# APPLICATION OF ARTIFICIAL NEURAL NETWORKS FOR PREDICTION OF AIR POLLUTION LEVELS IN ENVIRONMENTAL MONITORING

Małgorzata Pawul[1], Małgorzata Śliwka[1]

[1] AGH University of Science and Technology, Mickiewicza 30, 30-059 Krakow, Poland, e-mail: pawul@agh.edu.pl

**ABSTRACT**

Recently, a lot of attention was paid to the improvement of methods which are used to air quality forecasting. Artificial neural networks can be applied to model these problems. Their advantage is that they can solve the problem in the conditions of incomplete information, without the knowledge of the analytical relationship between the input and output data. In this paper we applied artificial neural networks to predict the PM 10 concentrations as factors determining the occurrence of smog phenomena. To create these networks we used meteorological data and concentrations of PM 10. The data were recorded in 2014 and 2015 at three measuring stations operating in Krakow under the State Environmental Monitoring. The best results were obtained by three-layer perceptron with back-propagation algorithm. The neural networks received a good fit in all cases.

**Keywords:** environmental monitoring, air pollution, artificial neural networks, prediction

## INTRODUCTION

High level of air pollution is currently a problem in many urban areas. Permissible concentrations of many pollutants are exceeded. In conjunction with the unfavorable weather conditions it leads to the formation of smog phenomena. It causes deterioration of the quality of life in urban areas and may cause many diseases. One of the aims of air quality monitoring carried out in these areas is detection of exceeding the permissible pollutant concentrations. In the event of exceedance, actions to improve air quality are undertaken. The knowledge of reliable forecasts of occurrence of high air pollution levels would allow to undertake preventive actions.

Therefore, in recent years, a lot of attention was paid to the improvement of modeling methods of phenomena occurring in the environment, including the creation of air quality forecasting. As we know, the forecasts generation may be based on statistical or deterministic models. There are analyzes based on modeling of meteorological fields and pollution dispersions in

deterministic models. Forecasting Air Pollution Propagation System is one of such models [Hajto et al. 2012]. The deterministic models are usually complex and complicated. The statistical models are simpler but also have disadvantages. They are created based on a long series of immission measurements, and they give the concentration forecasts only for measuring points. Artificial neural networks (ANNs) belong to the group of statistical models. It is worthwhile to remark that the possibilities of using neural networks are various. They are used not only in air protection, but also in other environmental science, business, medicine, industry, etc. [Tadeusiewicz and Dobrowolski 2004]. The ability to work with unknown relationships between the predictor variables and the predicted variables is the advantage of the mentioned method.

There are papers devoted to prediction of air pollution levels based on artificial neural networks. We can predict the concentration of the air pollutant or the occurrence of a certain range of concentrations (class of air quality) [Skrzypski and Jach-Szakiel 2008]. It should be remembered

that we cannot create one neural network model, which will work well for each pollutant and for each region. The choice of its architecture including the number and type of neurons, and the selection of a learning algorithm can significantly affect its performance and have to be studied individually for each case.

In this paper we focus on forecasts of the concentration levels of particulate matter PM 10 as factors determining the occurrence of smog phenomena. The permissible level of this pollutant is often exceeded in Krakow and the municipal authorities undertake various actions to reduce PM 10 emission.

## THE CAUSES OF SMOG FORMATION IN KRAKOW AND ITS IMPACT ON THE HEALTH OF RESIDENTS

Air pollution is the cause of a number of unfavorable effects, among which smog is regarded as particularly hazardous to human health. Smog is formed as a result of combination of smoke, fog and water vapor in large urban areas, during the occurrence of a high level of air pollution, especially: sulfur and nitrogen oxides and particulate matter. Depending on the type of the chemical pollutants and climatic conditions, there are two basic types of smog formation:

- photochemical smog, is formed in strong sunlight, low humidity, with the participation of such pollutants as: carbon monoxide, nitrogen oxides, aromatic and unsaturated hydrocarbons, ozone and industrial particulates,
- sulfuric smog (industrial smog), is formed in large urban areas in low temperate climates, mainly as a result of burning coal, in still air and high humidity of air. It is formed in the autumn and winter months as a result of temperature inversion, the main pollutants are sulfur oxides, nitrogen oxides, carbon oxides and particulate matter [Sobczyk et al. 2014].

