CONCERTED ACTION COMMITTEE ON INFORMATION SYSTEMS,
22nd/23rd OCTOBER 1998 IN BRUSSELS

Needs for Modelling Purposes

Airport Choice & Competition – a Strategic Approach

Benedikt Mandel[1] and Werner Rothengatter[2]

[1] MKmetric Gesellschaft für Systemplanung mbH, Karlsruhe
email: *mandel@mkm.de*
[2] Institute of Policy Research (IWW) at the University of Karlsruhe
email: *rothengatter@iww.uni-karlsruhe.de*

Abstract Basically the following article refers to the fact that data needs for modelling generally depend on policy needs.

Keywords Travel demand, transport policy, policy needs, queries, aggregate and disaggregate data, consumer behaviour, system approach, generation, distribution, mode choice, four-step approach, data needs.

TABLE OF CONTENTS

ANNEX - PRESENTATION 22ND & 23RD OF OCTOBER IN BRUSSELS

1. POLICY NEEDS AND MODELLING

To verify which data are needed for modelling purposes first investigations have to be done in the underlying requests. As models are representations of parts of the real world modellers are urged to specify the intended use of the models to develop.

Focussing on transport policy one can distinguish between two scopes of interest:

– Describing current and/or past situations in transport
– Explaining and forecasting travel demand

While the first scope could often be modelled with aggregate data, i.e. global statistics and less detailed figures, the second one requires individual based information, i.e. disaggregate data that in most kinds is provided through surveys. It is a question of detail of the requests by decision makers that decides whether aggregate or disaggregate sources are necessary for appropriate modelling. To a certain extent forecasting could also be done using aggregate data; when considering time series of indicators in the past trend based approaches might give a first rough outlook for future developments. Decision makers should be aware, that such approaches apply the developments of the past – often uncorrected – to coming periods, therefor these methods will give constrained predictions about the future.

The main drawback of such an approach is its limited consideration of interdependency of the underlying impacts that influence travel demand. The absence of information about the driving forces and – what counts most - their true interrelationship, makes it impossible that models based on aggregate time series could anticipate future developments as detailed as required. The absence of harmonised measures for e.g. different years can only lead to snapshots of transport patterns at a certain point in time and will only allow to fill existing gaps on the base of a priori assumptions. If the degree of aggregation is lowered from global statistics – e.g. annual mileage - to e.g. link counts, more detailed insights will be provided by the models developed. But such a degree of detail will only allow to analyse bottlenecks or general demand figures for a given supply of transport services. If sensitivity of models is required with respect to the individual consumer behaviour on changing supply, an approach is necessary that enables to compute micro-economical measures like elasticities, value-of- time, thresholds of travellers reactions.

Such a more sophisticated instrument will capture the interdependency of supply and demand on a disaggregate level. If decision makers are e.g. interested to evaluate ex ante the consumer reaction concerning transport policies like a planned fuel tax increase, stricter speed limits on highways, subsidies on intermodal transport or large infrastructure investments very sensitive model instruments to measure effects appropriately should be applied. Of course one can also apply time series models on aggregate data to produce elasticity figures and to analyse reactions ex ante, but due to the structure of these data (for example neither information on the true origin and destination of the trip nor individual elasticities are available) which differ a lot concerning the trip purposes, e.g. consumer segments, in general, models on aggregate data are by far not as refined as the ones based on disaggregate data and therefore the latter ones are more likely to be used to analyse the consumers' sensitivity to policy actions.

From a scientific point of view, modelling shall result in a homomorphic representation of the empirical object. This leads to the challenge to model transport behaviour on a micro scale with disaggregate data to explain the decisions behind the behavioural responses. The scientific issue of constructing a micro foundation of demand analysis results in high standards for data bases covering the socio-economic clustering, the decision environment (alternatives, constraints, etc.) and observable individual traffic behaviour. Hence, disaggregate data sources are compelling for a scientifically sound and sophisticated model approach while more aggregate data are necessary for validation of model results and calibration of structurally appropriate specifications.

When attempting to develop transport models covering Europe as a whole, model input from all possible sources in Europe will be necessary to ensure a representative instrument for all regions to consider. While some harmonised aggregate data is available across Europe from EUROSTAT and further material can be obtained from different national statistical offices, a real lack of harmonised disaggregate data exists. Household surveys are available in a number of countries but these seldom cover trips abroad – and if so the regionalisation in the foreign countries is often much too coarse to be indicative.

With some work one can combine the aggregate data to get the basic structure but for example one can not apply elasticities based on a German national survey to explain travel behaviour in Greece or one can not use the values-of-time based on a Cross-Channel survey to explain consumer reactions in Italy. One has to build up an European database if one wants to establish a real European model. Such a multi-country database is capable to reflect the different consumer behaviours so that models can work out the underlying decision structure and existing barriers. In case a homogeneous representative multi-country survey database does not exist a lot of econometrical effort and a number of exogenous hypotheses are needed to construct synthetic models which are by far not as sensitive and precise as comparable disaggregate models would be.

