



TRAFFIC POLLUTION ASSESSMENT USING ARTIFICIAL NEURAL NETWORK AND MULTIVARIATE ANALYSIS

Mario De Luca¹✉, Daiva Žilionienė², Saulius Gadeikis³, Gianluca Dell'Acqua⁴

^{1,4}Dept of Civil, Construction and Environmental Engineering, University of Naples Federico II,
Via Claudio 21, I-80125 Naples, Italy

²Dept of Roads, Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania

³Dept of Hydrogeology and Engineering Geology, Vilnius University, M.K.Čiurlionio g. 21/27,
LT-03101 Vilnius, Lithuania

E-mails: ¹mario.deluca@unina.it; ²daiva.zilioniene@vgtu.lt; ³saulius.gadeikis@gf.vu.lt;
⁴gianluca.dellacqua@unina.it

Abstract. The work addressed a study on pollution caused by traffic on the highway. In particular, it was considered the concentration of pollutant, resulting from the passage of vehicles on the freeway. Five different stations (sensors and samples) used to collect data. The data collection period around six months. Also, the following parameters were detected: wind speed and direction, temperature and traffic flow rate. Data processed with Multivariate Analysis and Artificial Neural Network approach. The best model it obtained with Artificial Neural Network approach. In fact, this model presented the best fit to the experimental data.

Keywords: Artificial Neural Network, concentration of pollutant, Multivariate Analysis, traffic flow rate, wind speed and direction, temperature.

1. Introduction and literature review

The level of pollution by vehicular transport is a problem that affects the entire globe. Laws and regulations for the control of this phenomenon have been produced in many countries in the world.

Kyoto Protocol is the international reference for the evaluation of air pollution. In the Europe Union, the rules governing air pollution are based on *Directive 2008/50/EC*, which provide guidelines and criteria for the control of pollution (also from motor transport). For this reasons, the scientific community is very interested in this topic. Many researchers in this field have produced in models and procedures useful for the control of air pollution.

Pasquill (1961) and Briggs (1967) conducted the first studies. In particular, Pasquill (1961) suggested a relationship that takes account of parameters characterizing the dispersion of pollutants and it identifies the relationships for calculating the standard deviations of the Gaussian distribution. Briggs (1967) uses the same formula employed by Pasquill (1961) for the calculation of the concentration, however, propose different relations for calculating the standard deviations of the Gaussian distribution. Subsequently, many studies have been proposed, e.g. many based on standard analytical techniques, other – more

innovative, on more advanced techniques such as ANN (Artificial Neural Network) and FLT (Fuzzy Logic Theory).

Sharan *et al.* (1996) developed a model for in low wind situations. This study based on the parameterization of the diffusivity coefficients expressing those regarding standard deviations of the crosswind and vertical distribution of the concentration Gaussian. Singh and Yadav (1996) proposed an evaluation model for the dispersion of air pollutants in conditions of light winds taking into account the spread in all directions. Raimondi *et al.* (1997) reported a study based on FLT, which allows taking into account model uncertainties and describes daily dynamics of a variable Dosage Area Product (DAP) – representative of ground level pollution produced by vehicular transport in urban areas of complex topography. Air pollution being an imprecise and variable event, the application of FLT an ultimate tool for improving the description of air pollution and is an area worth venturing in for future research. Tanaka *et al.* (1992) have reported other applications of FLT in air pollution. Moseholm *et al.* (1996) argued that neural networks be an effective and efficient method for exploring complex relationships between traffic, the wind, and short-term CO (carbon monoxide) concentrations near intersections. A neural network-based model for the analysis of CO pollution in the urban areas of Rosario due

to motor transport developed by Drozdowicz *et al.* (1997). Gualtieri and Tartaglia (1997) used a street canyon model to estimate NO₂ (nitrogen oxide) concentration due to traffic. Gardner and Dorling (1998) developed multi-layer perceptron model for forecasting hourly NO₂ (nitrogen oxide) levels in an urban area of London city. Tao and Xinmiao (1998) considered environment assessment methods. It involved applying multistage fuzzy clustering analysis, wherein after an initial setting up of an evaluation system, evaluation criteria, formulae for the subordination function, allocating weights and modelling design programme have been established. Johansson (1998) show levels of future specific vehicle emissions and the energy efficiency required to match long-term environmental. It appears to be possible to achieve sufficiently large reductions in both nitrogen oxide (NO₂) and non-methane volatile organic compound (NMVOC) emission to meet long-term Swedish environmental requirements even with continuing transport growth. Stockie (2011) described in detail the basic mathematics that lies behind the modelling of atmospheric dispersion, assuming a Gaussian dispersion type and point source. Chart-asa and Gibson (2015) showed a new approach for quantifying health impacts of traffic-related particulate matter air pollution at the urban project scale. Shorshani *et al.* (2015) showed a review of the status of the relationships between traffic, emissions, air quality, and water quality models, to recommend modelling approaches and to propose some directions for improving the state of the science.

