WCTR 2013 RIO

SELECTED PROCEEDINGS

TRANSPORT DEMAND PROJECTIONS: A BAYESIAN NETWORK APPROACH

SULE ONSEL, DOGUS UNIVERSITY, DEPARTMENT OF INDUSTRIAL ENGINEERING, ISTANBUL, TURKEY, SONSEL@DOGUS.EDU.TR

FUSUN ULENGIN, DOGUS UNIVERSITY, DEPARTMENT OF INDUSTRIAL ENGINEERING, ISTANBUL, TURKEY, FULENGIN@DOGUS.EDU.TR, CORRESPONDING AUTHOR

OZGUR KABAK, ISTANBUL TECHNICAL UNIVERSITY, INDUSTRIAL ENGINEERING DEPARTMENT, ISTANBUL, TURKEY, KABAK@ITU.EDU.TR

OZAY OZAYDIN, DOGUS UNIVERSITY, DEPARTMENT OF INDUSTRIAL ENGINEERING, ISTANBUL, TURKEY, OOZAYDIN@DOGUS.EDU.TR

This is an abridged version of the paper presented at the conference. The full version is being submitted elsewhere. Details on the full paper can be obtained from the author.

ISBN: 978-85-285-0232-9

13th World Conference on Transport Research

www.wctr2013rio.com

15-18 JULY 2013
Rio de Janeiro, Brazil

unicast

TRANSPORT DEMAND PROJECTIONS: A BAYESIAN NETWORK APPROACH

Sule ONSEL, Dogus University, Department of Industrial Engineering, Istanbul, Turkey, sonsel@dogus.edu.tr

Fusun ULENGIN, Dogus University, Department of Industrial Engineering, Istanbul, Turkey, fulengin@dogus.edu.tr, *Corresponding Author*

Ozgur KABAK, Istanbul Technical University, Industrial Engineering Department, Istanbul, Turkey, kabak@itu.edu.tr

Ozay OZAYDIN, Dogus University, Department of Industrial Engineering, Istanbul, Turkey, oozaydin@dogus.edu.tr

ABSTRACT

In this paper, a decision support system is proposed based on Bayesian Net to support policy makers in their analysis of the impact of socio-economic, environmental and transport-related variables on the future passenger and freight transportation demand. Different what-if analyses are conducted to guide transportation policy makers in their future strategic decisions. By this way, it is possible to facilitate analysis of the possible consequences of a specific policy on changing the share of transportation modes for both passenger and freight transportation and to highlight in detail the conditional relationships among variables that are considered relevant in the transportation system analyzed. What-if analyses can also be used to show the impact of a change in any variable on the whole system.

Keywords: Bayesian Networks, transport demand projection, transportation policy, what-if analysis

1. INTRODUCTION

European Union (EU) countries recently admitted that, in terms of environment, business efficiency, health and extending road capacity to levels that can keep pace with predicted growth in traffic, their current trends in traffic are unsustainable.
Highways have the highest share of passenger (79%) and freight transport (44%). The share of railways in freight transport decreased from 21% in 1970 to 8.4% in 1998, while its share in passenger transport is currently only 6%. Transport problems are expected to worsen due to the fact that automobile ownership tripled between 1970 and 2000 (from 62.5 million to 175 million) and the movement of goods is projected to increase by 50% by 2010 (European

Commission, 2001). In 2007 the percentage of EU-25 countries are 84.9 and 43.5 for highway passenger transportation and highway freight transportation respectively (Republic of Turkey, Ministry of Transport, 2007).

In September 2000, the European Commission accepted the "European Transport Policy for 2010: Time to Decide" White Paper and thus accepted the importance of an integrated multimodal system which will provide a balance between the transport modes by 2010. The aim is to provide high quality and safe transportation that will support sustainable growth. An annual budget of 30 million Euros has been allocated to the Marco Polo Program for this purpose, in order to provide the integration of railways and maritime lines. In terms of infrastructure, in accordance with the TEN, the commission is concentrating on investments that will permit the transfer the shipment of goods to railways (European Commission, 2001).

The European Commission 2011 White Paper also admits the strategic importance of transportation for both the economies and societies and provides a roadmap to a Single European Transport Area with a competitive and resource efficient transport system (European Commission, 2011). The entire above-given trend shows that the transport investment decision should be realized with great care in a way to reduce the imbalances among the transport modes. This also necessitates a detailed analysis of demand projections in order to see under which conditions an unbalanced transportation demand projections will occur and what are the countermeasures can be taken for this purpose.

There has been a considerable increase in transport demand modeling since the late 1970s. This has been due mainly to the developments, which have taken place in the field of discrete choice modeling, together with the availability of better microeconomic data sets. Understanding and predicting traveler behavior is a complex activity.

Conventional models transport mode demand is based on the maximization of the random utility, which is the rule of discrete choice modeling (Ben Akiva and Lerman, 1985; Cascetta, 2001). Recently, a different approach to choice analysis, based on Artificial Neural Network (ANN) models has been proposed. ANN is a non-linear black box model and it seems to be a useful alternative for modeling the modal choice behavior with a travel demand model in the presence of complex non-linear relationship and is recently used by several researchers (Cantarella and Luca, 2005; Vythoulkas and Koutsopoulos, 2003; Tsai et al., 2009; Yuehua et al. 2008).

In the literature, discrete choice models are generally used to investigate which transport mode will be selected in a specific journey (Filippini and Molini, 2003; Egger, 2000). This information is usually gathered through a survey conducted with the travelers in order to reveal their selection of a specific mode in a specific origin-destination (O-D). Therefore, the related model is a micro econometric model. In the transport mode selection for a specific O-D, the data concerning the travel time and cost as well as the demographic and socioeconomic characteristics of the surveyed people will be available. In particular cases, even the impact of environmental impacts of transportation modes can also be revealed. Then, the discrete choice model, that will based on these data will forecast and test how and in which way, the travel time, cost as well as the characteristics of the passengers will influence the selection of particular transportation modes. However, the transportation demand model that will be developed in our research, are based on macro data. The basic reason of this is the incompleteness of an O-D matrix as well as the lack of a data concerning the information about the travel behavior of the individuals.

Despite the increased interest in disaggregate demand modeling; aggregate demands models continue to be useful. The main advantage of disaggregate models is that their econometric specification can be directly obtained from microeconomic consumer theory. Moreover, they allow for an accurate definition of the variables at an individual level, which makes it possible to obtain precise policy conclusions. The exogenous variables employed usually show larger variability, therefore allowing for more precise estimation of the model's coefficients. Despite these advantages, aggregate models have other characteristics, which make them play an important role in transport demand analysis. Their main advantages are related to the costs involved in obtaining the required data, Procuring individual data is usually the main problem when trying to estimate a disaggregate model. If only the results of disaggregate models are available, their aggregation requires that certain assumptions are made, which reduces the initial advantages of disaggregate models (Asensio, 2000).

As can be seen in the above-given researches, in many national forecasts, such as those for Germany and the methods adopted in United Kingdom until 1998, the basic strategy is "predict and provide" (Schafer and Victor, 2000). However the assumption underlying this strategy is that the future will be similar to the past and this cannot provide a useful guide to the EU in the attempt to adopt the measures toward a balanced multimodal system. In accordance with EU policies, the British Government issued a white paper on its future transport strategy, which proposes abandoning the "predict and provide" strategy in order to make way for "pragmatic multi-modalism", a more integrated transport system better suited to tackle the problems of congestions and pollution (DETR, 1998). In fact, the same arguments are recently underlined by Piecyk and Mc Kinnon (2010).

Prediction of passenger and freight demand for the distant future is critical to the planning of long-lived transport infrastructures and to assessing its consequences in order to guide transport planners in the specification of policies to be used and to avoid an undesirable growth of any transport mode. Such distant future predictions necessitate large-scale, long-term models of the transportation system but that pressing need contrasts sharply with the capabilities of existing, traditional forecasting and modeling techniques. These multivariate methods therefore deteriorate rapidly as projections for the future. It is well known that the "qualitative" or "technological" approaches to forecasting techniques are more suitable for long-term prediction. In the short-term, the assumption that the future will be similar to the past can be more easily defended. However, when the period of analysis is the medium- or long-term, it becomes very difficult to accept this principle.

