

Nijkamp, Peter; Reggiani, Aura; Sabella, E.

Conference Paper

A comparison of the performance of Discrete Choice Models and Biocomputing Models in Transport Systems Analysis

39th Congress of the European Regional Science Association: "Regional Cohesion and Competitiveness in 21st Century Europe", August 23 - 27, 1999, Dublin, Ireland

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Nijkamp, Peter; Reggiani, Aura; Sabella, E. (1999) : A comparison of the performance of Discrete Choice Models and Biocomputing Models in Transport Systems Analysis, 39th Congress of the European Regional Science Association: "Regional Cohesion and Competitiveness in 21st Century Europe", August 23 - 27, 1999, Dublin, Ireland, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/114212>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A COMPARISON OF THE PERFORMANCE OF DISCRETE CHOICE MODELS AND BIOCOMPUTING MODELS IN TRANSPORT SYSTEMS ANALYSIS

P. Nijkamp¹, A. Reggiani², E. Sabella²

¹ Department of Spatial Economics, Free University of Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam (The Netherlands); e-mail: pnijkamp@econ.vu.nl;

² Department of Economics, Faculty of Statistics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna (Italy); e-mail: reggiani@economia.unibo.it; sabel100@economia.unibo.it

Abstract

The present paper aims to offer a comparative empirical study of the performance of various spatial flow models with a view to their applicability and relevance in freight transport analysis.

Two classes of models will be considered in more detail, viz. *discrete choice models and biocomputing models*. From the class of discrete choice models the conventional logit model will be employed. Next, from the class of biocomputing models we will address in particular two types, viz. *neural network models* and *evolutionary algorithms* (in particular, *genetic algorithms*). Also a combined version of neural networks and genetic algorithms will be developed and compared with the previous classes of models.

A mutual comparison of all these types of models will be carried out on the basis of an extensive data set of European transport flows for various commodity classes, based on a rather detailed regional subdivision of European countries.

The results will also be used to investigate the sensitivity of transport movements vis-à-vis changes in transport costs, by developing and testing several transport policy scenarios on user charges for the freight transport sector.

1. ANALYSIS OF COMPLEX NETWORKS: AN INTRODUCTION

Spatial competition is the result of unequal production opportunities and efficiency conditions in a open spatial market. Consequently, there will be a ‘natural’ (self-organising) tendency towards a case of losers and winners. The efficiency differences between regions are inter alia reflected in interregional import and export flows. And hence, transportation networks (at the supply side) and transportation flows (at the demand side) are critical in assessing spatial dynamics. This paper will address the issue of spatially interwoven trade flows by offering new modelling perspectives.

The modelling of spatial-economic systems already has a history of several decades and the increasing popularity of spatial network models is noteworthy (see Reggiani *et al.*, 1999). Spatial networks (transport networks, interacting life-style groups in local communities, systems of cities or regions, etc.) are in general, open multi-layered, multi-faceted, multi-actor and dynamic systems with a rich variety of (direct and indirect) interconnections¹. It should be noted at the outset that the analysis of complex spatial networks – with interactions among diverse components and layers – is in principle no different from the research task in other disciplines such as physics or biology which address network phenomena (Nijkamp and Reggiani, 1998). Clearly, spatial systems have an additional degree of complexity compared to natural science systems, viz. human activity at both an individual and a collective level. The collective level of human action, which is basic and far-reaching, necessitates the search for general analytical principles helpful in better comprehending the degree of variation and stochasticity in spatial systems.

An important methodological step forward in this context has been offered by synergetic theory (see amongst others Haken, 1983a, 1983b). Synergetics studies the complex relationships between interactive components and their consequences for evolving macrostructures by seeking simple universal principles that govern the dynamics of flows and morphology in the organisation of a system. Such dynamic forces may generate qualitatively new structures in space and time, while new functions of these systems may also be generated (see Nijkamp *et al.*, 1997).

The synergetics concept provokes many new research issues, e.g. regarding resilience, robustness, equilibrium, or sustainability of such systems. Synergetics maps such changes and their implications by focusing on qualitative changes in a complex system. Its main assertion is that the complex nature of a macro-system can be characterised by only a few parameters which govern the movement or behaviour of the micro-components comprising particular (sub)systems. The search for simple, dynamic and non-linear equations of motion able to adequately describe and predict the behaviour of a multi-component multi-actor system is apparently based on a synergetics concept (see also Domanski, 1992, 1994, and Weidlich and Haag, 1983).

It is conceivable then that complexity and synergetics have found significant appeal in contemporary spatial science research. Examples are the application of self-organising models, evolutionary models, cellular automata, fractal growth analysis, or neural networks to complex spatial systems. Clearly, such concepts do not require entirely new classes of models, but can instead build upon existing spatial models and add to the new logic of complex systems (Nijkamp and Reggiani, 1998).

Evidently, the analysis of complex networks has become an important research issue in many disciplines. It has intensified the search for universal principles driving non-linear dynamics systems with a particular interest in methodological frameworks and instruments. Similarly, we also observe emphasis on network evolution, since networks are prominent examples of interacting dynamic multi-component systems in

¹ “Networks are not only specific, organized spatial structures (i.e. based on node-edge interconnectivity); they offer user functions designed to improve the efficiency of spatial interactions. Thus, morphological structure and spatial interaction are dual phenomena. This is also echoed in the recent conceptual interest in networks as complex socio-temporal systems (see e.g. Batten *et al.*, 1994, and Reggiani and Nijkamp, 1996a). In this literature, networks are interpreted as spatial-economic or socio-economic systems characterized by evolutionary processes governed by a multi-layer, multi-component organization of interdependent sub-systems. Clearly, a hierarchical system is only a special case of a more general network ramification” (Nijkamp *et al.*, 1997).

which flow dynamics and morphological evolution play a crucial role. In spatial sciences these systems also have a strong human behavioural and policy aspect (flow behaviour, regulatory regime).

It should be added that many real-world network systems emerging from the complex interaction between slow and fast dynamics of the system's components not only reveal the complexity of dynamic systems, but could also mirror self-organising powers which may be depicted by various typologies of evolutionary non-regular behaviour (Reggiani and Nijkamp, 1996b).

An interregional trade or transport system is an illustrative example in this context, whereby its non-linear evolution may induce a (more or less stabilising) perturbation within the entire spatial network of regions and cities (including interregional connections). A popular concept in transport has become the term 'sustainability', which describes the continuity potential of a dynamic system under changing external conditions and countervailing powers and behaviours within the system (Nijkamp, 1991).

Clearly, further investigation into sustainable development of networks requires greater insight into the level of complexity – as a result of the dynamic interaction among subsystems – and the level of uncertainty associated with the dynamic behaviour of such systems. This will be further highlighted in the present paper.

The paper is organised as follows. Section 2 addresses recent important contributions to the spatial sciences in, for example, neurocomputing models. Next, we illustrate in Section 3 the potential for a new analytical departure, in particular by means of evolutionary algorithms, since these methodologies belong to the general field of neurocomputing and have gained popularity in the spatial/transport literature. Section 4 explores the potential and applicability of a hybrid approach based on evolutionary algorithms and neural network analysis. The new approach is applied to the modelling and forecasting of European freight transport flows. The paper concludes with suggestions for future research (Section 5).

2. NEW APPROACHES: NEUROCOMPUTING MODELS

2.1 Neural Network Models

Neural networks (NNs) are becoming popular tools for analysing complex systems. Complex choice problems are increasingly analysed via NNs in conjunction with their general similarity to functions of the human brain. This may also be seen in bio-computing and artificial intelligence.

Contemporary scientific literature has dedicated much attention to NNs as alternative models of information processing (see for a review, Reggiani *et al.*, 1998b). NN approaches stand apart from the aforementioned models in their capacity to generalise from experience without fixing – *a priori* – any behavioural rule/model among variables (see e.g., Maren *et al.*, 1990, and Rumelhart *et al.*, 1986). Applications of NNs are abundant. In the spatial sciences this approach has gained much popularity, e.g. in transport and spatial-economic interactions (see for a review, Himanen *et al.*, 1998, and Reggiani *et al.*, 1998b). When using NNs it is necessary – as with complex networks – to adopt tools able to map connectivity, communication, adaptivity, control and prediction patterns (see also Nijkamp *et al.*, 1997).

