Ann Reg Sci (1997) 31:411–429

The Annals of Regional Science

Spatial choice behaviour: logit models and neural network analysis

Peter Nijkamp[1], **Aura Reggiani**[2], **Tommaso Tritapepe**[2]

[1] Department of Economics, Free University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
[2] Department of Economics, University of Bologna, Piazza Scaravilli 2, I-40126 Bologna, Italy

Received: June 1996 / Accepted: February 1997

Abstract. Neural networks are becoming popular analysis tools in spatial research, as is witnessed by various applications in recent years. The performance of neural network analysis needs to be carefully judged, however, since the theoretical underpinning of neuro-computing is still weakly enveloped. In the present paper we will use the logit model as a benchmark for evaluating the result of neural network models, based on an empirical case study from Italy. The present paper aims to assess the foreseeable impact of the high-speed train in Italy, by investigating competition effects between rail and road transport modes. Two statistical models will then be compared, viz. the traditional logit model and a new technique for information processing, viz. the feedforward neural network model. In the study two different cases – corresponding to a different set of attributes – are investigated, namely by using only 'time' attributes and by using both 'time' and 'cost' attributes. From an economic viewpoint, both models appear to highlight the advantage of introducing the high-speed train system in that they show high probabilities of choosing the improved rail transport mode. The feedforward neural net model seems to provide reasonable predictions compared to those obtained by means of a logit model. An important lesson however, is that it is important to define properly the neural network architecture and to train sufficiently the network during the learning phase.

1. Introduction

The analysis of spatial behaviour in changing networks (e.g., as a result of the introduction of a new technology or new mode) is fraught with many difficulties of an analytical and empirical nature. The purpose of this paper is to assess the intermodal substitution effect as a result of the high-speed train (HST) in Italy. This requires an investigation of the competition effects of two main transport modes, viz. rail and road. Competition effects between distinct transport modes are usually analysed by means of discrete choice models, e.g., a logit model. Their general purpose is to assess the choice behaviour based on a utility function related to each transport mode in relation to relevant explanatory attributes (e.g. time, cost, distance) associated with that mode. Clearly, there are also alternative approaches, such as spatial interaction models and neural network (NN) models. Since our purpose is to estimate the new rail – and consequently road – probabilities for the entire network in Italy after the introduction of the HST, two methodologically different models will be tested against one another, viz. the conventional behaviourally-based logit model and the neurocomputing-based feedforward NN model.

NN analysis has a fairly recent history, although its roots can be traced back to McCullough and Pitts (1943) who modelled simple neurons as binary threshold units. After several decades of desperate research attempts, the real breakthrough took place approximately a decade ago, when Rumelhart et al. (1986) showed the scientific potential of back-propagation NNs. More recently, recurrent back-propagation methods have been introduced in NN analysis (see Taber 1995). NN research has been applied in a wide variety of fields, such as telecommunication, pattern recognition, network forecasting etc. Various applications of NN models in the transport sector do in the meantime exist, particularly in relation to engineering issues like incident detection, pavement performance or traffic control (see e.g. among others, Dougherty 1995; Himanen et al. 1997; Reggiani et al. 1997). Regrettably, however, thus far only a few applications related to transport economic behaviour or spatial behaviour do exist (see e.g. Fischer and Gopal 1994; Reggiani and Tritapepe 1997). There is certainly a need to test results from NN models on their behavioural contents. This is the principal goal of the present article. The comparative application of our two competing models depends principally on two background factors: 1) the necessity to provide a broader range of predicted values (rather than a single point prediction); 2) the need to investigate the potential of the new technique for information processing, i.e. the feedforward NN, compared to the benchmark of a traditional logit model.

The aim of the present paper is not to describe theoretical aspects of the models used, as we may refer here to previous works and publications cited therein (Nijkamp et al. 1996a, 1997). For the sake of completeness however, in Sect. 2 a brief introduction to neurocomputing will be offered. Section 3 will describe the basic characteristics and earlier results of our empirical application (in Subsects. 3.1 and 3.2); next, we will present in Subsects. 3.3 and 3.4 the predictions of our models as a result of the introduction of the HST in Italy. These predictions will then mutually be compared in Sect. 3.5. Finally, Sect. 4 contains some concluding remarks.

2. A brief introduction to neurocomputing

In recent years – and mainly in the last ten years – there have been various efforts to emulate the human learning process by means of artificial machine learning (see Kodratoff and Michalski 1990). Computers have become so much a part of our lives that it is often forgotten that they are simply executors of procedures supplied to them. Of course, their computing power is very high, but they do not self-improve with experience. This is the main difference with respect to the human learning process, which, while performing any kind of activity, is able to improve the way it performs its tasks. Computers work on the basis of *algorithms* which are a series of instructions supplied to them in order to achieve a desired aim. Such algorithms are representations of human knowledge. In other words, humans are able to develop their behaviour by *learning*, while computer power is limited only to those tasks for which a human is capable to elaborate an algorithm (see Aleksander and Morton 1990). It should be added however, that in recent years it is increasingly recognized that machine learning is to some extent possible in relation to emergent computation and artificial life, where computers learn from their environment and modify their behaviour accordingly. Such new research directions hold a great promise to imitating the human brains. According to the above described primary difference with learning processes it is possible to define *neurocomputing* as the first alternative to programmed computing (see Hecht-Nielsen 1990). Neurocomputing technology is one of various technologies (such as genetic algorithms, fuzzy logic, fractal systems, cellular automata etc.) which are usually denoted by the common name *biocomputing*. “Biocomputing refers to biologically inspired approaches to creating software” (see Valdes 1991). The idea behind biocomputing is the attempt to explain complex phenomena by means of a few number of simple rules, according to the principle that intricate structures like living systems are made out of simpler components (cells).

Then, starting both from the need of emulating the human learning process – which is based on experience – and from the concept of biocomputing, *neural networks* represent a new technology for information processing based on current theories concerning the way the human brain works. In a human brain, nerve cells, called neurons, are the fundamental elements of the central nervous system. The central nervous system is made up of about 5 billion neurons; their simple cooperation generates a complex behaviour. The basic features of a neuron may be summarized as follows (see Davalo and Naim 1991):

- it receives signals coming from other neurons;
- it integrates these signals;
- it propagates the resulting signal to other neurons (with different intensities) by means of electrochemical connections.