Quantitative forecasting techniques analyze past data and make forecasts based on the relationship between the variables according to this data. In technological forecasts, however, although past data is important, the experts' opinions and their speculations also play a crucial role. In EU countries, depending on the wide spectrum of critical issues encountered in the transportation sector, there are several scenario-based analyses conducted, such as integrated transportation forecasting, profitability of high speed train usage and highway freight projections, etc. The database and scenarios for strategic transport report is a good reference for the variables and the related databases used in the scenarios conducted for long-term prediction of transportation demand in different EU countries. Generally, those variables are grouped as "socio-economic data", "transport economy data", "energy data", "foreign trade data", "environment data", "transportation mode price data" and "accident data" (APAS, 1999). As can be seen in the scenario models mentioned in the referred EU publication, in almost all of the scenarios developed to forecast the transportation demand of EU countries, "gross national product" is the basic variable included in the model. The chief reason for this is the high level of correlation between gross national product and transportation demand. This is followed by export, import and employment variables. The inclusion of other variables in the models depends on the nature of the research and the level of detail requested.

Ulengin et al. (2007) proposes an integrated Transportation Decision Support System (TDSS) is proposed to allow formulation of aggregate and long-term scenarios (countrywide, regional or global) in order to see the impact of different policies. However, Ulengin et al. (2007) do not take into account the impact of external costs in the analysis of transport demand projections and in scenario analysis. In this paper, we will overcome this problem by taking into account the shadow prices.

The basic aim of the proposed decision support system (DSS) is to support policy makers in their analysis of the impact of socio-economic, environmental and transport-related variables on the future passenger and freight demand. Developed as such, the proposed DSS is expected to guide transportation policy makers in their future strategic decisions; facilitate analysis of the possible consequences of a specific policy on changing the share of transportation modes for both passenger and freight transportation; highlight in detail the conditional relationships among variables that are considered relevant in the transportation system analyzed and finally show the impact of a change in any variable on the whole system. The analysis conducted in this paper is based on Bayesian Net (BN) approach. A what-if analysis is also conducted in order to see the impact of different strategies in the analysis of different transport mode demand. The proposed model provides a dynamic what-if analysis opportunity to policy makers in their attempt to reduce uncertainties and to specify a direction to pursue in the future based on EU-25 Countries' data. Additionally, using the proposed BN model, it will become possible to include into the analysis, the quality of infrastructure as well as environmental-related variables.

The second section highlights the basic features of the proposed aggregate model and underlines its basic steps while providing guidelines to transport policy makers. The third section presents what-if analyses designed for projecting transportation demands in different situations. Finally, conclusions and suggestions are given.

2. PROPOSED DECISION SUPPORT SYSTEM

In the proposed DSS, initially the relevant variables that exist in transportation system are specified. This is realized through a literature survey and experts' judgments. Then, a belief network is prepared based on the significant relationship among the variables. The resulting map is subject to what-if analysis to help transportation planners support policy makers in their analysis of the impact of socio-economic and environmental variables and transport-related variables on future passenger and freight transportation demand.

2.1. Bayesian Networks

Bayesian networks are graphical models that encode relationships among variables of interest (Nadkarni and Shenoy, 2004). They are especially useful in modeling uncertainty in a domain and have been applied particularly to problems that require diagnosis of problems from a variety of input data. A Bayesian network model can represented at two levels, qualitative and quantitative (Nadkarni and Shenoy, 2001). At the qualitative level, there is a directed acyclic graph in which nodes represent variables and directed arcs describe the conditional independence relations embedded in the model. At the quantitative level, the dependence relations are expressed in terms of conditional probability distributions for each variable in the network.

According to Nicholson et al. (2008), there are at least 4 reasons to convert existing regression models to BNs. First, BNs provide a clear graphical structure with a natural causal interpretation that most people find intuitive to understand. Second, they provide good estimates when some predictors are missing. Third, BNs separate prior distributions from other model parameters, allowing easy adaptation to new populations and fourth they can easily incorporate additional data (Nicholson et al., 2008).

BNs are representations of joint probability distributions (Korb and Nicholson, 2011). There is a fundamental assumption that there is a useful underlying structure to the problem. If there is a directed arc from a variable X_1 to a variable X_2, then X_1 is called as the parent of X_2 and X_2 as the child of X_1. Consider a BN containing n nodes, namely; X_1 to Xn. A particular value in the joint distribution is represented by $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$. The chain rule of probability theory allows factorizing joint probabilities as it is given in the following formula. By this formula, the answer that the system will give under some certain probability states can be calculated.

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) &= P(x_1, x_2, \ldots, x_n) \\ &= P(x_1) \cdot P(x_2 \mid x_1) \cdot \ldots \cdot P(x_n \mid x_1, \ldots, x_{n-1}) = \prod_i P(x_i / x_1, \ldots, x_{i-1}) \end{aligned} \tag{1}$$

The structure of a BN implies that the value of a particular node is conditional only on the values of its parent nodes, so the formula becomes as follows:

$$P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i / Parents(x_i)) \tag{2}$$

There are several strategies to induce BNs from data (Sebastiani and Perls, 2008). The underlying premise of a Bayesian model selection strategy is to assess each dependency model by its posterior probability and select the model with maximum posterior probability. The value of the posterior probability is computed by using Bayes' theorem to update the prior probability with the marginal likelihood.

Though predominantly used in the decision making context, BNs are also used for data mining purposes, especially after BN learning algorithms for creating BN from data were available (Cinicioglu et al., 2012). There exists a growing interest for BN because of its semantic clarity and understandability by humans, the ease of acquisition and incorporation of prior knowledge, and the ease of integration with optimal decision-making models. Furthermore, by the help of BNs "what-if" analyses, about the variables in a network can be performed easily.

When building a BN, there are a number of steps that must be undertaken (Korb and Nicholson, 2011). Initially the variables of interest have to be identified. This involves both to identify the nodes to represent and the values that they can take. Secondly, the structure of the network has to be constructed. This step includes capturing qualitative relationships between variables. And after specifying the structure of the net, the next step is to quantify the relationships between connected nodes.

2.2. Specification of Variables

The first step of constructing a BN is to identify the variables of interest. For this purpose, initially the relevant variables that exist in transportation system are specified. In the estimations conducted to predict the freight and passenger demands of a country, the variables that are thought to influence the transport demand projection are found based on the literature survey as well as on experts judgments which were derived in Ulengin et al. (2007). In addition to these explanatory variables where a historical data for period 1990-2009 are available; environmental related variables are also taken into account. These are especially the air pollution costs, accident costs, congestion costs, noise costs, and climate change costs where only 2004 year value is given for EU25 countries in IMPACT. BN model permit to analyze the whole system based on EU25 countries' related data. Besides, the quality of transport infrastructure provided by World Economic Forum (WEF) is also added to the variable list.
As a result the definitions of the potential explanatory variables as well as environmental related variables used in the model as well as the source of their related data are given in Table 1. The values of variables correspond to the yearly data in the period 1990 – 2009.

