Updating trip matrices for Copenhagen using multiple data sources

Otto Anker Nielsen (oan@ctt.dtu.dk) and Christian Overgård Hansen (coh@ctt.dtu.dk), Centre Traffic and Transport, DTU, Building 115, 2800 Kgs. Lyngby

ABSTRACT

Traffic planning in the Greater Copenhagen Region (have over the last 10 years mainly been supported by the so-called OTM traffic model. The behavioural models in OTM include advanced state-of-the-art utility-based formulations, which are combined with base-matrices in a pivot-correction procedure. Before and after studies of specific projects have shown, that these matrices might be the Achilles heal of the whole model system. Vuk and Hansen (2006) therefore validated the present version of the OTM traffic model (version 4.0 from summer 2000) and concluded that a major drawback of the model was indeed outdated base 1992 matrices. From January 2005 to March 2007 the OTM model has therefore been in a large-scale process of updating where creation of new trip matrices has been the main focus.

The article describes the undertaken methodology for constructing the travel matrices for Copenhagen, and how the improved matrices influence the performance of the model. A main focus in the work has been to utilise various data sources for estimation of the new matrices. This includes telephone survey data, cordon line surveys and existing transport surveys to construct the base matrices, traffic counts to adjust these, and digital network databases. The article also demonstrates a new approach to adjust public transport matrices to counts.

INTRODUCTION

The Orestad Traffic Model (OTM) is a tactical traffic model for the Greater Copenhagen Region (GCR). OTM consists of demand and assignment models for both passenger and freight transport. The demand models include trip production, trip distribution and mode choice models, all following a utility-based framework, and the road network model is based on Mixed-Probit formulations and equilibrium algorithms. The model system includes feedback cycles to take congestion into account. The behavioural functions have been estimated based on the combination of multiple Revealed and Stated Preference data-sets.

The first version of the model was developed in 1994. The matrices built however on older 1992-matrices which again used adjusted 1989-matrices from a prior model and data sources. Since then the model has been continuously improved, latest in summer 2000 (Jovicic and Hansen, 2003). The matrices describing the 1992 travel patterns built upon older travel analyses have been adjusted to the counted traffic numerous times since the first version of the model was built in 1995. A major re-estimation of the car matrices was carried out in the harbour tunnel model project in 1998 (Paag et.al., 2001).

With respect to planning of the alignment of Copenhagen's Metro's phase 4, the so called Metro City Ring, a group of clients headed by the Ministry of Energy and Transport wished to upgrade OTM 4.0 to a new version, version 5.0, where a number of improvements were proposed. Most importantly the OTM 5.0 includes new base 2004 matrices. The reasons for building new travel matrices for Copenhagen were following:

- Reduction of uncertainty in forecasting, since base matrices are applied in a pivot point procedure where the demand model is adjusted to fit the base-year matrices,
- The metro is now an existing mode (first phase opened October 2002, second phase in fall 2003, and third phase in fall 2007), whereby the behavioural models for metro can be based on Revealed Preference data, and
- The matrices can be applied in numerous types of travel analyses beyond modelling purposes, e.g. for the analyses of travel behaviour.

Improvement of the model was initiated in January 2005 and finalized March 2007. In the period three main types of improvements have been performed. First, base matrices were updated from 1992 to 2004. Second, the demand sub-models were re-estimated. Finally, improvements and updates of the model zone structure, road and public transport networks and zone data were made.

The article introduces the approach used for the estimation of new trip matrices and the data foundation. The main part of the article discusses the method used to construct the base matrices and adjust them to fits counts. The article is concluded with results and experiences gained from the study.

MAIN APPROACH

The main approach in the estimation of the matrices followed the steps;

1. Base matrices were estimated at an aggregated zonal level (90 zones) based on personal interview data. First, interview data was expanded to population stratified by socio economy and home location and adjusted for seasonal and weekly variations. Second, whilst mode was estimated directly on data, splitting by trip purpose was supported by a modelling approach due to the limited amount of data.
2. The matrices were then adjusted according to the postcard survey to take account of children and residents outside Copenhagen (e.g. tourists) and integrated with the postcard data.
3. The zone aggregated matrices were spatial detailed into the finer system of 818 zones by use of zonal generation and attraction rates and formulated as Generation-Attraction (GA) matrices.
4. Time of day factors computed from survey data with respect to travel purpose and distance were applied to split the all day matrices into seven time-of-day matrices.
5. Trips with one or two legs outside GCR defined by 17 port zones were added to the matrices based on counts.
6. The car matrices were re-estimated on the detailed zonal structure based on traffic counts. A modified Multiple Path Matrix Estimation (MPME) procedure where used (Nielsen, 1998), which assumes that car users choose routes according to a stochastic user equilibrium framework (Sheffi, 1985).
7. The public transport matrices were re-estimated based on traffic counts in busses and at stations. Since one zone may connect too many stops, and each stop may have zonal-connectors from several zones, a fairly complex procedure was developed for the estimation. The results where correction factors at zonal level, which were used in a modified Furness (1970) procedure.
8. After the matrix estimations, the matrices were reformulated into GA-based tours.