A very significant problem that reduces the quality of life in Krakow is sulfuric smog, that settles over the city during the autumn and winter months. The reason for this situation is high air pollution (PM 10, PM 2.5, BAP, absorbed heavy metals, sulfur and nitrogen oxides) and unfavorable location of the city. Krakow, which is located in the valley, is not sufficiently aired due to unfavorable urban layout and lack of channels needed for airing of the city. Also, local climatic conditions, often occurring phenomenon of temperature inversion and a large number of windless days effect on the smog formation in Krakow. Burning solid fuels (coal, wood) in household stoves and traffic pollution should be mentioned among the major sources of air pollution. The pollution coming from the villages near Krakow has also a significant impact on air quality in the city. Polluted air has a negative impact on human health, and is the cause of many diseases. It causes e.g.:

- diseases of respiratory system (asthma, chronic obstructive pulmonary disease, allergies, runny nose, cough, sore throat),
- cardiovascular diseases (atherosclerosis, hypertension, heart failure, myocardial infarction),
- diseases of nervous system (problems with memory, concentration, more frequent depressive behavior, faster aging of the nervous system, increased risk of Alzheimer's disease),
- cancers (particularly of lung cancer, throat and larynx cancer) [Manahan 2006, Miśkiewicz 2014].

Polluted air adversely affects children before their birth (low birth weight, susceptibility to respiratory disease, lower intelligence quotient).

Frequent, multiple exceeding the permissible levels of pollutants [Dz.U. 2012 poz. 1031] has a significant impact on health and life expectancy in Krakow. About 400 people die prematurely due to the bad air quality every year here. Within the ad hoc action, the municipal authorities implement free of charge city communication for people who leave their cars at home (on the basis of a proof of the vehicle registration) when the pollution levels of PM 10 exceed 150 µg/m$^3$ (the results of five measuring stations) or one station exceeds the level of 200 µg/m$^3$ (currently in Poland for PM 10: alarm level > 300 µg/m$^3$, the information level > 200 µg/m$^3$ [Uchwała RMK 2015].

## ARTIFICIAL NEURAL NETWORKS AND THEIR APPLICABILITY

Recently, intensive development of algorithms in artificial intelligence is observed. Artificial neural networks were among the first algorithms of this type. Their characteristic feature is that they can be used to solve the problem in

the conditions of incomplete information, without the knowledge of the analytical relationship between the input and output data. This feature causes that ANNs are very important tool for modelling complex unknown relationships between the variables. They are widely used in various spheres of life e.g. in classification, analysis and processing of images, time series prediction, analysis of production problems, forecasting stock prices, forecasting weather events etc. In environmental protection, they can be used to providing missing data from environmental monitoring, predicting air and water pollution levels and sound levels, automatic image analysis and interpretation of biological monitoring results, environmental impact assessment, determination of ore lithological composition and many other issues [Haupt et al. 2009, Kwiecień and Pawul 2012, Krawczykowski et al. 2009, Singh et al. 2009, Szczepańska and Kmiecik 2001, Tadeusiewicz and Dobrowolski 2004].

There are many papers in which neural networks are discussed. For a comprehensive review on the artificial neural networks, one can refer to [Tadeusiewicz 1993]. There are also a number of papers devoted to prediction of the air pollution levels based on artificial neural networks. Most researchers focused on short- and long-term forecasts of the concentration levels of nitrogen oxides ($NO_x$) and particulate matter PM 10 as factors determining the occurrence of smog phenomena. There are also papers devoted to the predictions of the concentration levels of other air pollutants, for example sulfur(IV) oxide. These predictions were made based on meteorological data, data of air pollutants emissions, etc. Other studies are these dedicated to filling gaps in monitoring data, based on existing data [Abderrahim et al. 2016, Chellali et al. 2016, Kukkonen et al. 2003, Russo et al. 2013, Skrzypski and Jach-Szakiel 2008].

