SAMPERS – Erfarenheter av komplexa modeller i ett komplext utvecklingsprojekt

Forfatter(ere):

Staffan Algers, Muriel Beser och Gunnar Johansson Sveder, Transek AB. Address: Transek, Solna Torg 3, S-171 45 Solna, Sweden, Phone +46 8 735 20 20, Fax +46 8 735 2030, Email: staffan@transek.se, gunnarjs@transek.se

Background

The planning process

Swedish transportation authorities have a long tradition of developing traffic demand models. The first generation of national models was developed in the beginning of the 1980s, a second generation during the first half of 1990s. These models have also been frequently used in a large number of projects but also as a part of the regular national strategic transport investment plan. The national planning process has been a four-year cycle of revising a ten-year investment scheme. The first step in this process is to undertake an analysis to decide on a general policy (like promoting accessibility or focus on environmental protection). Here the models are used to analyse a few main alternatives, representing major differences in transportation policy and economic development (such as heavily increased petrol taxes to reduce carbondioxide emission).

Based on the decision on the general policy, taken by the parliament, the next step is to perform a more detailed analysis on what projects to include in the ten-year investment plan. The outcome of this process also contains tradeoffs between rail and road investment, which makes it essential to base the analysis on the same forecasting tool and the same assumptions on economic development, land use etc.

The actors in the process are the sector authorities (notably the road administration and the rail administration), and a co-ordination authority (The Swedish Institute for Communication Analysis SIKA). The forecasting work is carried out by the different actors, and co-ordinated by SIKA. The cost benefit score is a major assessment criterion in establishing the investment plan. Thus it is vital also in this respect that projects in different sectors can be compared on equal grounds.

Finally, the next ten-year investment plan is approved by the parliament. As can be expected, the political process does affect the outcome of the investment scheme.

Cause for new models and a new forecasting system

The previously used models covered car ownership, trip frequency, destination choice, mode choice and route choice for long distance, regional and local trips. The trip frequency, destination and mode choice models were nested logit models divided into private and business long distance trips (>100 km) and seven trip purposes for regional and local trips. For route choice, and for choice of public transport submodes for regional and local trips, the assignment package EMME/2 was used. The models were rather comprehensive but not integrated into one single system, as a result of them being developed sequentially over the years by different organisations.

One of the main problems with the old models was the lack of integration and their user unfriendliness. A new national travel survey had also been carried out, which made it possible to update, improve and extend the performance of the travel demand models.

SIKA, the Communications Research Board and the national planning authorities, all of which forming a clients group headed by SIKA developed an extensive project specification. After an international tendering process, the project was commissioned to Transek with subcontractors. The main subcontractor for the modelling work is the Hague Consulting Group, the tasks of which included general advice and development of the international models.

A major innovation in the system is a model for long distance trips, extended to include departure time and ticket type choice. This extension is addressed in the paper, which also reports on the other models developed in the project and on the integration of these models into one system.

Scope of new system

The general scope of the Sampers project was i) to develop a user-friendly computer traffic forecasting system, ii) to develop new models covering all trips in Sweden.

All trips

By the notion all trips is meant trips having at least the origin or the destination in Sweden. Through trips between for example Finland and Denmark is not modelled. As in previous systems, this means that domestic long distance trips, regional and local trips are modelled. The previous models that had been used by the national planning authorities did not contain a model for international trips, so this is a new element in the model system.

Level of detail

It is obvious that different applications require different level of detail in the forecasting system. An analysis of a road link in an urban environment requires a higher geographical resolution than an analysis of a high-speed train service. Therefore, local and regional trips are handled with a higher resolution than long distance and international trips. For local and regional trips, Sweden is divided into 6000 zones, which would imply very large matrices and corresponding problems if no further breakdown were made. Therefore, 5 regions are defined, which are run separately.

For domestic long distance trips, 670 zones are defined. The same zones are used in Sweden for international travel, and the world outside Sweden (Europe) is divided into 200 zones, coarser as distance increases.

Integrated system

In order to make the models system user friendly, the demand models, the databases and the EMME/2 system had to be integrated into one software under the Windows NT operating system. An important task for the system is also to make it possible to add car trips from the 700-zone level to the regional 6000 level, and to add up train trips from the regional, long distance and international levels.

To get a travel forecast is seldom the final step in an investment or policy analysis. To make successive steps in the analysis easier, an effects module and a cost benefit module were required to be included in the system, as well as a module for accessibility analysis.