The principal inspiration of NNs is the human brain whose structure is comprised of billions of connected neurons. Every neuron is a cell body which receives electrochemical input signals through the dendrites² and transmits the resulting electrochemical output signal through the axon³; information transmission (electrochemical signals) is then spread by spatial connections, the 'synapses', between axons and dendrites of different neurons. When the combined signals are strong enough the neuron is 'activated' and will produce an output signal.

In computational sciences, the structure of NNs is represented by logical units ('neurons'), connected by channels of communication ('synapses') which intercompute independently, since each unit cooperates in the transmission of information by means of a different 'weight'⁴. This differentiation in the weights thus

² Dendrites are usually defined as 'tree-like networks of nerve fibres' (see Hertz *et al.*, 1991).

³ An axon is a 'single long fibre' (see Hertz *et al.*, 1991).

⁴ Here weight is a real number assigned to a connection between two units.

corresponds to different values of the synapses. This occurs particularly during the ‘learning’ phase in order to allow adaptation to new conditions. In fact, as with cerebral behaviour, NNs are able to recognise previously unobserved patterns. This characteristic of ‘generalising’⁵ identifies the behaviour of the system as ‘intelligent’. In other words, since ‘real’ events never repeat themselves precisely, intelligent systems are able to observe, by means of past experience, the continuity and similarity of real events, by offering the possibility of correctly predicting future events.

It should be added that within the context of Artificial Intelligence, NNs are distinguished by their ability to elaborate and create information by means of Parallel Distributed Processing systems. This is an important aspect of NNs, since massive parallelism provides, on the one hand, the possibility of significantly increasing computer speed (see Kosko, 1992) and, on the other hand, offers a great ‘fault tolerance’ (since inter-connection among units is essentially local).

It is difficult to define NNs in formal terms; in principle, they can be considered as non-linear dynamic systems with many freedom degrees as well as ‘free’ models of estimation (see Kosko, 1992). The common element is the concept of ‘freedom’; in other words, there is ‘free biological behaviour’ within NNs which cannot be subjected to any mathematical model (usually creating logical boundaries between output and input). And so in contrast to the necessity to programme computers that require knowledge of the mathematical model which represents a real context, NNs are trained; that is, they learn from a set of ‘examples’ with input and output. In other words, NNs can be defined as connected systems that can solve a wide range of complex problems which are not able to be formalised by logical mathematical models.

Moreover, NNs are suitable in a forecasting context, given their ability to generalise, i.e., to elaborate new situations for different scenarios. In this framework, it should be noted that this peculiarity strictly depends on the chosen training set and the architectural configuration of the network (number of hidden levels, number of units on these levels, etc.) (see e.g. Fischer and Gopal, 1994). The first point presupposes the choice of a representative sample in order to get an ‘unbiased’ distribution of observations, while the second point implies a ‘careful’ mapping of NN architecture, since the optimal configuration is reached only by means of experimental methodologies (see also Malliaris and Salchenberger, 1992).

Details on NN structures and typologies can be found in Nijkamp and Reggiani (1998), Nijkamp *et al.* (1997), and Reggiani *et al.* (1998b). Several applications on high-dimensional complex networks, like the Italian passenger transport network (Nijkamp *et al.*, 1996) or the European freight transport networks (Reggiani *et al.*, 1998a), have shown a high performance level for NNs. However, NNs are still not easily interpretable from a behavioural viewpoint, although recent results show a compatibility between NNs and binary logit models emerging from micro-economic theory (Schintler and Olurotimi, 1998).

In this research endeavour towards complex network modelling a great potential is certainly offered by evolutionary computations, particularly evolutionary algorithms that are able to reach the optimal solution by means of natural selection and genetics. This innovative tool will be further discussed in the next sections.

2.2 Evolutionary Algorithms

There has been an avalanche of contributions to evolutionary thinking in spatial economics as well as the development of several ecologically-based model experiments which have stimulated the use of evolutionary algorithms (EAs) in social science research.

EA is a generic term used to describe computer-based problem solving systems which use computational models of evolutionary processes and structures as key elements in their design, specification and implementation. In other words, EAs are – usually stochastic – search methods of human behaviour that aim to mimic the metaphor of biological evolution used in social science research issues. They normally operate on a population of potential solutions to problems by applying the principle of ‘survival of the fittest’ to produce increasingly better approximations to a final

⁵ ‘Generalizing’ is the capacity of a system to create new patterns in accordance with previously studied examples.

equilibrium solution. For each relevant generation of solution types, a new set of solution approximations is created through a selection process of individual approximations according to their level of ‘fitness’ in the problem domain, and by combining them by means of operators borrowed from genetics. This repetitive process leads to the evolution of sets of individual solutions which are better suited to their choice environment than the individuals they were originating from by means of a process of natural adaptation. EAs try to simulate three main characteristics generally belonging to a natural dynamic system: (a) adaptivity; (b) stochasticity; (c) parallelism (see Colomi *et al.*, 1994). The first property refers to the possibility – for a system – to modify its solution results by feedback effects; the second allows the system to find ‘good’ solutions in a short time by using property (a); and the third property outlines the possibility of using high parallel computer power as a consequence of property (b).

In other words, EAs map out, by means of the above properties, a fundamental characteristic of networks in natural systems, i.e., the synergy effect, characterising also the functioning and operation of socio-economic and spatial networks (see Nijkamp and Reggiani, 1996). EAs thus seem able to overcome several limitations inherent so far in conventional social science and spatial-economic models. It is therefore worthwhile to explore the potential characteristics and scientific contributions of this new tool in greater detail.

3. EVOLUTIONARY ALGORITHMS FOR THE ANALYSIS AND MODELLING OF COMPLEX DYNAMIC NETWORKS

3.1 The Basic Structure of Evolutionary Algorithms: an Introduction

EAs are based on the imitation of processes found in the natural evolution of species. Their origin, as mentioned earlier, is in biology rather than in computer sciences.

Evolution as a concept in dynamic biology and population dynamics is governed by chromosomes⁶: organic information carriers which contain the exact characteristics of a living being. The living being can be ‘constructed’ by decoding its chromosomes. How this is done is not yet fully known, but the following properties are commonly accepted (see also Goldberg, 1989):

- (i) Evolution is a process working on chromosomes instead of the living beings they represent.
- (ii) Natural selection is the dynamic relationship between chromosomes. In other words, it is the successful performance of their decoded structure which will more often reproduce.
- (iii) Evolution occurs while reproducing. Mutation can, for instance, be the reason why chromosomes of the children sometimes differ from the ones of their parents in certain places. The chromosomes of the parents are combined in a certain way so as to create new and different chromosomes for the children.
- (iv) Biologic evolution has no memory. All it knows about individuals that perform well in their environment is stored in the set of chromosomes of the present individuals and in the way these chromosomes are encoded.

EAs work on populations of individuals represented by chromosomes instead of single solutions⁷ (see Figure 1 displaying the basic structure of an EA).

In an EA structure a number of individuals (the ‘population’) is randomly initialised (initial generation) in order to start the related computational algorithm. The objective function (OF) is then

⁶For a definition of the biological terminology we refer to Mitchell (1996, p.5): ‘All living organisms consist of cells, and each cell contains the same set of one or more *chromosomes* – strings of DNA – that serve as a ‘blueprint’ for the organism. A chromosome can be conceptually divided into *genes* – functional blocks of DNA – each of which encodes a particular protein. Very roughly, one can think of a gene as encoding a trait, such as eye color. The different possible ‘settings’ for a trait (e.g., blue, brown, hazel) are called *alleles*. Each gene is located at a particular *locus* (position) on the chromosome’.

⁷In EAs, the term ‘chromosome’ refers typically to a candidate solution to a problem, often encoded as a bit string.

evaluated for these individuals. If the optimisation criteria are not met, the creation of a new generation begins (see also Table 1).

Individuals are selected according to their fitness (i.e., contribution to the optimal solution) for the production of offspring⁸. All offspring will be mutated with a certain probability. The fitness of the offspring can afterwards be computed. Next the offspring are inserted into the population replacing the parents, thus producing a new generation. This cycle is continued until the optimisation criteria are reached (see also next section).