It should also be stated that the availability of data decreases in general due to the established common EU-market which had a big impact on the national data collection (see border-crossing freight statistics) and the ongoing privatisation where former public companies are no longer obliged to publish any transport related information (see e.g. the German railway company). Therefore it soon will happen that also more global statistics will vanish or have to be substituted by synthetic information due to a missing legal basis for a reporting system.

Coming back to the point of view of transport policy, models are needed which are transparent, easy to understand and policy sensitive such that the models can produce detailed information on the expected impacts of policy actions. For policy applications it is also important that models include the dynamic dimension and consider the basic feedback mechanisms in a broader system (e.g. transport and land-use). While transparency and inclusion of dynamic feedbacks would advocate to construct simple transport models for direct decision support the issue of policy sensitivity would imply a higher sophistication. This is because different policy actions may affect different levels of decision making and result in different impact mechanisms. For example large infrastructure investments are influencing the accessibility of regions and therefore their economic or touristic attraction and herewith the choice of destination, the mode choice, the route choice and in addition they might induce traffic and cause other socio-economic and environmental as well as transport specific synergetic effects.

As a consequence modelling approaches are preferred which have a scientifically based and sophisticated foundation, as well as allow for an easy aggregation to give decision support to the transport policy maker. Practically this means that demand modelling should be based on detailed traffic survey data allowing for disaggregate econometrics which has to be embedded in a GIS environment to show the spatial impact dimension and a macro–surface to generate aggregate indicators which can be interpreted on a cross section or dynamic base.

2. QUERIES AND DATA NEEDS FOR MODELLING

Referring to the well-known structure of the classical 4-step transport modelling procedure examples for queries a transport policy maker might have and the required data needs for modelling purposes is outlined.

2.1. Generation

The first step considers the generation and attraction factors of transport of a zone. Usually the socio-economic and land-use factors as well as zone specific attractions like environmental and cultural elements determine the attractivity of a region. These generating and attracting moments can be influenced directly by economic policy and indirectly by transport policy due to a better accessibility of the region.

For transport planer questions mobility rate and time budget for example are of interest. Variables could be seen as overall indicators or with respect to the specific consumer segments and their behaviour, e.g. the influence of a highway on the economic prosperity of the region versus a possible reduction of economic activities. On the freight side it is interesting to know what kind of economic structure leads to which transport activity and how the development of the region in the long term will be influenced.

In this step it is crucial to have a homogeneous database consisting of a maximum of socio-economic and land-use factors as detailed as possible. In addition environmental and cultural attractivities should complete the description of a region. These information are the base to analyse why and how much transport demand is generated in or attracted from a region, they define the market potential. For modellers, transport demand depends on such independent information. While the factors can be covered by the regional statistics, the demand can be based on counts or representative household surveys. The first has the advantage to be quite cheap and easy to be installed as permanent automatic procedure which reflects to a certain time a specific situation as aggregated network load, but only the latter information type can provide the modellers with knowledge about the number of trips, immobile persons, time spent for travelling, relation to the modes, trip purposes, etc. by consumer segments.

If decision makers only want to have a rough idea about the relationship of demand and factors then link counts can be used to answer the question. If they want to know something about the process of transport, the interdependencies and synergies, why and how transport is taking place, etc. more generally, if they want to ask questions about the background to base their transport policy decisions on that knowledge then one needs to have disaggregate information provided by representative household surveys.

2.2. Distribution

Based on the transport potential and the attractivity of a region, transport will be distributed in space according to the different attractivities of the regions considered. The different attractivities reflect the competitive situation of regions and therefor determine the distribution of demand. Influencing a given accessibility structure by a transport policy measure like a large infrastructure investment directly changes the competitive situation of the regions. Transport flows can be shifted from one zone to another which is equal to a substitution of destinations, in addition new demand can be induced because now the accessibility of a region just moves the zone into economic valuable distance so that the destination becomes interesting, and even complementary effects will occur due to the time savings to go to the region (e.g. the seaside) one can combine two trips (e.g. visiting the marine museum near by). Of course such effects will not happen from one day to another and not all time savings will be spent for additional trips but all the stated effects will take place and have somehow to be quantified if one wants to know the influence of transport policy decisions completely.

Beside a homogeneous database consisting of the factors and attractivities as stated above information about the accessibility of the regions are needed to describe the effects. Therefor network information are needed to compute the mode specific impedance between the regions. This defines a competitive transport situation among the relations of the considered regions, e.g. a matrix of competitive relations (origin-destination pairs). The transport demand potential which is now distributed among the different competing origin-destination pairs, e.g. specific transport flows, can be derived from link counts but these information do not allow to withdraw direct information concerning the distance distribution of the trips because the true origin and destination of a trip is just not available at this aggregate level. Applying assumptions about the basic urban network load and hypothesis about the mixture of trip distances will enable modellers to construct a distribution model but it is obvious that effects of complementarity, substitution and induced transport can only be analysed on a synthetic base. Information about different consumer segments and their behaviour due to policy actions can not be worked out by models based on aggregate data.