Model overview

The models that have been estimated in the project and will be more detailed described below are the following:

- Regional models
- Long distance models
- International models

All models are of the discrete choice logit type, except for an ordinary least squares trip frequency model for foreigners travel to Sweden. Experience from earlier national model studies are also brought into the project (Lundqvist and Mattson, ed 2000)

Data

The main data source for the travel behaviour to be modelled was the national Swedish travel survey, RiksRVU 94-98, which is a continuous travel survey containing 30,000 interviews for the entire interview period. The travel survey contains a one-day diary including all trips, supplemented by trips over 100 km made the last month, and trips over 300 km made the second last month.

This data set is however not sufficient for international trips, and data collected in other major infrastructure investigations were also used, e.g. the Öresund survey and the Fehmarn Belt survey.

The client using the EMME/2 system supplied the transport supply data. Specifically for long distance trips, data not only for different seasons, days and time of day but also for different years was supplied in order to match the development of the infrastructure over time (such as the introduction of the high speed train X2000).

Statistics Sweden produced Land use data. This data was produced at the Small Area statistics level, and then aggregated to the zoning system used in the different segments mentioned above. The number of zones used were as follows:

- local and regional trips 6 000 zones;
- domestic long distance trips 670 zones;
- international trips 180 zones outside and 670 zones inside Sweden.

Regional Models

General

The task for the regional models is to produce estimates of trips for the following modes: car as driver, car passenger, bus, commuter train, bicycle and walk. The models work on a tour basis. Work tours are defined as home based tours, having a model structure as in Fig. 1.

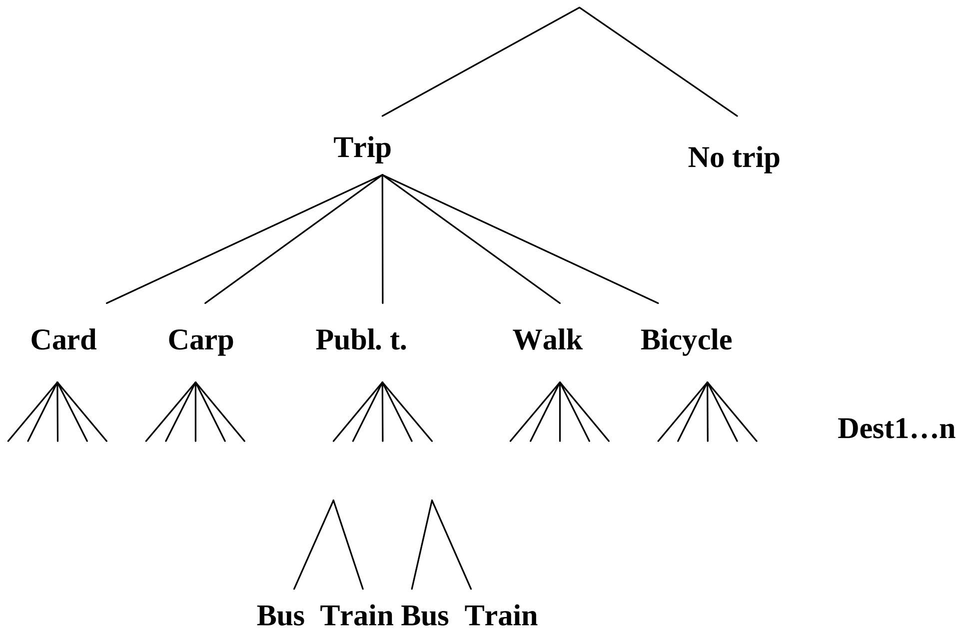

Figure 1 Model structure for home based tours

In order to better reflect similarities of the bus and train alternatives, a structure containing two mode choice levels was adopted, as can be seen from the figure. Public transport mode choice is now handled at two levels in the model. First, there is a general public transport mode at the mode choice level. Then there is a bus - train mode choice at the lowest level of the model. This mode structure gave a better fit for all trip purposes except for business trips.

There are 5 other trip purposes defined, for which home based tour models have been defined, with the same structure as in Fig. 1 (except for business trips, for which too few train trips were reported, forcing the train mode to be omitted form the business trips model). These trip purposes are:

- Business
- School
- Social
- Recreation
- Other

The aim of the model structure is to also capture trip chaining when fulfilling this task. This has been done by conditioning secondary destinations and work based tours on the work tour. Thus, if a person has made a work tour, he/she can choose to make for example a shopping trip not only as an ordinary home based trip, but also as a work based tour or as an intermediate stop on the way from work to home. This will of course not capture all types of trip chaining, but a fair share. The structure for non work tours will therefore take the form shown in Figure 2 (the work is currently ongoing):