The resulting OTM 5.0 base travel matrices describe an average weekday in 2004 in the GCR, which is split between 818 internal zones and 17 port zones. An average weekday is defined as Monday-Friday for a year where June, July and August are excluded.

The base 2004 matrices are segmented into travel mode, travel purpose and time periods. There are five main travel modes: Walk, Bicycle, Car driver, Car passenger and Public transport. Compared to the prior model, the split between car drivers and car passengers is a new feature of the model.

Further, there are six travel purposes: Home-Work (HW), Home-Education (HE), Home-Shopping (HS), Home-Leisure (HL), non Home based Leisure (nHL) and Business (BS). The split of “leisure trips” into the 3 purposes HS, HL and nHL is also an improvement of the model.

Finally, the time periods defined in the model are the following: 5 am to 7 am, 7 am to 8 am, 8 am to 9 am, 9 am to 3 pm, 3 pm to 6 pm, 6 pm to 9 pm and 9 pm to 5 am; in total seven time periods, compared to the three in the prior model.

The OTM 5.0 works therefore with 5 x 6 x 7 = 210 base 2004 travel matrices. The all day home-based matrices are formulated as tour matrices (GA matrices), whilst the nHL and BS matrices are trip matrices (OD matrices).

DATA

Traffic surveys

Data sources used to develop the travel matrices are TU data (the Danish annual national transport survey) for the years 1997-2003 and 2005 (newly completed interviews), and 2005 postcard data.

Since OTM 5.0 is applied for planning of the Metro's phase 4 (Metro City Ring) it was decided that the stratification of the 2005 TU interviews should mainly cover the alignment of the new metro line, i.e. most new interviews were completed with respondents living in the Copenhagen and Frederiksberg municipalities. In total, the 2005 TU data includes 16,285 interviews and 60,542 records. The 1997-2003 TU data includes 16,794 interviews and 51,960 records.

The postcard analysis was completed in March 2005 where over 61,000 postcards were handed out to train-, bus- and metro-passengers, car users and bicyclists travelling across "Sø-snittet" (main cordon north of Copenhagen City Centre). 18,376 postcards were successfully coded in a data file. When expanded to an average weekday (based on traffic counts), the survey gave in total 853,663 personal trips across the corridor.

Count data

For the purpose of matrix adjustment of car matrices 2,193 counts were colleted from the Danish Road Directorate, counties and municipalities across CGA. A lot of efforts were put into making count data consistent and reliable with respect to vehicle classification, time segments etc.

The bus, train and metro counts were collected from the transit operators. The bus count data were collected for November 2004 based on a sub-set of 5% of counting busses that runs in a stratified schedule during the entire year. The train counts were a train postcard analysis on all passengers completed the first Thursday of November 2004. The metro passengers counted by infrared equipment with a 100% sample referred to November 2004. The side-rail lines had only very imprecise counts (traffic estimates) collected by contacts to the companies, but they only carry a relatively small number of passengers. Bus, train and metro counts were adjusted to account for the average weekday in 2004.

All the automatic counting systems had the problem, that the sum of boarding and alighting passengers along a run did not necessary add up to 0 during the run due to the uncertainty of the equipment. Therefore a method was used, which corrected the deviations relatively along the run to secure balance and non negative values.

The matrix estimation method adjusts passengers entering and exiting the system. Therefore it was necessary to distinguish between passengers who transferred between lines, and passengers who entered or exited the system. This was supposed to be estimated based on survey data. As they turned out to be fairly old and quite imprecise, transfer patterns were instead estimated by use of a route choice model (Nielsen & Frederiksen, 2006).

MATRIX ESTIMATION BASED ON SURVEY DATA

Trips represented in the TU data gives an average of 0.15 trips for each zone pair between the 818 internal zones. If decided to build an 818x818 zone matrix directly from this data, this would result in a large number of matrix cells with zero trips. It was therefore decided to aggregate the 818 zones into 90 larger zones. With only 90x90 cells there were in average 13 trips per zone combination. When this was done, it was observed that 7% of zone combinations had zero trips.