Table 1 – Variable definitions used in the model

Short name	Description	Source
AIP	Air transport of passengers	Eurostat
AIG	Air transport, freight	World Bank - International Civil Aviation Organization, Civil Aviation Statistics of the World and ICAO staff estimates.
SEP	Sea transport of passengers	Eurostat
SEG	Sea transport of goods	Eurostat
MOT	Motorization rate	Eurostat
KLD	People killed in road accidents; Killed	EU Commission, DG Energy and Transport - CARE database
CAR	Car share of inland passenger transport	Eurostat
RDG	Goods transport by road; Thousands of tones	Eurostat
RAG	Rail transport of goods	Eurostat
RAP	Rail transport of passengers	Eurostat
MST	Modal split of passenger transport; Trains	Eurostat
MSR	Modal split of freight transport; Railways	Eurostat
GDP	GDP per capita (current US$) (natural logarithm of the data is used)	World Bank national accounts data, and OECD National Accounts data files.
IMP	Imports of goods and services (natural logarithm of the data is used)	International Monetary Fund, Balance of Payments Statistics Yearbook and data files.
EXP	Exports of goods and services (natural logarithm of the data is used)	International Monetary Fund, Balance of Payments Statistics Yearbook and data files.
EMP	Employment rate by gender Total	Eurostat

Short name	Description	Source
POP	Total population	Eurostat
INF	Inflation	World Bank, Indicators
POP1564	Population ages 15-64 (% of total)	World Bank, World Development Indicators
URBPOP	Urban population	World Bank, World Development Indicators
ED3	School enrollment, tertiary	UNESCO Institute for Statistics.
ED2	School enrollment, secondary (% gross)	UNESCO Institute for Statistics.
EDUEXP	Adjusted savings: education expenditure	World Bank staff estimates using data from the United Nations Statistics Division's Statistical Yearbook, and the UNESCO Institute for Statistics online database.
PET	Unleaded 95 Petrol Prices	The Automobile Association Limited, Petrol and diesel price archive 2000 to 2010
DIE	Diesel 7000PPM Prices	The Automobile Association Limited, Petrol and diesel price archive 2000 to 2010
QAI	Quality of air transport infrastructure	World Economic Forum, Global competitiveness Report, Data sheets
QOI	Quality of overall infrastructure	World Economic Forum, Global competitiveness Report, Data sheets
VFT	Volume of freight transport relative to GDP	Eurostat
VPT	Volume of passenger transport relative to GDP	Eurostat
QPO	Quality of port infrastructure	World Economic Forum, Global competitiveness Report, Data sheets
QRO	Quality of roads	World Economic Forum, Global competitiveness Report, Data sheets
MWL	Total length of motorways	Eurostat
QRA	Quality of railroad infrastructure	World Economic Forum, Global competitiveness Report, Data sheets
RLL	Total length of railway lines	Eurostat
plrdFRE	Pollution road freight	HEATCO
plrdPAS	Pollution road passenger	HEATCO

2.3. Determination of the Relations between the Variables for Freight and Passenger Transportation

As to specify the relations between variables and learn the associated BN; the data of Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Portugal, Slovenia, Spain, Sweden, Turkey and United Kingdom are used. Initially, the data is normalized and is divided into three states, namely “low” “medium” and “high”. Since the values of the variables correspond to the yearly data in period 1999-2009; there are more than 1 data set for each country which makes 151 data sets in total. The yearly data for which all the data values are available, are used in the analysis process. That is why there are no missing values. Subsequently, based on the software developed by WinMine, the relationships among the variables based on conditional probabilities are tried to be specified. 80% of 151 data sets are used for training and 20% for testing purposes. During the learning phase, kappa (a factor that is used to set the granularity of the BN) is determined as 1. As kappa approaches closer to 1 the model become denser. The Bayesian network learnt on the training set is tested using the test data where the accuracy of the learned model was evaluated using the log score (Cinicioglu, 2007). Log-score is a quantitative criterion to compare the quality and performance of the learned BNs. The formula for calculating the logscore is given as follows.

$$Source(x_1, x_2, ..., x_N) = \sum_{i=1}^{N} \log_2 p(x_i \mid model) / nN \quad (3)$$

where n is the number of variables, and N is the number of cases in the test set.

In the model developed at the first stage it is seen that several of the environmental variables do not affect or are not affected by the system. As a result, these unrelated variables are taken out of the system. That is why, in the finalized model, only the environment variables “Air Pollution Cost (road-freight)”, “Air Pollution Cost (road-passenger)” and “Climate Change Cost” which are in interaction with the other variables are included.

According to Logscore value of -0.42, using the Bayes Net, the log probability that each variable assigns to the given value in the test case is 75%, given the value of other variables. Using WinMine, it is possible to analyze the differences between the developed and marginal model. A positive difference means that the developed model provides better solution compared to the results of the marginal model, and, hence, is a required situation. The value named as “lift over marginal” is important to show the extent to which the model explains the given data. In the analyzed model, it is found as 0.45, meaning that the improvement rate we obtained with the provided model compared to the marginal model is about 20%. The results indicate that the provided BN model in this study outperforms the marginal model, signifying that the BN created effectively represents the dependency relations of fundamental factors of transportation system.

2.4. Constructing the Final Bayes Net

As the next step, the network structure determined from the Win Mine is transferred to Netica software (www.norsys.com). In this way, the probabilistic relationship between parameters that determine the freight and passenger transportation of Turkey can be easily analyzed by entering evidence fort variables and observing the change in posterior probabilities consequently. The model is valuable since it allows experts to visually investigate the causal dependencies between various states of variables and detect immediately what kind of a change or any modification in the state of a given decision variable would lead to consequent changes in the variables that would be influenced. The BN created using the Netica software and the marginal probabilities of the variables in the network can be seen in Figure 1.

