EXPERIMENTAL AND ARTIFICIAL NEURAL NETWORK APPROACH FOR FORECASTING OF TRAFFIC AIR POLLUTION IN URBAN AREAS: THE CASE OF SUBOTICA

by

Bogdana B. VUJIĆ[a*], Srđan M. VUKMIROVIĆ[b], Goran V. VUJIĆ[b], Nebojša M. JOVIČIĆ[c], Gordana R. JOVIČIĆ[c], and Milun J. BABIĆ[c]

[a] Provincial Secretariat for Environmental Protection and Sustainable Development, Novi Sad, Serbia
[b] Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
[c] Faculty of Mechanical Engineering, University of Kragujevac, Kragujevac, Serbia

Original scientific paper
UDC: 551.588:551.506(497.1)
DOI: 10.2298:/TSCI100507032V

In the recent years, artificial neural networks have been used to predict the concentrations of various gaseous pollutants in ambient air, mainly to forecast mean daily particle concentrations. The data on traffic air pollution, irrespective of whether they are obtained by measuring or modeling, represent an important starting point for planning effective measures to improve air quality in urban areas. The aim of this study was to develop a mathematical model for predicting daily concentrations of air pollution caused by the traffic in urban areas. For the model development, experimental data have been collected for 10 months, covering all four seasons. The data about hourly concentration levels of suspended particles with aerodynamic diameter less than 10 μm (PM10) and meteorological data (temperature, air humidity, speed and direction of wind), measured at the measuring station in the town of Subotica from June 2008 to March 2009, served as the basis for developing an artificial neural networks based model for forecasting mean daily concentrations of PM10.

Key words: *PM10, meteorological parameters, forecasting PM10 concentrations, artificial neural network*

Introduction

Information on traffic-induced air pollution, irrespective of whether the pertinent data are obtained by measurement or by numerical modeling, as well as the information about spreading of the pollutants, is an important starting point for planning effective measures to improve air quality in urban areas. Such information may be used by the authorized institutions in the urban planning of the endangered area, planning traffic regime, as well as in

[*] Corresponding author; e-mail: bogdana.vujic@vojvodina.gov.rs

establishing the appropriate monitoring systems for warning people about the adverse effects to their health.

First measurements of the concentrations of particulate matter and studies of the effects on human health can be dated to the early 1970s, starting from large size particulates to those of very small diameters (<10 μm) [1]. Measurement of the concentrations of air pollutants such as sulfur dioxide, nitrogen dioxide, black smoke (BS) and sedimenting matter in the Republic of Serbia began just in the middle of the 1990s. Measurements of the concentration of total suspended particles (diameter $\geq$10 μm) have been carried out only for the purpose of pollution detection, with minimal data coverage of 14% at an annual level [2].

Systematic measurements of PM10 (diameter <1 μm) on the territory of Vojvodina Province has not been realized in full yet, although a network for automatic monitoring of air quality, consisting of seven automatic measuring stations with a total of four analyzers for monitoring PM10 concentration was established in 2008.

In view of the requirements of the European Directives, defining mean daily tolerable concentration of PM10 with a limited number of days exceeding the limit in a year, mean annual values of PM10 concentration, as well as margin of tolerance, we are faced with new tasks in the area of air quality management that would enable forecasting of the concentration of polluting matter [2, 3].

In view of this, the aim of the present study was to develop a mathematical model based on the measurement data and application of artificial neural network (ANN) for predicting daily concentrations of traffic caused air pollution in an urban area, in this case of the town of Subotica. This location has been chosen because it is a site of heavy traffic-induced pollution, as an important international road is crossing the town center

Experimental determination of air quality parameters

The measurements of the air quality were carried out in a residential-business zone, at the main crossroad of the Subotica town.

