SUSTAINABLE TRANSPORTATION – SYSTEM DIAGNOSIS USING NEURAL NETWORKS

By Dalve A. Soria ALVES[**], Koshi YAMAMOTO[***], Yaeko YAMASHITA[***] and Eizo HIDESHIMA[****].

1 – Introduction

The last century was known for technological advances. One of the main developments was that of the automobile industry, a response to the problems of urban displacement. Important achievements were made in this field, allowing a degree of urban growth unthinkable years before. At the same time these benefits were accompanied by disadvantages such as traffic jams and air pollution which have compromised the urban quality of life.

For this reason, international organizations started to consider environmental problems based on the implementation of a Sustainable Development actions concerning social, environmental and economical interests. Experience, however, has shown that no Sustainable Development is possible without Sustainable Transportation, i.e., improving transport's benefits while reducing its environmental impact to sustainable levels [1)]. Hence governmental policies on Sustainable Transportation have focused on reorganising the system while integrating it into the urban planning process, and trying to minimize the problem of vehicular pollution by various means. Consequently, instruments such as monitoring stations have been used to measure pollution and its impact on the environment. Such activity, however, is very costly and the complexity of the pollution dispersion in the atmosphere has made it, so far, impossible to achieve a precise diagnosis of its origins within the urban system. Most research conducted, which has tried not only to model air pollution behaviour from mobile sources but mainly its association with other components of the urban and transportation system, has gotten rarely good results.

However, technological advances and the development of computational intelligence systems have helped problems related to urban and transportation planning. In particular, the use of tools such as Neural Networks (NN) has made possible the representation of the non-linear relationship between variables derived from both urban space and transportation systems, such as land use, and those variables' relation to the Pollution Indexes. This aids in the solution of such complex problems as how to model air pollution in an urban context, something very difficult to do by traditional statistics methods. It is also believed that NN could contribute actively to the strategic planning process at all stages from data collection to the systemic evaluation of the implementation of Urban and Transportation Sustainability policies.

Thus, the purpose of this research is to present the 'System Diagnosis Step' based on the development of a Neural Network Model for Air Pollution (NNM-AP) as contributing to urban and transportation sustainable development within the concept of Strategic Planning. The NNM-AP tries to analyse how some urban, transportation, and meteorological intervening factors affect vehicle emission dispersion by establishing non-linear relationships between variables.

This paper initially discusses the concepts of Sustainable Transportation and Strategic Planning, and presents the 'System's Diagnosis Step' and the intervening factors of vehicle emissions in the urban context. Next, the description of the NNM-AP mathematical formulation on the modelling of vehicle emissions is presented. Based on this model a case study has been conducted in Nagoya City, Japan, in order to evaluate its efficiency. At the end of the paper, the results of the study are discussed and some proposals are made for future improvements on the model

2 – Sustainable Transportation and Strategic Planning

The idea of Sustainable Transportation could be understood as a particular extension of international initiatives to react against the continuous process of degradation of the urban quality of life, systematized by actions of Sustainable Development in order to secure the quality of life of future generations. Such actions, however, cannot be disassociated from interventions in the transportation system, one of the main agents related to urban activity, in an effort to improve urban displacements while minimizing the environmental impact. Thus, based on the principles of Sustainable Development, discussed at great length at the United Nations Conference on Environment and Development (UNCED) held in Rio de Janeiro in 1992, Sustainable Transportation was defined by the Centre for Sustainable Transportation as transportation which, "allows the basic needs of individuals and societies to be met safely and in a manner consistent with human and ecosystem health (...); is affordable, operates efficiently, offers choice of transport mode, and supports a vibrant economy, and; limits emissions and waste within the planet's ability to absorb them (...)"[2)].