Such an EA structure – which refers to a single population – performs well on a broad class of problems. This process has a similarity to many real-world dynamic choice processes. However, better results can be obtained by introducing many populations (multipopulations). Each micro-population is then called subpopulation. Every subpopulation evolves independently for a few generations (like the single population EA); next, one or more individuals are exchanged between the subpopulations. Consequently, the multipopulation EA models the evolution of a species in a way more similar to nature than the single population EA.

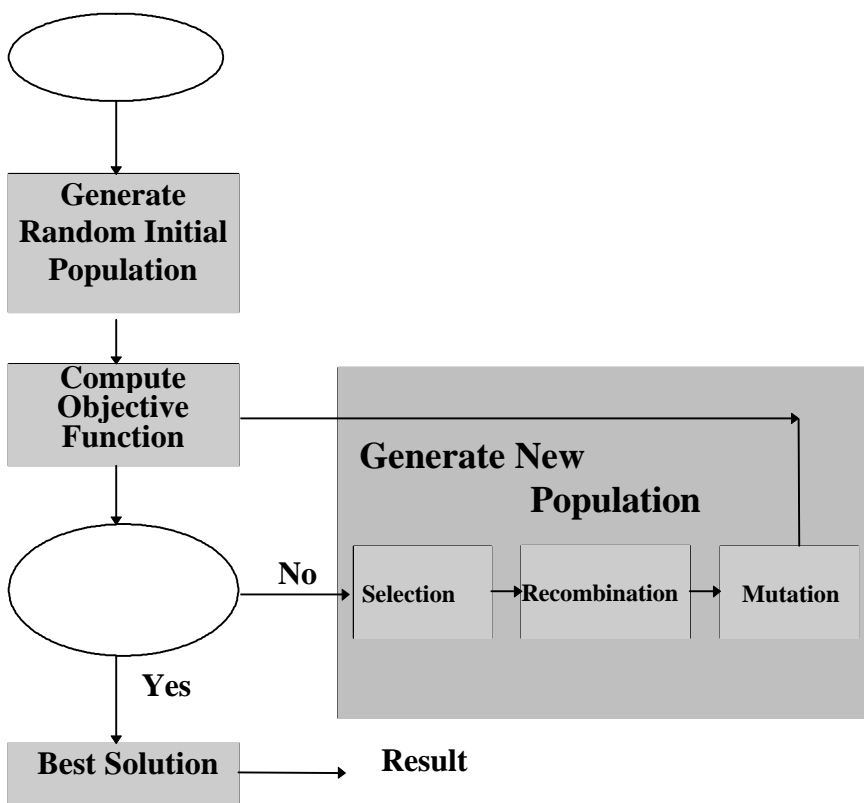


Figure 1: The basic structure of evolutionary algorithms

We will now illustrate one of the most relevant classes of models belonging to EAs: genetic algorithms (see also subsequent Sections 3.2).

3.2 Genetic Algorithms

Genetic algorithms (GAs) are computational models inspired by principles of natural evolution and population genetics . In scientific research, GAs have mainly been used as function optimisers for solving a wide range of problems (see for a review, Mitchell 1996). Their strength is essentially due to

⁸Offspring can be defined as the result of reproduction. In biological evolution offspring are children.

their ability to update an entire population of possible solutions during each iteration round; this allows for a parallel investigation of the search space (see Holland, 1975, and Bertoni and Dorigo, 1992). We will now offer a concise introduction to the GA computational paradigm on the basis of the work by Colorni *et al.* (1994) and Maniezzo (1994).

According to these authors, a GA evolves as a multiset of elements, called *population of individuals*. Each individual x_i ($i=1\dots n$) of the population X is represents a trial solution of the problem to be solved. Individuals are usually represented by strings of variables called *chromosomes*, every position of which is called a *gene* and its value is called its *allelic value*. The allelic value varies over a range on a set called *allelic alphabet* and it is usually restricted to a binary form (0,1).

The standard genetic algorithm proceeds as follows: an initially random population of such strings is generated. During each iteration round, the performance of each individual (solution) is evaluated according to a predefined *fitness function*. A new generation of individuals is then produced by applying a set of *genetic operators* to selected individuals from the previous generation (see Table 1).

A GA is capable of maximising a given FF computed on each individual of the population. If the problem is to minimise a given *objective function*, then it is required to identify and map the increasing FF values; this can be achieved by a monotonically decreasing function.

Table 1: The basic structure of genetic algorithms (source: Maniezzo, 1994)

- | |
|--|
| <p>Step 1: Generate randomly an initial population $X(0):=(x_1,\dots,x_n)$;</p> <p>Step 2: Compute the fitness $FF(x_i)$ of each individual x_i of the current population $X(t)$;</p> <p>Step 3: Generate an intermediate population $Xr(t)$ by applying the <i>reproduction</i> operator;</p> <p>Step 4: Generate $X(t+1)$ by applying other operators (subsequently defined) to $Xr(t)$;</p> <p>Step 5: $t:=t+1$; if not (<i>end - test</i>), go to Step 2.</p> |
|--|

The letter n denotes the population dimension.

The *end test* is usually a test on the number of generalisations or on the time length of the run, but it may also be based on more subtle indicators such as search stagnation.

In the work of Maniezzo (1994, p.40) the commonly used operators are described as follows:

i) *Reproduction*: This operator selects individuals from the current population according to their fitness function value. These individuals – which will make up the next generation – are selected randomly using a scheme with favours the more fit individuals (*Roulette wheel selection*). The extraction probability $p_r(x_i)$ of each individual x_i is proportional to its fitness $FF(x_i)$ as a ratio with respect to the average fitness of all individuals in $X(t)$:

$$p_r(x_i) = \frac{FF(x_i)}{\sum_{i=0}^n FF(x_i)} \quad (1)$$

ii) *Crossover*: This operator is applied with a certain probability p_c to exchange biologic information between two individuals. To apply the standard crossover operator the individuals of the population are randomly paired and cut their chromosome strings at some randomly-chosen position. Each pair is then recombined to produce *offspring* by the juxtaposition of the first part of one parent and the last part of the other parent.

iii) *Mutation*: This operator standard flips some of the bits in a chromosome with a fixed probability, where the mutation probability is a system parameter, p_m . Usually, the new allelic value is randomly chosen with a uniform probability distribution. This operation guarantees the possibility of exploring the whole search space independently from the initial population.

iv) *Local Search*: The need for of this operator in optimisation problems is still under debate. Local search is usually a simple gradient-descent heuristic that carries each solution to a local optimum (see, among others, Muhlenbein, 1989 and Colomi *et al.*, 1992a, b).

Crossover is generally considered as the principal search mechanism, with mutation relegated to a background operator whose exclusive role is to maintain diversity in the population and ensure that every point in the search space has some chance of being visited. By iterating the processes of selection, recombination and mutation, the population accumulates information about the distribution of fitness in the search space. One of the regions in which GAs perform quite well is optimisation. GAs are normally very *robust*, which means that they operate on a broad range of problems.

Given these characteristics of GAs, as well as their performance in existing applications (see, for a review, Colomi *et al.*, 1994), it seems now worthwhile to further explore this tool also in, for example, the modal split problem in a complex high-dimensional network (e.g. the European freight transport network). In previous works by the authors (Nijkamp and Reggiani, 1998; Reggiani *et al.*, 1998a, 1999), this problem was explored through a comparative analysis between logit and NN models. The results were quite interesting, but in the forecasting analysis sometimes rather significant differences emerged in the results of these two categories of models.

Herein, we follow-up previous research endeavours, since – in addition to logit and NN models – we investigate for the same European spatial network the ‘power’ of EAs, particularly by combining GAs with NNs. The results of to this empirical analysis are illustrated in Section 4.

4. EVOLUTIONARY NEURAL NETWORKS FOR EUROPEAN FREIGHT TRANSPORT MODELLING

4.1 Introduction

We now turn to the performance of EAs discussed in the previous sections to highlight the potentials/limitations of these new approaches. We will consider – as a case study – the European freight transport network with reference to the modal split problem between rail and road transport modes. Different NN models will be investigated and compared mainly in combination with EAs, giving rise to so-called evolutionary neural networks models (ENNs). The class of ENN models adopted here comprises four categories:

A: A Neural Network model using a genetic algorithm for the learning procedure [NN (GA)].

