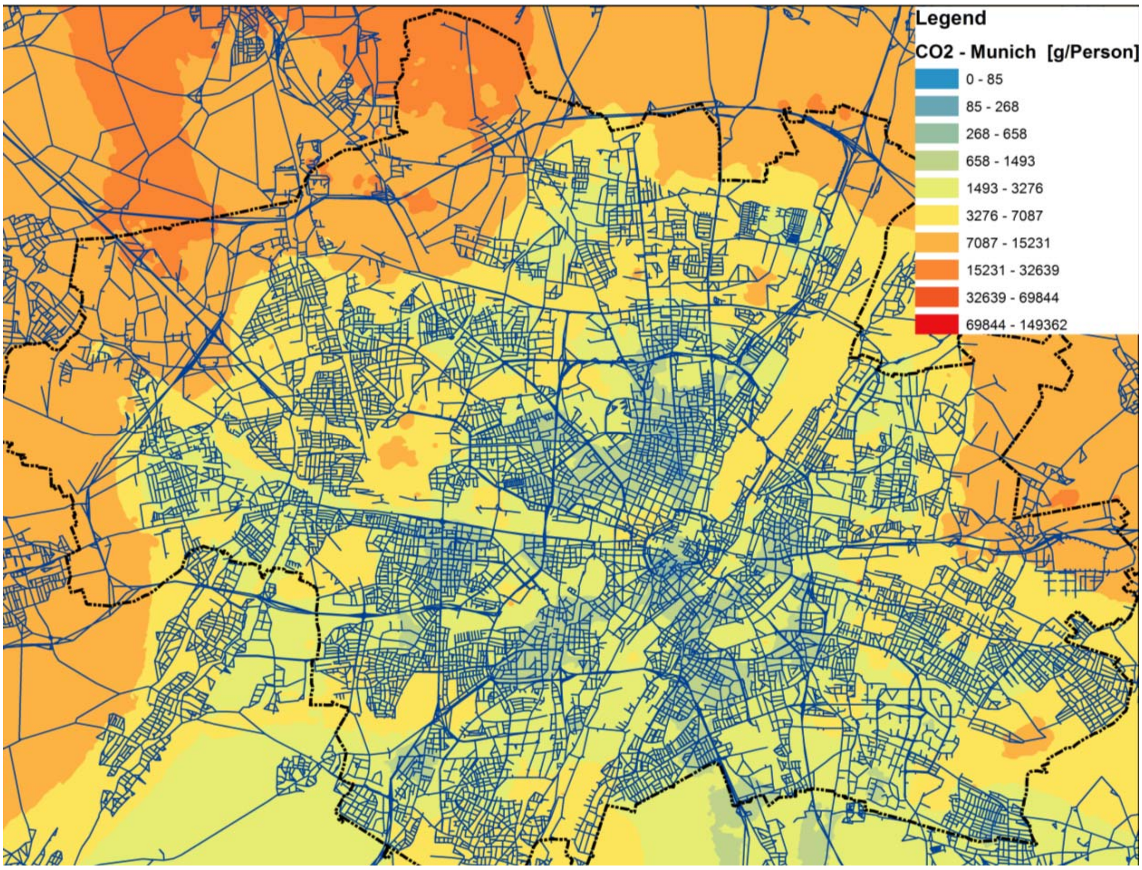


Figure 3: Spatial distribution of CO_2 emissions for the municipality of Munich [gram per person and day]; emissions are mapped to the home locations of their producers.

[7] The following parameters are used for IDW: {power = 1}, {search_radius = {distance = 20 km}{minNumberOfPoints = 10}}, and {cell_size = 1 m}.

5. DISCUSSION

This paper first provides aggregated information about the mobility behavior of two zones corresponding to an urban or a suburban life style, respectively. As expected, suburban life styles come along with substantially higher car use, less trips per person and day, and longer distances. Urban life styles tend to substitute car trips by taking public transit, by walking or biking. In consequence, overall emission levels are higher for suburban life styles, both per person and day and per car trip.