On the other hand, the practical realization of those goals is only possible through a general and systemic comprehension of actual conditions that determines strategic policies which could be implemented in different time periods, i.e. within the Strategic Planning Process. Hence Strategic Planning could be briefly defined as a process which provides planners with policy options and implementation strategies to help them make informed decisions and strategic agreements[3)]. Linking the idea of Sustainable Transportation to the Strategic Planning Process means allowing planners to choose from among different scenarios according to their principles based on urban and transportation sources, such as, land use, transportation system and pollution indexes. It could be understood methodologically through five cyclical stages described below: 1)*System's Diagnosis*

---

[*] Keywords: Sustainable Transportation, Air Pollution, Neural Network
[**] Student Non-Member of JSCE, BSc, Dept. of Civil Eng., Nagoya Institute of Technology
(Nagoya Institute of Tech., Gokiso-cho, Showa-ku, Nagoya, Japan, TEL.FAX. 052-735-5484, dalve@keik1.ace.nitech.ac.jp)
[***] Member of JSCE, Dr. Eng., Dept. of Civil Eng., Nagoya institute of Technology (e-mail yama@doboku2.ace.nitech.ac.jp)
[***] Non-Member of JSCE, PhD, Dept. of Civil Eng., University of Brasilia (e-mail yaeko@unb.br)
[****] Member of JSCE, Dr. Eng., Dept. of Civil Eng., Nagoya institute of Technology (e-mail hideshima@ace.nitech.ac.jp)

and Monitoring – Data collection through different means (questionnaire, monitoring, remote sensing, etc.); system diagnosis based on the relation of intervening variables; 2) *Evaluation and Formulation of Strategy* – Establishing principles for urban and transportation planning policies based on the system diagnosis; 3)*Future Scenarios Evaluation* – Creating and evaluating future scenarios based on the defined principles; 4) *Decisions* – Choosing the most appropriate scenario based on the defined principles; and, 5) *Implementation, Management and Control*.

In this paper the System Diagnosis step will be analysed.

2.1. – System Diagnosis Step

The concept of the NNM-AP is to establish a diagnosis of air pollution derived from mobile sources in the urban system, as the first stage in Strategic Planning on the way to develop a more sustainable city. The model will take into account the influence of urban dynamics such as land use and the transportation system on demand for travel and consequently on the dispersion of air pollutants.

The idea of urban dynamics could be understood as the way in which the population displaces within the urban area in between different patterns of land use in order to attend their economic or personal needs. Most urban environmental problems, however, are a consequence of the way these displacements are effected. A high volume of cars and buses on the roads, an aged fleet and long traffic jams are some of the main sources of pollution derived from mobile sources. With this in mind, studies have been trying to establish a relation between land use, transportation systems, traffic volume and pollution in order to minimize the effect of vehicles on the environment. Most of them have, however, been inefficient. The complex interaction of air pollution variables, which depend on various exogenous factors, makes accurate measurement difficult to achieve.

Against this background, Neural Networks (NN) modelling attempts to fill the gaps in the traditional studies of urban transportation and environment. NN has come to be thought of as, “a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use”[4], i.e., it works as a powerful instrument of parallel processing with various degrees of liberty, which is able to learn and to generalize the “acquired knowledge” and so become a dexterous and flexible modelling tool.

The main purpose of this research is to model the non-linear relation between urban, transportation and environmental conditions, using the NNM-AP, in order to evaluate its possibility of use as an instrument of diagnosis of urban system. This is the first step in defining the aims of the strategic planning process with the intention of developing urban and transportation sustainability.

