SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 230

FORECASTING PM_{10} CONCENTRATION USING ANN NARX FROM ROAD TRANSPORT IN SZCZECIN, POLAND

Wojciech KONICKI[1*], Milena BOJANOWSKA[2], Anna Eliza WOLNOWSKA[3], Karolina NADOLSKA[4], Urszula GABRIEL-PÓŁROLNICZAK[5], Anna KOKORNIAK[6]

[1] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; w.konicki@pm.szczecin.pl, ORCID: 0000-0002-5396-1257
[2] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; m.bojanowska@pm.szczecin.pl, ORCID: 0000-0003-0149-2281
[3] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; a.wolnowska@pm.szczecin.pl, ORCID: 0000-0003-4074-5747
[4] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; k.nadolska@pm.szczecin.pl, ORCID: 0000-0002-5801-9287
[5] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; u.polrolniczak@pm.szczecin.pl, ORCID: 0000-0002-7371-2231
[6] Maritime University of Szczecin, Faculty of Engineering and Economic of Transport; a.kokorniak@pm.szczecin.pl, ORCID: 0000-0001-8091-1805
* Correspondence author

Purpose: The aim of the article is to examine the possibility of using artificial neural networks to predict the concentration of PM_{10} from road transport on the example of the urban agglomeration in Poland, Szczecin.
Design/methodology/approach: Forecasting was carried out using measurement data for the concentration of particulate matter PM_{10} and meteorological parameters in the form of average daily air temperature, average daily atmospheric pressure, average daily relative humidity, average daily wind speed and daily precipitation (rainfall and snowfall). The data covered the period from January 1, 2020 to December 31, 2023. For the predictive analysis, NARX neural networks with the Levenberg-Marquardt gradient algorithm with four different architectures were used (5-5-1, 5-10-1, 5-20-1, 5-50-1), which differed in the number of neurons in the hidden layer. The Pearson's linear correlation coefficient (R) and mean squared error (MSE), were used to assess the models.
Findings: It was found that architecture 5-20-1 with 20 neurons in the hidden layer obtained the highest value of correlation coefficient for overall regression, equal to 0.7379. At the same time, the 5-20-1 architecture achieved also the best performance in the learning process, obtaining the lowest MSE value of validation equal to 0.00596. Experimental results demonstrated that NARX neural network with the Levenberg-Marquardt algorithm for tested architectures, showed a short time in the learning stage.
Research limitations/implications: In order to improve the quality of the model in subsequent studies, it would be advisable to introduce additional independent variables as chemical factors. A future research to improve the model may focus on periodic real-time data updates, and to use multiple measurement stations.

Social implications: The research will allow local authorities to manage air quality in the city by making decisions in the area of legal regulations and health safety of the local community. The forecasting of PM_{10} concentration can provides the public with the information with which they can make lifestyle decisions to protect their health.
Originality/value: The originality of the article lies in the possibility of using artificial neural networks to forecast the concentration of PM_{10} from road transport in in urban agglomerations in Poland.
Keywords: PM_{10} forecasting, air quality/pollution monitoring, artificial neural network, road transport, environmental management.
Category of the paper: Research paper.

1. Introduction

For many years, particulate matter pollution is a serious threat to the health and life of the population in urban areas. The topicality of the problem is evidenced by the WHO guidelines released in 2021, which indicate new levels of pollutant concentrations, including PM_{10} particles, below which the risk of negative effects on human health is minimal, i.e. 15 μg/m^3 for the average annual concentration and 45 μg/m^3 for daily concentration. The WHO justifies its actions by the fact that since the last update of the guidelines in 2005, the number of evidence on the adverse impact of air pollution on health, and above all on premature deaths, has increased (World Health Organization, 2021).

The source of particulate and gaseous air pollution in urban areas in Poland are mainly anthropogenic emissions from the municipal and housing sector, industry and means of transport. Although fuel combustion for heating residential, commercial and institutional buildings is the main source of particulate matter in Poland (in 2019, these sources accounted for approximately 67.3% of the national PM_{10} emissions). Automotive sources annually account for about 10-12% of total PM_{10} emissions in Europe and about 7% in Poland (European Environment Agency, 2021; Główny Urząd Statystyczny, 2021).