Obviously, similar results could be obtained by using aggregated models that link (average) trip distances to emissions for the different subpopulations. However, the underlying data from the emission tool offers much more disaggregated information. Figure 3 shows how this information can be visualized in a better way than averaging over predefined zones: it is possible to map emissions back e.g. to the home locations of individuals. Their emission level depends on *all* trips over time of day and includes cold start emissions derived from holding times at the activity locations. This is not possible when simply using origin-destination flows. Furthermore, the traffic flow simulation used in this paper accounts for the effect of congestion on emission levels. The model is, thus, in principle sensitive to policy measures. A much more detailed picture can be provided by depicting the estimated impacts of planned policy measures on a spatially very detailed level. In order to implement this, an estimation of the perception of travel times (and travel costs) for the different modes needs to be done, similar to Grether et al. (2009a). Then, time adaption and mode choice can be switched on for the learning cycle described in the Appendix.

A general drawback of the emission modeling with HBEFA 3.1 is that emission factors are only based on distance. Emissions on every link therefore have a maximum value that is lower than emission peaks calculated based on time-velocity profiles (Hülsmann et al., 2011). This would have to be solved within HBEFA or by using other emission factors as input data. A future enhancement of the presented approach will make it possible to obtain more realistic results: Population data contains detailed vehicle information for every household. This adds more heterogeneity to the emission modeling process and differences among individuals will become more important.

An open issue at this point is how to include public transit emissions, which is as of now assumed to run emission free. One could implement a detailed public transit simulation as Rieser and Nagel (2009a) already did for Zurich, Switzerland. A shortcut could be to approximate public transit emissions in an appropriate way.

6. CONCLUSION

In this paper it was shown that a mapping of warm and cold start emissions back to individuals is possible while still being applicable for large-scale scenarios. Emissions are differentiated by type and depend on traffic states that are deduced from road categories and actual speed. Furthermore, it was shown for the municipality of Munich that the average emission level is higher for people with a suburban life style in terms of emissions per person and day as well as per car trip. The underlying output data allows much more detailed analysis: it provides information *where* (on which link) and *when* (exact time of day) emissions are produced. Thus, the approach will help to analyze various effects of transport policies. Finally, a detailed visualization for such analyses has been proposed that does not require any averaging over predefined zones. This will add valuable information to the transport planning process, not only for planners and decision makers but also for the public.

Acknowledgements

This work was funded in part by the German Research Foundation (DFG) within the research project “Detailed evaluation of transport policies using microsimulation”. Important data was provided by the Municipality of Munich, more precisely by ‘Kreisverwaltungsreferat München’ and ‘Referat für Stadtplanung und Bauordnung München’. Our computer cluster is maintained by the Department of Mathematics at Berlin Institute of Technology.

REFERENCES

a) Books and Books Chapters

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.

D. Grether, Y. Chen, M. Rieser, and K. Nagel. Effects of a simple mode choice model in a large-scale agent-based transport simulation. In A. Reggiani and P. Nijkamp, editors, *Complexity and Spatial Networks. In Search of Simplicity*, Advances in Spatial Science, chapter 13, pages 167–186. Springer, 2009b. 10.1007/978-3-642-01554-0.

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*. Routledge, London, 2006.

b) Journal Papers

R. Arnott, A. de Palma, and R. Lindsey. Economics of a bottleneck. *Journal of Urban Economics*, 27(1): 111–130, 1990.

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

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, 2009b. 10.3141/2132-06.

W.S. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2): 251–260, 1969.

c) Papers Presented in Conferences

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

G. Flötteröd, Y. Chen, and K. Nagel. Behavioral calibration of a large-scale travel behavior microsimulation. Technical Report 11-2890, Washington D.C., 2011. Also VSP WP 10-07; see www.vsp.tu-berlin.de/publications.

C. Gawron. *Simulation-based traffic assignment*. PhD thesis, University of Cologne, Cologne, Germany, 1998.

D. Grether, B. Kickhöfer, and K. Nagel. Policy evaluation in multi-agent transport simulations considering income-dependent user preferences. In *Proceedings of the Meeting of the International Association for Travel Behavior Research (IATBR)*, Jaipur, India, December 2009a. Also VSP WP 09-13; see www.vsp.tu-berlin.de/publications.

M. Hatzopoulou and E. Miller. Linking an activity-based travel demand model with emission, dispersion, and population exposure models for appraisal of strategic transport plans. In *Proceedings of the Meeting of the International Association for Travel Behavior Research (IATBR)*, Jaipur, India, December 2009.

F. Hülsmann, R. Gerike, B. Kickhöfer, K. Nagel, and R. Luz. Towards a multi-agent based modeling approach for air pollutants in urban regions. In *Proceedings of the Conference on Air Quality*. Bundesanstalt für Straßenwesen, May 2011, forthcoming. Also VSP WP 10-15; see www.vsp.tu-berlin.de/publications.