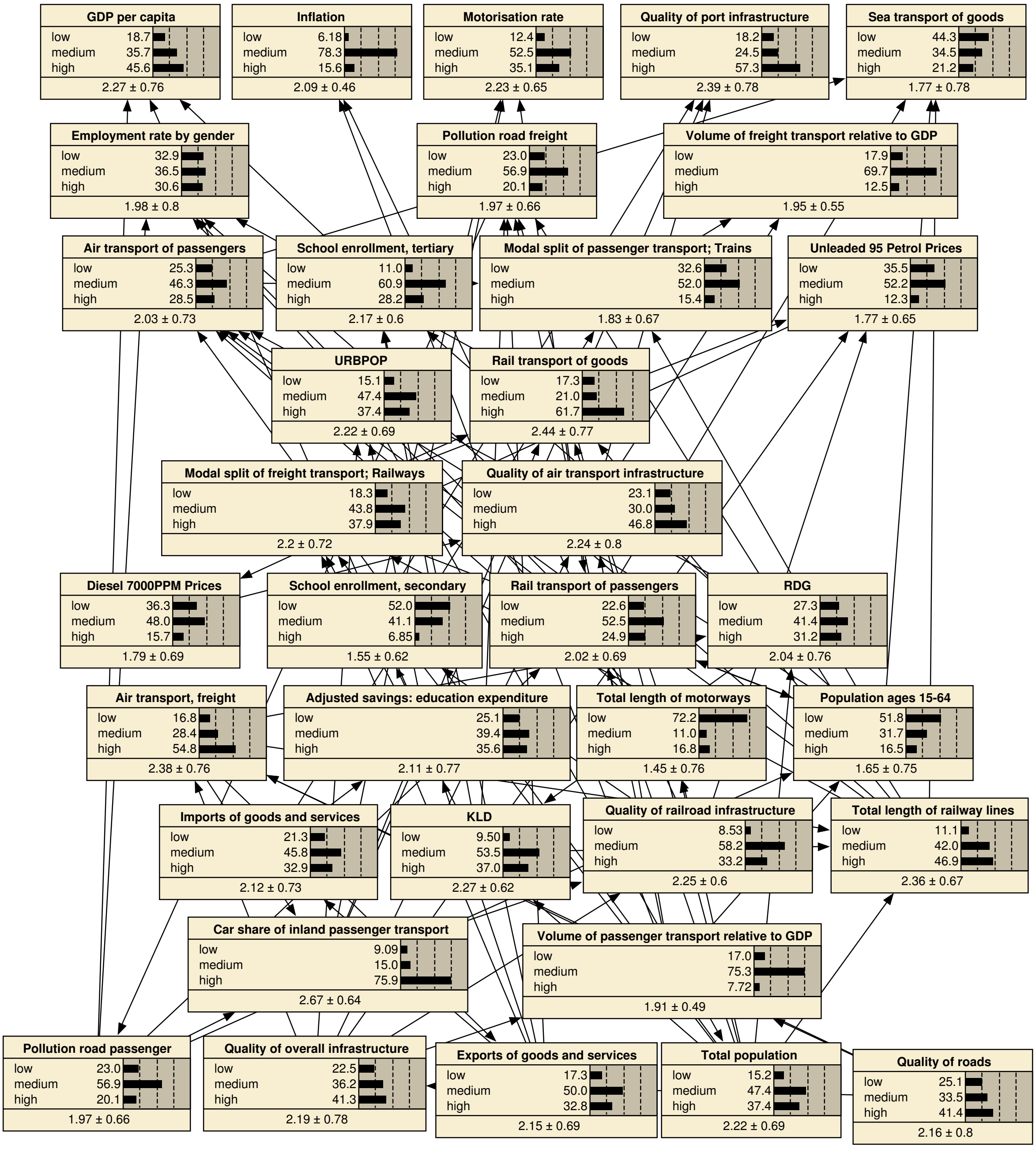
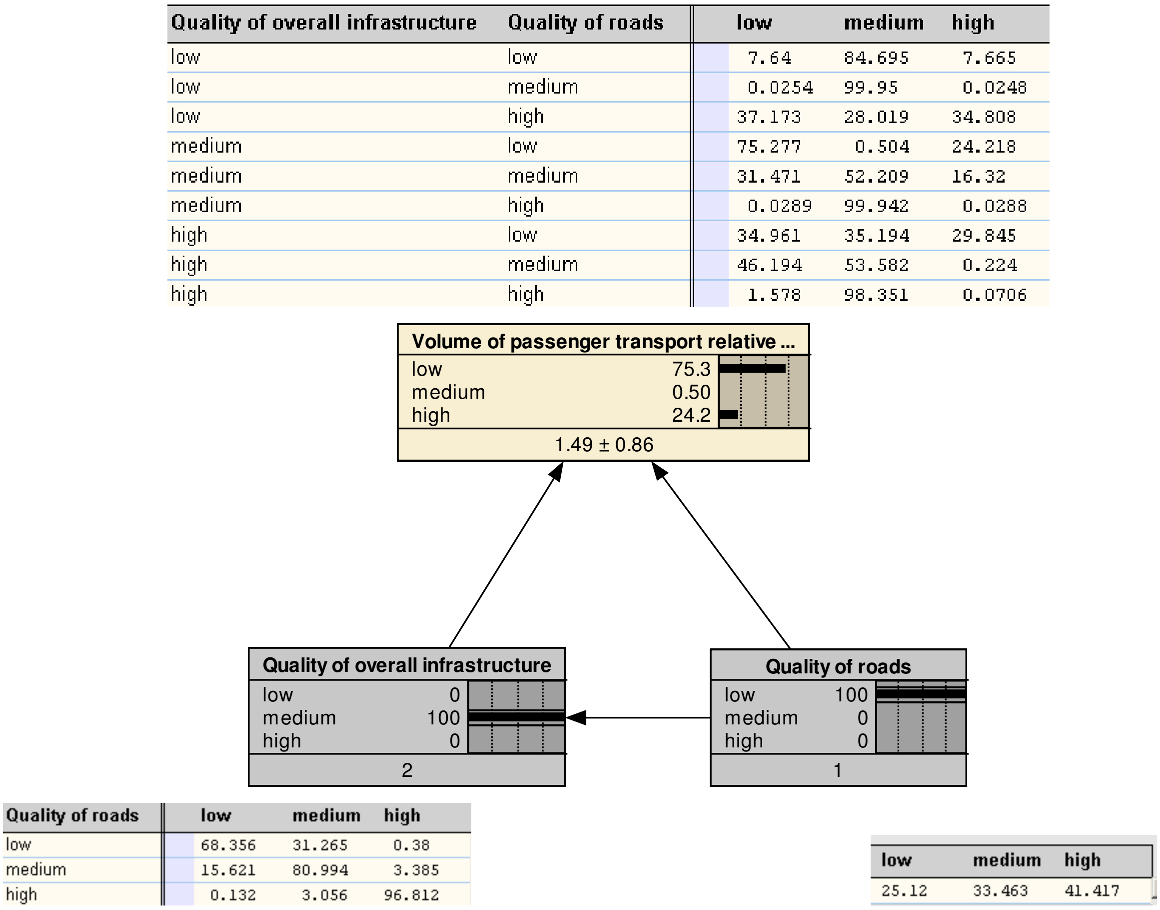


Figure 1 – Transportation BN

To give an idea of the analyses that can be conducted, Figure 2 shows that if the "Quality of overall infrastructure" is medium and "Quality of roads" is low, then "Volume of passenger transport relative to GDP" will be low with 75.3% probability, medium with 0.5% probability, and high with 24.2% probability. In practice, such an approach is computationally intractable when there is an extensive number of variables since the joint distribution will have an exponential number of states and values. Although BNs create an efficient language for building models of domains with inherent uncertainty, it may be time consuming to calculate conditional probabilities, even for a very simple BN.



Quality of overall infrastructure	Quality of roads	low	medium	high
low	low	7.64	84.695	7.665
low	medium	0.0254	99.95	0.0248
low	high	37.173	28.019	34.808
medium	low	75.277	0.504	24.218
medium	medium	31.471	52.209	16.32
medium	high	0.0289	99.942	0.0288
high	low	34.961	35.194	29.845
high	medium	46.194	53.582	0.224
high	high	1.578	98.351	0.0706

Quality of roads	low	medium	high
low	68.356	31.265	0.38
medium	15.621	80.994	3.385
high	0.132	3.056	96.812

low	medium	high
25.12	33.463	41.417

Figure 2 – Simple example of Bayes' rule

Forecasting methods, such as trend extrapolation and regression, are seen to be too dependent upon a projection of the past into the future to be useful for anticipating changes (Eden and Ackermann, 1998). Similarly they suggest a single view of the future. In contrast, scenario planning put forward a number of different alternative futures, each of which is possible. Scenarios focus less on predicting outcomes and more on understanding the forces that would eventually compel an outcome. BN permits to conduct what-if analyses to analyze these types of scenarios in a dynamic manner.

To give an idea, Figure 3 provides an example of a what-if analysis and shows the situation in whole system when the “Car share of inland passenger transport” variable is at “low” state. As can be seen from the figure, under this situation, for example “motorization rate”, “air pollution cost (road-freight)” or “air pollution cost (road-passenger)” variables also are at “low” state (with 1.24, 1.007 and 1.007 values respectively) while “air transport of passengers” or “rail transport of passengers” are at the “high” state (with 2.41 and 2.38 values respectively).