2. Techniques used in data analysis

Two different types of techniques used for the analysis in this study: Multivariate Analysis (MVA) and Artificial Neural Network (ANN). For the first, the description is omitted because it is present for many years in the technical literature. For the second, much more recent, the following basic principles described.

2.1. Artificial Neural Network

The inspiration for the structure of the ANN is taken from the structure and operating principles of the human brain. It is made of interconnected artificial neurons that mimic some properties of biological neurons. The function of a biological neuron is to add its input and produce an output. This output is transmitted to subsequent neurons, through the synoptic joints, only if the transmitted signal is high (more than a predetermined value). Otherwise, the signal is not transmitted to the next neuron. In the network, therefore, a neuron calculates the weighted sum, using Eq (1) (considering the input x_i and weights w_i) and compares it with a threshold value; if the sum is more than the threshold value, the neuron lights up, and the signal is transmitted. Otherwise, the neuron does not turn on, and the flow stops (Žilioniene *et al.* 2014).

$$I = \sum_{i=1}^n w_i x_i \quad (1)$$

where I – weighted sum, dimensionless; w_i – weight, dimensionless; x_i – input, dimensionless.

The activation value U_j rather than U_j , connected to weight w_{ij} , is a function of the weighted sum of the input. This function may take various forms. In this study, a function of type Eq (2) used (Žilioniene *et al.* 2014):

$$u_j = \frac{1}{1 + e^{-(\sum(i)w_{ij}u_i + \theta_j)}}, \quad (2)$$

where θ_j – bias unit, dimensionless; u_i – degree of sensitivity of u_j when it receives an input signal from u_j , dimensionless; w_{ij} – weight between the connection of the neuron i with the neuron j , dimensionless.

2.2. Multi-Layer Perceptron and the Back Propagation algorithm

In this study, a neural network with Multi-Layer Perceptron (MLP) architecture used. Training carried out using the Back Propagation (BP) algorithm. The neurons (or units) that comprise this type of network organised into layers: an input layer, an output and some intermediate layers between input and output referred to as hidden, defined by the user. Initially, the weights values are assigned random (normalized in the range [0, 1] or [-0.5, +0.5]); moreover, there is input vector $X_p = (X_0, X_1, X_2, \dots, X_{n-1})$ with $X_0 = 1$ and an output vector $T_p = (T_0, T_1, T_2, \dots, T_{m-1})$ (Žilioniene *et al.* 2014).

In this way, the network will consist of $n-1$ input neurons and $m-1$ output neurons. The weighted sum of the inputs for each layer calculated using Eq (1) and its value of activation, e.g. output using Eq (2). Then, the weights changed so that the output of the network (e.g. the output of the last layer of neurons) increasingly approximates the target set by the user. It defined a function error (Eq (3)) proportional to the square of the difference between the output and target for all output neurons:

$$E_p = \frac{1}{2} \sum_j (T_{p_j} - O_{p_j})^2, \quad (3)$$

where T_{p_j} – target (dimensionless); O_{p_j} – output (dimensionless).

Subsequently, BP is applied, e.g. the weights varied, so that error E_p tends towards zero (starting from the last layer to the first). It is defined, for the current pattern p , a variation Δw_{ij} of weight w_{ij} between the neuron i and j that given by Eq (4) (Žilioniene *et al.* 2014).