Out of the total 18,376 postcards available in the project, 11,358 postcards (62%) describe trips between zone relations comparable to the TU based matrices. When these trips were expanded by applying traffic counts a total of 503,899 trips across the Sø-snittet (cordon line north of the city centre) were achieved. As expected, the postcard total was larger than the TU total because it includes trips made by children under 8, trips made by tourists and temporarily residents, and because the business trips are better described in the postcard data than in TU data.

The statistical error in the two data sources was judged to be about the same. It was therefore decided to combine the two sources when calculating relevant zone-to-zone trips over Sø-snittet based on the following assumptions: a) walk trips were judged to be correct in the TU data, b) trip totals for other travel modes than walk were judged to be correct in the postcard data, c) the ratio between the travel mode totals for postcard data and TU data (factors are over 1.0) were applied for all cells (though not for walk mode).

Table 1 shows GCR travel matrices based on a combination of TU and postcard data. To take account for tourists etc. the zone-to-zone TU trips are expanded by applying the experience from the postcard survey as explained above, and external trips added based on traffic counts.

TABLE 1 Weekday trips including the external traffic, 2004

Travel mode	HW	HE	HS	HL	nHL	BS	Total Trips
Walk	51,285	89,059	310,678	349,350	162,704	11,852	974,927
Bicycle	252,207	184,082	178,620	288,696	151,981	25,360	1,080,946
Car, driver	559,776	37,545	393,361	649,925	355,994	168,199	2,164,799
Car, passenger	135,330	71,324	200,103	465,599	162,936	45,567	1,080,859
Public transport	298,929	137,678	121,552	219,528	111,130	32,434	921,251
Total	1,297,527	519,688	1,204,313	1,973,097	944,745	283,412	6,222,782

The matrix adjustment was then applied on car and public transport matrices based on the existing traffic counts for 2004. There are three reasons for that. First, the TU data applied in the project relates to the period 1997-2005 while the counts are collected for 2004, which is the model base year. Second, there was not enough information in the TU data and postcard data to split the 90x90 zone matrices into 818x818 zone matrices without introducing additional uncertainty. Finally, the TU data does not include information about trips within GCR made by persons living outside GCR. The adjustments for car and public transport respectively are described in the following two sections.

MATRIX ADJUSTMENT, CAR TRANSPORT

Initial preparation

Van and truck matrices originated from OTM 4.0 were simply adjusted to the new zone system and time of day periods. The car passenger matrices as shown in Table 1 were assigned onto the road network together with van and truck matrices in order to compare with the available road traffic counts. Some general deviations were noticed. For instance, the traffic was underestimated across the island of Amager as well as on the corridor along the Motorring 3 (motorway around Copenhagen), whilst it was overestimated on the main access roads towards the city. Therefore, before the matrix adjustment was started the existing matrices were corrected for the above described tendencies.

The Multiple Path Matrix Estimation method

The matrix adjustment applied in the project is a so called Multiple Path Matrix Estimation method (MPME), developed by Nielsen (1998). The MPME is a heuristic method, which simultaneously re-estimates the matrix while the car assignment model iterates.

The method ensures that the estimated link load minimised the weighted square average deviation relative to the available link counts. All paths for each zone-pair are used for the estimation relative to the likelihood of the path being used, and all counts along each path are used.

When zone-to-zone traffic is calculated, the original matrix is adjusted in a heuristic way where the square deviation between the original matrix and the new matrix is minimised. The method therefore adjusts the matrix to fit the available counts as good as possible, and conditional to this changing the original matrix at least as possible.

The more iterations the better the matrices fits to the observed traffic. However, if the number of counts is not optimal (i.e. not enough counts for a sub-area) the matrices can be adjusted wrongly if running with many iterations. In cases with no counts, a Furness-like method is used, where the relative correction of other trips to/from the specific OD-pair is used as an proxy for the estimation (refer to Nielsen, 1998). Naturally, both this and zonal pairs with only one or few counts are much more imprecisely estimated than pairs with many count.

The matrix estimation method was adjusted in order to keep the totals in the port zones, since these usually had very precise counts (e.g. the bridge to Sweden with a toll station).