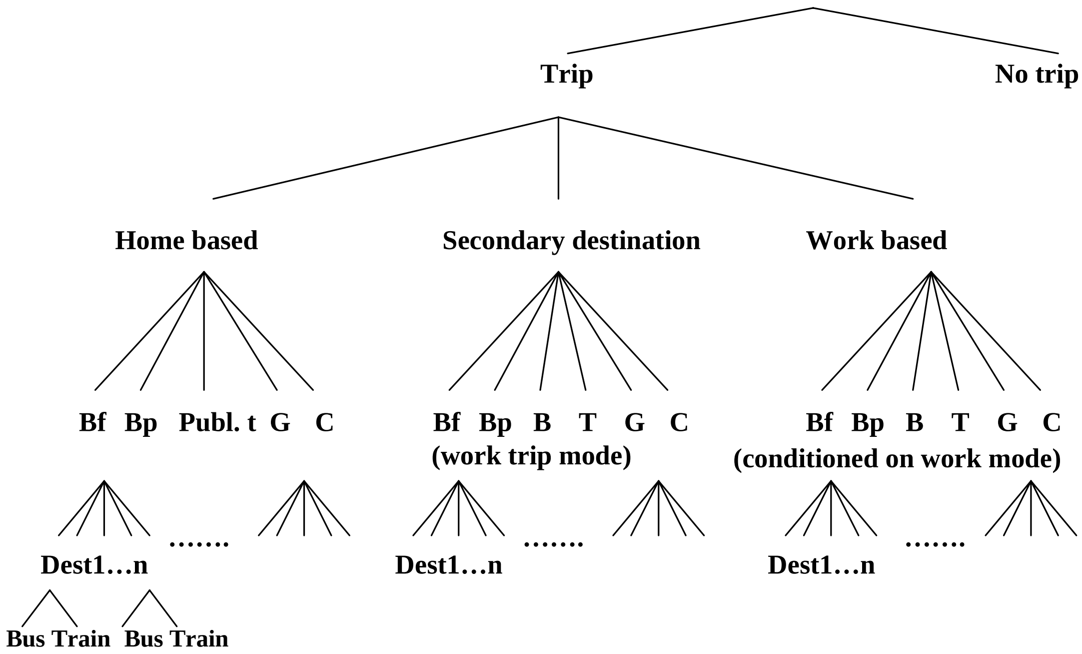

Figure 2 Preliminary model structure for non-working tours

Work Tours

A base model was first developed on the total regional data set. Then, the same specification was used on the different data sets for each of the five subregions. A major finding is that the cost and in-vehicle time parameters are very similar in the different regions. Therefore, a model including a number of region specific constants was formulated. In this way, the major part of the regional differences could be accommodated in one single model.

The model was simultaneously estimated, and contains significant logsum parameters between the different choice dimensions.

Concerning travel time components, it has been found that a piecewise linear formulation of (first) waiting time and auxiliary time gives significant improvements. It has also been found that the in vehicle time parameter is very similar for the different modes. An important part of the cost variable is related to Swedish taxation rules, making it possible to deduct work trip travel cost in some cases. The implementation of the model allows for analysing effects of changes in these rules.

The number of destination variables has been very limited. The only size variable used was the log of the total number of employees. Another variables that turned out to be significant was a dummy variable indicating the central area in each county. This variable was defined to be mode specific, and was supposed to account for omissions of parking costs, parking search times etc.

Socio-economic variables have also been included, such as car ownership, license holding, gender and type of employment.

The value of time for work trips amounts to 50 SEK per hour, which is about 40 percent higher than what was found in the 1994 Swedish Value of Time Study, based on Stated Preference data (Algers et al., 1995).

Non Work Tours

For home based tours, models including mode, destination and frequency choice have been estimated. As for work tours, regional differences were captured by region specific constants rather than estimating one model for each subregion, the main reason being scarcity of data for the train mode.

As for work tours, cost and in vehicle time parameters get significant values. Also, the resulting values of time are substantially higher than those found in the 1994 Swedish Value of Time Study.

Generally, destination choice is a week part in travel demand models. In previous Swedish models, size variables have normally defined as the number of employed persons in different economic sectors, because of the general availability of this information. Such variables give of course a very parsimonious characterisation of a destination zone. In this project, an extra effort was made to improve this part of the models. This was done by collecting information on a number of zone attributes, such as number of university students, number of beds in hospitals, number of summer house square meters, dummy variables for supply of winter and summer activities (like skiing facilities and camping facilities) etc. These variables improved the model significantly.