There are many types of the artificial neural networks which differ in structure and principle of operation, e.g. the fully connected feedforward networks known as multi-layer perceptron (MLP) or the radial basis function networks (RBF). The basic structure of the artificial neural network consists of three types of layers of neurons (interconnected nodes). The first is the input layer where data are introduced. The second is the hidden layer where data are processed in order to extract the intermediate data required to determine the final solution. The hidden layer

may be one or more than one. The third type of layers is the output layer, where the results are produced. In the process of defining the multi-layer neural network, first we must specify the number of layers and the number of neurons in each layer. The number of neurons in the input layer is equal to the number of feature vector components. One hidden layer is sufficient to solve most of the classification problems. The number of neurons in the hidden layer depends on the complexity of the problem. While applied ANNs to more complicated problems, more neurons in hidden layer are required. The number of neurons in the output layer is equal to the number of predefined classes (in the classification problem) or the number of the output data (in the prediction problem).

The data passing through the neurons are modified by weights and transfer functions, so to define the neural network we also must specify the type of neuron activation function (non-linear relationship between the signal of the total neuron stimulation and its response), the learning algorithm (used to determine the best weights) and the size of subset of the learning, validation and testing data. In many cases, the activation function takes a form of sigmoid (logistic) function or hyperbolic tangent (often works better than the logistic function) [Tadeusiewicz 1993, Bishop 1995]. Other activation functions may be linear, exponential, sinusoid or Gaussian (used in RBF networks).

There are several learning algorithms used to determine the best weights. The most popular algorithm of MLP learning is the back-propagation algorithm. This simple algorithm is quite slow, but very effective. It works by adjusting the weights to minimize the error between the actual and the desired outputs (by propagating the error back into the network) [Tadeusiewicz 1993]. The sum squared error function is often used in this algorithm.

In some cases, the weight values in MLP are modified in the learning process during the Conjugate Gradient or Quasi-Newton algorithms.

The neural network overfitting is a negative phenomenon that can occur during learning process. The validation process should be used to avoid it. In order to increase the reliability of the final network model, the testing data are used. All relevant cases should be represented in these three data subsets: learning, validation and testing [Tadeusiewicz 1993].

The huge popularity of artificial neural networks and their wide applicability have made that there is software for modeling neural networks. Package Neural Networks in STATISTICA is recommendable. It allows the use of different neural networks, learning methods, activation functions and error functions.

## THE DATA SOURCE USED TO CREATE THE NEURAL NETWORKS THAT PREDICT CONCENTRATION LEVELS OF AIR POLLUTIONS

The artificial neural network learns by comparing the input and output data, so the correct selection of the data set is very important in building a model of neural network. It is worth noting that the input data that affect the output data should be used. The introduction of input data not associated with forecasted data to the model degrades the network working. The input data set containing the concentrations of selected air pollutants and meteorological data is normally used for predicting concentration level of the other air pollutant. Additionally, we can use the data on the emission of air polluting substances in a given area and traffic flow. To create the neural network model we can use temporary data, average hourly or average daily data.

In this paper, to create the artificial neural network that predicts concentration of particulate matter PM 10, which is the main cause of the occurrence of the phenomenon of sulfuric smog, we used meteorological data and data of concentrations of PM 10. They are average daily concentrations during the period from 1 January 2014 to 31 December 2015, which were recorded at three measuring stations operating in Krakow under the State Environmental Monitoring. At the station at Krasinski Avenue the level of traffic pollutions are measured. At the second station Nowa Huta (located at Bulwarowa street) the level of industrial pollutions are measured. The Kraków-

Kurdwanów Station (located at Bujaka street) is a background station. The measurement of PM 10 is done automatically. There are measured concentration levels of other air polluting substances, besides PM 10, at above mentioned measuring stations. These air polluting substances are: particulate matter PM 2.5, nitrogen(II) oxide, nitrogen(IV) oxide, sulfur(IV) oxide, carbon(II) oxide, benzene, ozone (measured in air) and lead, cadmium, nickel, arsenic, benzo(a)pyrene (measured in particulate matter PM 10). The data collected during the measurements are continuously displayed on the website of the Voivodship Inspectorate of Environmental Protection in Krakow (www.krakow.pios.gov.pl).