The route choice model

The route choice model is an integrated part of the matrix estimation procedure, since MPME allow for this as long as it is solved by the Method of Successive Averages (Nielsen, 1998). The route choice model was estimated in a Mixed Logit framework, which beside an error term describes heterogeneity of preferences represented by stochastic coefficients (Nielsen, et.al. 2002). The route choice model was hereafter calibrated compared to the road network data. The following utility function was used based on work in Nielsen (2004);

$$U = k \bullet l + c + \beta_{free} \bullet (t_{free} + \beta_{con} \bullet t_{con}) + \varepsilon$$

The normal driving costs (petrol, etc) are assumed to be proportional *k* with the length *l*. k is assumed to be 0.7 DKK per km (0.13 USD). β_{free} is the Value of Time (VoT) for free flow time, t_{free}. β_{free} is estimated from data to be logarithmic normal distributed. It was estimated for four different segments: commute, business, other (leisure, shopping etc.) and van/trucks. β_{con} is the extra VoT due to congestion relatively to β_{free}. Based on AKTA (Nielsen, 2004) this was also estimated to be logarithmic normal distributed $(1+\ln(\mu, \sigma^2))$ i.e. $\log(\beta_{con} -1) = N(\mu,\sigma^2)$.

Applied adjustment procedure

Because the passenger car matrices estimated from survey data were judged to be quite reliable and the number of counts rather limited, the matrix adjustment was reduced to few iterations.

Whilst the MPME-procedure only increased the number of trips by 0.7%, the average trip length was reduced by 7%. Since the time-of-day factors were estimated on some older travel survey data, a major benefit of MPME was adjustments of the time distribution of trips. For instance, the number of trips in the time segment from 5 am to 7 am was increased by 30% due to a rapid growth in road congestions in the morning peak not existing in older survey data.

The output matrices were made symmetrical over the day and finally they were modified to be GA matrices for the model segments which were home-based. The number of car passengers was computed based on passenger loads in the input matrices.

The number of van trips was increased by 4% in the MPME-procedure, while truck trips were reduced by 21%. The strong correction of trucks is contributed by two factors. First, the vehicle classification scheme has been changed by the road administration. Second, buses were preloaded from bus timetables to the road network and subtracted from truck counts more accurately than in previously matrix estimations. Trip lengths were only marginal changed by the MPME-procedure since it decreases by 4% for vans and increases by 4% for trucks.

The van and trucks matrices would have benefited from a larger number of iterations to improve the quality of the matrices. However, this was not possible without changing the passenger car matrices in the simultaneous assignment and matrix adjustment procedure and priority was given to passenger car matrices.

MATRIX ADJUSTMENT OF THE PUBLIC TRANSPORT

The public transport matrices were also assigned onto the network and adjusted according to the available counts.

Mabit and Nielsen (2006) describe in details the public transport network, which was also used in the matrix estimation project. The model has 3,951 zonal connectors which represent walk/bicycle access to the public transport stops, 270 public transport lines (with 1,170 variants, that has different alignments or stopping patterns) and 5,023 stops (bus-stops, train- and metro stations). The day schedule includes 17,744 runs, where a run represent a given bus or train running from the start to the end of a line-variant.

The timetables were imported from the official timetable database (the organization behind www.rejseplanen.dk), and linked to a digital map (KRAKS Geodatabase, www.krak.dk) in ArcGIS (www.esri.com) using the Traffic Analyst software package in a modified version (www.rapidis.com). This means that a very high accuracy have been obtained concerning the network data.

There is one important methodological difference between car and public transport matrix adjustments, as the public transport matrix adjustments are based on counts on the stop level (boarding and aligning passengers) whilst the car matrix adjustment are based link counts.

In the applied public transport matrix adjustment method we calculated adjustment factors, which afterwards were corrected manually taking specific local traffic conditions and land use into consideration. A common reason for need for manual corrections is that the number of transfers when travelling from O to D by public transport is not known in counts.

The public transport assignment method applied in the matrix adjustment was based on the actual bus and train timetables. It is a stochastic assignment model, which considers distributed values of travel time (Nielsen & Frederiksen, 2006).