3 – Definition of the Neural Network Model

Thus, in NNM-AP, Air Pollution (*AP*) is obtained from a non-linear relation between Land Use patterns (*LU*), Transportation System (*TS*), Spatial Location (*SL*), Demography Conditions (*DC*) and Environmental Conditions (*EC*) as defined in equation (1).

$$AP = f\left(LU;TS;SL;DC;EC\right) \tag{1}$$

where, *f* is a non-linear function establishing a set of weights *W* between the dependent variables (*LU, TS, SL, DC and EC*). These weights are determined through NN computation based on a Multi-Layer Perceptron (MLP) architecture. Similar to standard regression modelling, independent (input vector *I*) and dependent (output vector *O*) variables have to be defined in MLP architecture. Mathematically, an urban area, which is divided into *nz* unit of analysis (zones, macro zones, traffic zones, census borders, etc). For each unit of analysis *I* (zone), a vector $\overrightarrow{I_i}$ is defined as shown in equation (2).

$$\overrightarrow{I_i} = \left\{\overrightarrow{TS}_i, \overrightarrow{LU}_i, \overrightarrow{SL}_i, \overrightarrow{DC}_i, \overrightarrow{EC}_i\right\} \tag{2}$$

where: $\overrightarrow{TS}_i$ vector containing transportation system features of *i*; $\overrightarrow{LU_i}$ vector representing land use characteristics of *i*; $\overrightarrow{SL_i^t}$ vector for describing spatial location of *i*; $\overrightarrow{DC_i^t}$ vector with demographic conditions of *i*; and $\overrightarrow{EC_i^t}$ vector with environmental conditions of *i*. On the other hand, AP_i expressing vehicular pollution of the unit of analysis *i* is assigned to the output vector $\overrightarrow{O_i}$.

In order to obtain the set of weights *W*, a training algorithm might be selected. In this specific case of NNM for air pollution modelling, a well-known approach is the adoption of a back-propagation algorithm, which is based on the error-correction learning rule.

Using training data sets, training process might be conducted until the NN reaches the convergence status. This happens if after a number *q* of iterations, additional increments on $\Delta w_{jh}(q)$ do not compute considerable improvement on the results ($y_j(q)$). Based upon the weights *W* obtained for the trained NN, testing has to be conducted in order to evaluate the generalization power of the modelling function. One of the measurements that is usually employed is the Minimum Square Error (MSE), which is computed as presented in equation (3).

$$MSE = \sum_{i=1}^{nt} \sum_{j=1} \left(d'_j - y'_j\right)^2 \Big/ nt \tag{3}$$

where: *nt* is the number of samples in the testing data set; d'_j is the desired output in the testing data set for neuron *j*; and y'_j is the calculated output in the testing data set for neuron *j*.

4 – Case Study

The case study was conducted in Nagoya City, Japan. In 1991, when the statistics used here were gathered, the population was about 2,000,000 distributed over 326.35 Km^2. Nagoya has a radial and concentric urban structure divided into 16 wards and 248 traffic zones related to the predominant land use. The transportation system is formed by the interaction of the approximately 1,003 km of roads and 75 km of urban express highways and the public system, creating a high dependency on cars (around 70%). In order to control the pollution dispersed in the air, 26 monitoring station have been established at important spots, roadsides and rivers, and densities of different types of pollutant are constantly measured.

Concerning the data used for this research, the land use, transportation system and demographics were obtained from Nagoya City Bureau, while pollution, wind direction and velocity were obtained from Nagoya City Environmental Affairs Bureau. In order to measure the pollution dispersion in function of distance, the distance of the centre of each traffic zone to the closest monitoring station was measured using a Geographical Information System (GIS) platform. The pollutant used for the modelling was NO_x, collected in almost all stations, and the general average emission for vehicular source in Nagoya is around 65%, varying according to the ward's main land use.

4.1. – NN simulations for obtainment of NNM-AP

Activities of this phase of the application of NNM-AP comprise the creation of input and output vectors through pre-processing of data, definition of training and testing data sets, training and testing.

Initially, transportation system vector $\overrightarrow{TS}_i$ is defined as shown in equation (4).

$$\overrightarrow{TS}_i = \{PT_i, RT_i, ET_i\} \tag{4}$$

where: PT_i is the total extension (Km) of public transportation for zone *i; RT_i* is the total extension (Km) of road transportation for zone *i*; and, ET_i represents the existence of an express railway in the zone *i*.