D.E. Kaufman, K.E. Wunderlich, and R.L. Smith. An iterative routing/assignment method for anticipatory real-time route guidance. Technical Report IVHS Technical Report 91-02, University of Michigan Department of Industrial and Operations Engineering, Ann Arbor MI, May 1991.

B. Kickhöfer, M. Zilske, and K. Nagel. Income dependent economic evaluation and public acceptance of road user pricing. In *Proceedings of the Kuhmo Nectar Conference on Transport Economics*, July 2010. Also VSP WP 10-03; see www.vsp.tu-berlin.de/publications.

N. Lefebvre and M. Balmer. Fast shortest path computation in time-dependent traffic networks. In *Proceedings of Swiss Transport Research Conference (STRC)* str (). See www.strc.ch.

M. Rieser and K. Nagel. Combined agent-based simulation of private car traffic and transit. In *Proceedings of the Meeting of the International Association for Travel Behavior Research (IATBR)*, Jaipur, India, December 2009a. Also VSP WP 09-11; see www.vsp.tu-berlin.de/publications.

d) <u>Other Documents</u>

J. Bottom. *Consistent anticipatory route guidance*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 2000.

S. Böhme and L. Eigenmüller. Pendlerbericht Bayern. Technical report, IAB, 2006.

Y. Chen, M. Rieser, D. Grether, and K. Nagel. Improving a large-scale agent-based simulation scenario. VSP Working Paper 08-15, TU Berlin, Transport Systems Planning and Transport Telematics, 2008. See www.vsp.tu-berlin.de/publications.

R. Follmer, U. Kunert, J. Kloas, and H Kuhfeld. Mobilität in Deutschland – Ergebnisbericht. Technical report, infas/DIW, 2004. See www.kontiv2002.de.

D. Guth, C. Holz-Rau, V. Killer, and K. Axhausen. Räumliche Dynamik des Pendelverkehrs in Deutschland und der Schweiz: Das Beispiel München und Zürich. 2010.

INFRAS. Handbook Emission Factors for Road Transport 3.1. Technical report, 2010. See www.hbefa.net.

ITP/BVU. Prognose der deutschlandweiten Verkehrsverflechtungen 2025. Technical report, 2005.

MVV – Munich Local Transport Provider. Regionaler Nahverkehrsplan für das Gebiet des Münchner Verkehrs- und Tarifverbundes, Munich, 2007.

RSB. Municipality of Munich: Referat für Stadtplanung und Bauordnung, 2005.

APPENDIX: SIMULATION DETAILS

The following paragraphs are meant to present more information about the MATSim simulation approach that is used in this paper. Every step of the iterative loop in Sec. 2.1 is now illustrated in more detail.

Plans generation

An agent's daily plan contains information about his planned activity types and locations, about duration and other time constraints of every activity, as well as the mode, route, the desired departure time and the expected travel time of every intervening trip (= leg). Initial plans are usually generated based on microcensus information and/or other surveys. The plan that was reported by an individual is in the first step marked as "selected".

Traffic flow simulation

The traffic flow simulation executes all selected plans simultaneously in the physical environment and provides output describing what happened to each individual agent during the execution of its plan. The *car traffic flow simulation* is implemented as a queue simulation, where each road (= link) is represented as a first-in first-out queue with two restrictions (Gawron, 1998; Cetin et al., 2003): First, each agent has to remain for a certain time on the link, corresponding to the free speed travel time. Second, a link storage capacity is defined which limits the number of agents on the link; if it is filled up, no more agents can enter this link.

The *public transit simulation* simply assumes that traveling takes twice as long as traveling by car on the fastest route in an empty network[8] and that the travel distance is 1.5 times the beeline distance between the activity locations. Public transit is assumed to run continuously and without capacity restrictions (Grether et al., 2009b; Rieser et al., 2009b). In this paper, *all other modes* are modeled in a similar way: travel times are calculated based on the distance between activity locations and a mode specific velocity as presented in Sec. 0.

The output of the traffic flow simulation is a list that describes for every agent different events, e.g. entering or leaving a link, arriving or leaving an activity. These events include agent ID, time and location (link or node ID). It is therefore quite easy to grab very detailed information and to calculate indicators such as travel time or costs per link (which is used by the router), trip travel time, trip length, percentage of congestion, and many more.