Method

A pure MPME-method for public transport would be too time-consuming and difficult to implement – especially due to the time aspect. A "simplified" approach was therefore adapted, where only boarding and exiting passengers in each zone represented by zonal connectors was considered. The principle in this approach was;

1. The route choice model estimates the traffic flows.
2. The route choice model was used to estimate transfer patterns. This was validated on available surveys on transfer patterns. The share of transfers and new boarding or exiting passengers at each stop could hereby be estimated. The counts were then adjusted so that they only represented the boarding and exiting passengers – NOT transfers.
3. Modelled and counted passenger volumes could then be compared at each stop (station or bus-stop), and the relative deviation be calculated.
4. The zonal connectors to the specific stop could then be assigned the same relative deviation (note that connectors from several zones could be connected to the same stop).
5. The relative deviations for all connectors for each zone could then be compared. If the deviations have different signs, this could mean that there were some problems with the Level of Service (LoS) or the connectors to the zone, or the location of the zonal centroid (e.g. if the activity density centre differs from the calculated one from the population and work place data). The model was then adjusted manually to secure the same sign on the deviations. After this a new route choice model was run, and point 4 and 5 repeated until satisfactory results obtained.
6. A weighted correction factor could then be calculated for each zone. A matrix adjustment procedure was then run by the Furness (1970) method. However, to be conservative (i.e. not changing the matrices too fast and too much) all adjustment factors were evaluated manually before this adjustment was made. After this, a new route choice model was run, and the workflow returned to point 4.

The main challenge in this respect was that a trip starting in a given time-period, may first reach its destination in the next (or even following) periods. And a count at a given bus-stop may consist of trips from the same as well as prior time-periods. In principle the same problem exists for car trips. The public transport trips are however often slower due to the slower speed of public transport, why the problem is larger than for the car trips.

To make the adjustment feasible, a method had to be developed which took this into accounts. As this has not been published before, it is being described in mathematical and algorithmic terms in the following.

The matrix adjustment is described by using the following notation;

T_{ijkxy} Modelled traffic *T* from zone *i* to *j* with purpose *k* i beginning in time-interval *x* and arriving in time-interval *y*. The absent of an indices indicates that the matrix has been summarised over this dimension.

E_i Traffic as it ought to have been modelled (Expected) from zone *i* or *j* to fit the counts.

T_{is} Modelled traffic along zonal a zonal-connector between zone *i* and stop *s* in the network

E_{is} Traffic as it should have been modelled along the zonal-connector between zone *i* and stop *s* to obtain the correct stop-volume compared to counts

R_{is} Relative share of traffic from zone *i*, which uses the zonal-connector to stop *s*

v_{sap} Observed traffic volume at stop *s*. Indices *a* for exiting passengers and *p* for boarding passengers

T_{sap} Modelled traffic at stop *s*

R_{sap} Relative deviation between observed and modelled traffic at a stop

R_j Relative deviation between modelled and estimated traffic from zone *i* or *j* (matrix factor)

(n) Indices for iteration number.

The algorithm for the matrix adjustment is described in pseudo-code in the following:

0. Initialisation

All traffic volumes are read:

$T_{iskxy(0)}$, $T_{sikxy(0),}$ V_{sp}, and V_{sa}

Expected volumes are set equal to the modelled (to at initialise volumes for stops and zonal-connectors for which not counts exists).

$E_{iskx(0)} := T_{iskx(0)}$
$E_{siky(0)} := T_{siky(0)}$
$R_{sax(0)} := 1$
$R_{spy(0)} := 1$
$R_{ix(0)} := 1$
$R_{jy(0)} := 1$

The share of traffic on each zonal-connector is estimated

$R_{iskx} := T_{iskx(0)} / T_{ikx}$
$R_{siky} := T_{siky(0)} / T_{jky}(j=i)$

The iteration number is set to *n*=0

1. Updating, s

The traffic at all stops are calculated as it should have been, if it should equal the weighted sum of counts at the end of the zonal-connectors. The corresponding traffic on the zonal connectors is calculated as it should have been to fit counts:

If $V_{sp} <> \text{Null}: R_{spx} := V_{spx} / T_{spx(n)}$
If $V_{sa} <> \text{Null}: R_{say} := V_{say} / T_{say(n)}$
$E_{iskx(n)} := T_{iskx} \bullet R_{spx}$
$E_{siky(n)} := T_{siky} \bullet R_{say}$

2. Updating, ij

The expected traffic for all zones are calculated as the sum of the traffic along all zonal-connectors. Matrix adjustment factors can hereafter be calculated as follows:

$E_{ikx} = \Sigma_{\text{over zonal connector s}} (\Sigma_{n=\#\text{ time intervals } \in[x;y]} T_{iskyn} R_{spn})$
$E_{jky} = \Sigma_{\text{over zonal connector s}} (\Sigma_{n=\#\text{ time intervals } \in[x;y]} T_{sikxn} R_{san})$ (j=i)
$R_{ix} = E_{ix} / T_{ix}$
$R_{jy} = E_{jy} / T_{jy}$

3. 'Assignment'

n:=n+1

For all zonal-connectors (the traffic is assigned onto the zonal-connectors after the constant relative ratios adapted from the full assignment of the base-matrix from the prior adjustment round):

$T_{iskx(n)} = R_{iskx} \bullet E_{ikx}$
$T_{siky(n)} = R_{siky} \bullet E_{jky}$ (j=i)

For all zones (stations volumes are calculated as the sum of zonal-connectors to each stop):

$T_{spkx(n)} = \Sigma_{\text{over zonal connector is}} T_{iskx(n)}$
$T_{saky(n)} = \Sigma_{\text{over zonal connector si}} T_{siky(n)}$

4. Stop criterion

If a stop criteria is not reached (e.g. number of iterations as the simplest one), then go to step 1.