If decision makers accept the risk of a non validated model with all its estimation biases one can stick to a link count system but if they request validated answers on the base of true observations a policy sensitive model based on surveys is needed. Especially in the context of an analyses of large infrastructure effects concerning the economic welfare of a region a detailed modelling is needed which requires detailed data. For example, if the policy wants to increase the accessibility of peripheral regions by building new highways. Now two extremes can happen, on the one hand industry will settle down to the benefit of this region and on the other hand the industry of other regions will withdraw employees from the now better accessible region and people of that region go shopping into the other regions so that just a housing area with reduced economic activity will be left over. To grasp such effects one needs detailed information covering a couple of years.

2.3. Mode choice

Within this modelling step the transport flows are split among the available modes of each origin-destination relation. This requires detailed information about the supply side including the transport relevant policy issues as well as any information about the consumers' choice of the mode. Some models also incorporate socio-economic information to capture the different behaviour of consumer segments. It is obvious that aggregate data like link counts neither contain consumer specific information like trip purposes nor can they reflect the origin - destination specific competitive situation among modes based on the different impedances, e.g. time - cost - service relationship. If one does not use hypothesis it is nearly impossible to define explicitly the consumer reaction on travel time, price or service changes which can be directly influenced by transport policy decisions.

This classical problem of multi-modality, competition of modes, where one mode is substituted by another reflects only one area of interest. Other questions concerning the inter-modality, co-operation of modes, and the intra-modality, competition within one mode segment, imply a more complex modelling and therefor more complex data which can not be provided by aggregate information. If the decision makers are interested in inter-modality the data for modelling have to contain information about the different modes used for a trip or transport action. Depending on the requested level of detail this can be broken down to an exact description of a transport chain which covers the modes used at the specific route segments of a trip. In any case, to model inter-modality one needs at least information about the main mode and the access/egress mode used, e.g. for a trip from Leeds to Heidelberg the chain of modes used like car - plane - rail and the place of mode transfer like Birmingham - Frankfurt should be asked in the survey. This information doesn't identify the exact chosen route in detail but on European level this information satisfies the needs for modelling purposes concerning the problem of inter-modality. A higher level of detailed data is requested if questions about intra-modality should be answered. Here exact information about the mode specific product chosen on a certain route is requested. For example in air the distinction between non-stop and via flights versus a certain hub, in rail whether a high speed train or an inter-regional train was used and in road which route was driven to receive any information about the behaviour concerning tolls.

Of course the Europe wide collection of such information is very expensive and can hardly be financed but a more general survey stating the points of interconnection of the modes will at least cover the European level of the inter-modality problem without increased costs. Link count data are not capable to provide modellers with data concerning inter-modal consumer behaviour. In addition the sensitivity of travellers to changes of mode characteristics depends on which origin - destination pair the trip takes place on e.g. travel time reductions on a short trip have different effects on mode choice than on a long trip. The same lack of information in the aggregate data makes it impossible to compute thresholds where consumers start to react unproportionally to the mode characteristic changes or (the other way around) questions of transport policy makers like: if x % (or an absolute number of y) travellers should be moved from road to train on a specific (or a couple of) origin-destination pair(s) what type of service (travel time, fares, frequency, ...) is necessary, can not be answered. One can enrich this argumentation by a lot of other points of interest but in general all follow the same line. If the decision makers wants to know which type of consumer, why, when and how is reacting to transport policy actions, then there is a necessity to work out the consumer’s elasticity, captivity, threshold, marginal rate of substitution, e.g. value of time, etc. on the disaggregate data level. Based on this information the policy actions can be levelled to the extent of the consumer reaction the transport policy wants.

2.4. Assignment

The last step of the classical modelling approach assigns the mode specific transport flows to the different networks using well established operations research routines based on criteria like shortest path (time, cost, generalised costs, ...) or the idea of equilibrium. Capacity checks are involved to regulate the network loads and to give a feed back to the models in the previous steps when impedances between regions change due to the exceeding capacity. Of course the basic network loads on urban level have to be considered in addition and a calibration process can be attached to reflect the measured aggregate link count information.

On European level an in-time-simulation/assignment of traffic flows will exceed the financial possibilities but if the policy makers are focusing in detail on route choice problems information about the chosen path or product and the telecommunication services have to be provided. If they are interested in the inter-modality problem skipping the exact route choice, then just information about the mode transfer points is required. In both cases the algorithms have to be enriched by route choice models or at least adjusted to reflect the observed behaviour and in the case of forecasting to load the network in consideration of the consumer behaviour.

While the aggregate link count data just state at a certain link the load and do not carry any information about the users of the link or the structure of the load (despite some technical description concerning vehicles and time-of-day-curves) up to now one applies hypotheses, heuristics and a variety of assumptions to simulate this information. To shed a light on problems like 'which consumer uses the link for a specific origin - destination trip' regional observations are undertaken where licence plates of cars are filmed or cars were stopped for a short survey (from and to where are you going, why, number of persons, how often, which alternative routing they choose, etc.). For air, rail and ferry disaggregate data already exist in form of airport surveys, rail or ferry surveys and in addition ticket & coupon databases but unfortunately these data are private, e.g. owned by the ferry and railway companies and air carriers / airports, so that the access is restricted strongly. If transport policy makers want to have any of these information they could try to get on a honoris causa base at least old data, try to build a joint survey or buy the data from the private companies. Due to the ongoing liberalisation, deregulation and privatisation process in the transport sector and herewith the increasing competition between the modes and within the modes themselves it is doubtful that private companies will release or share any data. Anyhow, if the policy makers are interested in assignment problems like stated above then the survey should contain at least some routing information.