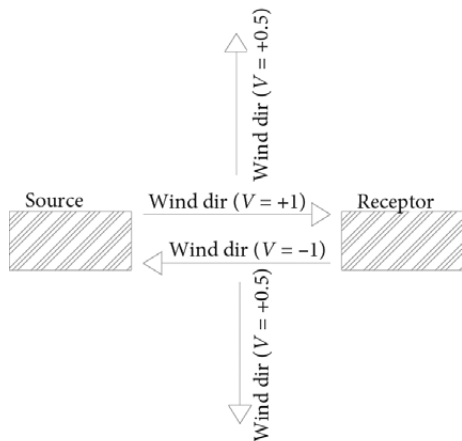
$$\Delta_{pw_{ij}} = -\alpha \frac{\partial E_p}{\partial w_{ij}} + \beta (\Delta_{p-1}) w_{ij}, \quad (4)$$

where α – the learning coefficient (learning rate); β – the momentum; $\Delta_{p-1} w_{ij}$ – the variation of the same weight calculated according to the previous model. Eq (5) gives the new weights (Žilioniene *et al.* 2014):

$$w_{ij}^{new} = w_{ij}^{old} + \Delta_{pw_{ij}} \quad (5)$$

Table 2. Dataset aggregation

Traffic class Cl , vph	Temperature T , °C	Wind speed u , m/s	Wind direction V , dimensionless	E total E_{top} , g/s	C observed, $\mu\text{g}/\text{m}^3$
25–50	21.3	1.1	–0.3	38 475	12.38
50–75	18.2	1.8	–0.1	58 101	12.59
75–100	17.6	1.7	–0.4	79 138	16.71
125–150	18.5	1.7	–0.5	104 364	19.72
150–175	19.3	1.0	–0.8	150 687	39.21
175–200	20.2	1.6	0.5	175 630	34.95
200–225	19.9	1.5	–0.3	195 606	47.83
225–250	16.5	3.0	–0.2	212 440	27.40
250–275	17.2	2.7	–0.3	233 499	31.82
275–300	17.6	2.7	–0.2	265 967	38.83
300–325	16.1	2.8	–0.2	282 082	38.44
325–350	18.8	2.6	–0.2	306 184	46.00
350–375	26.1	1.9	–0.5	334 345	47.66
375–400	25.3	1.9	–0.5	360 096	53.05
400–425	25.3	2.3	–0.4	381 465	49.47
425–450	23.9	1.6	–0.5	402 141	76.42
450–475	26.1	2.3	–0.7	426 860	37.74
475–500	25.4	2.5	–0.6	446 892	46.99
500–525	25.4	2.8	–0.5	442 879	49.24
525–550	22.9	1.8	0.0	438 846	97.40

**Fig. 3.** Wind value direction**Table 3.** Model parameters

Parameter	Estimate	Std. Error	Interval		Significance
			Lower Bound	Upper Bound	
b_1	4.692	1.234	1.385	6.091	>95%
b_2	3.131	0.759	0.834	4.529	>95%

3.1. Data analysis and results

The data reported in Table 1 aggregated into traffic classes (to 25 vpd). Subsequently (to build the columns of Table 2) for all the variables given in Table 2 it was considered the average value in each class.

For the wind direction variable, the criteria presented in Fig. 3, were considered.

4.1. Multivariate Analysis model

The technique of MVA has been applied to the data contained in Table 2 using C as the dependent variable and the other variables as a predictor (T , u , V , E). The better relation between the dependent variable and predictors it has been the Gaussian type. Following is shown the expression of the model obtained:

$$C = \frac{ET}{u\sqrt{2\pi}} b_1 \exp\left[-\frac{1}{2} b_2 V^2\right], \rho^2 = 0.94. \quad (9)$$

The model (9) was characterized by a coefficient of determination $\rho^2 = 0.94$ and a significance more than 95% (Table 3).

4.2. Artificial Neural Network model

The ANN technique has been applied, as in the case of the model (9), to the Table 2 data. The model has been obtained with the technique of ANN as given in Section 2.1. The variables that were considered and used in the model are listed in Table 1. The 70% of the data was used to train the network, and the remaining part of the data was used for verification. Different configurations were considered for the architecture of the neural network. Fig. 4 presents the best network architecture. Table 4 gives values of the estimated parameters.