## RESULTS OF PREDICTION OF PARTICULATE MATTER PM 10 CONCENTRATION LEVELS

In this section, we present the results to demonstrate the effectiveness of ANNs for solving the prediction of particulate matter PM 10 concentration levels. In our studies the input data set consists of meteorological data (maximum, minimum and average temperature, average wind speed, average temperature of the previous day) and average daily concentrations of particulate matter PM 10 of the previous day. The used neural network models have one output – predicted concentrations of PM 10. It is important to notice that for all experiments the data are randomly partitioned into three separate subsets: 75% for training subset, 15% for validation one and 15% for testing one.

Various types of neural network architectures were constructed and tested in order to find the best network for each measurement station. In all the tested cases, the MLP networks with one hidden layer achieved better results than the RBF networks. The best results were achieved when the MLP networks were trained by means of back-propagation algorithm (BP).

**Table 1**. The characteristics of the best neural networks

| Measurement station | Neural network type (number of neurons) | Training algorithm | Correlation coefficient | Average difference between the expected value and the predicted one [µg] |
|---|---|---|---|---|
| Krasinski Avenue | MLP (6–13–1) | BP | 0.908 | 12.64 |
| Nowa Huta | MLP (6–18–1) | BP | 0.921 | 9.92 |
| Kraków-Kurdwanów | MLP (6–15–1) | BP | 0.933 | 9.89 |

Detailed results of conducted tests are given in Table 1. The second column gives the best ANNs with the number of neurons in three layers (input, hidden and output). From Table 1 one can see that the use of BP algorithm gives the correlation coefficients above 0.9 for each measurement station. The smallest average absolute value of difference between the expected (real) value and the predicted value (for testing subset) was 9.89 µg for Kraków-Kurdwanów Station. The biggest one was equal 12.64 µg for Krasinski Avenue Station.

Figures 1–3 illustrate the ability of the MLP networks for the prediction of particulate matter PM 10 concentrations. The blue lines show the expected values of PM 10 concentrations, the red lines show predicted values. As we can see the neural networks have received a good fit in all cases.
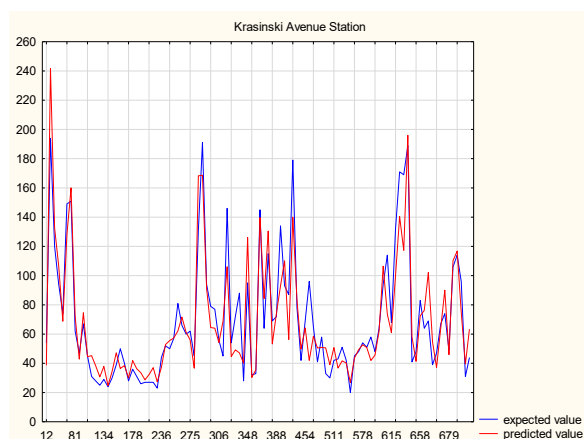
It can be observed that the distributions of the differences between the expected value and the predicted value are similar for each measurement station. The most of the differences are in the range of -10 and 10 µg. Experiments indicate that 57.7% of the results obtained for the Krasinski Avenue Station is within this range. These values are 73.3% and 61.5% for the Kraków-Kurdwanów Station and Nowa Huta Station respectively. The exemplary histogram of mentioned differences we can see in Figure 4.