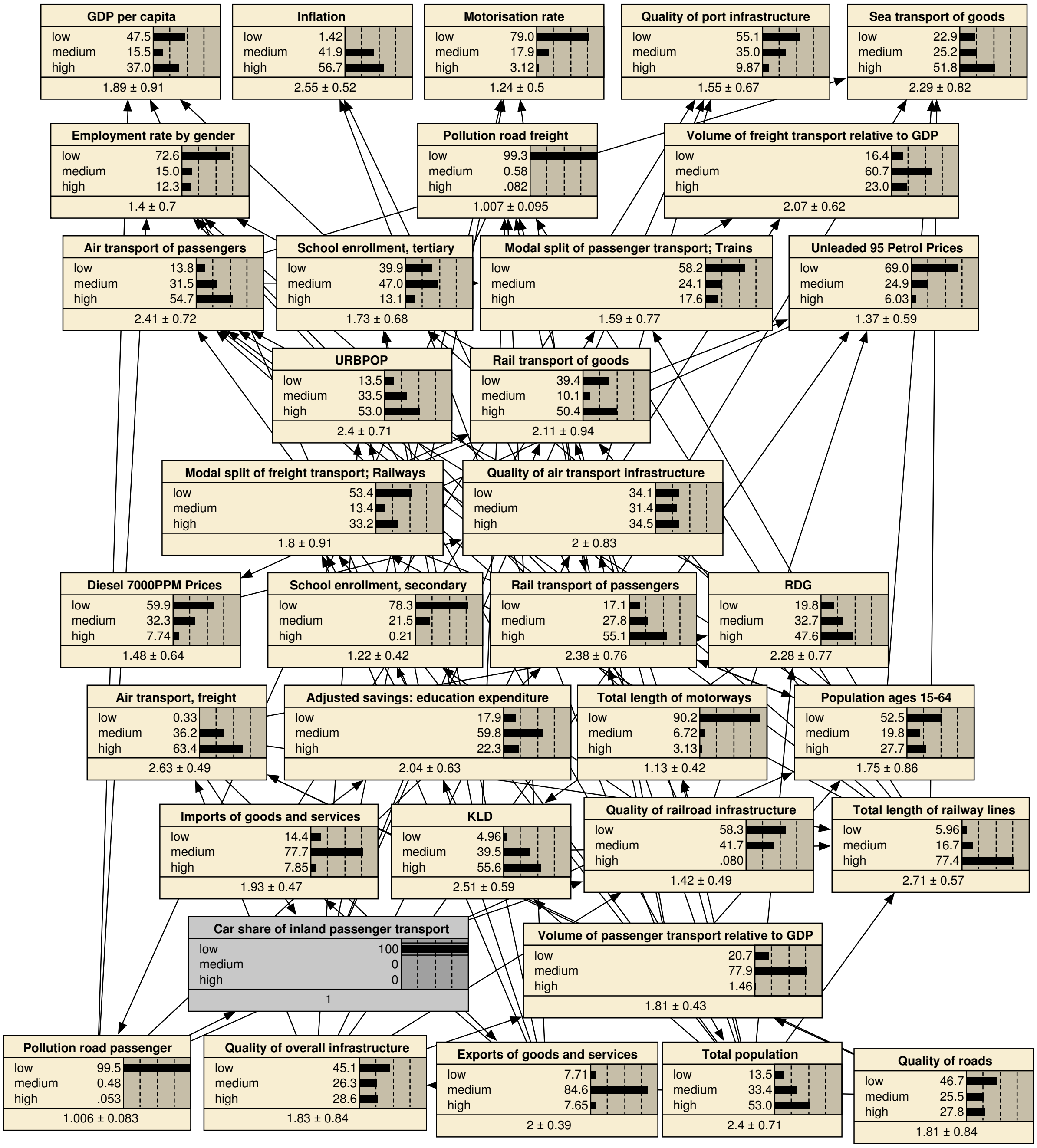


Figure 3 – Car Share Variable at "low" state

However, when "Car share of inland passenger transport" is at "medium" state, the situation of the resulting system is given in Figure 4. An increase from "low" to "medium" directly influence the state of "motorization rate", "air pollution cost (road-freight)" and "air pollution cost(road-passenger)" variables and resulted with an increase in these states(with values of 1.79, 1.55 and 1.55 respectively). Similar trends are experienced when the related variable increase from "medium" to "high" state (see Figure 5). Table 2 gives a summary about the outcome of the system when changes in the states "Car share of inland passenger transport" occur. As can be seen from Table 2, when "Car share of inland passenger transport" increases from low to medium and then medium to high, motorization rate and GDP increases consequently. Although air tranport of passengers and rail transport of passengers decrease initially, they both increase afterwards depending on the increased level of GDP. The main reason of this is that: as GDP goes up, the level of travel demand increases, and so air, rail and car use increase.

Table 2 – Summary table for the changes in the states of “Car share of inland passenger transport”

Car share of inland passenger transport		Motorisation rate	GDP	Air transport of passengers	Rail transport of passengers
	low	1.24	1.89	2.41	2.38
	medium	1.79	2.05	1.91	2.18
	high	2.43	2.36	2.01	1.95