B: A Neural Network model using a genetic algorithm + backpropagation for the learning procedure [NN (GA+BP)].

Furthermore, Models A and B will be compared with the standard NN model:

C: A Neural Network model using a backpropagation algorithm for the learning procedure [NN(BP)].

These three categories of neurocomputing models will be compared with a conventional choice model often used in transportation research, viz. :

D: A Logit model using a Newton-Raphson algorithm for the calibration procedure.

Finally, a sensitivity analysis will be conducted in order to investigate the results of the models under different policy scenarios on environmental costs of freight flows.

4.2 The Data Set

The data set used in our ‘complex’ transport network contains the freight flows and the attributes related to each link between 108 European regions for the year 1986 (see for details Reggiani *et al.*, 1998a). The attributes considered are *time* and *cost* between each link (ij) with reference to each transport mode. In particular, each observation of the data set pertains to variables related to each link (ij). Furthermore, the flow distribution in the matrices concerned refers to one particular kind of good: food.

Since 108 areas have been considered, the data set should ideally contain 11664 observations (108 × 108). However, our data set contains finally 4409 observations because of the following considerations:

- the intra-area freight flows are zero;
- for each link, only the transport movements towards one direction $i \rightarrow j$ have been considered;
- only the links where the flows and the attributes (of both road and rail) are different from zero have been considered (empty cells are excluded).

The data set has been randomly subdivided into three sub-sets:

- a *training set* containing 3000 observations, about 68% of the data-set;
- a *cross-validation set* containing 600 observations, 20% of the training set;
- a *test set* containing 1409 observations, about 32% of the data-set.

The training set is used for the learning/calibration phase of the models adopted. The cross-validation set is useful to determine when to stop training the NN for solving the overfitting problem in the learning phase (see for details, Fischer and Gopal, 1994). The test set is used for comparing the results of models adopted after the introduction of data not used in the learning/calibration phase.

4.3 Empirical Application: the Models Adopted

In our application we have essentially used the GA tool to improve the learning procedure of an NN approach. Two models have been adopted according to the classification presented in the Section 4.1.

Model A

Model A refers to an ENN approach combining an NN model with evolutionary search procedures.

We have concentrated on the most popular class of NNs (i.e. a two layer feed-forward totally connected NN) combined with GA (see Section 3.2), in order to define a (near)optimal connection weight, instead of using the conventional procedures for the weight optimisation, like the backpropagation algorithm (BP; see Section 2.1).

The methodological structure of the main steps involved in the application of a feed-forward NN consists of three stages: i) definition of the network architecture; ii) learning phase; iii) forecasting phase (see Reggiani *et al.*, 1998b).

In our application, the ENN architecture (see Figure 2) contains four inputs which correspond to the attributes *time* and *cost* related to each transport mode (rail and road), one output unit corresponding to the probability of choosing one mode⁹ (e.g., the road mode) and 10 hidden units

⁹ The choice probability of the other mode is just the complement.

defined by a trial and error procedure¹⁰. The representation of the connection weights is – from a GA perspective – a binary string of a certain length. The set of connection weights is simply represented by the concatenation of all the weights in the network in a binary form. The objective function used is the mean square error (MSE¹¹).

Figure 2: The structure of Model A

Note: W_{ij} ($i=1..I$; $j=1..J$) are the weights (real number/binary form) of the synaptic connections between the i -th neuron of the input layer and the neurons of the hidden layer linked to it, where I = number of neurons in the input layer and J = number of the neurons in the hidden layer.

As already mentioned in Section 3.2, the GAs are designed for maximisation problems. For the minimisation problems – in this study we have to minimise the MSE – the objective function (OF) must be distinguished from the fitness function (FF). For this purpose, FF is usually obtained by transforming OF through a decreasing monotonic function. According to Vitetta (1993), the transformed function utilised here is a negative exponential function. The value of the FF for a generic solution x_i can be express as follows:

$$FF(x_i) = e^{-aOF(x_i)} \quad \mathbf{a} > 0 \quad (2)$$

As consequence the reproduction probability p_r expressed by (1) for a generic individual x_i results as follows:

¹⁰ The number of hidden units has been defined by numerical experiments.

¹¹ For the definition of MSE indicator, see Annex 1.

$$p_r(x_i) = \frac{e^{-aOF(x_i)}}{\sum_k e^{-aOF(x_k)}} \quad (3)$$

Model B

The efficiency of the evolutionary training of Model A – with reference to the computational time – can be improved by incorporating a local search procedure into the evaluation process. In fact, evolutionary training is computationally slower than gradient descent-based training, such as BP.

The basic idea is to use the BP algorithm as a genetic local-search operator (see Section 3.2) into a GA approach. This combination leads to an efficient method of training an NN, because it takes both the advantages of GAs and BP (the fast initial convergence of GAs and the powerful local search of BP). The new standard structure of a GA combined with BP is then (see Table 2):

Table 2: Genetic algorithm combined with backpropagation

Step 1:	Generate randomly an initial population $X(0) := (x_1, \dots, x_n)$;
Step 2:	Compute the $FF(x_i)$ for each individual x_i of the current population $X(t)$;
Step 2:	$X(t+1) = \text{reproduction} [X(t)]$;
Step 3:	$X(t+1) = \text{crossover} [X(t+1)]$;
Step 4:	$X(t+1) = \text{mutation} [X(t+1)]$;
Step 5:	$X(t+1) = \text{backpropagation}^{12} [X(t+1)]$;
Step 6:	$t = t+1$; if not end-test go to Step 2.

The architecture (network configuration) for the NN model is the same as that of Model A.

These two models have also been compared with the traditional NN model described in Section 2 and defined here as model C.

Model C

Model C is a conventional NN(BP) model (see Section 2.1). This ‘basic’ model is considered here in order to be evaluated in comparison with the other two novel approaches (Models A and B).

The architecture (network configuration) adopted is the same as in Models A and B, while a BP algorithm is used as a learning procedure.

Finally, all these Models A, B and C will be compared with the conventional approaches derived from micro-economic theory, such as the well-known logit model (Model D).

Model D

A widely adopted approach for modal split analysis is the logit model (see e.g. Ben Akiva and Lerman, 1985). Recent experiments using logit models/spatial interaction models to map the freight transport in Europe have been carried out by Tavasszy (1996), who also showed the suitability of logit models for the goods transport sector (where data are more “fuzzy” and incomplete compared to the passenger sector).

Logit models are discrete choice models used for modelling a choice from a set of mutually exclusive and exhaustive alternatives. It is assumed that the decision-maker chooses the alternative with the highest utility among the set of alternatives. The utility of an alternative is determined by a

¹² Another possibility should be of using the BP algorithm after the GA iteration. It means that after the GA learning phase a BP algorithm as a local search operator is computed (see Kitano, 1990).

utility function, which consists of independent attributes of the alternative concerned and the relevant parameters.

Since in our case two discrete choices – rail and cost – have been considered, a binary logit model is adopted. The variables “time” and “cost” between the 108 regions for the two transport modes have been considered in our empirical application (see also Reggiani *et al.*, 1998a).

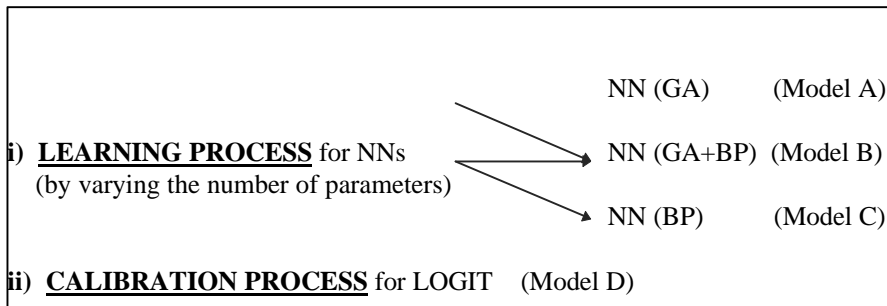
The results concerning the application of the models adopted to the European freight transport network will be presented and discussed in the next section.