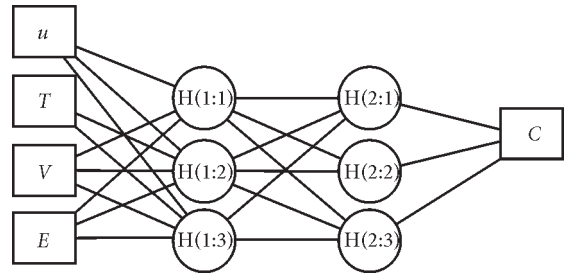


Fig. 4. Architecture of ANN model

Table 4. Parameters of ANN model

Predictor	Predicted						Output Layer Leq
	Hidden Layer 1			Hidden Layer 2			
	H(1:1)	H(1:2)	H(1:3)	H(2:1)	H(2:2)	H(2:3)	
Input Layer	(Bias)	-.827	-.366	.310			
	T	-.053	-.300	.268			
	u	-.157	-.538	-.099			
	V	.326	.266	-.112			
	E	-1.176	.403	.596			
Hidden Layer 1	(Bias)				.341	-.051	.097
	H(1:1)				.132	-.656	-.105
	H(1:2)				-.118	.711	.753
	H(1:3)				-.073	.330	.092
Hidden Layer 2	(Bias)						-.432
	H(2:1)						-.279
	H(2:2)						1.193
	H(2:3)						.757

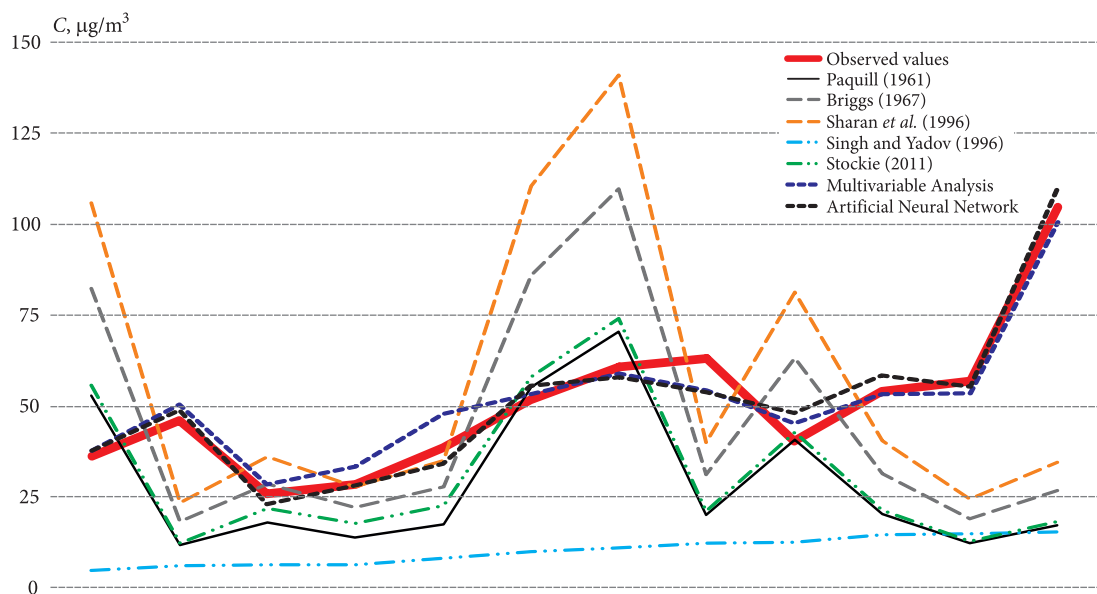
Table 5. Main models available in literature

No.	Model	Equation
1	Pasquill (1961)	$C = \frac{E}{u} \frac{1}{2\pi\sigma_y\sigma_z} \exp\left[-\left(\frac{y^2}{2\sigma_y^2} + \frac{z^2}{2\sigma_z^2}\right)\right], \sigma_y(x) = \frac{k_1 x}{\left[1 + \left(\frac{x}{k_2}\right)^2\right]^{k_3}}, \sigma_z(x) = \frac{k_4 x}{\left[1 + \left(\frac{x}{k_2}\right)^2\right]^{k_5}}$
2	Briggs (1967)	$C = \frac{E}{u} \frac{1}{2\pi\sigma_y\sigma_z} \exp\left[-\left(\frac{y^2}{2\sigma_y^2} + \frac{z^2}{2\sigma_z^2}\right)\right], \sigma_y = \sigma_z = \sigma = ax(1-bx)^c$
3	Sharan et al. (1996)	$C = \frac{q}{2\pi(K_y K_z)^{1/2} x} \exp\left[-\frac{U}{4x} \left(\frac{y^2}{K_y} + \frac{z^2}{K_z}\right)\right]$
4	Singh and Yadov (1996)	$C = \frac{Q}{U\pi\sqrt{\beta\gamma}x^2} \left[1 + \frac{\alpha}{x^2} \left(\frac{y^2}{\beta} + \frac{z^2}{\gamma}\right)\right]^{-((1/2\alpha)+1)}$
5	Stockie (2011)	$C = \frac{Q}{4\pi u \sqrt{r_y r_z}} \exp\left(-\frac{y^2}{4r_y}\right) \left[\exp\left(-\frac{(z-H)^2}{4r_z}\right) + \exp\left(-\frac{(z+H)^2}{4r_z}\right) \right]$