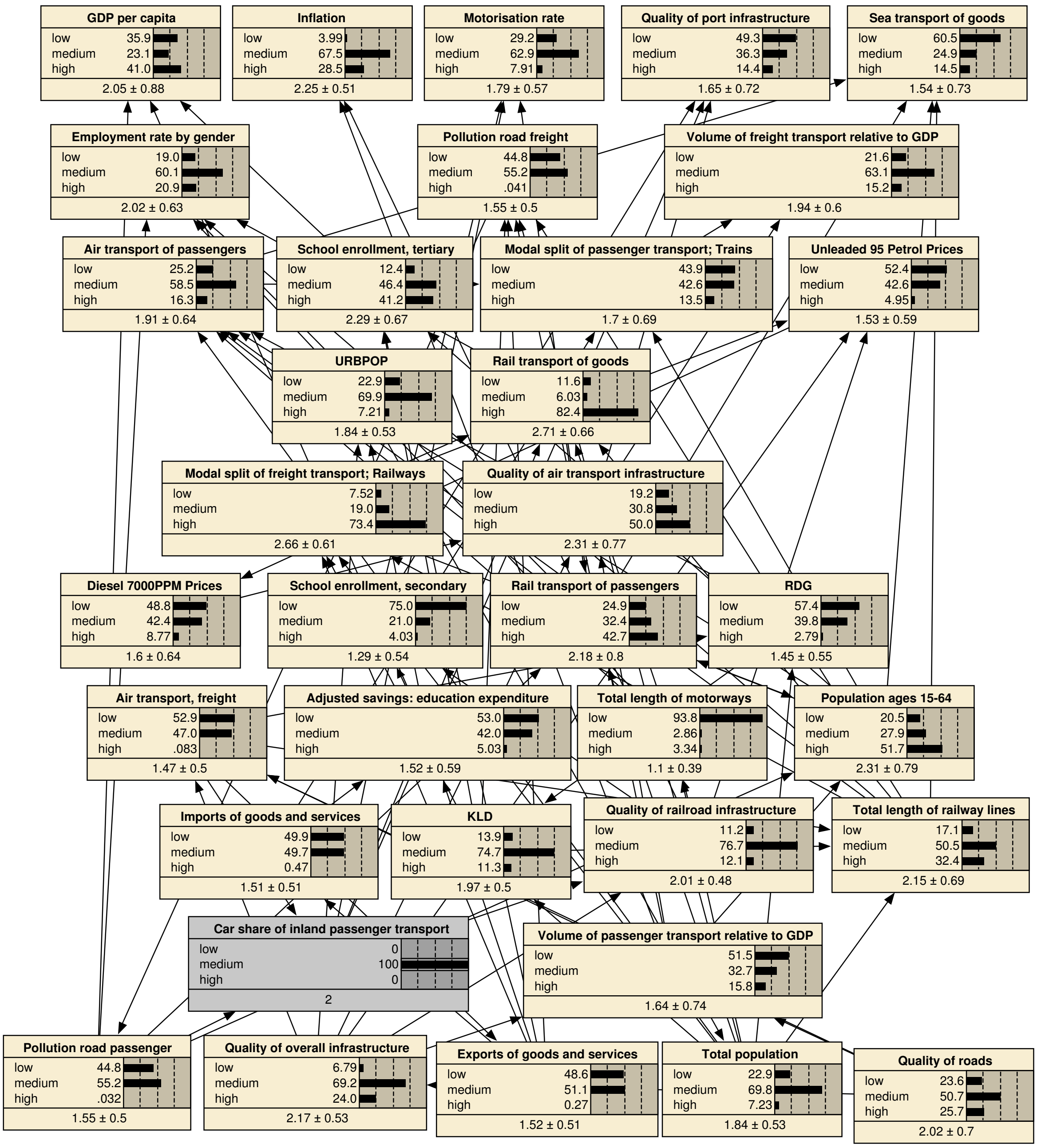


Figure 4 – Car share variable at “medium” state

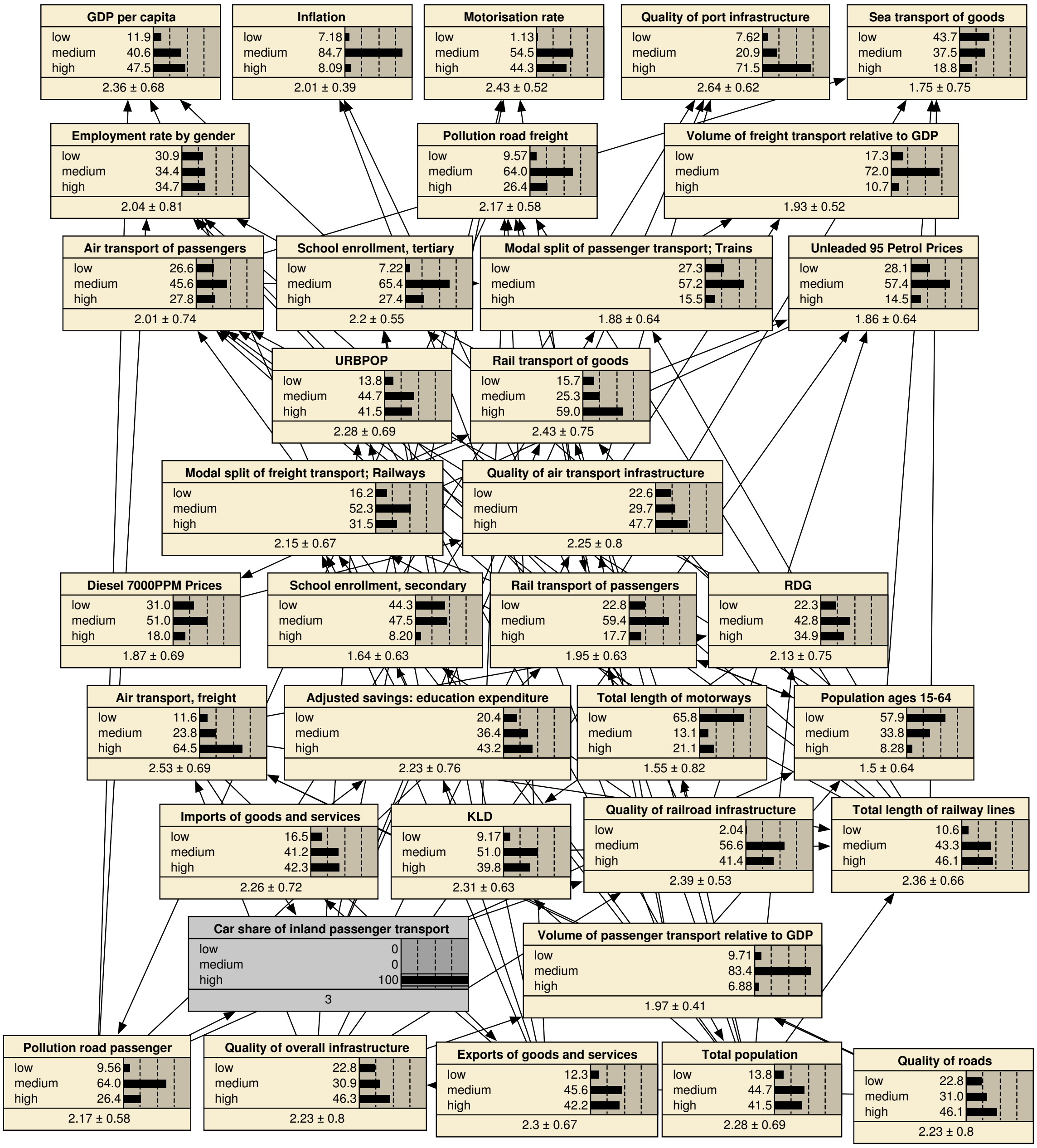


Figure 5 – Car Share Variable at "High" State

3. WHAT-IF ANALYSIS

In this section several what-if analyses are conducted in order to see the impact of different input parameters on freight and passenger transportation. The results of the conducted what-if analyses are given in Table 3. The variable names and their related definitions are given in Table 1 above.

According to results of BN-based what-if analysis it can be seen that, due the infrastructure deficits which results with low quality of railway infrastructure, the railway freight transportation (RAG) will lag far behind road freight transport (RDG). But the railway will experience the highest increase in demand if the quality of railway infrastructure is set at the "High" level. When the investment to other transportation modes are considered, it can be seen that a “Medium” level of investment made to other modes will also result with an increase in the railway freight demands due to the increase in the intermodal transportation potential. However, when the investment made to road and sea is at “high” level, this positive impact will be eliminated due to the switch of freight demand directly to these modes.

The demand for railway freight transportation shows an important increase, whenever the GDP is increased but the motorway length is increased at a relatively lower level. On the other hand, contrary to the expectations, when there is an increase in import and export levels, this results with a decrease in railway freight demand. This shows the continuing tendency of using road-based freight transportation in Europe despite the attempts of increasing the railway shares.

Table 3 – Results of BN-based What-if Analysis

			Transport of Goods			**Transport of Passengers**			
			Road (RDG) x million tkm	**Rail (RAG) x1000 tkm**	**Air (AIG) x1000 tkm**	**Rail (RAP) million pkm**	**Car share (CAR) %**	**Air (AIP) Passenger**	**Air pol. (-)**
GDP	<24159	L	26.715	6,888.2	44,27	5681,0	70,33	509428,46	0,1557
	24159-43098	M	93.165	9,404.0	657,92	5252,2	82,26	607699,05	0,2181
	>43098	H	31.122	9,404.0	104	4668,8	81,28	433952,28	0,2076
IMP	<466*10^8	L	14.506	9197,18	8,59	5050,1	77,07	413570,55	0,1625
	466*10^8-2700*10^8	M	36.255	8994,91	146,34	4761,4	77,49	363778,16	0,2076
	>2700*10^8	H	122.973	8229,38	1514,2	5462,4	83,52	936957,48	0,2154
EXP	<400*10^8	L	13.163	9197,18	3,78	5356,3	76,51	433952,28	0,1491
	400*10^8-2600*10^8	M	35.262	8994,91	162,14	4668,8	77,63	357991,32	0,2115
	>2600*10^8	H	124.691	8229,38	1529,8	5462,4	83,52	936957,48	0,2154
QOI	<3.84	L	88.134	9197,18	614,47	5793,6	79,59	658429,34	0,1989
	3.84-5.4	M	28.240	8994,91	34,85	4855,7	78,05	440967,01	0,1848
	>5.4	H	45.270	8603,63	343,78	4951,9	80,29	501324,67	0,2115
QRA	<3.08	L	47.195	6736,73	192,34	7625,3	65	772948,23	0,1352
	3.08-5.04	M	40.513	8603,63	100,5	4951,9	78,89	493349,8	0,1929
	>5.04	H	47.195	9615,45	332,24	4761,4	83,24	477778,62	0,2276
QRO	<3.57	L	53.474	8414,42	179,64	5570,6	78,19	579156,84	0,1882
	3.57-5.44	M	36.762	9403,99	69,02	4855,7	78,89	477778,62	0,1917
	>5.44	H	44.031	8603,63	299,87	4951,9	80,15	493349,8	0,2102
QAI	<3.05	L	49.889	9403,99	199,02	5252,2	78,75	534534,29	0,1953
	3.05-4.89	M	44.031	8229,38	123,37	5050,1	79,17	501324,67	0,1941
	>4.89	H	38.860	8994,91	156,69	4951,9	79,59	485501,79	0,2026
QPO	<4.12	L	29.441	6888,22	38,62	5570,6	72,44	517663,24	0,1566
	4.12-5.46	M	26.715	9615,45	52,52	4761,4	76,93	427049,14	0,1826
	>5.46	H	58.118	9197,18	368,09	4951,9	82,26	534534,29	0,2207

When the same types of analysis are developed for passenger transport, it can be seen that when the GDP level is increased, this result with a decrease with railway passenger demand. In fact, this result was also estimated in Muller (2007) and was explained by the fact that when GDP increases, the time value of the passengers also increases. Therefore, passengers do not prefer railways if high speed trains are not used. On the other hand, when the motorway length is increased from low to medium level, the railway passenger demand is decreased dramatically and an increase even further may have a positive impact on railway transportation due to the possibility of intermodal connections possibility.

Table 4 – Effects of Change in Level of Total Length of Railways (RLL) whenever the "overall infrastructure quality" is at "high" level

	Rail transport of goods(RAG)	**Rail transport of passengers (RAP)**	**Air pollution**
RLL "low"	5045.47	2915.31	0.2361
RLL "medium"	6443.68	4578.12	0.2480
RLL "high"	8603.63	4951.93	0.1941

Table 5 – Effects of Change in Level of Total Motorway Length (MWL) whenever the "overall infrastructure quality" is at "high" level

	RDG	**Car share of inland passenger transport**	**Air pollution**
MWL "low"	33358.11	80.15	0.2194
MWL "medium"	62294.42	82.68	0.2262
MWL "high"	91880.96	83.24	0.2262

Generally, it is expected that the increase in the quality of infrastructure of a transportation mode result with an increase in the passenger transportation of the related mode. However, the improvement in the quality is a subsection of the infrastructure. What is really important is to keep the high level of quality whenever the total route length is increased. Based on this perspective, the impact of increasing the length of the total motorway length and railway length on the passenger and freight demand while the "overall infrastructure quality of the country" variable is kept at the high level is evaluated. For example, as can be seen from Table 4, whenever the "overall infrastructure quality" is at "high" level, an increase in the total length of railways from "low" to "medium" and then to "high" will result with an important increase in rail transport of goods (RAG) and rail transport of passengers (RAP) and a reduction in air pollution.