It is also important to note that for each measurement station, the differences between the expected value and the predicted value are smaller for PM 10 concentrations of the lower ranges. Simultaneously, the cases with PM 10 concentrations below 100 µg predominate in the data set used in the learning process.
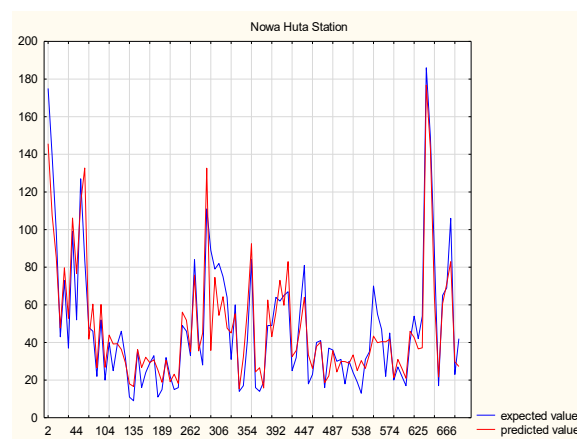


**Figure 1**. Real and predicted values of PM 10 concentrations obtained from the MLP model for Krasinski Avenue Station.



**Figure 2**. Real and predicted values of PM 10 concentrations obtained from the MLP model for Nowa Huta Station.
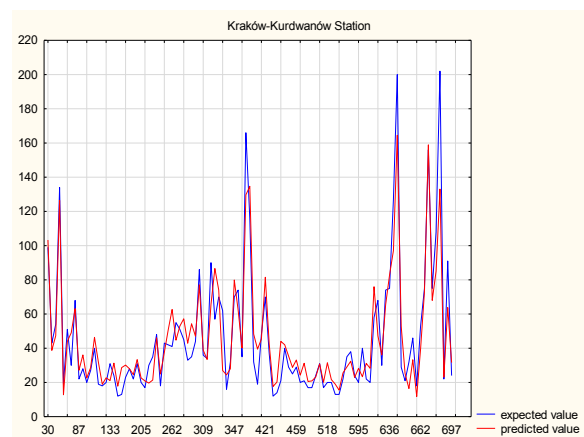


**Figure 3**. Real and predicted values of PM 10 concentrations obtained from the MLP model for Kraków-Kurdwanów Station.
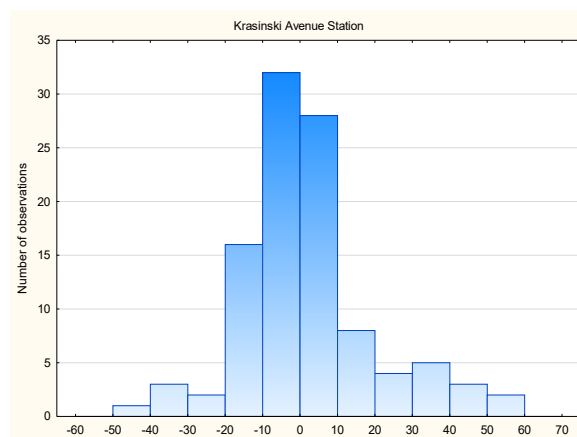


**Figure 4.** Differences between the expected value and the predicted value for Krasinski Avenue Station [µg]

## CONCLUSIONS

The construction of a simple and an effective tool for air quality forecasting is highly desirable. As our research has shown, the artificial neural networks are good tool in the forecasting of PM 10 concentrations. The issue of prediction of PM 10 and other air pollutants concentrations by means of ANNs was undertaken by many researchers. They used various parameters as input data adapted to output data and local conditions. Proper selection of input and output data with clear dependence between them is necessary to get good results. It is not possible to construct one universal model of the neural network, which will allow for prediction of various pollutants in different areas. The network must be designed and trained individually for each case.

We focused on forecasts of the PM 10 concentrations because of the permissible level of this pollutant is often exceeded in Krakow. It turned out that, for our conditions, the best results were obtained by means of three-layer perception with back-propagation algorithm. The neural networks received a good fit in all cases. The correlation coefficients were above 0.9 for each measurement station. The distributions of the differences between the expected value and the predicted value were similar for each measurement station. The average absolute values of mentioned differences were in the range of 9.89 μg and 12.64 μg. Referring to the presented results we can conclude that the performance of the MLP networks are satisfactory.