2.5. Interdependency of the modelling steps

As already mentioned in the assignment modelling step there is a feed back mechanism between models to reflect the decision process in transport in a realistic way and to incorporate a dynamic dimension in the system (e.g. transport and land-use). On the one hand such dynamic regulations take into account that a change of impedances by policy actions, e.g. increased tax on fuel, or just by given capacity constraints of the network, e.g. traffic jam, will affect all modelling steps (changing impedances influences route, mode, destination choice and the decision whether a trip is undertaken or not (urban activity)), on the other hand different policy actions may affect different levels of decision making and result in different impact mechanisms. For example large infrastructure investments are influencing the accessibility of regions and therefor their economic or touristic attraction and herewith the choice of destination, the mode choice, the route choice and in addition they might induce traffic and cause other socio-economic and environmental as well as transport specific synergetic effects.

It is obvious that complex structures of interdependencies can not be reflected by aggregate information, this requires an observation in form of a survey. Aggregate information can enrich the observed picture but can not substitute the survey on which the structure of the problem has to be worked out and modelled.

2.6. Other models

In the same way the classical four-step approach was discussed, one can go on elaborating other type of models for topics like assessment, environment and economy but this would just enrich the details of requested information and does not add additional value for the decision whether a Pan-European mobility survey is needed or not. In the same way other models depend on the policy needs and the resulting queries which the models should answer. Therefor further descriptions are skipped for the sake of clearness.

2.7. Supply side

All transport models need information about the supply side. It is not important to have an exact graphical representation of the different tracks like satellite pictures provide it is moreover crucial for modellers to have information about the physical restrictions, the service provided and the political framework. While a good GIS representation of the networks makes life easier to understand results, to make complex structures visible, to increase the analytical component and transparency of the models, the exact representation of the supply side by model relevant link attributes is essential for modelling because these information determine the attractivity of the modes or the accessibility of a region both in quantity and quality. In addition such attributes can directly be influenced by the transport policy. If they change it directly effects the consumer behaviour.

Therefore a second effort to install a network database containing physical link descriptions as well as service attributes and a database about transport policy regulations and general cost structures for the specific modes would complete the data needs for modelling. Of course the network information have to fit to the years of the survey.

3. RESUME

Following the above mentioned issues one can summarise that the data needs for modelling are generally depending on the policy needs. But transport scientists need high-quality harmonised disaggregate data to fulfil the informational needs of the decision makers. The continuous increase in complexity in the political process forces the researchers to develop new and better - i.e. more detailed – planning instruments. This must be accompanied by new methods that enrich the existing modelling worlds so that more questions can be answered in more detail.

Modellers are interested to have both kinds of data: disaggregate data for modelling and aggregate data for model validation purposes. While some sparse aggregate data exist on European level a homogeneous disaggregate database (i.e. household surveys) for policy sensitive modelling is missing completely. After negotiating with different ministries and private transport companies and consultants one could combine some of the already existing surveys but besides to overcome the licence problem still the data can not easily be combined to produce a homogeneous representative sample.

Although regional aggregate data is provided by EUROSTAT its consistency and degree of detail is lagging behind the material that e.g. are offered by the US Bureau of Transport Statistics for North America. These free-of-charge sources have a great impact on the transport research in the USA and also world-wide due to the free access to this harmonised and disaggregate data. We would be happy if such sources would be available for the European continent or the Community at least.

If an European database with the components:

- homogeneous socio-economic and land-use factors,
- environmental and cultural attractivities,
- Europe wide representative surveys,
- already existing link counts and
- network representations with model relevant information combined with information about transport policy regulations and general cost structures for the specific modes

is available all projects in research and consultancy could use a common database which ensures the compatibility between results, models and different research areas. A neighbouring effect would be the savings of costs of each transport project because up to now every consultant or researcher generates, implements and uses the own sources which now could be provided by the EC.

Additionally we would request the involvement of modellers as well as decision makers in case a survey is designed. Considering their needs and requirements in this early step is crucial to ensure an appropriate data framework that enables all parties to rely on information as perfect as possible.