4.4 The Results

4.4.1 The experiments

The structure of the experiments carried out in our empirical application is shown in Table 3. In order to compare Models A, B, C and D described in Section 4.3, we used the same data set (see Section 4.2) for the learning/calibration phase.

Table 3: The structure of the experiments



The learning phase incorporated in Models A and B showed the following results and findings:

Model A

- The best final result is obtained with high values of parameter α (reproduction parameter).
- The convergence to the final result improves with an increase of the population dimension in the GA.
- The final results seem to be better when the probabilities of crossover and mutation are 0.4 and 0.2, respectively.

Model B

- The GA+BP method – internal to the NN model – converges consistently faster than a ‘simple’ GA.

The results concerning the max-likelihood estimators - for the logit model - are shown in Table 4.

Table 4: The values of max-likelihood estimators for the logit model (Model C)

LOGIT MODEL			
Variable	Coefficient	Std. Error	t-ratio
Rail_time	0.49783E-0.3	0.4745E-0.3	1.049
Rail_cost	0.56234E-0.1	0.3251E-02	17.295
Road_time	-0.14995E-0.2	0.4944E-03	-3.033
Road_cost	-0.20208E-01	0.4692E-02	-4.307

For each Model A, B, C and D, the statistical indicators¹³ MSE, RMSE, R², ARV, MAPE have been considered to compare the results of the models adopted.

4.4.2 Comparative analysis of the models adopted

The results concerning the test phase of the adopted models are shown in Table 5. In order to calculate the values of the statistical indicator we have used a *test set* (see Section 4.2). The phase of model testing means to evaluate the performances of the four approaches under consideration by running for each model new observations (examples/patterns) for which the expected values are known.

Table 6 shows that the performance of the NN models combined with the GA (Models A and B) performs better than Models C and D. It is also noteworthy that Model A performs slightly better than Model B (but with higher computational time). Clearly, the results concerning these last Models C and D may also be considered as very good. There is also a ‘structural’ difference between the three ‘arche-types’ A, B and C in association with the NN model on the one hand and the logit model (Model D) on the other.

Table 5: The results of the testing phase after the calibration/learning phase for the adopted models

THE MODELS ADOPTED	MSE	RMSE	R ²	ARV	MAPE (%)
A) NN(GA) [*]	0.0386	0.1964	0.8456	0.1723	11.33
B) NN(GA+BP) ^{*,**}	0.0386	0.1964	0.8427	0.1741	11.34
C) NN(BP) ^{**}	0.0398	0.1995	0.8123	0.1927	12.11
D) LOGIT ^{***}	0.0464	0.2154	0.7507	0.2274	13.77

Note: ^{*}GA parameters:

crossover probability (p_c) = 0.4;
 mutation probability (p_m) = 0.2;
 dimension of the population = 30.

^{**}BP parameters:

learning rate (η) = 0.5;
 momentum factor (μ) = 0.

^{***} For the logit parameters see Table 4.

5. EMPIRICAL RESULTS FOR THE TRANSALPINE NETWORK

5.1 The Calibration/Learning Analysis

In order to better evaluate – for all our adopted models – the performance of our approach in a more specific setting, an interregional comparison of estimated data against the real data will be shown (see Tables 6-9). In this context, we have focused on the Transalpine area. Considering that the Alpine chain ‘virtually’ separates Europe from Greece and Italy, a closer comparison of these data has been carried out (see Tables 6 and 8). More precisely, Table 6 shows the estimated/real flows for the outflows from Greece and Italy towards the rest of Europe, while Table 8 illustrates the

¹³ For the definition of the statistical indicators MSE, RMSE, R², ARV, MAPE, see Annex A.

estimated/real flows for the inflows from Europe to Italy and Greece. We have also calculated (see Table 7 and Table 9):

- i) the relative prediction error from the real data of all Transalpine regions for each model A, B, C and D;
- ii) the mean value of the variations from the real data indicated by (M);
- iii) the mean of the absolute variations from the real data indicated by (MA).

The values illustrated in Tables 6-9 emerge from an estimation process conducted on a data set consisting of the sum of the learning and test set for all methodologies (see Section 4.2). It is evident from Tables 7 and 9 that the NN models combined with GAs (Models A and B) perform overall slightly better than the other models C and D (see the values of the relative prediction error for each link as well as the related mean values, M and MA), as already mentioned in the previous section.

In the next section we will illustrate a sensitivity analysis to investigate the results of the models under different policy scenarios on environmental costs. In this latter context we will display only the results related to models A [NN(GA)] and D (logit), since these two models are most representative from a methodological viewpoint (model A emerges from the most recent neurocomputing approaches, while model D derives essentially from micro-economic choice theory).

Table 6: Transalpine food transport flows by road from the Italian and Greek regions to the rest of Europe (flows in tons; year:1986)

Regions	Real flows	Model A NN(GA)	Model B NN(GA+BP)	Model C NN(BP)	Model D Logit
Thessaloniki	19343	18412	18457	18389	17141
Athens	24462	24341	25056	24695	23560
Patras	22251	20189	20300	20738	18240
Heraklion	17491	16374	16765	16175	14833
Turin	75349	74155	74030	75193	66045
Milan	305144	334545	330123	337618	296754
Venice	19715	24930	24621	25008	21203
Bologna	266350	263357	263215	263660	239286
Florence	28137	27465	27198	27341	26946
Ancona	24579	23987	23545	23314	22124
Pescara	29448	28357	28012	28364	24960
Rome	15864	15376	15750	15018	13624
Naples	18368	17119	17115	17161	14782
Bari	30197	30340	30432	30440	26576
Reggio C.	4435	4468	4434	4511	4129
Palermo	75710	76004	76005	76037	66934
Cagliari	6171	6387	6354	6545	6668

Table 7: The variations of the estimated values from the real data for the models adopted (transport flows by road from the Italian and Greek regions to the rest of Europe)

Regions	Model A Variation (%)	Model B Variation (%)	Model C Variation (%)	Model D Variation (%)
Thessaloniki	-4.81	-4.58	-4.93	-11.38
Athens	-0.49	2.43	0.95	-3.69
Patras	-9.27	-8.77	-6.80	-18.03
Heraklion	-6.39	-4.15	-7.52	-15.20
Turin	-1.58	-1.75	-0.21	-12.35
Milan	9.64	8.19	10.64	-2.75
Venice	26.45	24.88	26.85	7.54
Bologna	-1.12	-1.18	-1.01	-10.16
Florence	-2.39	-3.34	-2.83	-4.23
Ancona	-2.41	-4.21	-5.15	-9.99
Pescara	-3.70	-4.88	-3.68	-15.24
Rome	-3.08	-0.72	-5.33	-14.12
Naples	-6.80	-6.82	-6.57	-19.52
Bari	0.47	0.78	0.80	-11.99
Reggio C.	0.74	-0.02	1.71	-6.90
Palermo	0.39	0.39	0.43	-11.59
Cagliari	3.50	2.97	6.06	8.05
M	-0.05	-0.05	0.20	-8.91
MA	4.9	4.71	5.38	10.75

Note: M= the mean value of the variations from the real data.

MA= the mean of the absolute variations from the real data.