5. Matrix-adjustment

$T_{ijk(new)} := T_{ijk(0)}$

Rows and columns in $T_{ijk(new)}$ are adjusted iteratively to fit with E_{ik} and E_{jk} (equal to a common factor model or Furness method). This is done for each time-period separately.

The pseudo assignment in step 3 assumes that the trip-distribution is independent of the adjustment of the zonal-connectors (the matrix adjustment is in step 5). In principle a full matrix adjustment as in step 3 should be followed by a new full assignment. The assignment is – however – itself an iterative algorithm, which takes several hours, whilst the pseudo assignment takes few minutes. The overall calculation time with a “true” assignment procedure would therefore increase from minutes to weeks, if a full assignment were to be part of the matrix adjustment algorithm. This is why the “simplification” above was used, and why a full MPME-method could not be used for the large public transport network in the present model.

After the above algorithm has run, it is necessary to run the full assignment to evaluate the results. For each main iteration, too full assignments are therefore needed (before and after the adjustment).

Matrix adjustment

During the estimation procedure, it was realised, that it was inappropriate to have counts on bus-stops that were not connected to zones by zonal connectors, since trips hereby were overlooked. Bus stops which did not offer transfer options nor was connected to zones by zonal connectors was therefore located. Some stops were then additionally connected to zones by zonal connectors. Stops with transfers but no connectors were manually inspected to assess whether connectors should be added.

Following this, counts on bus stops with no transfers or connectors were moved proportionally to the nearby stops, and the non-modelled stops were removed from the network model. A special ArcGIS-based software was developed for this purpose. The problem mainly occurred in the outer parts of the Copenhagen region, where the zones are pretty big, as they only were used to describe the hinterland traffic to Copenhagen, but not local traffic in great details.

The core indicators suggested relative adjustments at the stop-level, at the connector level (which considered the average suggested adjustments of all zones using the stop at the end of the connector), and at the zonal level. Based on thorough manual inspections it was first examined whether justified (e.g. with uneven land use in the zone) relocation of Zonal centroids could improve the fit. This was typical the case where some stops had too much traffic and some too little. Then it was examined whether some connectors were missing, typically to stops with too little modelled traffic, whilst neighbouring stops had too much traffic.

The adjustments of zonal connectors, network and centroids was repeated 3 times before the matrix adjustments was begun. A major challenge in this respect was, that a connector factor had to be added for the long connectors, since the route choice model tend to assign too much traffic on these and too little on local busses to the train network (refer to Mabit & Nielsen, 2006).

The next two steps only adjusted the total daily traffic, while some fine-adjustments on connectors and centroids were still made.

Finally, the matrices were also adjusted with respect to the 7 time-periods during the day. The calculation of these adjustment factors needed also to use a bookkeeping of the start time of trips and when they reach the count. This fine-tuning with respect to time periods was done by using a spreadsheet mode. This fine-tuning of matrices was repeated twice.

Some observations on the public transport matrix adjustment

It is interesting to note, that very limited changes were made on the daily matrices, when the 818x818 matrices where aggregated to the 90x90 matrices that had been estimated based on the travel surveys. This basically indicates a consistency between the surveys and the traffic counts.

The main exception for this was the zones in the new urban development of the "Ørestad", where the transport surveys cover 2004 as well as prior years, whilst the counts covered 2004. Due to the fast development in this town area, the adjustment procedure increased flows to/from these zones.