Thus, analogously, the structure of NNs is generally represented by logical units (“neurons”) connected by channels of communication which intercompute independently, since each unit cooperates in the transmission of information by means of a different "weight"[1] . By changing the values of the weights such as to get the desired output, the learning process takes place; NNs are trained to output the desired results. Especially the back-propagation algorithm is able to assign back the mean-squared error signal from the output units to the input units. In this sense NNs can learn from experience; this is the key advantage of NNs over conventional algorithms.

The application areas of NNs are broad and widespread, although the main task is pattern recognition. In particular, in recent years, they have been adopted for image processing, speech synthesis, noise filtering, robotic control, financial modelling, etc.

The term 'neural networks' is used to describe a number of different models which are usually distinguished into two classes: NNs *without* and *with* supervisor. This difference is based on the difference in learning processes. In fact, the networks with *unsupervised training* do not need the target outputs and they modify the weights by means of competitive learning algorithms, in response to the input data. On the other hand, the *supervised training* implies the knowledge of input/output data in order to find, during the learning phase, the weights[2] of the network which minimize the error function of the target outputs and the network outputs. Although different training algorithms[3] exist, the most utilized one is the back-propagation algorithm.

In the application described in the next section, we will use the above mentioned NN model, viz. the feedforward NN model which is characterized by a back-propagation algorithm as learning procedure.

3. An application of neural networks in transport

3.1 Introduction

As mentioned above, NNs are well suitable for solving *pattern recognition* problems. They have been applied in various areas for solving this specific task and, in general, they have shown a very good performance.

Of course, much attention has also been paid to this new technology in the transportation area and, again, the majority of the applications has been developed for solving patterns recognition problems, in particular for traffic control (as already mentioned in Sect. 1).

However, NNs have at least two features which distinguish themselves from other methodologies. For example, they are suitable for exploratory

[1] Here a weight is a real number assigned to a connection between two units

[2] The weights are the values (adaptive coefficients) assigned to the links between the units of the network. The aim of the training phase is to find the values of the weights which will produce a reasonable output in response to input

[3] The aim of a training algorithm is to minimize the error function by adjusting the value of the weights

data analysis (see Nijkamp et al. 1996a) due to the following characteristics:

- NNs do not a priori require algorithms or rules' development; this feature may be very useful in cases of a large quantity of data (e.g., in a GIS context; see Fischer 1994) in which the knowledge of the exact statistical model for explaining the phenomenon examined is lacking.
- NNs can learn and then forecast even on the basis of incomplete, noisy and fuzzy information.

Furthermore, NNs may also be very useful in a forecasting context. In fact, in microeconomic applications, classical models, which are based on micro-variables, are often used by adopting aggregated variables, under the assumption that classes of individuals behave in the same way. Therefore, the statistical models often reflect the limits associated with this assumption. NNs seem to be able to overcome this limit by capturing the stochastic elements neglected in the previous assumption. In addition, NN models can capture nonlinearities that a traditional discrete choice network models cannot. Also the transparibility of NN results is an interesting potential, as results may be used to forecast behaviour in other regions or sectors. This is a major advantage compared to discrete choice models, where parameters must be re-estimated for each relevant region.

This latter peculiarity will be investigated and tested in the following application, where the standard logit model will be compared with the feed-forward NN model. It should be noted that also the logit model has several strong points. In particular, it gives information on the effect of each independent variable on the dependent variable, thus avoiding the 'black box' impression of NN models.

The experiments will be carried out by using aggregated data referring to a modal split problem for road-rail competition after the introduction of the foreseen HST in Italy.

3.2 Results from previous experiments

As mentioned in Sect. 1, the aim of the empirical application described in this study is to investigate the impact of the HST in Italy. The entire Italian territory (except the Sardinia Island) has been subdivided into 67 areas corresponding to both single provinces and an aggregation of two or three provinces; such a subdivision is shown in Fig. 1.

The introduction of the HST is studied by evaluating how the rail mode and the road mode flows vary, as a result of the implementation of the HST system in Italy. This is a typical *modal split problem* for an entire network. To test the sensitivity of the results, two methodologically different models are used, the classical logit model and the new NN model. An illustrative scheme of such a problem is depicted in Fig. 2.

As mentioned in Sect. 1, the present analysis refers to the experiments carried out in previous work (see Nijkamp et al. 1996a, 1997). In these