4. ANNEX - PRESENTATION 22ND & 23RD OF OCTOBER IN BRUSSELS 1998

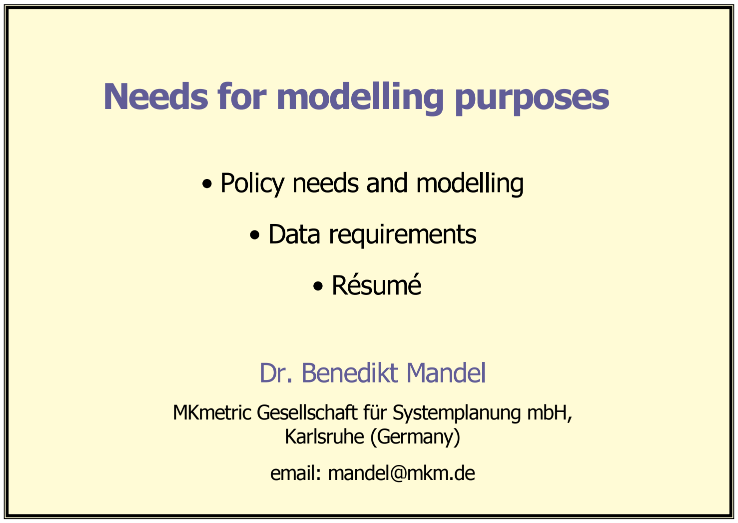

Fig. 1. Presentation Title

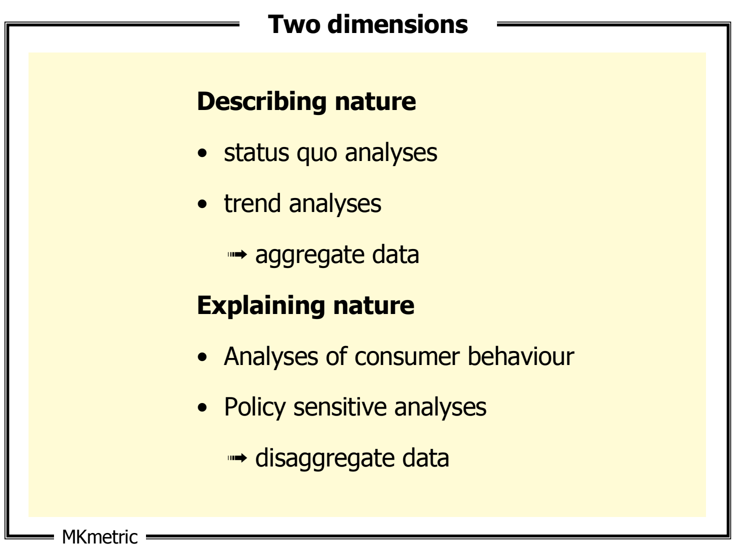

Fig. 2. Data types

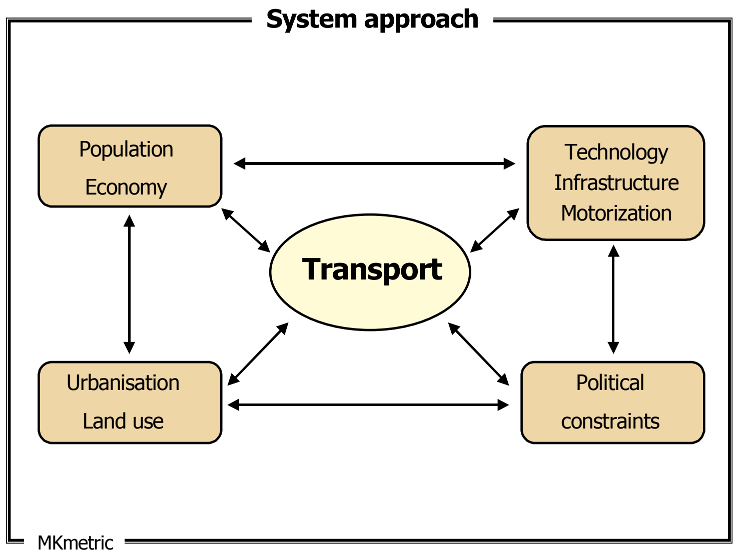

Fig. 3. System Approach

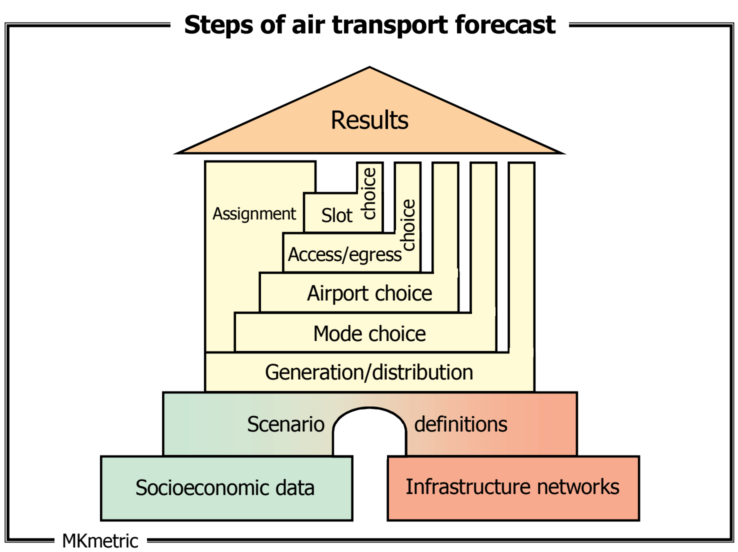

Fig. 4. Steps of air transport forecast

Modelling steps

Modelling step	Policy query	Principle data request
Generation	total demand potential mobility rate time budget	observed travel behaviour socio-economic land use culture and environment
Distribution	substitution complementarity induced traffic competition of regions economic effects	observed travel behaviour supply data attractivities/factors
Mode choice	multi-modality inter-modality intra-modality	observed travel behaviour supply data
Assignment	network capacity route choice	observed travel behaviour supply data