Similarly, as can be seen from Table 5 whenever the "overall infrastructure quality" is at the "high" level, the increase in the total motorway length from "low" to medium and then to "high" results with an important increase in road transport of goods (RDG) and car share (CARS) . This does not result with a significant increase in air pollution. However as the lowest level of the air pollution is already very high (0.2194), even very small affect can deteriorate the whole system dramatically.

However, in the current configuration, the "infrastructure quality" variable does not provide sufficient information alone. That is why; a more detailed analysis can be realized by adding to the model "transportation investment as a percentage of GDP" variable for each transport mode. By this way, it will be possible to analyze the impact of those investments on the overall system.

On the other hand, it can be seen that an increase in demand in a transportation mode results with an increase in air pollution, the only exception being passenger air transportation. In fact, when the passenger air transportation is increased from low to medium and high level, the air pollution is decreased. The most dramatic increase in air pollution occurs with an increase in car share of inland passenger transport. In fact, as can be seen from Figure 6, in passenger transportation, when the "Car share of inland passenger transport" is at "high" and "Air transport of passengers" and "Rail transport of passengers" at "low" level, the air pollution is the highest.

Contrarily, however, as can be seen from Figure 7, in passenger transportation when the "Car share of inland passenger transport" is reduced to the "low" level and "Air transport of passengers" and "Rail transport of passengers" is increased at the "high" level, the air pollution is reduced to "low" level with approximately 100% probability.

Similar to those given above, the policy makers can develop other type of what-if analyses based on BN in order to see the implication of a policy variable on the transport mode demands.

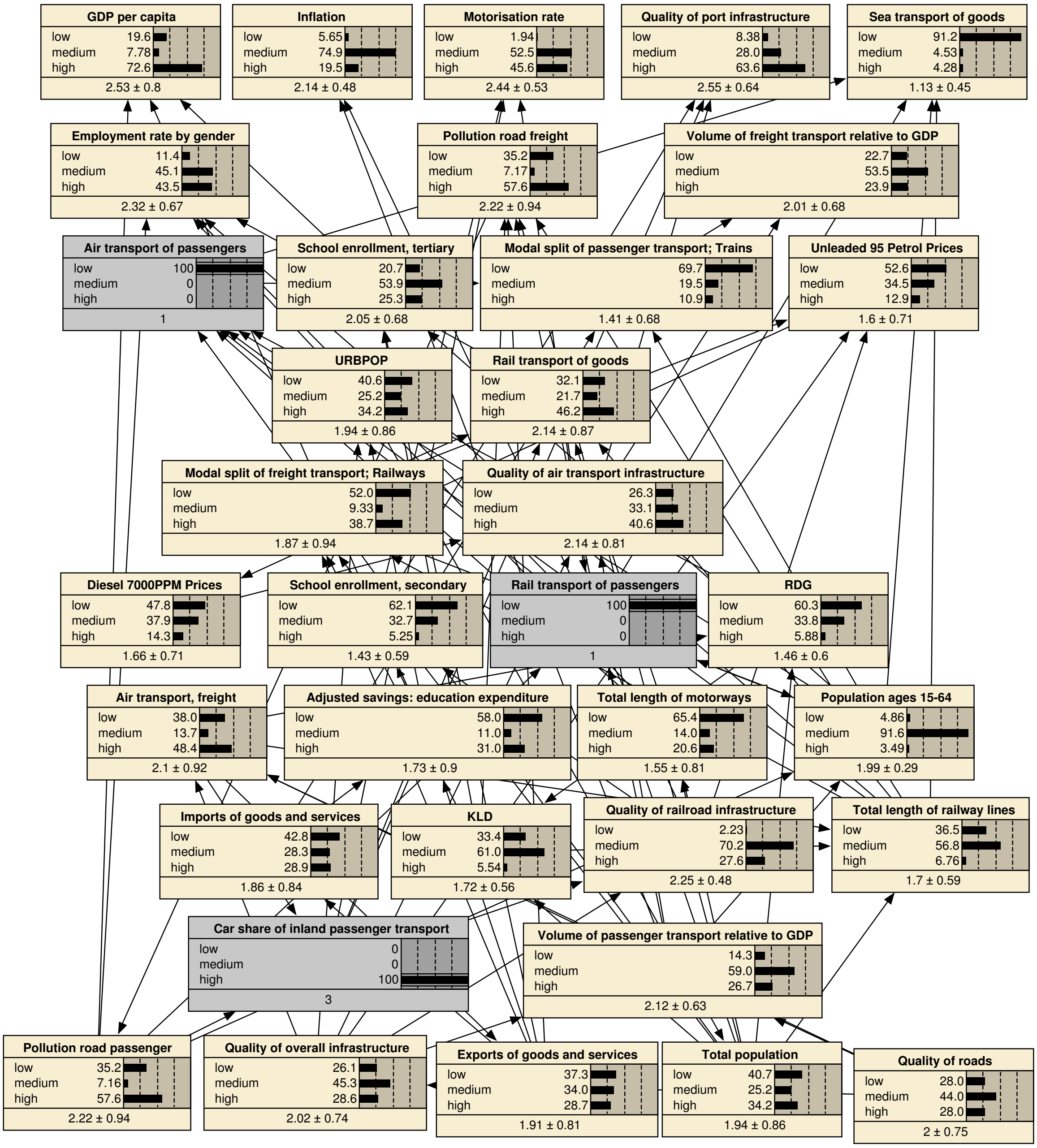


Figure 6 – The system configuration with “car” at “high” level and “air” and “rail” at “low” level in passenger transportation