Table 1. Characteristics of the analyzers and sensors in the measuring station

Analyzer of suspended particles PM10, TEOM 1400[a]	Sensors for meteorological parameters
Sampling system: filter papers placed on the oscillating microbalance. The decrease in the oscillation frequency caused by particle settling is directly proportional to the mass of the sample. Measurement principle: Filter sampling, direct measurements of the mass, continuous automatic measurement. Precision:± 5μg m^{-3} for 10-minute mean; ± 1.5μg/m^3 for 1-hour mean Measurement range: 0–5000μ μg m^{-3}	Sensors for wind speed and direction, Wind sonic Option3 Measurement principle: ultrasonic 2-D;Measurement range: speed: ±2%; direction: ±3°; Resolution: speed: 0.01ms^{-1}; direction: 1°; Air temperature sensor, Model DMA 575 Measurement principle: Thermo-hygrometer for measuring air temperature and humidity, thermometric sensor 1/3 DIN Pt100 with signal conversion; Measurement range: from –30 °C to +70 °C; Accuracy: 0.15 °C; Sensor for air humidity, Model DMA 575; Measurement principle: Thermo hygrometer for measuring air temperature and humidity, capacitive humidity sensors; Range: 0%-100% relative humidity (RH); Accuracy: 1.5% (for 5-95% RH, 23 °C)

The measurement station mainly serves to monitor the level of traffic-originated pollutants. It is equipped with the analyzers for measuring concentrations of nitrogen oxides, carbon monoxide, ozone, sulfur dioxide, PM10, BTEX. At the same time it is equipped with the sensors for measuring the meteorological parameters such as wind speed and direction, intensity of solar radiation, temperature, and air humidity (Table 1).

The automatic station is equipped with a total of 6 analyzers for monitoring the concentration of the pollutants and 5 sensors for meteorological parameters. The station characteristics are classified in accordance with the Directive on reciprocal exchange of information – EoI [3].

Analysis of the influence of meteorological parameters on the PM10 concentration

The data obtained by the measurements in the investigation period that lasted from June 2008 to March 2009 were subjected to a detailed analysis with the aim of determining the concentration levels, behavior of the pollutants in a particular time and space, as well as the influence of the other parameters affecting the behavior of the pollutants themselves.

Table 2. Statistical treatment of the PM10 values and meteorological data

	Mean value	Standard deviation	Maximum value	Minimum value
PM10 1-h mean [μgm^{-3}]	49.58	35.34	344.60	0.00*
PM10 24-h mean [μgm^{-3}]	49.27	21.85	144.4	0.00*
Wind direction [°]	169.64	102.67	360.00	0.00*
Wind speed [ms^{-1}]	1.18	0.84	6.38	0.00*
Temperature [°C]	11.05	9.91	35.30	–12.40
Pressure [hPa]	995.59	8.04	1021.00	962.00
Radiation [Wm^{-2}]	112.24	206.5	968.00	2.00
Relative air humidity [%]	62.01	16.7	94.00	16.00

By analyzing the meteorological parameters it was concluded that the PM10 concentration is influenced by the following parameters: temperature, wind speed, and air humidity.

Temperature – High temperatures are related to the slowing of air motion, clear and sunny weather, as well as to the stagnation and poor circulation of air masses in the higher atmospheric layers. The high temperatures and immobile air masses lead to the production and accumulation of the pollutants, so that temperature can be considered as one of the parameters that are powerful predictors of the high PM10 concentration levels. The temperature gradient is influenced by the solar radiation, which increases the kinetic turbulence energy and the height of the layer of mixing of air masses. If the height of air mixing layer above the urban area is low, the pollutants tend to accumulate in the surface layer of the air mass. The temperature and solar radiation influence the formation of the new, secondary, particles in the atmosphere [4]. The correlation between the PM10 concentration and the difference between the maximal and minimal daily temperatures (Δt) of r = 0.29 is considered as significant [5]. The analysis of the measurement data served as the basis for defining the correlation between the difference of maximal and minimal temperatures measured during a day and mean daily concentration of PM10 (PM10 D24), and the value r = 0.87 was considered as significant for the further numerical evaluation.