MKmetric

Fig. 5. Modelling steps

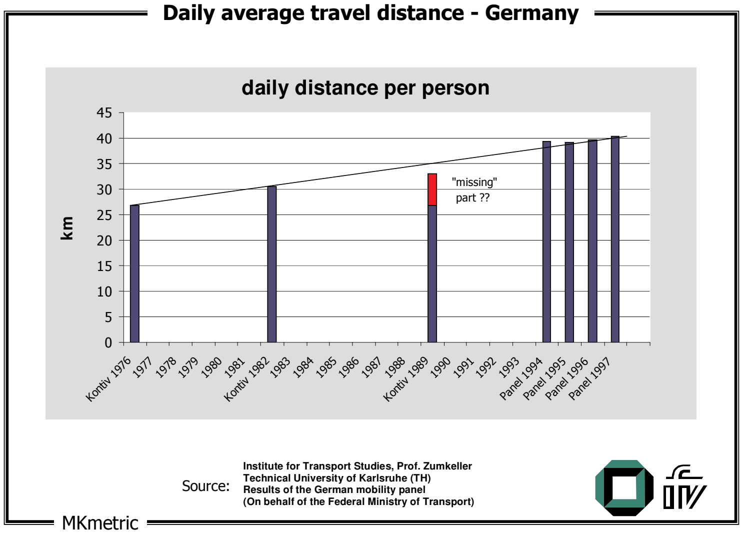

Fig. 6. Daily average travel distance - Germany

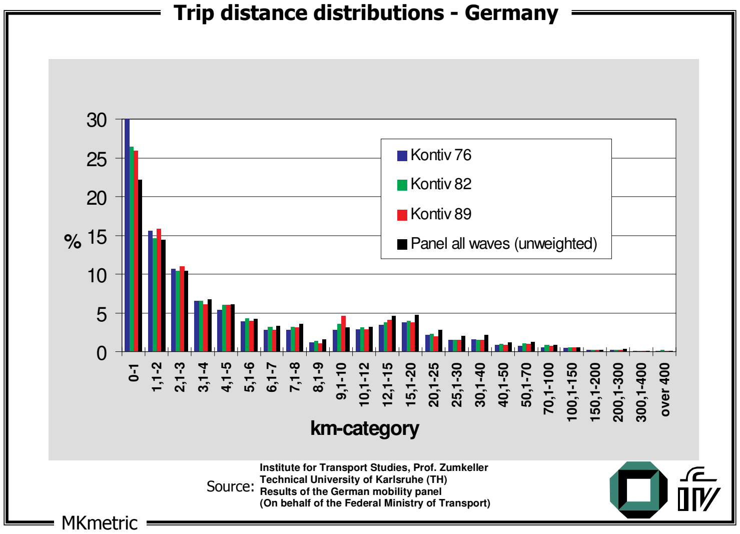

Fig. 7. Trip distance distributions – Germany

Mobility indicators - Germany

Indicator / Source	Panel 1994	Panel 1995	Panel 1996	Panel 1997	Kontiv 1989	Kontiv 1982	Kontiv 1976	BMV 1992
Trip making [%]	91,9	93,9	92,9	92,0	85,0	82,2	90,0	-
Daily trips per person	3,32	3,39	3,46	3,52	2,75	3,04	3,09	3,13
Daily trips per mobile person	3,61	3,61	3,73	3,82	3,24	3,70	3,43	-
Cars per inhabitant	0,502	0,467	0,511	0,518	-	-	-	0,508
Travel time budget [h:min]	1:19	1:20	1:21	1:22	1:01	1:12	1:08	-
Daily travel distance per person	39,3	39,2	39,6	40,4	26,9	30,5	26,9	33,8
Average trip length [km]	11,8	11,5	11,5	11,5	9,8	10,0	8,7	10,8

BMV 1992: „Verkehr in Zahlen", adjusted figures

Source: Institute for Transport Studies, Prof. Zumkeller
Technical University of Karlsruhe (TH)
Results of the German mobility panel
(On behalf of the Federal Ministry of Transport)