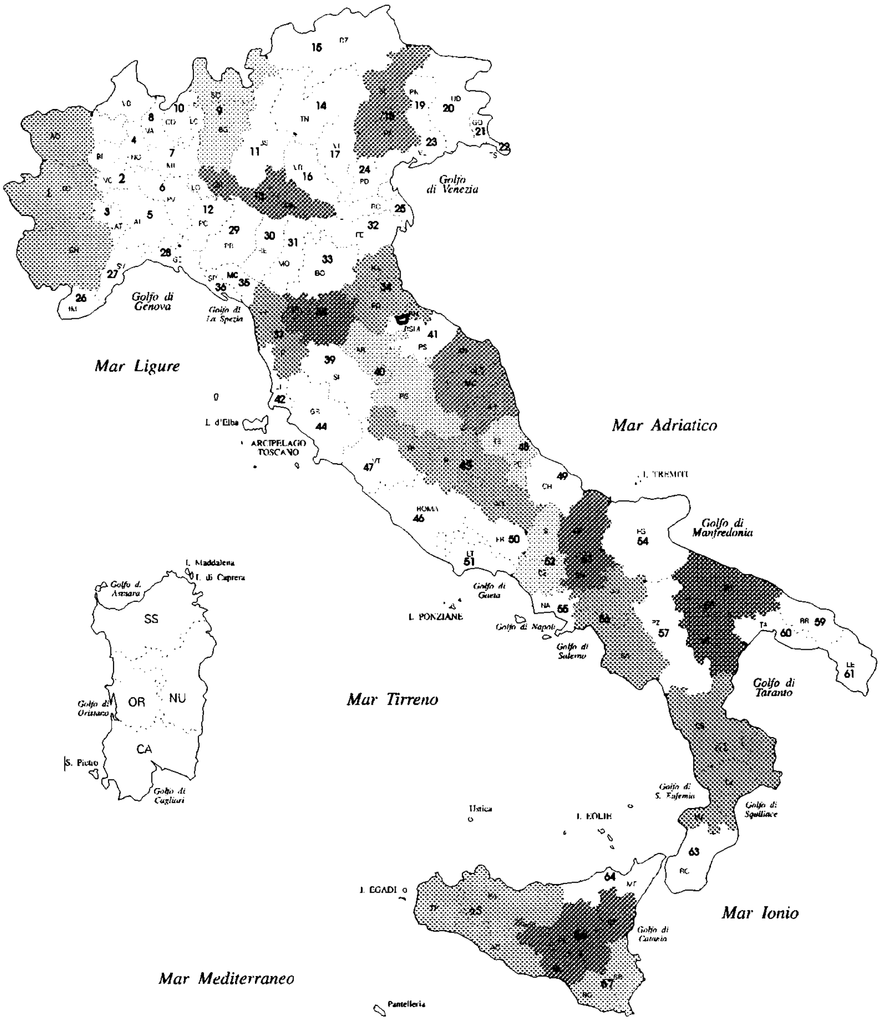

Fig. 1. Subdivision of the Italian land

studies the best configuration of the models has been investigated; they are consequently utilized in the present study. The approach adopted for this purpose consists of calibrating (training) the models by using the same data set[4] (containing 698 observations[5]) and in testing them by means of a test-set which contains 349 observations referring to other links (never used before; thus we may speak of *spatial forecasting*). The results are evaluated by means of the Average Relative Variance (*ARV*) statistical indicator (see

[4] The whole data set has been kindly provided by the Italian State Railways ('Ferrovie dello Stato') and it refers to Census data (1987)

[5] Each observation contains variables such as the attributes (e.g. 'distance', 'time' and 'cost') and the flows with reference to each transport mode.

The provinces					
2	Vercelli	21	Gorizia	41	Pesaro
3	Asti	22	Trieste	42	Livorno
4	Novara	23	Venezia	44	Grosseto
5	Alessandria	24	Padova	46	Roma
6	Pavia	25	Rovigo	47	Viterbo
7	Milano	26	Imperia	49	Chieti
8	Varese	27	Savona	50	Frosinone
10	Como	28	Genova	51	Latina
11	Brescia	29	Parma	54	Foggia
12	Piacenza	30	Reggio Emilia	55	Napoli
14	Trento	31	Modena	57	Potenza
15	Bolzano	32	Ferrara	59	Brindisi
16	Verona	33	Bologna	60	Taranto
17	Vicenza	35	Massa	61	Lecce
19	Pordenone	36	La Spezia	63	Reggio Cal
20	Udine	39	Siena	64	Messina

Aggregation of two provinces					
9	Bergamo	38	Firenze	56	Salerno
	Sondrio		*Pistoia*		*Avellino*
13	Cremona	40	Arezzo	58	Bari
	Mantova		*Perugia*		*Matera*
18	Treviso	48	Pescara	62	Cosenza
	Belluno		*Teramo*		*Catanzaro*
34	Forli	52	Caserta	67	Siracusa
	Ravenna		*Isemia*		*Ragusa*
37	Pisa	53	Benevento		
	Lucca		*Campobasso*		

Aggregation of three provinces					
1	Torino	45	Terni	66	Catania
	Aosta		*Rieti*		*Enna*
	Cuneo		*L'Aquila*		*Caltanissetta*
43	Ancona	65	Palermo		
	Macerata		*Trapani*		
	Ascoli Piceno		*Agrigento*		

e.g. Nijkamp et al. 1996a, 1997; Fischer and Gopal, 1994). This NN analysis is conducted in two phases:

1. by investigating three different possible NN architectures (see Nijkamp et al. 1996a);
2. by varying the attributes in both models (see Nijkamp et al. 1997).

Finally, referring to the questions raised in Sect. 1, the best configuration results in Case A illustrated in Fig. 3 (according to the *ARV* values for both models); it corresponds to the case with only two attributes, i.e. the rail and the road 'time'. In Fig. 3 Case B is also illustrated, corresponding to the

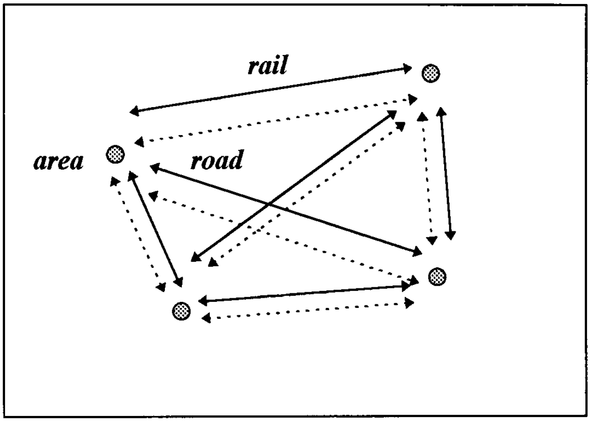

Fig. 2. Scheme of a modal split problem between 67 areas

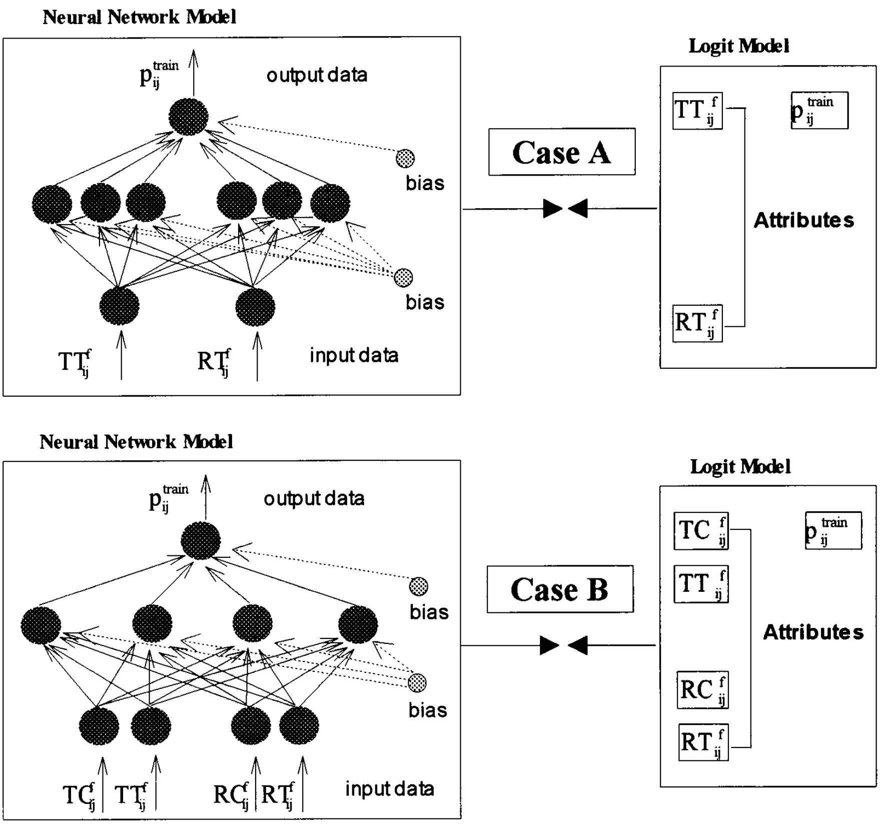

Fig. 3. Configuration of the two cases studied

case with 4 attributes, i.e. ‘time’ and ‘cost’ related to both transport modes. This case will be analysed here further, as – in order to evaluate the impact of the HST – it is important to take into account ‘cost’ attributes as economic choice criteria.