Scoring Plans

In order to compare plans, it is necessary to assign a quantitative score to the performance of each plan. In this work, a simple utility-based approach is used. It is related to the parameters of the Vickrey bottleneck model (Vickrey ,1963; Arnott et al. (1990), but is modified in order to be consistent with our approach that is based on complete daily plans (Charypar and Nagel, 2005; Raney and Nagel, 2006).The elements of our approach are as follows:

- The total score[9] of a plan is computed as the sum of individual contributions:

$$U_{total} = \sum_{i=1}^{n} U_{perf,i} + \sum_{i=1}^{n} U_{tr,i} \qquad (2)$$

  where U_{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 – both “home” – are counted as one); $U_{perf,i}$ is the (positive) utility earned for performing activity i; $U_{tr,i}$ is the (negative) utility earned for traveling during trip i.

[8] This is based on the (informally stated) goal of the Berlin public transit company to generally achieve door-to-door travel times that are no longer than twice as long as car travel times. This, in turn, is based on the observation that non-captive travelers can be recruited into public transit when it is faster than this benchmark (Reinhold, 2006).

[9] Note that the terms ‘score’ and ‘utility’ refer to the same absolute value.

- A logarithmic form is used for the positive utility earned by performing an activity:

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

  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. $t_{0,i}$ is a scaling parameter that is related both to the minimum duration and to the importance of an activity. As long as dropping activities from the plan is not allowed, $t_{0,i}$ has essentially no effect.

- The (dis)utility of traveling depends on the mode travel time $t_{tr,i}$ and on the transport mode specific marginal utility of time $\beta_{tr,\mathrm{mode}}$:

$$U_{tr,i} = \beta_{tr,\mathrm{mode}} \cdot t_{tr,i} \quad (4)$$

Arriving early to an activity could be punished. There is, however, no immediate need to punish early arrival, since waiting times are already indirectly punished by foregoing the reward that could be accumulated by doing an activity instead (opportunity cost). In consequence, the effective disutility of waiting is already $-\beta_{perf} t_{*,i} / t_{perf,i} \approx -\beta_{perf}$. Similarly, this opportunity cost has to be added to the time spent traveling.

Learning mechanism

A plan can be modified by various modules that correspond to different choice dimensions. These modules are customizable; they can be independently switched on or off or even be replaced by other modules. In this paper, only the first of the three standard choice dimensions is considered (see Sec. 0).

1. **Router module:** The router is a time-dependent best path algorithm (Lefebvre and Balmer, 2007), using for every link generalized costs of the previous iteration.
2. **Time allocation module:** This module is called to change the timing of an agent's plan. A simple approach is used which just applies a random "mutation" to the duration attributes of the agent's activities (Balmer et al., 2005).
3. **Mode choice:** This choice dimension was for a long time not represented by its own module, but instead by making sure that every agent has at least one car and at least one public transit plan (Grether et al., 2009b; Rieser et al., 2009b). New software implementations make it now possible that agents change the transport mode for single trips within their plans. It is assured that when using so called 'chain-based-modes' like car or bike, agents need to pick up their vehicle at the last parking position.

The modules base their decisions on the output of the traffic flow simulation (e.g.

knowledge of congestion) using feedback from the multi-agent simulation structure (Kaufman et al., 1991; Bottom, 2000). This sets up an iteration cycle which runs the traffic flow simulation with the selected plans for the agents, then uses the choice modules to generate new plans; these are again fed into the traffic flow simulation, etc., until consistency between the modules is reached. The feedback cycle is controlled by the agent database, which also keeps track of multiple plans generated by each agent. The probability to change from the selected plan to an existing, randomly chosen plan is calculated according to

$$p_{change} = \min(1, \alpha \cdot e^{\beta \cdot (s_{random} - s_{current})/2}), \tag{5}$$

where

- α : The probability to change if both plans have the same score, set to 1%
- β : A sensitivity parameter, set to 2 for the large-scale simulation
- $s_{\{random,current\}}$: The score of the current/random plan

In the steady state, this model is equivalent to the standard multinomial logit model $p_j = \frac{e^{\beta \cdot s_j}}{\sum_i e^{\beta \cdot s_i}}$, where p_j is the probability for plan *j* to be selected.

The repetition of the iteration cycle coupled with the agent database enables the agents to improve their plans over several iterations. As the number of plans is limited for every agent by memory constraints, the plan with the worst performance is deleted when a new plan is added to a person who already has the maximum number of plans permitted. 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.