MKmetric

Fig. 8. Mobility indicators - Germany

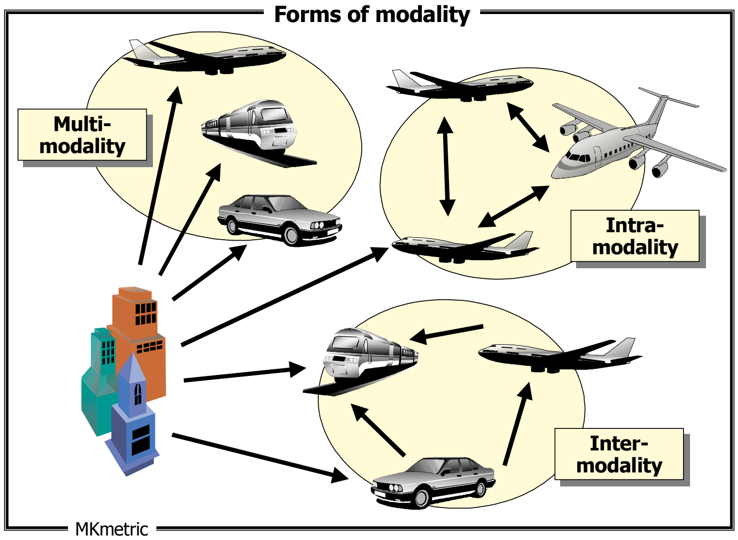

Fig. 9. Forms of modality

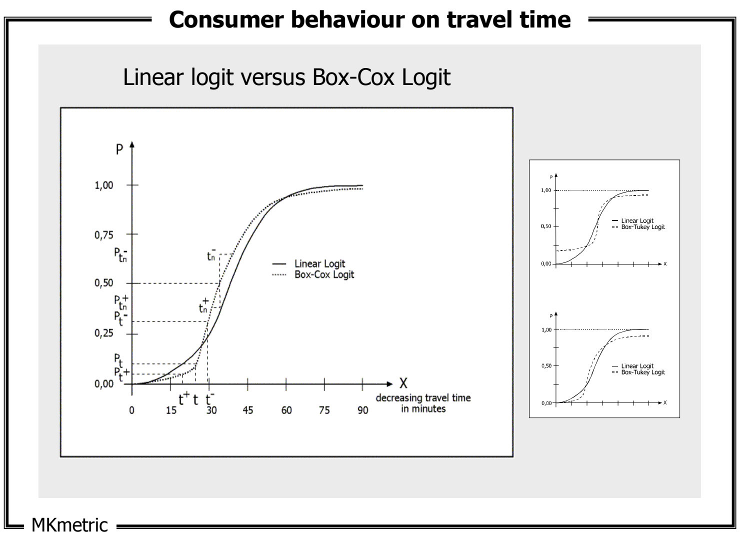

Fig. 9. Consumer behaviour on travel time

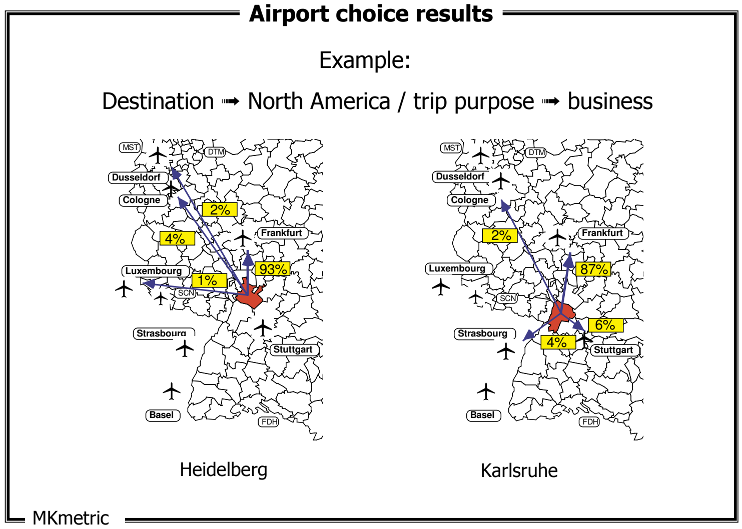

Fig. 10. Airport choice results

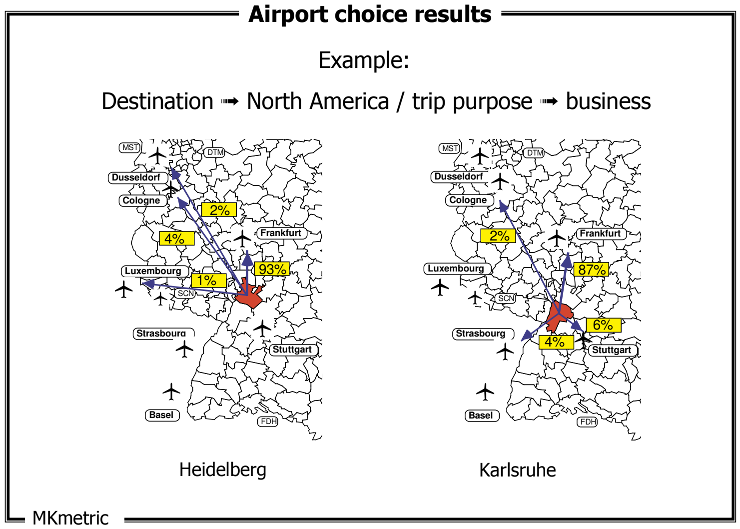

Fig. 11. Airport choice results

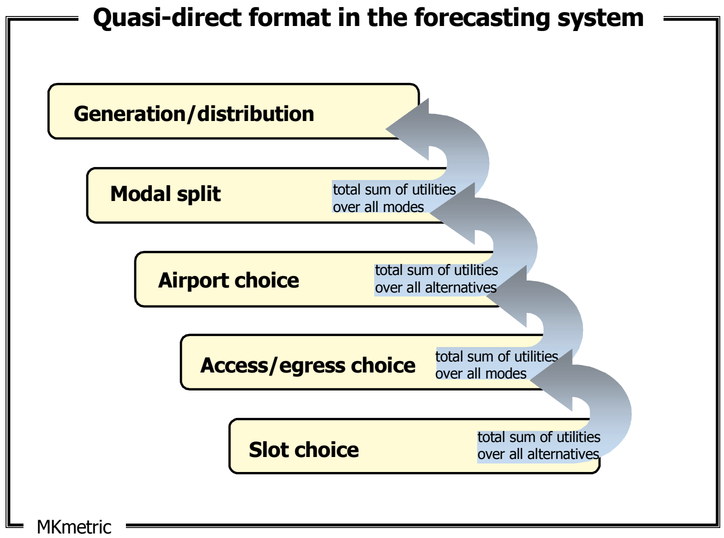

Fig. 12. Quasi-direct format in the forecasting system

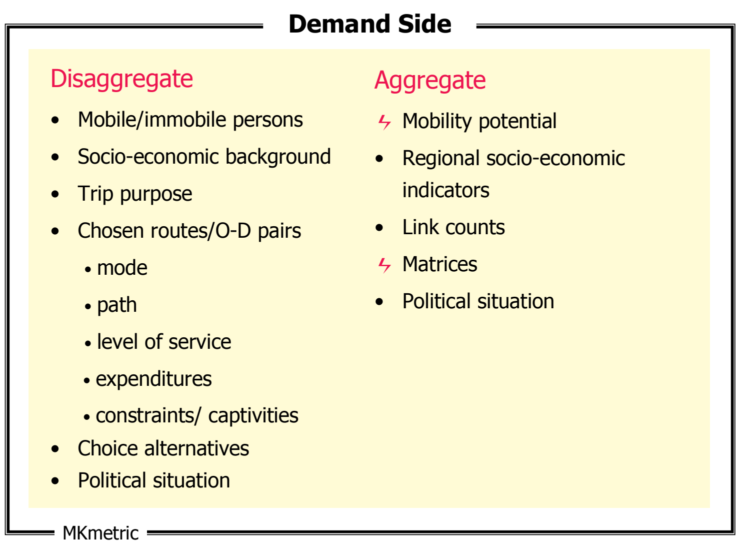

Fig. 13. Demand side

Supply Side

Infrastructure

- Network representation by mode
- Link-based information
  - time
  - speed
  - capacity
  - etc.
- Transport policy

Transport services

- Algorithms by mode and purpose
- Vehicle-based information
  - speed
  - capacity
  - level of service
  - fix/ variable costs
- Number of competitors
- Transport policy

MKmetric

Fig. 14. Supply side