Having described now two main background results to be used for our new empirical forecasting experiments, we will describe in subsequent sections our findings for the expected competition effects between road and rail in Italy. Air will not be included due to lack of spatial data, but might certainly be an interesting addition. Subsections 3.3 and 3.4 will describe the experiments related to Case A and Case B mentioned above. The models will be calibrated (trained[6]) by using the whole data set containing 1396 observations, while the impact of the HST will be evaluated only on the links shown in Fig. 4.

In fact, according to the new data provided by the Italian State Railways, which refer to the new values of the attributes (rail ‘time’ and rail ‘cost’) after the introduction of the HST, the links[7] which are supposed to be connected by the new line HST are the following:

- *Torino↔Bologna; Torino↔Roma; Torino↔Napoli; Milano↔Bologna; Milano↔Roma; Milano↔Napoli; Venezia↔Roma; Venezia↔Napoli; Genova↔Bologna; Genova↔Venezia; Genova↔Napoli; Bologna↔Roma; Bologna↔Napoli; Roma↔Napoli.*

We will now present in more detail the results for cases A and B.

3.3 Impacts of ‘time’ attributes: case A

In the previous subsection we have already referred to the two cases A and B, which will be analysed and compared in this paper. The choice of these cases was depending on the values of the statistical indicator evaluated on the test-set (see above), as can be easily seen from Fig. 5.

Thus, according to Case A, it seems that, on the basis of both the problem under investigation and the data set, the choice of the transport mode for the generic passenger does not depend on the ‘cost’ attributes.

By calibrating, respectively training, the logit model and the feedforward NN model (for a mathematical exposition on these models see Nijkamp et al. 1997) by means of a (training) set containing 1396 observations, it is then possible to make intermodal forecasts by using the new values of the attributes related to the links which are supposed to be connected by the HST (see above). Only for these links, in Table 1 the attributes *before* and *after* the HST are shown.

[6] Note that for the NN model – in order to take into account of the *overfitting problem* (see e.g. Fischer and Gopal 1994; Reggiani and Tritapepe 1997) – it is necessary to know the point (number of iterations) when stopping the learning-phase. Usually this stopping-point is determined by means of the Cross-Validating Technique (see e.g. again Reggiani and Tritapepe 1997); in the Cases A and B described here, the points which have been determined in the work of Nijkamp et al. (1997) will be used

[7] Each link has to be considered in both directions

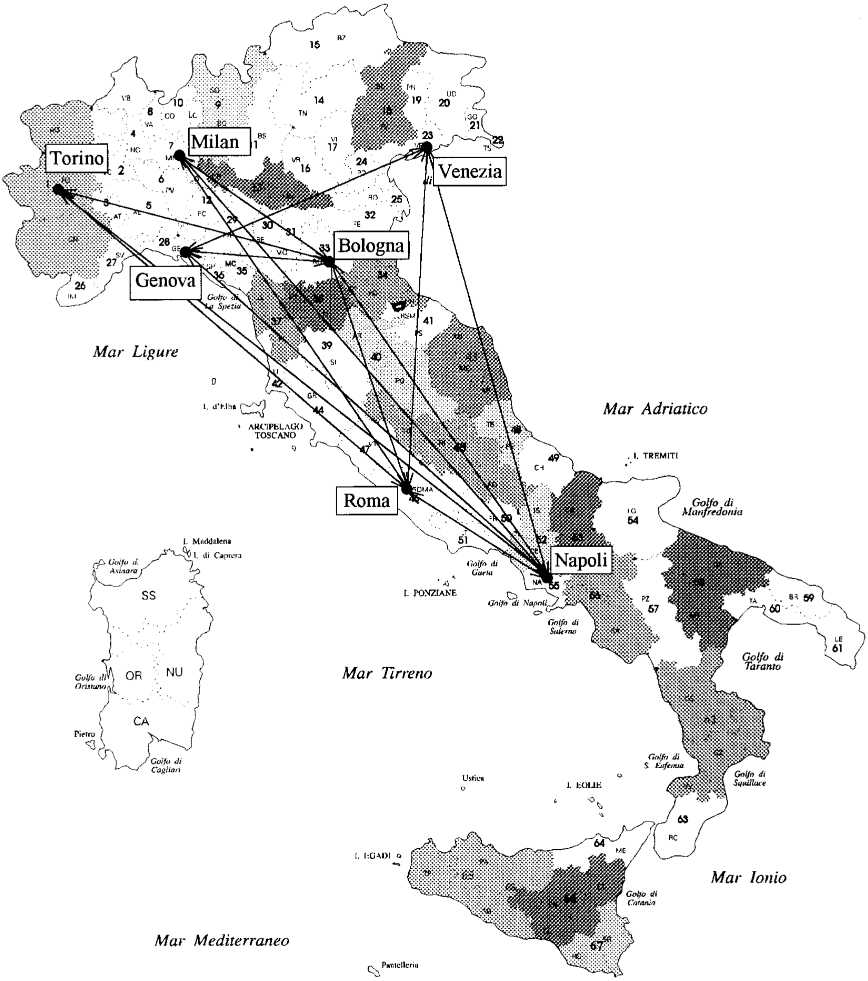

Fig. 4. Links which will be connected by the HST

The predicted flow probabilities obtained from the application of both models are compared to the flow probabilities without the new HST connection, and both are shown in Fig. 6 and in Table 2.