The adjustment did, however, make many changes of the "sub-zones" within each of the large base-zones. Some general observations was made of the corrections *after* the pure algorithmic approach had been made, i.e. it is observed changes not used criteria for the adjustments;

- More public transport journeys close to the fast and/or high frequency bus lines ("A" or "S" busses), train stations or the metro.
- Less public transport journeys in industrial areas with poor public transport service.
- More public transport in newly developed urban areas. This was due to the combination of several years in the transport surveys, while the adjustments were done using the 2004-counts.
- More journeys at stations that opened during the survey period, where the counts where after the opening (two new stations only).
- Much fewer journeys in rural areas. These areas typically only contain a small population share of the larger 90-zones, but are populated with car-owning population segments.
- Many more journeys to areas with summer houses. This could both be due to the use of population and workplaces for the initial split of the 90x90 matrices, or since some of these houses are used illegal for whole-year living, which is most likely not reported in e.g. telephone surveys (since the inhabitants have pro-forma addressed elsewhere)
- More journeys from zones with larger stations in towns in the outer part of the region.
- Various special cases, where sub-zones include special attractions such as museums, aquarium, etc,
- Zones with concentrated education activities (high schools, colleges, universities,…) or student hostels.

VALIDATING THE RESULTING MATRICES

Mode choice and trip length

Table 2 shows the final 2004 day matrix after the matrix adjustment.

TABLE 2 Final base day matrix 2004

Travel mode	HW	HE	HS	HL	nHL	BS	Total Trips
Walk	51,285	89,059	310,678	349,350	162,704	11,852	974,927
Bicycle	252,207	184,082	178,620	288,696	151,981	25,360	1,080,946
Car, driver	549,634	39,827	395,508	623,134	378,187	192,358	2,178,648
Car, passenger	132,697	77,366	201,808	447,111	172,905	52,081	1,083,968
Public transport	276,965	117,921	126,202	212,997	123,222	31,898	889,205
Total	1,262,788	508,255	1,212,816	1,921,288	988,999	313,549	6,207,694

To compare the new trip 2004-matrices with prior trip matrices, a forecast with the previous model - OTM 4.0 - was conducted to update the 1992-matrices to a 2004 reference year. The total number of trips in the new 2004-matrices fits fairly well with the updated 1992-matrices, since it is only 6% less.

While the number of car fits very precise, the divergence of walk and bike trips are large. In the 1992-basis matrices used in pivot-point corrections of OTM 4.0, walk and bike trips have only been roughly estimated and possible biased. In contrast passenger car matrices were adjusted in 1998 (Paag et. al. 2001) and later using MPME on the older matrices. The number of public transport trips is 6% less in than in updated 1992-matrices.

Analyses reveal, however, that the trip length distribution differs between the two set of matrices. The average length of public transport trips is 25% larger in the new 2004-matrices than in the updated 1992-matrices and 20% for passenger car trips. Explaining the differences it is likely;

- That persons travel longer distances today than in 1992 not completely captured in the model forecast.
- That the matrix estimation extensively applied in development of the 1992-matrices has changed the trips length distribution.
- That the respondents in the TU-survey forget to report short distance trips.

A consequence of the longer trips distances person km. by car and public transport is larger in the new 2004-matricer than prior.

Trip purposes

When looking at the trip purposes, a dramatic change of the matrices are evident. For the model, this is important, since the different trip purposes have different Value of Times, and these result in different changes of behaviour when new policies or services are introduced.

We need to recall the data foundation of the two set of matrices. The new 2004-matrices are based on TU-surveys and minor matrix adjustments of car and public transport. Therefore, TU is the main source for segmentation of the 2004-matrices, whilst segmentation of the 1992-matrices is estimated form a cordon survey downtown in 1994 and commuter statistics.

For instance, there are 40% less home-work trips in the 2004-matrices than in the 1992-matrices. While the 1992-matrices clearly overestimate commuter trips there may be underestimation in the 2004-matrices. Since the prior estimations of trip purpose distribution were based on commuter statistics which did not consider absence and part time employment, this is probably the main reason to overestimation of commuter trips. In OTM 5.0, home-work trips are modelled like simple tours with an out and return leg. Therefore, it does not allow combined trip chain modelling and combined home-work-shopping trips may be split into other trip purposes than commuting although efforts have been put into trip purpose segmentation of TU-data.

The number of business trips are much larger in the updated 1992-matricer than in the 2004-matrices. We conclude again that it must be due to major overestimations in the old matrices and underestimations in the new matrices. In the 1992-matrices, trip purpose split is based on a survey in an area downtown Copenhagen with a proportional high number of business trips compared to GCR as whole. While the 1992-matrices reflect quite well business trips in the City Centre of Copenhagen, larger divergences are likely in the suburban and rural areas of GCR. With respect to the 2004-matrices it is noticed to be difficult to capture business trips in household TU-interviews.