Table 6. Differences between models and observed values

Paquill (1961)	Briggs (1967)	Sharan <i>et al.</i> (1996)	Singh and Yadov (1996)	Stockie (2011)	Multivariate Analysis Model	Artificial Neural Network Model
16.76	46.27	69.71	-31.35	19.47	1.58	1.56
-34.37	-27.84	-22.64	-39.97	-33.77	4.30	2.90
-7.84	2.62	10.18	-19.63	-3.88	2.64	-2.90
-14.43	-6.23	-0.52	-21.97	-10.49	5.07	-0.06
-21.08	-10.71	-3.57	-30.42	-15.81	9.16	-4.54
3.55	34.34	58.80	-41.82	6.38	1.55	3.91
9.62	48.90	80.10	-49.88	13.23	-1.95	-2.98
-42.95	-31.80	-22.93	-50.68	-41.93	-8.72	-9.07
0.36	23.00	41.00	-27.70	2.44	4.87	7.92
-33.95	-22.68	-13.72	-39.60	-32.91	-0.88	4.28
-44.70	-37.85	-32.41	-42.19	-44.07	-3.39	-1.59
-87.48	-77.85	-70.20	-89.28	-86.59	-4.22	5.39
Total -256.51	-59.83	93.80	-484.49	-227.93	10.01	4.82

Note: minimum difference marked by grey colour.

**Fig. 5.** Comparison of models

4.3. Artificial Neural Network model and Multivariate Analysis model versus literature models

The reliability of the models obtained through ANN and MVA was compared to the main models available in the literature (Table 5).

Table 6 and Fig. 5 present the comparison among the models. In particular, the estimated values with the analysis ANN turn out to be closer to the observed values (more less residual).

5. Conclusions

1. The pollution by motor transport affects the entire globe. Laws and regulations for the control of this phenomenon have been produced in many countries in the world. The

most important document is Kyoto Protocol; it is the international reference for the evaluation of air pollution. In the Europe Union, the rules governing air pollution are based on *Directive 2008/50/EC*, which provide guidelines and criteria for the control of pollution (also from vehicular traffic). For this reasons, the scientific community is very interested in this topic.

2. Two models of prediction of the concentration of pollutant produced by motor transport on the highway were built. The data collection period six months and involved a section of 30 km. Also, detected wind speed and direction, temperature and flow rate for the traffic.

3. The data aggregated for traffic classes and processed by Artificial Neural Network and Multivariate

Analysis techniques. The results showed that the two models (Model 1 proceeds through Artificial Neural Network and Model 2 proceeds through Multivariate Analysis) are very reliable. In fact, they have residues less than the observed values.

4. The simulation capacity of the models compared to the main models available in the literature. The comparison showed that the Artificial Neural Network model is the most reliable because it presented the best fit to the experimental data (it presented a less sum of residuals, i.e. difference between observed values and estimated values).