By examining Fig. 6, some qualitative observations may be made:

- according to both the logit model and the feedforward NN model, the introduction of the HST gives rise to an increase in the probabilities of choosing the rail transport mode;
- in the NN model the increase in the rail probabilities is higher than in the logit model;
- NN forecasts seem to follow a stochastic pattern while logit forecasts tend to stabilize the predicted values.

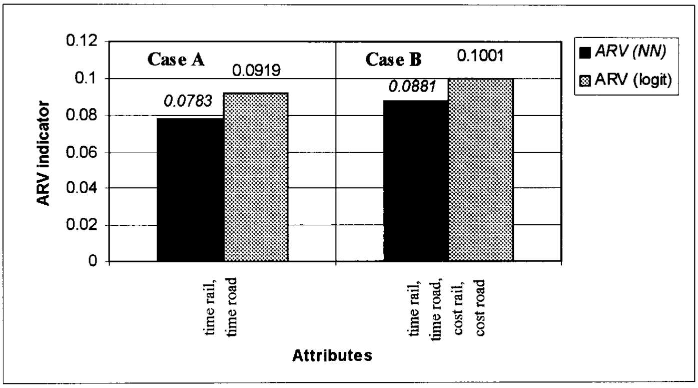

Fig. 5. The results of spatial forecasting for the two cases studied

Table 1. The attributes 'before' and 'after' the HST for case A (time in minutes) including the 'entrance' and 'exit' time (data provided by the Italian Railways Company)

Attributes case A							
Before high-speed train				After high-speed train			
O	D	tot rail time	tot road time	O	D	tot rail time	tot road time
Torino	Bologna	409	346	Torino	Bologna	392	346
Torino	Roma	549	589	Torino	Roma	517	589
Torino	Napoli	666	724	Torino	Napoli	590	724
Milano	Bologna	322	275	Milano	Bologna	291	275
Milano	Roma	488	533	Milano	Roma	452	533
Milano	Napoli	589	669	Milano	Napoli	509	669
Venezia	Roma	474	488	Venezia	Roma	462	488
Venezia	Napoli	679	624	Venezia	Napoli	609	624
Genova	Bologna	367	304	Genova	Bologna	357	304
Genova	Venezia	475	375	Genova	Venezia	456	375
Genova	Napoli	587	596	Genova	Napoli	565	596
Bologna	Roma	385	390	Bologna	Roma	376	390
Bologna	Napoli	532	526	Bologna	Napoli	473	526
Roma	Napoli	359	291	Roma	Napoli	307	291

These results emphasize the good 'nonlinear pattern' of an NN approach, in comparison with logit model. On the other hand, the logit predictions are surely helpful in order to better 'optimize' the NN results. Consequently, the 'combination' of both results seems quite suitable in a forecasting analysis where a slightly broader range of proper values may be preferred to 'unique' point values. Such a kind of sensitivity analysis is also methodologically interesting, as it may increase our understanding of the robustness of the predictions.

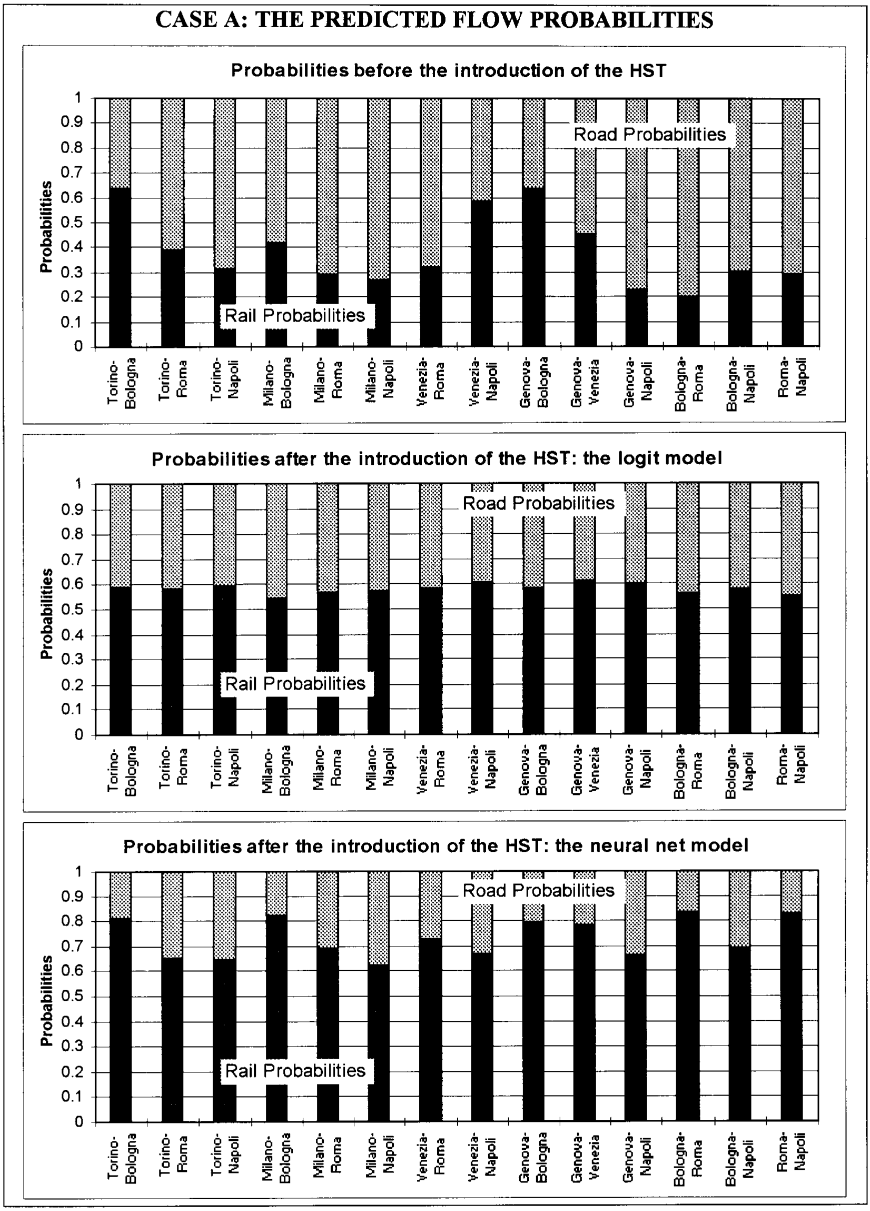

Fig. 6. Impact of the HST in the case with two attributes

3.4 Impacts of 'time' and 'cost' attributes: case B

In this case, the analysis of the impact of the HST is undertaken by means of four attributes, i.e. 'time' and 'cost' attributes for both models. These values, *before* and *after* the HST, are shown in Table 3.