Next, we defined the indexes related to land use patterns in the equation (5).

$$\overrightarrow{LU}_i = \{CL_i\} \tag{5}$$

where: CL_i is the occupied area (m^2) of commercial pattern for zone *i;*

Spatial location vector ($\overrightarrow{SL}_i$) is expressed by SD_i representing the distance (Km) from zone *i* to Sakae's TV Tower which is the main reference point located in the city centre as shown in equation (6).

$$\overrightarrow{SL}_i = \{SD_i\} \tag{6}$$

Next, as defined before the environmental conditions $\overrightarrow{EC}_i$ as presented by equation (7).

$$\overrightarrow{EC}_i = \{WD_i, WV_i, CD_i\} \tag{7}$$

where: WD_i is the wind predominant direction for zone *i; WV_i* is the average of the wind velocity (m/s) for zone *i*; and, CD_i is the distance between the centroid of the zone *i* and the closest monitoring station (Km).

Finally, $\overrightarrow{DC}_i$ was defined as the number of inhabitants in the zone *i*.

Then, once $\vec{I}_i$ was composed for all traffic zones, it was normalized for each of its component ς, in order to fit original values such as areas, extensions, etc into a limited interval. The interval between 0.1 and 0.9 was used by applying the following equation (8):

$$\overline{I_i}[\varsigma] = 0.1 + 0.8(I_i[\varsigma] - I_{\min}[\varsigma])(I_{\max}[\varsigma] - I_{\min}[\varsigma])^{-1} \tag{8}$$

where: $\overline{I_i}[\varsigma]$ is the normalized value of $I_i[\varsigma]$ for component ς; $\overline{I}_{\max}[\varsigma]$ is the maximum value of $I_i[\varsigma]$ for component ς, inside $\overrightarrow{IT_i}$*;* and $\overline{I}_{\min}[\varsigma]$ is the minimum value of $I_i[\varsigma]$ for component ς, inside $\overrightarrow{IT_i}$.

On the other hand, similar to the normalization procedure which was applied in equation 8, AP_i was processed by using equation (9).

$$\overline{AP_i} = 0.1 + 0.8(AP_i - AP_{\min})(AP_{\max} - AP_{\min})^{-1} \tag{9}$$

where: $\overline{AP_i}$ is the normalized value of AP_i inside $\overrightarrow{AT_i}$; AP_{max} is the maximum value of AP_i inside $\overrightarrow{AT_i}$; and AP_{min} is the minimum value of AP_i inside $\overrightarrow{AT_i}$.

Next, the NN was used to perform simulations towards the obtainment of modelling function capable to calculate the concentration of NO_x (AP) per zone of the year 1991 considering the vectors presented previously. In this sense, $\overline{I}[\varsigma]$ and $\overline{AP_i}$ were associated to $\vec{Y}$ $\vec{X}$ input and $\vec{Y}$ output vectors. Normalized vectors were randomly divided into training and test data sets. A distribution of 75% (186 vectors) and 25% (62 vectors) was used to obtain vectors $\vec{X}'$ $\vec{Y}'$ for training and $\vec{X}''$ input and $\vec{Y}''$ output for testing. Figure 1a shows the spatial distributions of training and test-related traffic zones.

In the sequence, a three-layer NN structure was established. It was configured with 15-input, 15-hidden and 1-output processing units. Applying a back-propagation algorithm with a learning rate of 0.001 (η =0.001) and using sigmoid activation functions (α=∞), the network was trained until the minimum *MSE* in the test set was reached. The most accurate results are shown in Table 1 and Figure 1b. In the Table 1, it can be verified that after 3,885,912 iterations the weights were defined as the best optimisation. The application of the trained NN on a testing data set ($\vec{X}''$ and $\vec{Y}''$) generated a maximum positive error in 43% and maximum negative error in -27%. It is also observed that the average error per zone (ξ), as defined in equation (10), 61 air pollution indexes per zone was reached.

$$\xi = \left(\sum_{i=1}^{nt} |AP_i'' - Y_i''|\right) \Big/ nt \tag{10}$$

where: AP_i" is the desired air pollution indexes for zone *i* on the testing data set; Y_i" is the measured air pollution indexes for testing data zone *i*; and *nt* is the total number of vectors for testing.