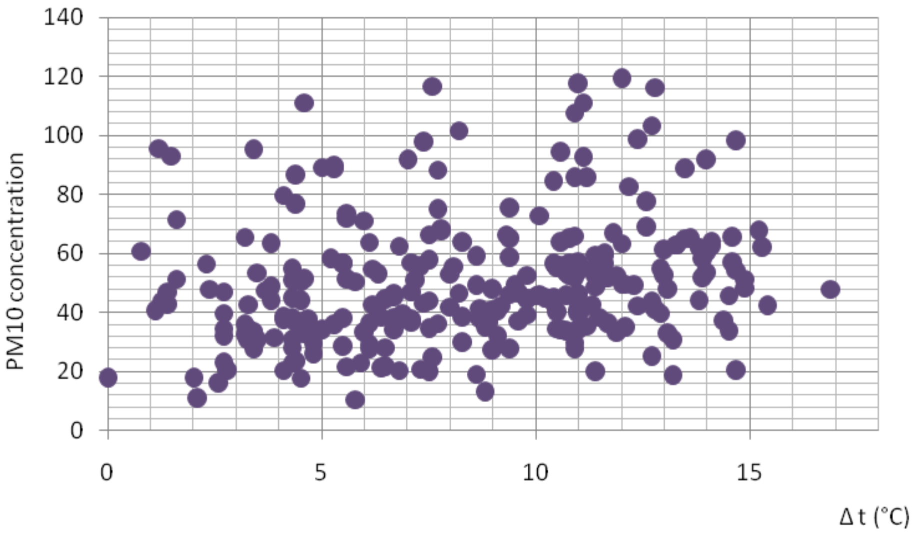

Figure 1. Dependence of the PM10 concentration [μgm^{-3}] on the difference between maximal and minimal daily temperatures (Δt)

Wind speed and direction – Wind speed appeared to be the most significant meteorological parameter that determines the horizontal transport and dispersion of PM10. It was found that low wind speeds result in elevated concentrations of the pollutants. As can be seen from figure 2, showing the dependence of the logarithm of the PM10 (D24) concentration on the wind speed, that at higher wind speeds the PM10 concentrations have a decreasing tendency (r = –0.28), which can be explained by increased dispersion of the pollutants in a larger air volume. The correlation between the PM10 concentration and wind speed is statistically significant.

Daily variations of the PM10 concentration – A detailed analysis of the data showed that the PM10 concentration varies during a day. An increase in the concentration is characteristic of the morning hours (morning peak) from 7 to 10 h, followed by the stagnation from 11th hour and a mild decrease, but with still high concentrations registered to 18 h, when another increase is observed to 22 h. The mean concentrations in the interval from 18 and 22 h are above the maximal tolerable daily value of 50 μgm^{-3}. During the night, the PM10 decreases to the early morning hours, when a new cycle begins (figure 3).

It is important to point out that the distribution of certain PM10 concentrations varies during a day. Statistical treatment of the available data leads to the conclusion that during 24 h the percentages of particular PM10 concentrations show daily variations. During the early morning hours (from 1 to 7 h) the highest percentage makes the concentration of PM10 up to 30 μgm^{-3}, during the major part of the day (from 7 to 18 h) the highest concentrations are of the particles of 30-50 μgm^{-3}, whereas later (after 18 h), an increase is observed in the percentage of high PM10 concentrations (figure 3).

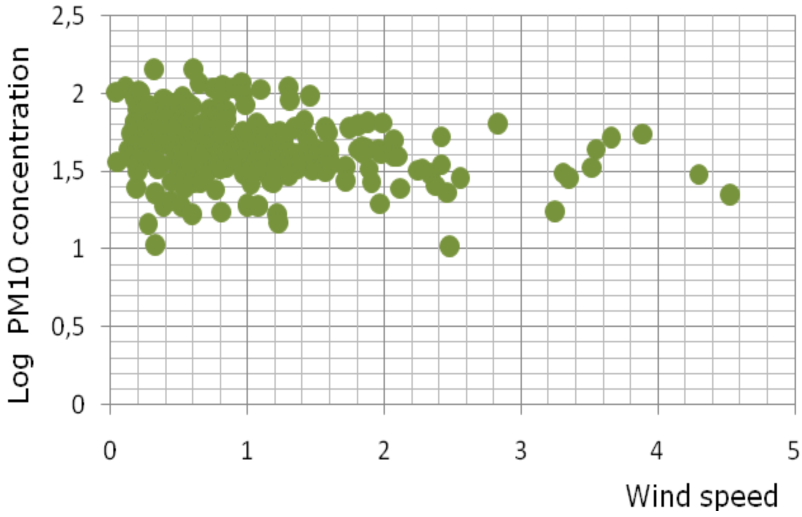

Figure 2. Dependence of the logarithm of the PM concentration [μgm^{-3}] on wind speed [ms^{-1}]

Architecture of the adopted ANN

The obtained experimental results were used to develop a mathematical model for predicting the mean daily concentration of PM10. The mathematical model is based on the application of the ANN in the frame of the software package Matlab R2006a [6]. Survey of the input parameters used for the ANN model is given in Table 3. These parameters have been collected at the measuring place.