Table 2. The rail mode flow probabilities 'before' and 'after' the HST for case A

Forecasted rail mode flow probabilities: case A							
Neural net model				Logit model			
O	D	before HST	after HST	O	D	before HST	after HST
Torino	Bologna	0.6353	0.8108	Torino	Bologna	0.6353	0.5903
Torino	Roma	0.3900	0.6551	Torino	Roma	0.3900	0.5869
Torino	Napoli	0.3141	0.6502	Torino	Napoli	0.3141	0.5944
Milano	Bologna	0.4209	0.8228	Milano	Bologna	0.4209	0.5460
Milano	Roma	0.2949	0.6920	Milano	Roma	0.2949	0.5699
Milano	Napoli	0.2695	0.6266	Milano	Napoli	0.2695	0.5758
Venezia	Roma	0.3198	0.7259	Venezia	Roma	0.3198	0.5828
Venezia	Napoli	0.5858	0.6713	Venezia	Napoli	0.5858	0.6068
Genova	Bologna	0.6354	0.7939	Genova	Bologna	0.6354	0.5869
Genova	Venezia	0.4564	0.7858	Genova	Venezia	0.4564	0.6119
Genova	Napoli	0.2301	0.6636	Genova	Napoli	0.2301	0.5999
Bologna	Roma	0.1983	0.8341	Bologna	Roma	0.1983	0.5630
Bologna	Napoli	0.3049	0.6922	Bologna	Napoli	0.3049	0.5802
Roma	Napoli	0.2944	0.8282	Roma	Napoli	0.2944	0.5523

In Fig. 7 and in Table 4, the probabilities predicted from these models, compared to the probabilities of choosing rail or road transport modes before the introduction of the HST, are presented. They will concisely be commented upon.

By examining Fig. 7, one can see that no drastic change occurs, with reference to the previous case. The general considerations underlined for case A (see the previous Sect. 3.3) still hold for case B. In particular, these last results confirm the 'robustness' of the logit model which seems to predict in case B the same values as those depicted in case A. The same predictions in both the two cases are likely caused by the independence from irrelevant alternative (IIA) feature underlying discrete choice models. On the contrary, the NN approach seems to be more sensitive to changes in the input information. Finally, in the next section, the predictions of both models for both Cases A and B will be mutually compared.

3.5 Comparison of different cases

Next, a comparison of the results of the models for both cases is depicted in Fig. 8 and in Table 5, by showing a homogeneity of results especially for logit models. NN show slightly different values in the two cases, by underlying their 'intrinsic' sensitivity to data variations.

A first qualitative observation, by looking at Fig. 8, is that the feedforward NN model forecasts in Case B lower rail probabilities than in Case A; on the other hand, the logit model forecasts in Case B higher rail probabilities than in Case A. Then, by observing in more detail Table 5, it is noteworthy that the above observation is valid for all links except for the

Table 3. The attributes‘ before’ and ‘after’ the HST for case B

Attributes case B											
Before high-speed train						After high-speed train					
O	D	tot rail time	tot road time	tot rail cost	tot road cost	O	D	tot rail time	tot road time	tot rail cost	tot road cost
Torino	Bologna	409	346	33.580	58.375	Torino	Bologna	392	346	36.672	58.375
Torino	Roma	549	589	50.814	115.982	Torino	Roma	517	589	70.503	115.982
Torino	Napoli	666	724	57.757	144.814	Torino	Napoli	590	724	82.091	144.814
Milano	Bologna	322	275	32.144	43.913	Milano	Bologna	291	275	39.640	43.913
Milano	Roma	488	533	58.496	96.953	Milano	Roma	452	533	79.111	96.953
Milano	Napoli	589	669	69.108	125.785	Milano	Napoli	509	669	95.771	125.785
Venezia	Roma	474	488	44.423	90.709	Venezia	Roma	462	488	60.277	90.709
Venezia	Napoli	679	624	53.455	119.541	Venezia	Napoli	609	624	75.208	119.541
Genova	Bologna	367	304	28.828	52.883	Genova	Bologna	357	304	31.498	52.883
Genova	Venezia	475	375	31.732	65.385	Genova	Venezia	456	375	31.732	65.385
Genova	Napoli	587	596	52.100	122.690	Genova	Napoli	565	596	58.814	122.690
Bologna	Roma	385	390	42.752	69.168	Bologna	Roma	376	390	55.871	69.168
Bologna	Napoli	532	526	55.775	98.000	Bologna	Napoli	473	526	76.219	98.000
Roma	Napoli	359	291	32.623	45.964	Roma	Napoli	307	291	39.949	45.964