### Acknowledgements

## REFERENCES

1. Abderrahim H., Chellali M. R., Hamou A., 2016, Forecasting PM10 in Algiers: efficacy of multilayer perceptron networks, Environmental Science and Pollution Research, 23, 1634–1641

2. Bishop C., 1995, Neural Networks for Pattern Recognition. Oxford University Press

3. Chellali M. R., Abderrahim H., Hamou A., Nebatti A., Janovec J., Artificial neural network models for prediction of daily fine particulate matter concentrations in Algiers, Environmental Science and Pollution Research (published online 4 .04.2016)

4. Hajto M. J., Godłowska J., Kaszowski W., Tomaszewska A. M., 2012, FAPPS Forecasting of Air Pollution Propagation System – assumptions, capabilities, development , Ochrona powietrza w teorii i praktyce. T. 2 (red. Konieczyński J.), 89–96

5. Haupt S. E., Pasini A., Marzban C., 2009, Artificial Intelligence Methods in the Environmental Sciences. Springer

6. Kukkonen J., Partanen L., Karppinen A., Ruuskanen J., Junninen H., Kolehmainen M., Niska H., Dorling S., Chatterton T., Foxall R., Cawley G., 2003, Extensive evaluation of neural network models for the prediction of $NO_2$ and PM 10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki, Atmospheric Environment, 37, 4539–4550

7. Kwiecień J., Pawul M., 2012, Application of artificial neural networks to spring water quality prediction, Polish Journal of Environmental Studies, 21, 5A, 271–275

8. Krawczykowska A., Trybalski K., Krawczykowski D., 2009, Application of neural networks models to lithological composition determination of copper ore., Górnictwo i Geoinżynieria 33 (4), 141–151 [in Polish].

9. Manahan S.E., 2006., Environmental toxicology. Chemical and biochemical aspects. PWN, Warszawa [in Polish]

10. Miśkiewicz P., 2014, Pollution impact on health, WHO Country Office in Poland.Warszawa (https://www.mos.gov.pl) [in Polish]

11. Rozporządzenie Ministra Środowiska z dnia 24 sierpnia 2012 r. w sprawie poziomów niektórych substancji w powietrzu (Dz.U. 2012 poz. 1031)

12. Russo A., Raischel F., Lind P. G., 2013, Air quality prediction using optimal neural networks with stochastic variables, 79, 822–830

13. Singh K.P., Basant A., Malik A., Jain G., 2009, Artificial neural network modelling of the river water quality – A case study, Ecological Modelling 220, 888–895

14. Skrzypski J., Jach-Szakiel E., 2008, Prediction of air quality class state as an instrument of ecological safety management in urban-industrial agglomerations, Current problems in atmospheric air protection, Musialik-Piotrowska A., Rutkowski J. (red), 151–154 [in Polish]

15. Sobczyk W. (red), Pawul M., Śliwka M. et al., 2014, Selected issue of environmental protection and environmental engineering, Wydawnictwa AGH, Kraków [in Polish]

16. Szczepańska J., Kmiecik E., 2001, Application of the neural network for the mine waste disposal impact on water environment, Współczesne Problemy Hydrogeologii 10 (2) [in Polish]

17. Tadeusiewicz R., 1993, Neural Networks. Akademicka Oficyna Wydawnicza, Warszawa 1993 [in Polish].

18. Tadeusiewicz R., Dobrowolski J.W., 2004, Artificial intelligence and primary prevention of health hazards related to changes of elements in the environment, Polish Journal of Environmental Studies 13 (3), 349–352

19. Uchwała Nr XXXIV/578/15 Rady Miasta Krakowa z dnia 16 grudnia 2015 r. w sprawie ustanowienia uprawnień do bezpłatnych przejazdów środkami Komunikacji Miejskiej w Krakowie na podstawie dowodu rejestracyjnego samochodu osobowego w sytuacji wystąpienia przekroczenia określonych poziomów szkodliwych substancji w powietrzu.