Table 8: Transalpine food transport flows by road from the rest of Europe to the Italian and Greek regions (flows in tons; year: 1986)

Regions	Real flows	Model A NN (GA)	Model B NN(GA+BP)	Model C NN (BP)	Model D Logit
Thessaloniki	44380	42144	42228	39696	42745
Athens	52047	49915	48461	45734	48798
Patras	53626	50382	49915	46839	51436
Heraklion	56930	54249	54133	50191	55948
Turin	275039	399557	397894	403826	372693
Milan	414190	415570	415184	421212	373309
Venice	53795	55026	55009	55709	46704
Bologna	365183	363457	361699	366398	338100
Florence	178632	186418	186016	186659	172325
Ancona	43653	42104	41871	41985	41413
Pescara	119774	114664	113206	112817	110799
Rome	35705	33358	33445	32993	31760
Naples	183553	193880	193657	190551	165469
Bari	105824	101171	98945	99434	90306
Reggio C.	29960	30768	30476	29554	28350
Palermo	126464	125287	123516	120470	108555
Cagliari	64435	62187	61560	60045	55122

Table 9: The variations of the estimated values from the real data for the models adopted (transport flows by road from the rest of Europe to the Italian and Greek regions)

Regions	<u>Model A</u> Variation (%)	<u>Model B</u> Variation (%)	<u>Model C</u> Variation (%)	<u>Model D</u> Variation (%)
Thessaloniki	-5.04	-4.85	-10.55	-3.68
Athens	-4.10	-6.89	-12.13	-6.24
Patras	-6.05	-6.92	-12.66	-4.08
Heraklion	-4.71	-4.91	-11.84	-1.72
Turin	45.27	44.67	46.82	35.51
Milan	0.33	0.24	1.70	-9.87
Venice	2.29	2.26	3.56	-13.18
Bologna	-0.47	-0.95	0.33	-7.42
Florence	4.36	4.13	4.49	-3.53
Ancona	-3.55	-4.08	-3.82	-5.13
Pescara	-4.27	-5.48	-5.81	-7.49
Rome	-6.57	-6.33	-7.60	-11.05
Naples	5.63	5.50	3.81	-9.85
Bari	-4.40	-6.50	-6.04	-14.66
Reggio C.	2.70	1.72	-1.36	-5.37
Palermo	-0.93	-2.33	-4.74	-14.16
Cagliari	-3.49	-4.46	-6.81	-14.45
M	1	0.28	-1.33	-5.67
MA	6.13	6.60	8.47	9.84

Note: M= the mean value of the variations from the real data.

MA= the mean of the absolute variations from the real data.

5.2 The Forecasting Analysis

It goes without saying that freight transport causes high social costs, which from an economic perspective would have to be charged to the transportation sector (see also Reggiani *et al.*, 1998a). We will now investigate the consequences of varying the transportation costs for freight flows. A sensitivity analysis of the previous results based on some economic scenarios will now be carried out in this section by again using both the NN (GA) model (Model A) and the binary logit model (Model D). We have ‘chosen’ Model A among the three categories of neurocomputing approaches since it offers the best performance in the learning phase (see Tables 5-9). Three policy scenarios based on different external costs assignments will be used in our experiment. It should be noted that the road mode represents the highest percentage of freight flows (82 %) in comparison to the rail mode. For this reason, as well as the well-known problems of congestion and environmental externalities of European road systems the policy scenarios have been focused on road transport mode.

In summary, we assume in all three scenarios that a uniform European tax policy for freight transport is adopted – in order to reduce, for example, congestion problems – and that the cost attribute related to the road mode is increased for all links (ij) to reduce road usage. In Scenario 1 the cost attribute is increased by a 5% increase and in Scenario 2 by 10%. If the high social costs of external effects (for example congestion, accidents, etc.) were to be included in the cost of road usage, the cost would significantly increase. Therefore, Scenario 3 is based on a ‘draconic’ scenario assuming the introduction of an ecotax 50%.

Tables 10-13 illustrate the results of the sensitivity analysis. More precisely, Table 10 and 12 display the results concerning the NN(GA) approach (Model A), for both flow directions, while Tables 11 and 13 illustrate the values for the Logit model (Model D), again for both flow directions. These tables indicate that the binary logit model is relatively more sensitive to changes in the cost attribute than the NN(GA) model, as in the results obtained from the previous experiments (see Reggiani *et al.*, 1998a).

5.3 Concluding Remarks

The conclusion from the above experiments is interesting in that the combined approach ‘EA-NN’ can certainly be considered as a valid tool for spatial forecasting (in the same year of the data used for the learning process). With temporal forecasting we should be more cautious, given the absence of recent data. Overall, however, the above results are plausible, even in the case of less updated information, particularly due to the ‘good’ results in the learning/calibration phase.

Table 10: The results of the sensitivity analysis for the model NN(GA) (transport flows by road from Italy+Greece @rest of Europe)

Model A - NN(GA) -							
REGIONS	Model A NN(GA)	Scenario 1 Cost+5%		Scenario 2 Cost+10%		Scenario 3 Cost+50%	
	Estimate flow	Estimate	Variation (%)	Estimate	Variation (%)	Estimate	Variation (%)
Thessaloniki	18412	18397	-0.08	18382	-0.16	18339	-0.40
Athens	24341	24323	-0.07	24305	-0.15	24244	-0.40
Patras	20189	20173	-0.08	20162	-0.13	20104	-0.42
Heraklion	16374	16361	-0.08	16352	-0.13	16297	-0.47
Turin	74155	74098	-0.08	74068	-0.12	73902	-0.34
Milan	334545	334346	-0.06	334224	-0.10	333583	-0.29
Venice	24930	24918	-0.05	24908	-0.09	24863	-0.27
Bologna	263357	263154	-0.08	263021	-0.13	262355	-0.38
Florence	27465	27453	-0.04	27442	-0.08	27374	-0.33
Ancona	23987	23967	-0.08	23957	-0.13	23898	-0.37
Pescara	28357	28341	-0.06	28326	-0.11	28267	-0.32
Rome	15376	15360	-0.10	15353	-0.15	15315	-0.40
Naples	17119	17106	-0.08	17097	-0.13	17055	-0.37
Bari	30340	30315	-0.08	30297	-0.14	30218	-0.40
Reggio C.	4468	4464	-0.09	4461	-0.16	4447	-0.47
Palermo	76004	75952	-0.07	75922	-0.11	75725	-0.37
Cagliari	6387	6385	-0.03	6381	-0.09	6365	-0.34
M			-0.07		-0.12		-0.37

Note: M= the mean value of the variations from the estimate value.

Table 11: The results of the sensitivity analysis for the Logit Model (transport flows by road from Italy+Greece@rest of Europe)

Model D - LOGIT -							
REGIONS	Model D LOGIT	Scenario 1 Cost+5%		Scenario 2 Cost+10%		Scenario 3 Cost+50%	
	Estimate flow	Estimate	Variation (%)	Estimate	Variation (%)	Estimate	Variation (%)
Thessaloniki	17141	16807	-1.95	16456	-4.00	13036	-23.95
Athens	23560	22998	-2.39	22404	-4.91	16596	-29.56
Patras	18240	17784	-2.50	17312	-5.09	13224	-27.50
Heraklion	14843	14463	-2.56	14062	-5.26	10255	-30.91
Turin	66045	65284	-1.15	64494	-2.35	57126	-13.50
Milan	296754	293263	-1.18	289643	-2.40	256275	-13.64
Venice	21203	20868	-1.58	20521	-3.22	17366	-18.10
Bologna	239286	236959	-0.97	234551	-1.98	212382	-11.24
Florence	26946	26651	-1.09	26342	-2.24	23365	-13.29
Ancona	22124	20695	-6.46	20420	-7.70	17817	-19.47
Pescara	24960	24547	-1.65	24116	-3.38	20078	-19.56
Rome	13624	13435	-1.39	13238	-2.83	11359	-16.63
Naples	14782	14457	-2.20	14122	-4.46	11178	-24.38
Bari	26576	25893	-2.57	25185	-5.23	18874	-28.98
Reggio C.	4129	4042	-2.11	3951	-4.31	3074	-25.55
Palermo	66934	65270	-2.49	63539	-5.07	47778	-28.62
Cagliari	6668	6593	-1.12	6513	-2.32	5692	-14.64
M			-2.08		-3.93		-21.15

Note: M= the mean value of the variations from the estimate value.