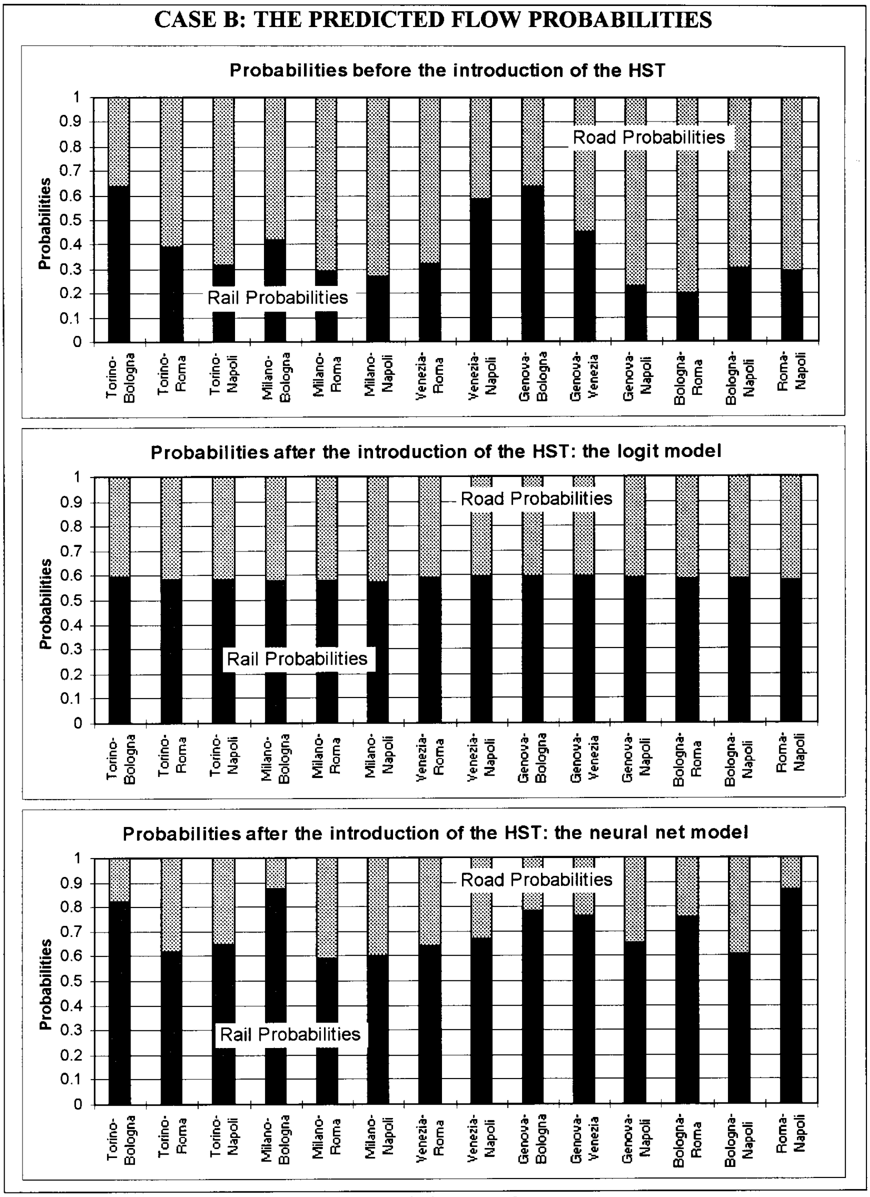

Fig. 7. Impact of the HST in the case with four attributes

link *Genova-Venezia* (in both models). A few relatively large differences are also to be ascribed to inaccuracies in some of the railway data used. Clearly, the model results are influenced by the assumptions made in cases A and B.

Of course, at least one question arises concerning the goodness of fit of the predictions and, in particular, concerning the question which model pro-

Table 4. The rail mode flow probabilities 'before' and 'after' the HST for case B

Forecasted rail mode flow probabilities: case B							
Neural net model				Logit model			
O	D	before HST	after HST	O	D	before HST	after HST
Torino	Bologna	0.6353	0.8242	Torino	Bologna	0.6353	0.5969
Torino	Roma	0.3900	0.6204	Torino	Roma	0.3900	0.5840
Torino	Napoli	0.3141	0.6492	Torino	Napoli	0.3141	0.5827
Milano	Bologna	0.4209	0.8753	Milano	Bologna	0.4209	0.5816
Milano	Roma	0.2949	0.5927	Milano	Roma	0.2949	0.5780
Milano	Napoli	0.2695	0.5998	Milano	Napoli	0.2695	0.5720
Venezia	Roma	0.3198	0.6432	Venezia	Roma	0.3198	0.5890
Venezia	Napoli	0.5858	0.6680	Venezia	Napoli	0.5858	0.5991
Genova	Bologna	0.6354	0.7821	Genova	Bologna	0.6354	0.5964
Genova	Venezia	0.4564	0.7605	Genova	Venezia	0.4564	0.5992
Genova	Napoli	0.2301	0.6555	Genova	Napoli	0.2301	0.5892
Bologna	Roma	0.1983	0.7543	Bologna	Roma	0.1983	0.5878
Bologna	Napoli	0.3049	0.6061	Bologna	Napoli	0.3049	0.5873
Roma	Napoli	0.2944	0.8715	Roma	Napoli	0.2944	0.5819

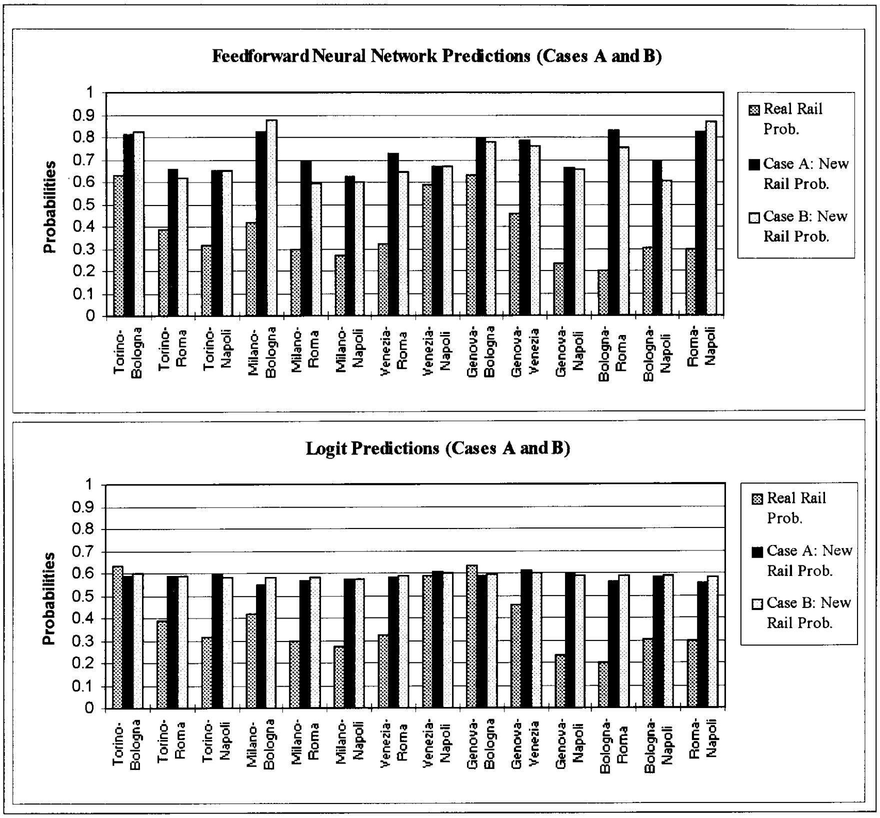

Fig. 8. Comparison of the predictions in the cases A and B

Table 5. Summary of the predictions of the experiments carried out

Predicted rail mode flows						
Before the HST			Neural net model after the HST		Logit model after the HST	
Origin	Destination	rail flows	Case A	Case B	Case A	Case B
Torino	Bologna	385	491	499	358	362
Torino	Roma	411	690	654	619	616
Torino	Napoli	103	213	213	195	191
Milano	Bologna	1585	3099	3296	2056	2190
Milano	Roma	659	1547	1325	1274	1292
Milano	Napoli	325	756	723	694	690
Venezia	Roma	158	359	318	288	291
Venezia	Napoli	99	113	113	103	101
Genova	Bologna	291	364	358	269	273
Genova	Venezia	110	189	183	147	144
Genova	Napoli	49	141	140	128	126
Bologna	Roma	259	1089	985	735	768
Bologna	Napoli	125	284	248	238	241
Roma	Napoli	3141	8837	9299	5893	6209

vides better results. In this respect, the following concluding considerations may be relevant:

- according to the values of the *ARV* indicators, which evaluate the ability of the models for spatial forecasting, the feedforward NN model tends to forecast slightly better than the logit model;
- although a conclusive ‘test’ concerning our forecasting analysis is not entirely possible, a ‘global view’ on both the logit and NN results teaches us that our findings are in general quite plausible;
- despite a global similarity on results of the two different approaches, the relatively modest discrepancies suggest the validity – for each link – of a slightly broader range of predicted values. Such a sensitivity analysis of results from different methodologies may be appealing for planners or decision-makers, as this focuses the attention more on the recognition of the patterns of results than on individual point estimates which may be hampered by data flows on separate links.

4. Conclusions

In the present paper the impact of the HST on modal split in Italy has been studied. The analysis has been carried out by means of two methodologically different models in order to compare the performance of the models and to provide a range of values predicted.

First of all, the best configuration of the models investigated and their results showed that the 'distance' attributes may lead to a distorsion in both models, probably because it is correlated with other attributes in our analysis.

Next, the introduction of the HST has been investigated; from an economic viewpoint, both models have shown the benefits of the new HST, in particular in regard to the rise in the probabilities of choosing the rail transport mode on most links in Italy.

From a methodological viewpoint, some discrepancies were observed by using the logit model; the feedforward NN model seems to provide acceptable predictions, although it is important to well define the NN architecture and to well train the network during the learning phase (see Nijkamp et al. 1996a). In this framework, a comparative analysis among different architectures/algorithms as well as different software packages could be certainly useful, while also the development of a better data base with a view on monitoring new developments seems to be a sine qua non.

Clearly, NN models take sometimes much training time and this may limit their application for quick large scale problems, while also their sensitivity to the preprocessing of data may be a cause for concern, as this may lead to fluctuating results. A fascinating methodological issue is still the question on the behavioural contents of NN models. It seem to be possible to use the weights inside a NN to derive coefficients of a discrete choice model (see Schintler and Olurotimi 1997) by designing a specific NN formulation and architecture. It is also worth mentioning that the sigmoid function used in an NN analysis for training the network is essentially a logistic function related to a binary choice model. This may lead to another correspondence between NN models and discrete choice models. It is clear that the application of NN analysis in transportation behaviour still deserves more careful methodological attention.

Acknowledgements. The authors wish to acknowledge helpful comments given on a previous version of their paper by J. Poot and L. Schintler. Also constructive comments of two anonymous referees have to be acknowledged. The second author gratefully acknowledges the CNR "Finalizzato Trasporti 2" Project, contract no. 96.00098.PF74. The third author acknowledges a Human Capital and Mobility Research Grant of the EC (Strategic Infrastructure Networks in Europe, Contract no. ERBCHBGCT920162), as well as the good facilities to carry out this study at the host institution (Economic and Social Institute, Department of Economics, Free University, Amsterdam).

References

Aleksander I, Morton H (1990) An Introduction to Neural Computing. Chapman & Hall, London

Ben Akiva M, Lerman SR (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, Massachusetts

Davalo E, Naim P (1991) Neural Networks. Macmillan Education LTD, London

Dougherty M (1995) A Review of Neural Networks Applied to Transport. Transportation Research C, 3: 247–260

Fischer MM (1994) Expert Systems and Artificial Neural Networks for Spatial Analysis and Modelling. Geographical Systems, 1:221–235

Fischer MM, Gopal S (1994) Artificial Neural Networks: a New Approach to Modelling Interregional Telecommunication Flows. Journal of Regional Science 34:503–527

Hecht-Nielsen R (1990) Neurocomputing. Addison-Wesley Publishing Company, Inc, USA

Himanen V, Nijkamp P, Reggiani A (eds) (1997) Neural Networks in Transport. Avebury, Aldershot (forthcoming)

Kodratoff Y, Michalski R (1990) Machine Learning an Artificial Intelligence Approach. Morgan Kaufmann Publishers, Inc, San Mateo, California

Nijkamp P, Reggiani A, Tritapepe T (1996a) Modelling Inter-Urban Flows in Italy: a Comparison between Neural Network Approach and Logit Analysis, in: Reggiani A (ed) Special Issue on Design and Management of Transport Networks: New Approaches and Experiments. Transportation Research C, 4:323–328

Nijkamp P, Reggiani A, Tritapepe T (1997) Analysis of Complex Networks: An Overview of Methodologies and a Neural Network Application to International Transport in Italy, in: Capineri C, Rietveld P (eds) Policy Aspects of Networks. Avebury, Aldershot (forthcoming)

McCullough W, Pitts WH (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 5:113–115

Reggiani A, Romanelli R, Tritapepe T, Nijkamp P (1997) Neural Networks: an Overview and Applications in the Space-Economy, in: HimanenV, Nijkamp P, Reggiani A (eds) Neural Networks in Transport Systems. Avebury, Aldershot (forthcoming)

Reggiani A, Tritapepe T (1997) Neural Network and Logit Models to Commuters' Mobility in the Metropolitan Area of Milan, in: Himanen V, Nijkamp P, Reggiani A (eds) Neural Networks in Transport Systems. Avebury, Aldershot (forthcoming)

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning Representations by Back-Propagating Errors. Nature, 323:533–536

Schintler L, Olurotimi OO (1997) The Use of Neural Networks in Planning for Demand-Responsive Transit, in: Himanen V, Nijkamp P, Reggiani A (eds.), Neural Networks in Transport Systems. Avebury, Aldershot (forthcoming)

Taber RA (1991) Matter of Degree. Helix 3:50–55

Valdes R (1991) What Is BioComputing? Dr. Dobb's Journal 46:108–109