For predicting the mean daily concentrations of PM10 use was made of the feed-forward type of ANN. It is a non-recursive network, so that the higher layers do not return information to the lower layers. Optimal ANN structure is obtained by the numerical experiments, by repeating the teaching in several cycles that is by repeating it until a network is obtained that shows the smallest quadratic error on the data set that is taken for the validation. The ANN used is a multilayer network which, apart from the input and output layers, has a hidden layer at the medium level. In the hidden layer use is made of a transit activation function which allows the attaining of the non-linearity in the network (figure 4). The input data for the ANN simulation were the forecasted meteorological data (for $N = 1$) and mean PM10 concentration in the interval from 0 to 10 h on the day to which the forecast is related ($N = 1$) (Table 3).

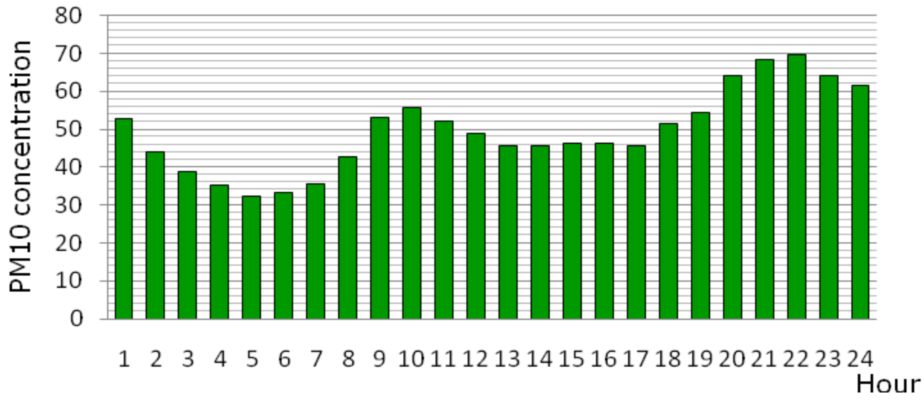

Figure 3. Mean PM10 concentration levels (μgm^{-3}) during 24 hours

Table 3. Description of the input parameters for forecasting mean PM10 concentrations

Parameter	Parameter description
Day (0)	Day when the forecast was made
Day (N)	Day relative with respect to the day 0 (N=...-1,0,1,..)
PM10 (D1), PM10 (D24)	Mean values of PM10 concentrations
WS (D1), WS (D24)	Wind speed
WD (D1), WD (D24)	Wind direction
T(D1), T(D24)	Temperature values
RH (D1), RH (D24)	Relative air humidity
DOW	Day of the week
cos(WD(D24)$_1$), sin(WD(D24)$_1$)	Cosine and sine of the wind direction…

***** D1– for hourly values, D24 – for 24-hour values

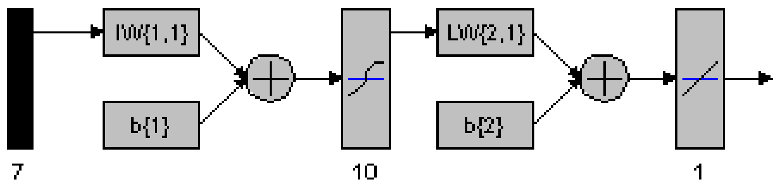

Figure 4. Topological neural network used for training and optimization, where: LW{1,1}, LW{2,1} – the matrices of weighting factors between the input layer and hidden layer, hidden layer and output layer, respectively; b{1}, b{2} – bias matrices of the weighting factors between the input layer and hidden layer, hidden layer and output layer, respectively

Table 4. Comparative survey of calculated errors

Error type	Developed model	[4,5,7 – 17]
RMSE	13.64	4.53 – 75.78
MAE	8.24	2.77 – 8.59
MAPE	25.36	25.57
D	0.84	0.87 – 0.94
r	0.71	0.31 – 0.97

To evaluate the precision of the ANN model use was made of the following errors: root mean square error (RMSE), mean absolute error (MAE), coefficient of correlation (r), mean absolute percentage error (MAPE), and mean deviation error.