Table 12: The results of the sensitivity analysis for the model NN(GA)
(transport flows by road from the rest of Europe @ Italy+Greece)

Model A - NN(GA) -							
REGIONS	Model A NN(GA)	Scenario 1 Cost+5%		Scenario 2 Cost+10%		Scenario 3 Cost+50%	
	Estimate flow	Estimate	Variation (%)	Estimate	Variation (%)	Estimate	Variation (%)
Thessaloniki	42144	42077	-0.16	42046	-0.23	41824	-0.76
Athens	49915	49856	-0.12	49814	-0.20	49502	-0.83
Patras	50382	50344	-0.08	50307	-0.15	50130	-0.50
Heraklion	54249	54186	-0.12	54146	-0.19	53935	-0.58
Turin	399557	399137	-0.11	398889	-0.17	398036	-0.38
Milan	415570	415184	-0.09	414894	-0.16	412906	-0.64
Venice	55026	54968	-0.11	54930	-0.17	54774	-0.46
Bologna	363457	363211	-0.07	362919	-0.15	361561	-0.52
Florence	186418	186260	-0.08	186135	-0.15	185313	-0.59
Ancona	42104	42073	-0.07	42042	-0.15	41833	-0.64
Pescara	114664	114564	-0.09	114480	-0.16	113893	-0.67
Rome	33358	33334	-0.07	33309	-0.15	33131	-0.68
Naples	193880	193740	-0.07	193612	-0.14	192749	-0.58
Bari	101171	101083	-0.09	101041	-0.13	100427	-0.74
Reggio C.	30768	30745	-0.07	30730	-0.12	30583	-0.60
Palermo	125287	125153	-0.11	125070	-0.17	124680	-0.48
Cagliari	62187	62131	-0.09	62086	-0.16	61812	-0.60
M			-0.093		-0.16		-0.60

Note: M= the mean value of the variations from the estimate value.

Table 13: The results of the sensitivity analysis for the Logit model (transport flows by road from the rest of Europe @ Italy+Greece)

Model D - LOGIT -							
REGIONS	Model D LOGIT	Scenario 1 Cost+5%		Scenario 2 Cost+10%		Scenario 3 Cost+50%	
	Estimate flow	Estimate	Variation (%)	Estimate	Variation (%)	Estimate	Variation (%)
Thessaloniki	42745	42388	-0.84	41933	-1.90	37095	-12.95
Athens	48798	48206	-1.21	47546	-2.57	39357	-10.51
Patras	51436	50989	-0.87	50486	-1.85	43856	-11.10
Heraklion	55948	55537	-0.73	55072	-1.57	48701	-14.07
Turin	372693	369475	-0.86	366111	-1.77	333533	-9.21
Milan	373309	369740	-0.96	366044	-1.95	331866	-12.81
Venice	46704	46126	-1.24	45530	-2.51	40133	-8.52
Bologna	338100	335494	-0.77	332780	-1.57	306963	-12.76
Florence	172325	170437	-1.10	168467	-2.24	150256	-13.08
Ancona	41413	41133	-0.68	40838	-1.39	37885	-22.84
Pescara	110799	109726	-0.97	108582	-2.00	96663	-18.40
Rome	31760	31432	-1.03	31085	-2.13	27607	-21.37
Naples	165469	162287	-1.92	158956	-3.94	127675	-24.46
Bari	90306	88966	-1.48	87544	-3.06	73693	-17.63
Reggio C.	28350	27867	-1.70	27354	-3.51	22293	-15.12
Palermo	108555	106366	-2.02	104060	-4.14	82007	-13.22
Cagliari	55122	54322	-1.45	53483	-2.97	45405	-19.35
M			-1.17		-2.41		-14.74

Note: M= the mean value of the variations from the estimate value.

6. CONCLUSIONS AND NEW RESEARCH DIRECTIONS

We have addressed the use of evolutionary computation (and particularly of GAs combined with NNs) in order to measure ‘evolutionary activity’, i.e. the ‘spontaneous generation of innovative functional structures’ (see Bedan and Packard 1992). Like connectionism (i.e. the study of computer programmes inspired by neural systems), evolutionary computation is a ‘bottom-up’ paradigm in which humans write only very simple rules, while complex behaviour emerges from the massive parallel application and interaction of these simple rules. However, while in connectionism these rules are typically simple ‘neural’ thresholding, i.e., the activation and strength of connections, in evolutionary computation the rules are ‘natural selection’ with variation due to crossover and/or

mutation (see again, Mitchell 1996). Evolution is thus seen as a method for designing innovative solutions to complex problems and inspiring computational search methods based on the simple rules in which the 'fittest' solution tends to survive and reproduce.

Clearly, from a social science perspective, it also be interesting to investigate whether the selection process described above can be interpreted in terms of a utility maximisation process (or, in general, as a behavioural paradigm; see e.g., Ben-Akiva and Lerman, 1985) and hence to undertake further research on the theoretical compatibility between EAs and utility maximising (behavioural) models (such as logit models), as well as between EAs and NNs, particularly in the light of recent studies also offering a 'behavioural' framework for NNs (see Section 3.1). Fischer and Leung (1998), in particular, argue that "*Neural spatial interaction models are termed neural in the sense that they have been inspired by neuroscience. But they are more closely related to conventional spatial interaction models of the gravity type than they are to neurobiological models. They are special classes of general feedforward neural network models...*"

In this context it would be not surprising if even EAs were to show a 'compatibility' with spatial interaction models (and consequently with logit models), if we investigated under which conditions these 'conventional' models can be considered as a 'powerful' class of universal approximators for spatial/social interaction. Consequently, EAs may offer another interesting conceptual research question: can natural selection be interpreted in the framework of economic utility theory? Apart from further theoretical/methodological research, this would also require more rigorous empirical tests. We leave such theoretical/empirical experiments on this subject open for future research.

ACKNOWLEDGMENTS

The first author gratefully acknowledges the Italian CNR project No. 97.000264.PF74. The third author thanks the Faculty of Statistics (University of Bologna) and particularly its dean, Prof. Paola Monari, for the fellowship assigned to him. The authors also wish to thank Dr. Vittorio Maniezzo and Dr. Giovanni Russo for their kindly providing information.

REFERENCES

- Batten D., Casti J. and Thord R. (eds.) *Networks in Action*, Springer-Verlag, Berlin, 1994.
- Bedan M. A. and N. H. Packard, "Measurement of Evolutionary Activity, Teleology and Life", *Artificial Life II* (C.G. Langton, C.Taylor, J.D. Farmer, and S.R. Rasmussen ,eds.), Addison Wesley, Reading, MA, 1991.
- Ben-Akiva M. and S.R. Lerman, *Discrete Choice Analysis. Theory and Application to Travel Demand*, MIT Press, Cambridge, Massachussets, 1985.
- Bertoni A. and M. Dorigo, *Implicit Parallelism in Genetic Algorithms*, Technical Report, 92-102, Dipartimento di Elettronica, Politecnico di Milano, 1992.
- Colomi A., M. Dorigo and V. Maniezzo, "ALGODESK: An Experimental Comparison of Eight Evolutionary Heuristic Applied to the AQP Problem", *European Journal of Operational Research*, 1992a..
- Colomi A., M. Dorigo, and V. Maniezzo, "Scheduling School Teachers by Genetic Algorithms", *Combinatorial Optimization, New Frontiers in Theory and Practice* (M. Akgul, H.W. Hamacher and S. Tfekci, eds.), NATO ASI Series, Series F, Vol. 82, Springer-Verlag, Berlin, 1992b.
- Colomi A., M. Dorigo, and V. Maniezzo, "Introduzione agli Algoritmi Naturali", *Rivista di Informatica*, 3, 1994, pp. 179-197.
- De Jong K.A., *Analysis of the Behavior of a Class of Genetic Adaptive System*, Ph.D. dissertation, Univ. of Michigan, 1975.