Most of the models contain individuals from 7-74 years. This is of course a fairly heterogeneous group, and the models contain many variables to account for that. It may well be that further segmentation with regard to age would turn out to be efficient.

All models, except for school trip frequency, contain significant logsum parameters integrating the different choice dimensions.

Domestic Long Distance Models

Model Structure

Long distance trips are modelled as tours, longer than 100 km in one direction. The modes for these trips are defined to be car, bus, normal Inter-city (IC) train, X2000 (high-speed) train and air.

A two-phase approach has been defined in this case. The first phase concerns the development of a nested disaggregate logit model with frequency, destination and mode choice, which had been used before (Algers, 1993). In the second phase, departure time, class and access/egress mode choice is added to the choices modelled in the first phase. At this time, only the first phase model has been developed.

The model structure of the first phase model is such that frequency choice is at the highest level, mode choice at the middle level and destination choice is at the lowest level. In this phase, access/egress is treated in a simple way, like adding a mode specific access/egress distance variable in the main mode choice model.

Two trip purposes have been defined – private trips and business trips.

Estimation Results

Both for business trips and private trips, mode and destination choice models have been estimated simultaneously. It turns out that for private trips there is some heterogeneity related to the number of days at the destination. Therefore a further segmentation according to this criterion was made. The first segment model concerns those staying away for more than 5 days, and the second segment model concerns those staying away 5 days or less. The latter model contains a further partial segmentation, where time components are segmented on those who make a day trip, and those who stay away up to 5 days. The specifications are also different in structure – it turns out that the first segment (duration > 5 days) requires the mode choice to be at the lowest level, and that the other segments require destination to be at the lowest level. Thus it seems that we explain destination choice relatively less well for trips with longer duration, which are often vacation trips.

For the destination choice, size variables were first defined as the number of employed persons in different economic sectors. As for the regional models, the extra variables for destination choice improved the model significantly.

Frequency models have been estimated using disaggregate data, but taking the observation period to be less than the maximum trip frequency, thereby allowing the use of observed frequencies in terms of probabilities to make a trip in the defined period (Daly, 1997). As an example, if no one makes more than 10 trips in a month, then a tenth of a month can be taken as the observation period. The probability to make a trip in such a period will range from one for the observed maximum frequency, to zero for those not having made a trip, and values in between for the rest. The advantage of this approach is that the choice becomes a binary choice.

Values of time for domestic long distance private trips for those staying away up to 5 days have been calculated based on the mode and destination choice model. The values are higher for trips with a shorter duration, probably related to a sharper time constraint. The value of time for trips exceeding 5 days was substantially lower. The in vehicle values of time in the mode and destination model are on the average reasonably close to those found in the 1994 Swedish Value of Time study (which was not segmented in the same way) (Algers et al., 1995).

Also for business trips, values of time have been calculated. As for private trips, trips with shorter duration have higher values of time. As can be expected, time values for business trips are substantially higher than for private trips.

International Models

International trips are classified into two main models – one for Swedes travelling to and from other countries, and one for non-Swedes travelling to and from Sweden. There is one important exception from this rule; namely the short trips being made between the very south of Sweden and the Danish island Sjælland (including Copenhagen). These trips are handled as an extension of the regional model for the southern region.

Model Structure

The model structure contains three choice dimensions – trip frequency, mode choice and route choice, where routes are classified according to the ferry connection. This structure is depicted in Figure 3 below.