References

- Briggs, G. A. 1967. Plume Rise: a Critical Survey, in *Proc. of the USAEC Meteorological Information Meeting Held at Chalk River Nuclear Laboratories*. 11–14 September, 1967, Chalk River Ontario, Canada. 2787: 1–21.
- Chart-asa, C.; Gibson, J. M. 2015. Health Impact Assessment of Traffic-Related Air Pollution at the Urban Project Scale: Influence of Variability and Uncertainty, *Science of the Total Environment* 506–507: 409–421. <https://doi.org/10.1016/j.scitotenv.2014.11.020>
- Drozdowicz, B.; Benz, S. J.; Santa Cruz, A. S. M.; Scenna, N. J. 1997. A Neural Network Based Model for the Analysis of Carbon Monoxide Contamination in the Urban Area of Rosario, *Transactions on Ecology and the Environment* 15: 677–689. <http://dx.doi.org/10.2495/AIR970641>
- Gardner, M. W.; Dorling, S. R. 1998. Artificial Neural Networks (the Multilayer Perceptron) – a Review of Applications in the Atmospheric Sciences, *Atmospheric Environment* 32 (14–15): 2627–2636. [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0)
- Gualtieri, G.; Tartaglia, M. 1997. A Street Canyon Model for Estimating NOx Concentrations due to Road Traffic, in *International Conference on Measurements and Modelling in Environmental Pollution*. 211–220.
- Johansson, B. 1998. Will New Technology be Sufficient to Solve the Problem of Air Pollution Caused by Swedish Road Transport?, *Transport Policy* 5(4): 213–221. [https://doi.org/10.1016/S0967-070X\(98\)00023-7](https://doi.org/10.1016/S0967-070X(98)00023-7)
- Moseholm, L.; Silva, J.; Larson, T. 1996. Forecasting Carbon Monoxide Concentration Near a Sheltered Intersection Using Video Traffic Surveillance and Neural Networks, *Transportation Research Part D: Transport and Environment* 1(1): 15–28. [https://doi.org/10.1016/S1361-9209\(96\)00002-8](https://doi.org/10.1016/S1361-9209(96)00002-8)
- Pasquill, F. 1961. The Estimation of the Dispersion of Windborne Material, *Meteorological Magazine* 90(1063): 33–49.
- Raimondi, P. M.; Rando, F.; Vitale, M. C.; Calcara, A. M. V. 1997. Short-Time Fuzzy DAP Predictor for Air Pollution due to Vehicular Traffic, *WIT Transactions on Ecology and the Environment* 19. <http://dx.doi.org/10.2495/MMEP970191>
- Sharan, M.; Yadav, A. K.; Singh, M. P.; Agarwal, P.; Nigam, S. 1996. A Mathematical Model for the Dispersion of Air Pollutants in Low Wind Conditions, *Atmospheric Environment* 30(8): 1209–1220. [https://doi.org/10.1016/1352-2310\(95\)00442-4](https://doi.org/10.1016/1352-2310(95)00442-4)
- Shorshani, M. F.; André, M.; Bonhomme, C.; Seigneur, C. 2015. Modelling Chain for the Effect of Road Traffic on Air and Water Quality: Techniques, Current Status and Future Prospects, *Environmental Modelling & Software* 64: 102–123. <https://doi.org/10.1016/j.envsoft.2014.11.020>
- Singh, M. P.; Yadav, A. K. 1996. Mathematical Model for Atmospheric Dispersion in Low Winds with Eddy Diffusivities as Linear Functions of Downwind Distance, *Atmospheric Environment* 30(7): 1137–1145. [https://doi.org/10.1016/1352-2310\(95\)00368-1](https://doi.org/10.1016/1352-2310(95)00368-1)
- Stockie, J. M. 2011. The Mathematics of Atmospheric Dispersion Modeling, *SIAM Review* 53(2): 349–372. <https://doi.org/10.1137/10080991X>
- Tanaka, K.; Sano, M.; Watanabe, H. 1992. Identification and Analysis of Fuzzy Model for Air Pollution – an Approach to Self-Learning Control of CO Concentration, in *Proc. of the International Conference on Industrial Electronics, Control, Instrumentation, and Automation on IEEE, Power Electronics and Motion Control*, 13 November, 1992. 1431–1436. <https://doi.org/10.1109/IECON.1992.254391>
- Tao, Y.; Xinmiao, Y. 1998. Fuzzy Comprehensive Assessment, Fuzzy Clustering Analysis and Its Application for Urban Traffic Environment Quality Evaluation, *Transportation Research Part D: Transport and Environment* 3(1): 51–57. [https://doi.org/10.1016/S1361-9209\(97\)00021-7](https://doi.org/10.1016/S1361-9209(97)00021-7)
- Žilionienė, D.; De Luca, M.; Dell'Acqua, G.; Lamberti, R.; Biancardo, S. A.; Russo, F. 2014. Evaluating Freeway Traffic Noise Using Artificial Neural Network, in *Proc. of the International Conference “Environmental Engineering”*, 22–23 May, 2014, Vilnius, Lithuania. 9 p. <http://dx.doi.org/10.3846/enviro.2014.182>

10 October 2016; accepted 27 February 2017