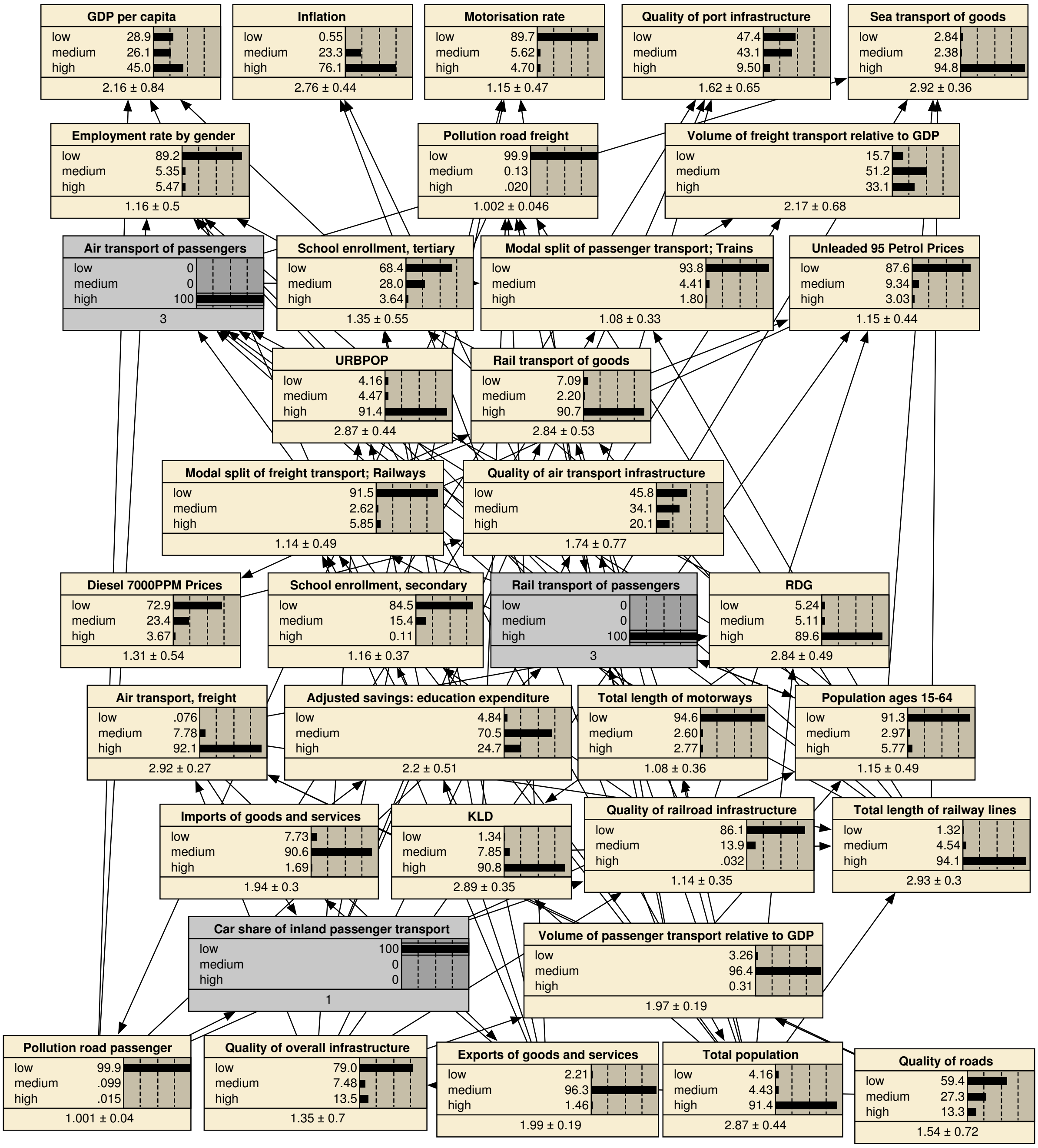


Figure 7 – The system configuration when passenger transportation with “car” is at “low” level and “air” and “rail” at “high” level

4. CONCLUSIONS AND FURTHER SUGGESTIONS

In this study, the basic aim is to provide an input for the transportation policy makers in their way to realize those strategic aims.

For this purpose, a BN model is developed that recognizes the need to look at the transportation problem as a whole, not in its separate components. In this way, each and each affecting and affected variables are considered in relation to its effect on other modes and other variables.

Subsequently, different what-if analyzes are performed using the BN. These analyses are just examples to check the validity of the model and observe the implications of certain strategies. Using this model and Netica software, policy makers can analyze any change of policy variables in order to see the impact of different policies. The what-if analysis based on BN showed that especially passenger transportation by car inevitably causes serious traffic congestion on the roads and cause environmental pollution. Another striking result of analysis is that, whenever there is an increase in GDP, there is a decline in the railway passenger demand. This finding is also in parallel with that of Muller (2007)'s traffic forecast which underlines that the share of rail transport will decrease to 2.2% in 2020 from 2.3% (2004 statistics) domestically, if high-speed train investment is not made. The basic reason of this is that, whenever the GDP level is increased, the time value of the people is also increased and people want to reach their destination much more quickly.
As can be seen in the previous section results, the BN permits to conduct what-if analysis not only by changing the value of just one variable but based on the changes of more than one variable at a time. This helps to see the overall changes in the system and facilitate to develop policies.
This proposed model can be used as a reference for the transport policy makers of EU as well as EU candidate countries in their attempt to improve the balance in the transport modes, taking also into account the environmental concerns.

Acknowledgements

The authors would like to express their gratitude to Andreas Kopp and Mediha Ağar (World Bank) as well as Turkish State Railways, General Directorate of Highways, Turkish Airways and Turkish Maritime Lines authorities for their invaluable contribution during data acquisition as well as for their invaluable comments during the preparation of the report. Our special thanks are for İbrahim Çelik, Ömer Çevik, Yaşar Rota (Turkish State Railways) and Kemal Yıllıkçı (Turkish Airways) for their great support during the data collection phase.

REFERENCES

APAS. (1999) *Strategic transport, databases and scenarios for European Transport.* Office for Publications of the European Commission, Luxembourg.
Asensio, J. (2000) The success story of Spanish suburban railways: determinants of demand and policy implications, *Transport Policy,* 7, 295-302.
Ben Akiva, M., and Lerman, S. (1985) *Discrete choice analysis theory and application to travel demand.* MIT Press, Cambridge, MA.
Cantarella, G., and Luca, S. (2005) Multilayer feed forward networks for transportation mode choice analysis: An analysis and comparion with random utility models. *Transportation Research Part C: Emerging Technologies,* 13(2), 121-155.
Cascetta, E. (2001) *Transportation Systems Engineering: Theory and Methods.* Springer.
Cinicioglu, E., Onsel, S., and Ulengin, F. (2012) Competitiveness Analysis of Automotive Industry in Turkey Using Bayesian Networks. *Expert Systems with Applications* (in press)
Cinicioglu, E., Shenoy, P., and Kocabasoglu, C. (2007) Use of radio frequency identification for targeted advertising: a collaborative filtering approach using Bayesian Networks,

Symbolic and Quantitative Approaches to Reasoning with Uncertainty ed K. Mellouli, pp. 889-900, Springer Berlin / Heidelberg.

DETR, (1998) *A New Deal for Transport: Better for Everyone, Cmnd 3950.* UK Government, Department of the Environment, Transport and the Regions. London: The Stationary Office.

Eden, C., and Ackermann, F. (1998) *Making Strategy: The Journey of Strategic Management.* Sage, London.

Egger (2000) A note on the proper econometric specification of the gravity equation, *Economics Letters*, 66(1), 25-31.

European Commission (2001) *European Transport Policy for 2010: Time to Decide,* Office for Official Publications of the European Communities, Luxembourg.

European Commission (2011) *Roadmap to a Single European Transport Area: Towards a Competitive and Resource Efficient Transport System.* Office for Official Publications of the European Communities Brussels.

Filippini, C., and Molini , V. (2003) The determinants of East Asian trade flows: a gravity equation approach, *Journal of Asian Economics*, 14(5) 695-711.

Korb, K., and Nicholson, A. (2011) *Bayesian Artificial Intelligence.* CRC Press, Taylor and Francis Group.

Nadkarni, S., and Shenoy, P., (2001) A Bayesian network approach to making inferences in causal maps, *European Journal of Operational Research*, 128, 479–498.

Nadkarni, S., and Shenoy, P., (2004) A causal mapping approach to constructing Bayesian networks, *Decision Support Systems*, 38, 259–281.

Nicholson, A., Twardy, C., Korb, K., and Hope, L. (2008) Decision support for clinical cardiovascular risk assessment, *Bayesian Networks: A practical guide to applications bayesian networks* eds O. Pourret, P. Naim, and B. Marcot, pp. 33-52, John Wiley and Sons, Cornwall.

Piecyk, M.I., and McKinnon, A.C.(2010) Forecasting carbon footprint of road freight transport in 2020, *International Journal of Production Economics*, 128, 31-42.

Republic of Turkey, Ministry of Transport. (2007) *Transport Operational Programme 2007-2009.* Ministry of Transport, Ankara.

Schafer, A., and Victor, D. (2000) The future mobility of the World population, *Transportation Research Part A, 34*, 171-205.

Sebastiani, P. and Perls, T.T., (2008) Complex genetic models, *Bayesian networks: A practical guide to applications bayesian networks* eds O. Pourret, P. Naim, and B. Marcot, pp. 33-52. John Wiley and Sons, Cornwall.

Mueller, R. (2007) *Technical assistance to transportation infrastructure needs assessment for Turkey*, Final Report (TINA, Turkey) Joint Venture, 2007

Tsai, T.-H., Less, C., and Wei, C. (2009) Neural Network based temporal feature models for short-term railwat passenger demand forecasting. *Expert Systems with Applications,* 36, 3728-3736.

Ulengin, F., Onsel, S., Topcu, Y.I., Aktas, E., and Kabak, O. (2007) An integrated transportation decision support system for transportation policy decisions: The case of Turkey, *Transportation Research Part A*, 41, 80-97.

Vythoulkas, P., and Koutsopoulos, H. (2003) Modeling discrete choice behavior using concepts from fuzzy set theory, approximate reasoning and neural networks. *Transportation Research Part C,* 11, 51-73.

Yuehua, Z., Yanyan, C., Xue, G., and Lina, L. (2008) Transport modal split of commercial sites based on artificial neural network. *Journal of Transportation Systems Engineering and Information Technology*, 8(1), 86-91.