General data needs for modelling

- Observed demand
- Harmonised socio-economic indicators
- Land use information
- Environmental and cultural characteristics
- Mode specific networks
- Transport policy regulations
- etc.

MKmetric

Fig. 15. General data needs for modelling

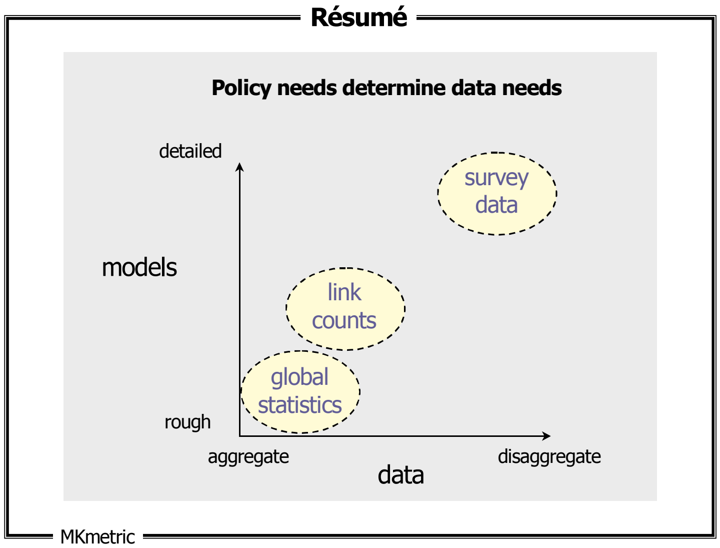

Fig. 16. Résumé

Résumé

Determinants

- Policy needs
- Trends in data availability
- Evaluation of large infrastructure investments (TEN)

Model improvements

- Pan-European model
- Policy sensitive
- Higher accuracy of results
- Consumer segment specific
- Decisions support system

MKmetric

Fig. 16. Résumé

Pan-European survey

- Household survey (mobile/immobile)
- Trip length minimum: 100 km
- Trip characteristics (detailed)
  - expenditures
  - route/destinations
  - purpose
  - etc.
- Socio-economic background
- Panel of interviewees

MKmetric

Fig. 17. Pan-European survey