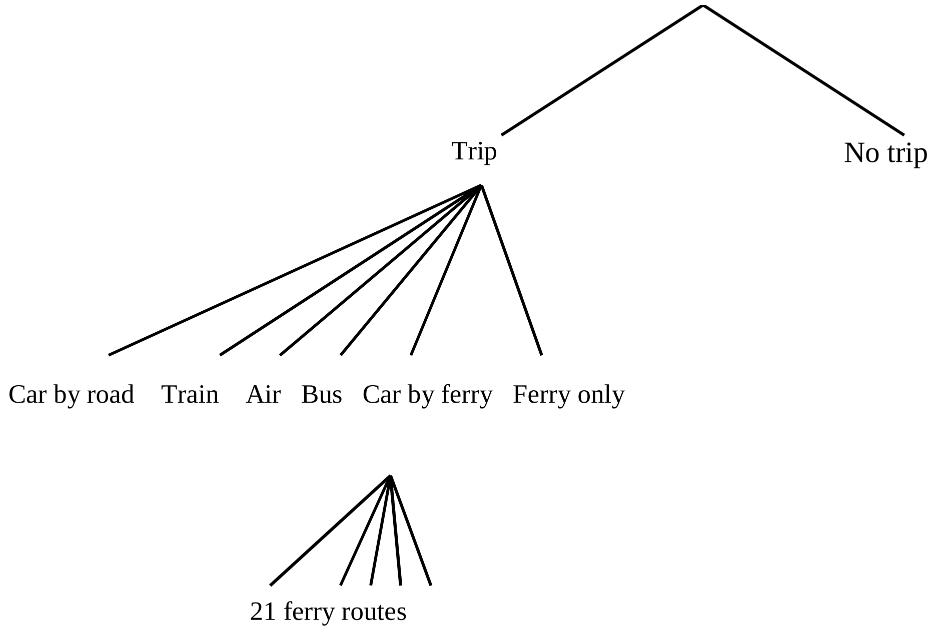

Figure 3 Structure of international trips generated in Sweden

The RiksRVU travel survey did not contain information on the route choice, so other data sources were needed to get the route choice part of the model. A joint estimation of the model is therefore deployed in order to get parameter estimates for the full model structure.

The estimation has resulted in significant estimates for time and cost variables for the mode and rout choice parts, and in significant estimates of main variables such as income and accessibility (logsum variables from the mode and route choice parts).

Car Ownership Model

As required by the client, the car ownership model implemented in the system is a previously developed cohort based model for car ownership, based on individual entry and exit probabilities for car ownership. The model was developed by the Swedish National Road and Transport Institute. The model gives zonewise car ownership levels. The main variables in this model are income, fuel price, age and company car.

Validation

The currently implemented models are now being validated. The validation of the models is made in different ways. In the estimation phase, the ability of the model to replicate the choices actually made is tested for different classifications of the data. After implementation, the models are compared with the base information from the travel survey. Finally, the model predictions for the base year are compared to other sources of information, mainly traffic counts. Also, elasticities are calculated before and after model implementation.

System design

Windows menu system

The SAMPERS system is built up as a Windows menu system. The software is developed using Visual Basic as the basic program language. The system contains a number of basic features, that can be put together to forecasting scenarios, which in turn can form a forecasting project. Scenarios are given properties that relate all runs within the scenario to a certain database for zonal data, forecasting year etc. In the same way, projects are given properties that relate all scenarios to the same level of aggregate result output. For forecasting, the following features may be invoked by the user:

- Car ownership model
- The EMME/2 system, directly or by macros
- Regional models
- Domestic long distance models
- International models
- Disaggregation of trips from national to regional level
- Iterations (such as car assignment and a regional model run)

The main menu is shown in Figure 4, where the user (by clicking on menu buttons) has put together two scenarios, in order to forecast the effects of an extended high-speed train network. The first scenario is the base scenario, and the second scenario is the extended high-speed train network. In the base scenario (which might of course have been run already), the first steps are to calculate supply matrices for the different modes. This is done by invoking EMME/2 macros. Input and output are defined by setting the properties of the macros.

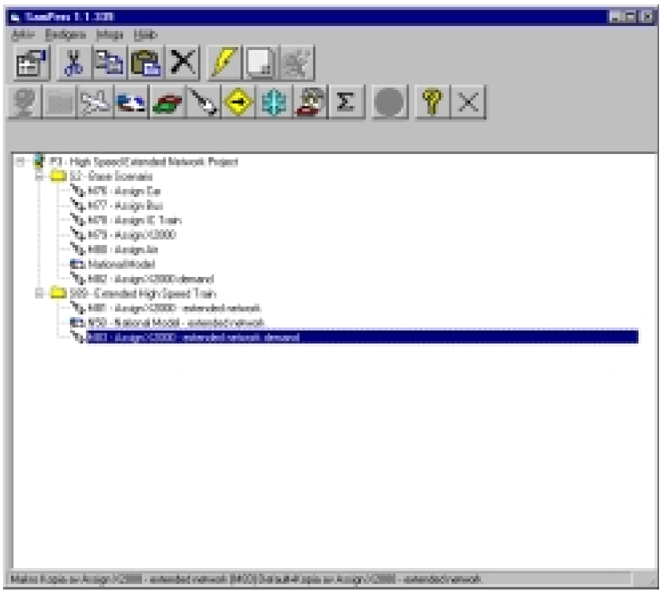

Figure 4 SAMPERS main menu

The next step is then to run the domestic long distance model, using the supply data created by the previous macro steps. As for the macros, input and output is defined by setting the properties of the domestic long distance model (as is shown in Figure 5). The last step is then to assign the resulting trips to the network, again by invoking EMME/2 macros for the different modes.