The number of home-education trips differs only slightly. This, however, cover a mix of divergences by mode. While the 2004-matrices contain more education trips by walk, car, and public transport than the 1992-matrices, the number of bike trips is less. The cordon survey in 1994 used in estimation of the 1992-matrices may include a higher proportion of cyclist compared to walk due to the distance to the nearby universities, whereas the use of public transport and car in connection with educational trips may be higher in the rural areas.

Finally, the number of private trips is larger in the new matrices partly caused by some misclassifications of commuter trips and business trips.

Time of day distribution

The number of time periods has been increased from three in OTM 4.0 to seven in OTM 5.0. A comparison of the time segments reveals quite large moves of traffic from the morning rush hours to out of peak periods. This tendency of earlier commuting to Copenhagen compared to 1992 confirms congestion measurements from the AKTA-project (Nielsen, 2004). It should be recalled that the trip length is considerable longer in the 2004-matricer than in the updated 1992-matrices and therefore impose more traffic and road congestion even though the number of trips are less in the morning peak.

CONCLUDING REMARKS

The article describes the applied matrix estimation procedures for OTM 5.0 and demonstrates the appraisals of the new matrices. We believe that quality of modelling depends on the data foundation as documented in Vuk & Hansen (2006) and the hybrid approach has been cost efficient.

The approach using surveys to estimate matrices at a large-zonal level (90 zones), and then to split these to sub-zones using survey data and with an adjustment (fine-tuning) using traffic counts, turned out to be quite successful.

The resulting matrices are very detailed including 6 trip purposes, 7 time-of-day periods and built on a zonal structure with 818 internal zones and 17 port zones. The old matrices built on surveys from the 1980'ies have been adjusted several times since then primarily based on traffic counts. Comparing the new matrices with the old revealed major differences mainly contributed to biases in the 1992-matrices and data definitions. However, there has been a dramatic change of time-of-day and trip distribution that the prior adjustments have not captured.

It is our belief that the new matrices improve the data foundation for OTM considerably. The car matrices will be much better to describe congestion. And the more correct split on purposes will improve forecasts on policy initiatives. The public transport matrices have been changed and hereby improved even more, and this will provide a strong basis for decision making the coming years.

REFERENCES

Furness K.P. Time Function Interaction. *Traffic Engineering and Control* Vol 7, No 7, pp19-36, 1970.

Jovicic, Goran and Hansen, Christian Overgaard. A passenger travel demand model for Copenhagen. *Transportation Research Part A*, vol. 37, pp. 333-349, Elsevier Science Ltd., 2003.

Mabit, Stefan & Nielsen, Otto Anker. The effect of correlated Value of Time Savings in public transport assignment. Article presented at European Transport Conference (ETC), September 2006.

Nielsen, Otto Anker. *Two new methods for estimating Trip Matrices from Traffic Counts". Chapter in Travel Behaviour Research: Updating the state of play*. Edited by Ortúzar, H. D., Hensher, D & Jara-Díaz, D. Elsevier Science Ltd., pp. 221-250, 1998.

Nielsen, Otto Anker; Simonsen, Nikolaj & Frederiksen, Rasmus Dyhr. Stochastic User Equilibrium Traffic Assignment with Turndelays in Intersections. *International Transactions in Operational Research*, Vol. 5, No. 6, pp. 555-568. Pergamon, Elsevier Science Ltd, 1998.

Nielsen, Otto Anker; Frederiksen, Rasmus Dyhr & Daly, Andrew. A stochastic multi-class road assignment model with distributed time and cost coefficients. *Networks and spatial economics. No 2*, pp. 327-346. Kluwer, 2002.

Nielsen, Otto Anker. Behavioural responses to pricing schemes: Description of the Danish AKTA experiment. *Journal of Intelligent Transportation Systems***,** Vol. 8(4). Pp. 233-251. Taylor & Francis, 2004.

Nielsen, Otto Anker & Frederiksen, Rasmus Dyhr. Optimisation of timetable-based, stochastic transit assignment models based on MSA. *Annals of Operations Research,* Vol. 144, Issue 1 pp 263-285. Kluwer, 2006.

Paag, H., Daly, A. & Rohr, C. *Predicting use of the Copenhagen Harbor Tunnel. Travel behaviour Research: The Leading Edge*. Chapter 36 in Book edited by David Hensher. Pergamon press, Elsevier, 2001.

Sheffi, Y. *Urban Transport Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, New Jersey. 1985.

Vuk, G.V. & Hansen, C.O. Validating the passenger traffic model for Copenhagen. *Transportation*, Volume 33, Issue 4, Page 371-392, Springer, 2006.