Figure 1b shows the relative error ξR_i for each zone i as defined by equation (11). The values of AP_i" and Y_i" were divided into five classes of results, which are distributed as presented in Table 2.

$$\xi R_1 = \left(AP_i'' - Y_i''\right) / AP_i'' \tag{11}$$

Table 1: Best Test results.

	Max positive error (%)	Max negative error (%)	Average error (%)	*MSE*
Test	43	-27	14	1.72

Table 2: Distribution of Air Pollution Indexes

	Classes of Results (air pollution indexes)				
	[49;61[	[61;73[	[73;85[	[85;97[	[97;109]
AP_i"	25	16	7	9	5
Y_i"	18	33	6	0	5

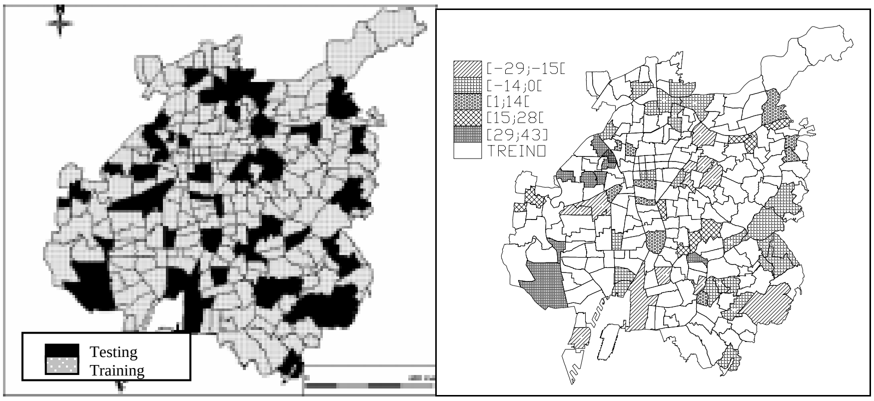

Figure 1a: Distribution of zones for testing and training; Figure 1b: Relative errors ξR_1(%).

5 - Conclusion

The development of the NNM-AP aims to contribute to the process of strategic planning within a sustainable transportation policy in different ways. Firstly, it works as an important and flexible instrument of diagnoses for the first stage of the strategic planning process, showing the relation between urban, transportation and environmental conditions in the production of vehicular air pollution, as presented in this study. Next, it is also believed it can contribute to the other phases of the strategic planning process within the sustainable transportation concept.

With particular regard to NNM-AP's formulations, the results of the simulations showed the model's efficiency in relating the urban, transportation and environment conditions to the indexes of vehicular air pollution collected in the monitoring stations. Some errors, however, can be credited to the data collection representation, mainly the pollution dispersion one, which should be in the next studies analysed more properly.

Finally, we intend to improve the model through greater integration of the whole process of strategic planning within the concept of Sustainable Development, in order to contribute to the realization and consolidation of a more functional and clean – Sustainable – city.

Acknowledgements – The first author wishes to thank Japan International Cooperation Agency (JICA) for the scholarship that supported this research.

References

[1)] OECD – Organization for Economic Co-operation and Development, *EST, Environmentally Sustainable Transport (Synthesis Report)*, Conference: Vienna, Austria, October 2000.
[2)] Centre for Sustainable Transportation, *Sustainable Transportation Monitor – Canada*, Mississauga, Ontario: No. 5, November 2001.
[3)] SPO – Strategic Planning Office, Internet Page: Seattle, WA: In: http://www.ci.seattle.wa.us/planning/ ;last update 12th, February, 2001.
[4)] Haykin, S., *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Company: New York, 1994.