The first step in the extended high-speed scenario is the supply macro for high-speed train, which in this scenario has an extended network. The supply for other modes is not changed, and does not have to be recalculated. The next step is to run the domestic long distance model, now with properties set to match the new high-speed train supply. In a final step, a new assignment is made of the resulting demand matrix for the high-speed train mode.

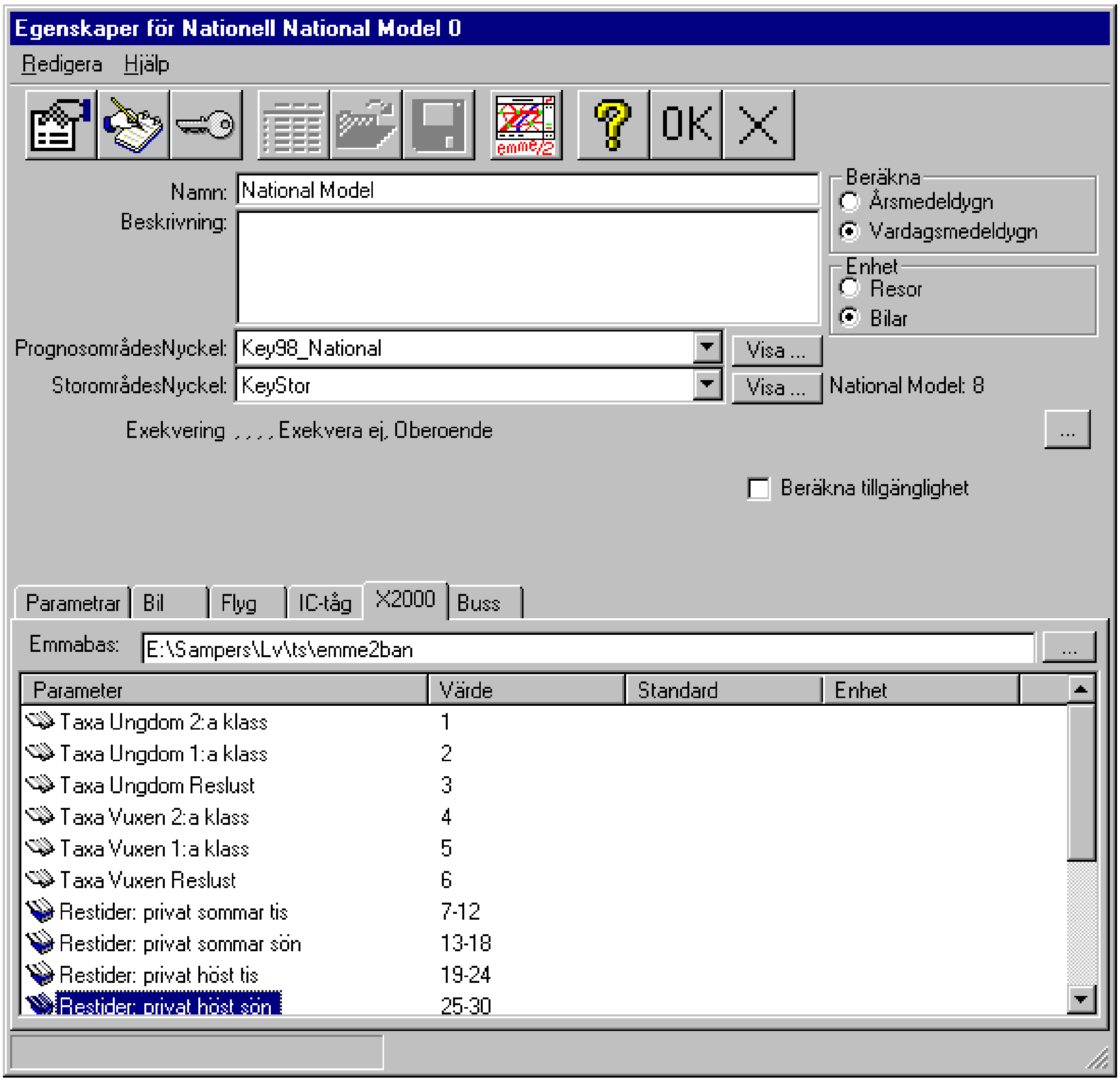

Figure 5   Example of properties sheet for the national model

The definition of the scenarios can be made separately from the actual running of the macros and models, to make the work as efficient as possible. Different steps can also be assigned to different computers, to give complicated runs (such as running many regional models) a shorter turnaround time. Dependencies between different steps can be introduced to ensure that steps are carried out in the right sequence.