The reason for choosing the above errors is the comparative analysis of the model developed in the present work with the literature models [4,5,7-17] (Table 4). For example, Papanastasiou *et al.* [7] developed a model for forecasting mean concentrations of PM10 with the following values of the errors: RMSE=11.37 $\mu g/m^3$, r = 0.78, d = 0.87. On comparing our model with the other models (Table 4) it can be seen that the errors obtained by evaluating of this model are within the limits of the literature data.

Numerical simulation of mean daily PM10 concentration

In the numerical simulation of the assessment of the daily PM10 dose using ANN the input data were the forecasted meteorological data (N = 1) and mean PM10 concentration for the previous day – PM10 $(D24)_0$. The features of the input parameters for the numerical finding of the mean daily PM10 concentrations are given in Table 5.

Table 5. Features of the input data for numerical finding of mean daily concentrations by simulation

cosWD $(D24)_1$ [–]	sinWD $(D24)_1$ [–]	WS$(D24)_1$ [ms^{-1}]	T$(D24)_1$ [°C]	RH $(D24)_1$ [%]	PM10 $(D24)_0$ [μgm^{-3}]	DOW
0.866752	0.498739282	0.46	18.5	77	30.7	5
– 0.37554	0.926807186	0.23	19.9	65	45.8	5
0.110417	– 0.993885326	0.36	23.2	51	66.9	5
0.987344	– 0.158592906	0.19	24.5	48	64.4	5
– 0.88797	– 0.459903491	0.8	25.3	49	57.9	5
0.517796	– 0.855504371	0.39	26.4	44	50.1	5
0.424179	– 0.905578362	1.22	25.9	52	49.1	5
– 0.96945	– 0.245281209	0.56	21.8	69	52.6	5
– 0.37554	0.926807186	0.7	22.8	49	64.8	5

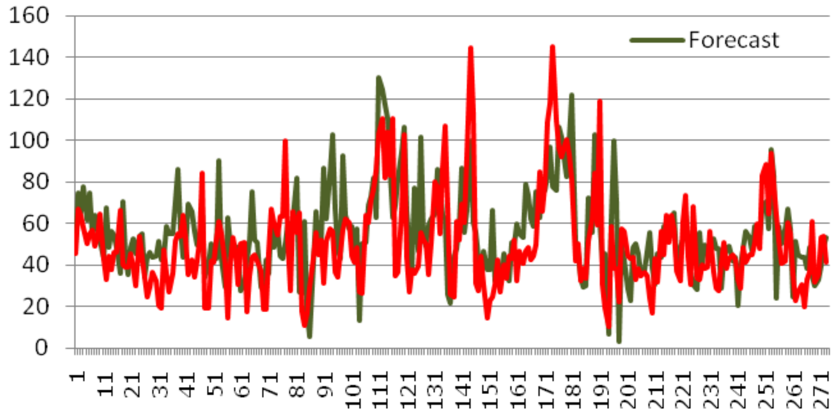

Figure 5. Comparison of the measured and numerical modeled PM10 concentrations

Figure 5 shows a comparison of the measured mean daily PM10 concentrations and those predicted by applying the ANN. The model evaluation was carried out on the basis of the comparison of the experimentally measured values with those obtained from the mathematical model. The mean values of measured daily concentration of PM10 is 50.05 μg/m^3; the MBE value of the developed model from the mean daily concentration is 0.42 μg/m^3, while the standard deviation of the measured values of PM10 concentration is 19.46 μg/m^3, which at the same time indicates that the model satisfies to a great degree the variability of the measured values.

Conclusions

Based on the measurement data for the concentration levels of PM10 and meteorological conditions obtained in the measurement station of the town of Subotica in the period June 2008 – March 2009, a model has been developed for predicting mean daily concentrations of PM10. The results showed that the mean daily concentrations may be obtained using ANN with a defined level of deviations. On the basis of the evaluation of the RMSE, MAE, MAPE, d and MBE it was concluded that the developed model shows a satisfactory forecasting performance.

The previous studies using a number of different methods to forecast the pollutants concentrations proved that standard modeling methods could not forecast sudden pollution episodes (*e. g.* 142nd day, figure 5), so that the developed model has to be further upgraded and adapted to new situations. In addition to standard meteorological data (temperature, wind speed and direction, air humidity) the parameters such as the amount of precipitation, height of the atmospheric mixed layer, traffic density, *etc.* play a significant role in the process of transport of the different particulate matter in the air, which may be subject of a future research.