- Domanski R., "The Changeability of Man-Environment Systems far from Equilibrium", *Space Economy in the Transition* (R. Domanski, ed.), Polish Academy of Science, Warsaw, 1992, pp. 85-104.
- Domanski R., "Transformation of Urban Systems in Terms of Synergetics", *Changes in the Regional Economy in the Period of System Transformation* (R. Domanski and E. Judge, eds.), Polish Academy of Science, Warsaw, 1994, pp. 11-23.
- Dorigo M. and V. Maniezzo, "Parallel Genetic Algorithms: Introduction and Overview of Current Research", *Parallel Genetic Algorithms: Theory and Applications* (J. Stenders, ed.), IOS Press, Amsterdam, 1992.
- Fischer M.M., "Computational Neural Networks", *Neural Networks in Transport Applications* (V. Himanen, P. Nijkamp and A. Reggiani, eds.), Ashgate, Aldershot, 1998, pp.3-20.
- Fischer M.M. and S. Gopal, "Artificial Neural Networks: a New Approach to Modelling Interregional Telecommunication Flows", *Journal of Regional Science*, Vol. 34, No. 4, 1994, pp. 503-527.
- Fischer M.M. and Y. Leung, "A Genetic-Algorithms Based Evolutionary Computational Neural Network for Modelling Spatial Interaction Data", *The Annals of Regional Science*, Vol. 32, No. 3, pp.437-458, 1998.
- Goldberg D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, MA, 1989.
- Haken, H., *Synergetics*, Springer-Verlag, Berlin, 1983a.
- Haken, H., *Advanced Synergetics*, Springer-Verlag, Berlin, 1983b.
- Hertz, J., A. Krogh and R.G. Palmer, *Introduction to the Theory of Neural Computation*, Addison Wesley, Redwood City, CA, 1991.
- Himanen V., P. Nijkamp, A. Reggiani (eds.), *Artificial Neural Networks in Transport Applications*, Ashgate, Aldershot, 1998.
- Holland J.H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.
- Kitano H, "Empirical Studies on the Speed of Convergence of Neural Network Training Using Genetic Algorithms", *Proceedings of the Eight National conference on Artificial Intelligence*, 1990, pp. 789-795.
- Kosko B., *Neural Networks and Fuzzy Systems*, Prentice Hall Inc., Englewood Cliffs, N.J., 1992.
- Malliaris M. and L. Salchenberger, "A Neural Network Model for Estimating Option Prices", (mimeo), 1992.
- Maniezzo V., "Genetic Evolution of the Topology and Weight Distribution of Neural Networks", *Proceedings of the IEEE International Conference on Transactions on Neural Networks*, Vol. 5, No. 1, IEEE Computer Society Press, 1994, pp. 39-53.
- Maren C., A. Harston. and R. Pap, *Handbook of Neural Computing Application*, Academic Press, San Diego, 1990.
- Mitchell M., *An Introduction to Genetic Algorithms*, The MIT Press, Cambridge, MA, 1996.
- Muhlenbein H., "Parallel Genetic Algorithms, Population Genetic and Combinatorial Optimization", *Proceedings of the Third International Conference on Genetic Algorithms* (J.D. Schaffer, ed.), Morgan Kaufmann, San Mateo, CA, 1989, pp. 416-421.
- Nijkamp P. (ed.), *Sustainability of Urban Systems*, Avebury, Aldershot, UK, 1991.
- Nijkamp P. and A. Reggiani, "Modelling Network Synergy: Static and Dynamic Aspects", *Journal of Scientific & Industrial Research*, Vol. 55, No. 5, 1996, pp. 931-941.
- Nijkamp P. and A. Reggiani, *The Economics of Complex Spatial Systems*, Elsevier, Amsterdam, 1998.
- Nijkamp P., A. Reggiani and T. Tritapepe, "Modelling Inter-Urban Flows in Italy: a Comparison between Neural Network Approach and Logit Analysis", *Transportation Research C*, Vol. 4., No. 6, 1996, pp. 323-338.
- Nijkamp P., A. Reggiani and T. Tritapepe, "Analysis of Complex Networks: An Overview of Methodologies and a Neural Network Application to Intermodal Transport in Italy", *Policy Aspects of Networks* (C. Capineri and P. Rietveld, eds.), Ashgate, Aldershot, 1997, pp. 285-305.
- Reggiani A. and P. Nijkamp, "Towards a Science of Complexity in Spatial Economic Systems", *Recent Advances in Spatial Equilibrium Modelling* (J.C.M. van den Bergh, P. Nijkamp and P. Rietveld, eds.), Springer Verlag, Berlin, 1996a, pp. 359-378.
- Reggiani A. and P. Nijkamp, "Modelling Network Synergy in a Sustainable Perspective", *Paper Presented at The International I.G.U. Conference*, The Hague, 1996b.

Reggiani A., P. Nijkamp and E. Sabella, "New Advances in Spatial Network Modelling: Towards Evolutionary algorithms", *European Journal of Operational Research*, 1999 (forthcoming).

Reggiani A., P. Nijkamp and W.-F. Tsang, "European Freight Transport Analysis Using Neural Networks and Logit Models", *Accessibility, Trade and Locational Behaviour* (A. Reggiani, ed.), Ashgate, Aldershot, 1998a, pp. 355-373.

Reggiani A., R. Romanelli, T. Tritapepe and P. Nijkamp, "Neural Networks: an Overview and Applications in the Space-Economy", *Neural Networks in Transport Applications* (V. Himanen, P. Nijkamp and A. Reggiani, eds.), Ashgate, Aldershot, 1998b.

Rumelhart D.E., G.E. Hinton and R.J. Williams, "Learning Internal Representation by Error Propagation", *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (D.E. Rumelhart and J.L. McClelland, eds.), MIT Press, Cambridge, MA, 1986.

Schintler L.A. and O. Olurotimi, "Neural Networks as Adaptive Logit Models", *Neural Networks in Transport Applications* (V. Himanen, P. Nijkamp and A. Reggiani, eds.), Ashgate, Aldershot, 1998, pp. 131-160.

Tavasszy E.T., *The Economics of Regulating Road Transport*, Aldershot: Edward Elgar, 1996.

Vitetta A., "The Urban Network Design Problem Solved through Genetic Algorithms", *Proceedings of the Second Meeting of the EURO Working Group on Urban Traffic and Transportation*, Paris, 1993.

Wegener M. and E. Graef, *Estimating Multiattribute Spatial Choice Models*, User's Manual, University of Dortmund, 1982.

Weidlich W. and G. Haag, *Concepts and Methods of a Quantitative Sociology*, Springer-Verlag, Berlin, 1983.

ANNEX A. Definition of Test Statistics

The Average Relative Variance (ARV)

The statistical indicator ARV(N) is to compare the estimates calculated from the NN and the Logit model on the basis of the data set of N data (see Fischer, 1998) and it is defined as:

$$ARV(N) = \frac{\sum_{i \in N} (y_i - \hat{y}_i)^2}{\sum_{i \in N} (y_i - \bar{y})^2} \quad (a.1)$$

where y_i is the target/real value, \hat{y}_i = predicted value by the adopted model, \bar{y} = the average of the target/real values belonging to the set of data N. Assuming that all the events are 'equiprobable' the relationship (a.1) becomes :

$$ARV(N) = \frac{1}{\hat{s}^2} \frac{1}{M_N} \sum_{i \in N} (y_i - \hat{y}_i)^2 \quad (a.2)$$

In the relation (a.2), M_N is a number of samples belonging to the set N and \hat{s}^2 is the variance of the expected value of entire data set. ARV(N) will tend to zero, the more estimate is exact.

The Correlation Coefficient (R^2)

The statistical indicator R^2 (N) is defined in analogous manner of the ARV(N):

$$R^2(N) = \frac{1}{\hat{s}^2} \frac{1}{M_N} \sum_{i \in N} (\bar{y} - \hat{y}_i)^2 \quad (a.3)$$

where the variables are defined as in equation (a.1-a.2). R^2 will tend to one, the more estimate is exact.

The Mean Squared Error (MSE)

The MSE indicator is applied for all estimates and is independent of the underlying methodology. In fact, it is specified as average (where s denote the number of the patterns) of the squared difference between the observed/real values y_i and the predicted values \hat{y}_i :

$$MSE = \frac{1}{s} \sum_{i=1}^s (y_i - \hat{y}_i)^2 \quad (a.4)$$

This indicator denotes a good performance of the calibrated model when its value is approaching zero.

The Root Mean Squared Error (RMSE)

This indicator is obtained as the root of the MSE indicator (see equation (a.4)):

$$RMSE = \sqrt{\frac{1}{s} \sum_{i=1}^s (y_i - \hat{y}_i)^2} \quad (a.5)$$

The Mean Absolute Percentage Error (MAPE)

The analytical expression of this indicator is the following (see equation (a.4) for the specification of the variables):

$$MAPE = \frac{1}{s} \sum_{i=1}^s \frac{|y_i - \hat{y}_i|}{y_i} \% \quad (a.6)$$

According to some authors (see e.g. Wegener and Graef, 1982), the MAPE values may be interpreted as follows:

- $MAPE < 10\%$; \Rightarrow extremely good forecast
- $10\% < MAPE < 20\%$; \Rightarrow good forecast
- $20\% < MAPE < 50\%$; \Rightarrow reasonable forecast
- $MAPE > 50\%$. \Rightarrow bad forecast