Emme/2 integration

The supply data needed for the models to be run are created in the EMME/2 system. Normally a number of travel time component matrices need to be exported from the EMME/2 databank to be accessed by the forecasting software. Then, result matrices need to be imported back into the EMME/2 databank. This may be very time- and storage consuming, and in order to avoid this a more direct process was implemented. Thus, in the SAMPERS system the EMME/2 databank is directly accessed from the SAMPERS modules, for reading as well as for writing.

Results

When the scenarios in the example are run, the user may want to look at the results. For EMME/2 macros, results are stored in the EMME/2 databank, and can be viewed by invoking the EMME/2 system directly from the SAMPERS menu. For model runs, results are always produced at an aggregate level in the form of tables. If the user wants, then also the Cognos PowerPlay system can be invoked from the SAMPERS system, allowing results to be graphically displayed for a rich variety of categories. In figure 6, the travel demand of a scenario displayed with regard to mode and trip length.

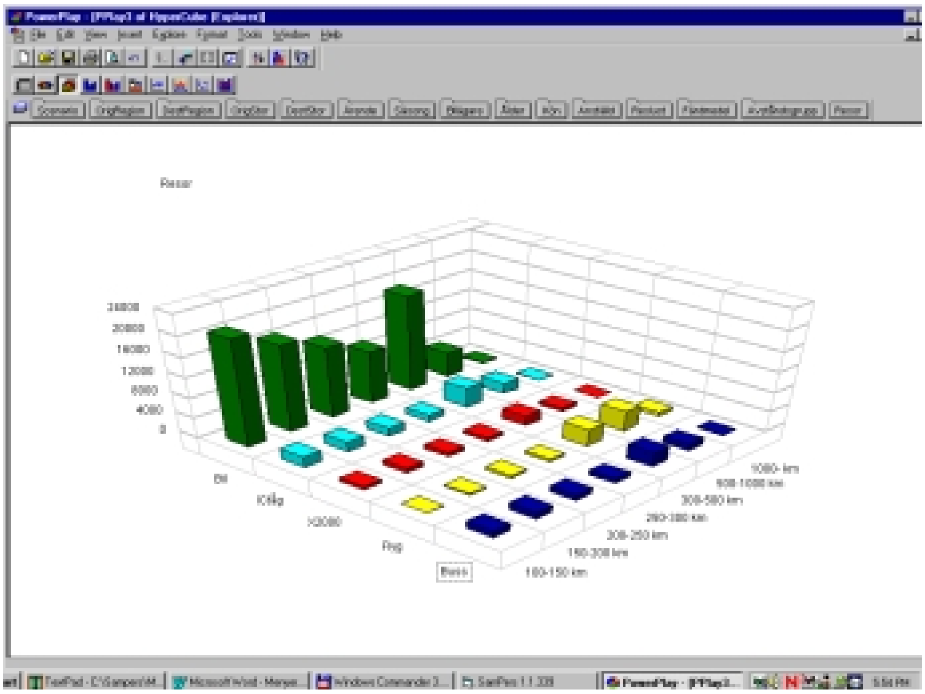

Figure 6 Cognos PowerPlay graph showing results from the national model

Cost benefit calculations

The demand forecast is often only a partial result in the planning process. Environmental impact and cost benefit calculations are further steps that are required by decision makers. The SAMPERS system contains two modules to match these requirements. In the effect module, emissions and accident rates are calculated. Then, in the cost-benefit module, these effects are economically evaluated together with the direct costs and benefits in terms of travel time and vehicle cost changes. As the other SAMPERS modules, the effect modules and the cost benefit module is invoked from the menu, and can also be included in the scenario calculations as separate steps like macros and model runs.

Accessibility

Yet another result dimension is formed by different accessibility measures. SAMPERS also contains an accessibility module, where a number of different accessibility measures can be analysed in a GIS-oriented way using ESRI MapObjects. (Figure 7). Some of these measures are related to single scenarios, whereas other measures relate to scenario differences. The are three different types of accessibility measures:

- Impedance measures, i.e. travel time or generalised cost to reach certain areas
- Closeness, i.e. how many work places can be reached within a certain time
- Model based data, i.e. passenger distance travelled, or logsum measures