According to the fact that information about ambient air pollution, which originated from traffic, is a significant base for air quality management in urban area, the model could be used as important tool to decision makers for urban planning such as transport modes planning as well as development of warning system for health protection in this area.

References

[1] Zickus, M., Greig, A., J., Niranjan, M., Comparison of Four Machine Learning Methods for Predicting PM10 Concentrations in Helsinki, Finland, *Water, Air, and Soil Pollution: focus* 2: pp. 717-729, 2002

[2] ***, Directive 2008/50/EC of the European Parliament and of the Council 0f 21 May 2008 on Ambient Air Quality and Cleaner Air for Europe, *Official Journal of the European Communities*, No. L 152/1. EC, 2008

[3] ***, Council Decision 97/101/EC of 27 January 1997 Establishing a Reciprocal Exchange of Information and Data from Networks and Individual Stations Measuring Ambient Air Pollution within the Member States, *Official Journal of the European Communities*, No. L 35/14. EC, 1997

[4] Boy, M., Kulmala, M., Influence of Spectral Solar Irradiance on the Formation of New Partiles in the Continental Boundary Layer, *Atmos. Chem. Phys. Discuss.*, *2* (2002), 5, pp. 1317-1350

[5] U. Brunelli, V. P., L. Pignato, F. S., Vitabile Two-days Ahead Prediction of Daily Maximum Concentrations of SO_2, O_3, PM10, NO_2, CO in the Urban Area of Palermo, Italy, *Atmospheric Environment, 41 (*2007), 14, pp. 2967-2995

[6] ***, The Mathworks, http://www.mathworks.com/products/ (15.07.2009)

[7] Papanastasiou, D. K., Melas, D., Kioutsioukis, I., Development and Assessment of Neural Network and Multiple Regression Models in Order to Predict PM10 Levels in a Medium–Sized Mediterranean City, *Water Air Soil Pollution*, *182* (2007), 1-4, pp. 325-334

[8] Asha, B. C., Devotta, S., Air Quality Forecasting Using a Hybrid Autoregressive and Non-linear Model, *Atmospheric Environment, 40* (2005), 10, pp. 1774-1780

[9] Luis, A. D-R., *et al.*, A Hybrid ARIMA and Artificial Neural Networks Model to Forecast Particulate Matter in Urban Areas: The Case of Temuco, *Chile Atmospheric Environment, 42* (2008), 35, pp. 8331-8340

[10] Goyal, P., Chan, T., Neeru, J., Statistical Models for the Prediction of Respirable Suspended Particulate Matter in Urban Cities, *Atmospheric Environment, 40* (2006) 11, pp. 2068 – 2077

[11] Casseli, M.,Trizio, L., de Gennaro, G., Ielpo, P., A Simple Feed-forward Neural Network for the PM10 Forecasting: Comparison with a Radial Basis Function Network and a Multivariate Linear Regression Model, *Water Air Soil Pollutant, 201* (2008), 1-4, pp. 365-377

[12] Perez, P., Reyes, J., An Integrated Neural Network Model for PM10 Forecasting, *Atmospheric Environment*, *40* (2006), 16, pp. 2845-2851

[13] Kolehmainen, M., Martikainen, H., Ruuskanen, J., Neural Networks and Periodic Components Used in Air Quality Forecasting, *Atmospheric Environment, 35* (2001), 5, pp. 815-825

[14] Hubbard, M., Cobourn, G., Development of a Regression Model to Forecast Ground-Level Ozone Concentration in Louisville, KY, *Atmospheric Environment, 32* (1998) 14-15, pp. 2637-2647

[15] Gardner, W., Dorling, R, Statistical Surface Ozone Models: An Improved Methodology to Account for Non-Linear Behavior, *Atmospheric Environment, 34* (2000), 1, pp. 21-34

[16] Zang, P., Qi, M., Neural Network Forecasting for Seasonal and Trend Time Series, *European Journal Research*, *160* (2005), 2, pp. 501-514.

[17] Corrani, G., Air Quality Prediction in Milan: Neural Network, Pruned Neural Networks and Lazy Learning, *Ecological Modeling*, *185* (2005), 2-4, pp. 513-529

Paper submitted: May 7, 2010
Paper revised: June 25, 2010
Paper accepted: July 1, 2010