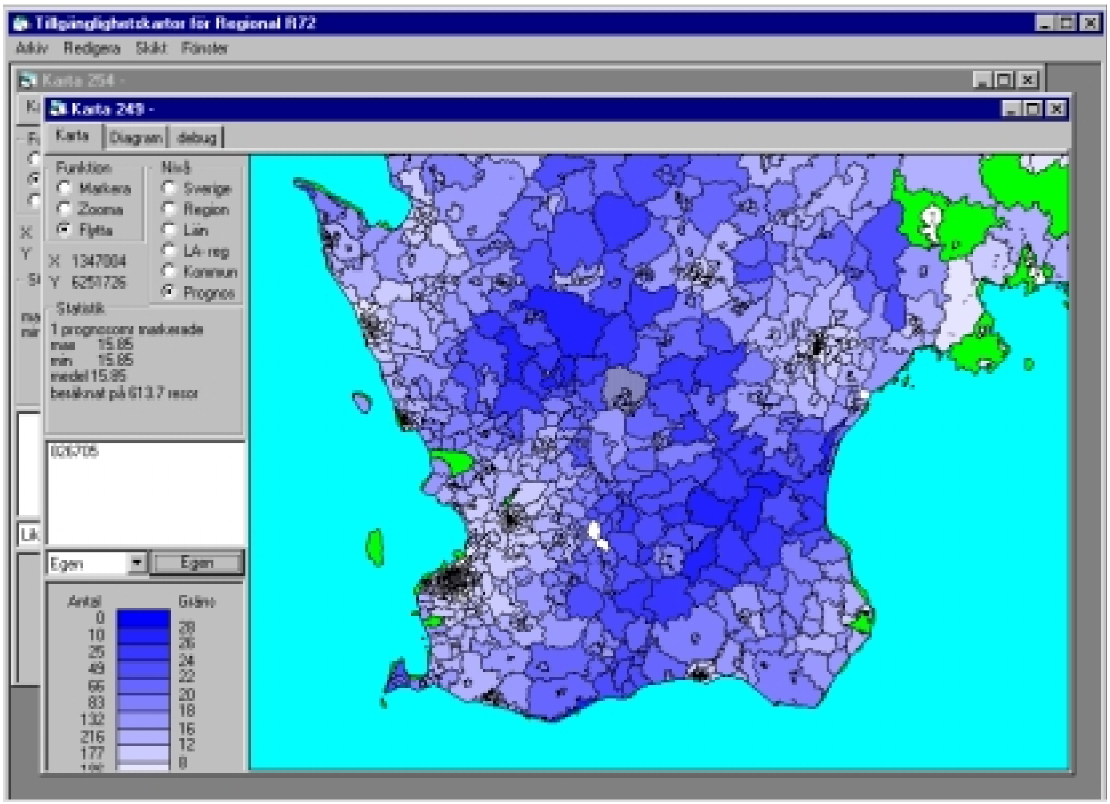

Figure 7 MapObject application in SAMPERS accessibility module

Discussion

The SAMPERS model system contains many improvements over previous Swedish model systems:

- It covers all personal trip making
- It is based on recent data covering all trips
- The model structures are enhanced
- All models, including assignment is handled within the same user-friendly software
- Effects, cost benefit and accessibility analysis can be carried out in the system

However, as can be expected from a system containing a high level of geographical and socio-economic detail, run times are long. On a standard PC, run times for the regional models (which take the longest time to run), vary from 4 to about 30 hours, depending on the size of the region. If a full set of models is to be run on one computer, several days are needed, especially if many iterations between demand and assignment are needed to capture road network capacity restraints. The possibility to run a project on different computers can however reduce the run times substantially.

Therefore, it is currently being discussed to create a sketch version of the system, in which some of the socio-economic detail is given up in favour of increased speed. For example, gender and age groups may be reduced, which may cut run time by a factor of at least four. The price is some aggregation error and less socio-economic detail in the results, which may not be too important until the end of the analysis process.

There are still some parts in the system that are not yet developed. These are regional trip chaining, long distance access/egress mode choice and departure time and ticket type choice for domestic long distance trips. They will not only add to the capability of the system, but will also make the system more time consuming to run.

References

Algers, S, (1993). An Integrated Structure of Long Distance Travel Behaviour Models in Sweden, Transportation Research Record 1413, Washington D.C. 1993.

Algers, S., Lindqvist-Dillén, J. and Widlert, S. (1995), 'The National Swedish Value of Time Study', Proceedings of Seminar F, PTRC, 1995.

Daly, A, (1997) Improved methods for trip generation, Proceedings of Seminar F, PTRC, 1997.

Lundqvist L, and Mattson, L-G, editors, National transportation models - recent developments and prospects, PTRC conference in Stockholm 1998, forthcoming 2000.