In the environment of urban-industrial agglomerations, the coarse dust fraction (PM_{10}) is emitted into the atmosphere as primary pollution released by stationary and mobile sources in the form of particulates, dust, soot and road dust undergoing resuspension. The automotive sources of particulate matter emission, in addition to particulate matter from internal combustion engines, are also solid materials formed as a result of tire friction with the road surface, as well as in the braking system. The emission of particulate matter, next to the emission of nitrogen oxides, is currently considered the most serious ecological problem of the automotive industry and a significant cause of air pollution in urban centers. In the context of urban air pollution, the factors influencing the particulate matter concentration from automotive sources include fuel and engine type, as well as vehicle type in fleet classification. A large part of the vehicles plying in the city corridors i.e. buses, light and heavy duty vehicles etc., are equipped with diesel fuel engines, which are known to emit more particulate matter than spark-ignition gasoline engines (Zhang et al., 2009; Srimuruganandam, Shiva Nagendra, 2012). Combustion engines, especially those with self-ignition, emit solid particles with dimensions below 10 μm, consisting of a carbon matrix (soot), organic and inorganic fractions, both containing compounds such as: benzo-a-pyrene, benzo-a-anthracene, benzo-f-fuloranthene and heavy metals (Canagaratna et al., 2004; Polkowska et al., 2007), which have detrimental effect on the environment and human health. High mutagenicity of particulate matter in the air has been demonstrated, mainly due to the content of Polycyclic Aromatic Hydrocarbons (PAHs), including mono- and bicyclic ones, aliphatic hydrocarbons, cycloalkanes and organic compounds containing metals, sulphur, chlorine and oxygen in their structure (Kuchcik, 2020; Costa, 2017). Inhalation of particulate matter has harmful effects on both the lungs and the cardiovascular system, causing or exacerbating diseases such as asthma, atherosclerosis, chronic obstructive pulmonary disease, ischemic heart diseases and arrhythmias. This could also cause brain damage (Chrabołowska et al., 2024; Li et al., 2022). Studies indicate that combined exposure to particulate matter ($PM_{2.5}$ and PM_{10}) and polycyclic aromatic hydrocarbons (PAHs) may have a negative synergistic effect on human health, including increased cancer incidence and mortality and increased oxidative stress (Wu et al., 2021; Akhbarizadeh et al., 2021; Bae et al., 2010).

The problem with maintaining local PM_{10} concentrations at a given level is related to the significant impact of factors that go beyond the primary sources of emissions that can be controlled. Such uncontrollable factors include meteorological parameters such as air temperature, relative humidity, wind speed and direction, radiation, precipitation and pressure, that have an extremely significant impact on the distribution of pollutant concentrations in the air. Consequently, for various places in the world many authors addressed the subject of the influence of atmospheric factors on the level of immission of particulate matter (PM) in the atmosphere. Ganguly and colleagues assessed the impact of NO_2 and SO_2 pollutants, as well as meteorological parameters such as wind speed and temperature, precipitation and relative humidity on the concentration of PM_{10} particles. Research was conducted to assess long-term PM_{10} trends based on data from 2011-2017 for two city monitoring stations in Shimla, India (Ganguly et al., 2019). Kirešová and Guzan (2022) conducted comparative studies on the concentration of PM particles in two locations in Slovakia, in Košice (city) and in a small village located 35 km from Košice. On the basis of PM concentration measurements from January, March and May 2022, an analysis of the correlation between PM_{10} particles and meteorological factors such as temperature, humidity and pressure was performed. For Poland, the impact of meteorological conditions on the variability of PM_{10} and $PM_{2.5}$ concentrations in the Poznań, Łódź, Kraków and Tricity agglomerations, for the period from 2006 to 2016, was studied by Jędruszkiewicz et al. (2017). Statistical analysis was performed on an hourly, monthly, annual and daily scale, considering the frequency of exceedances of permissible concentration thresholds. Correlations between PM and selected meteorological variables such as mean daily dew point, planetary boundary layer, sea level pressure, mean temperature, low cloud cover, total cloud cover, wind speed and precipitation were investigated.

One of the purposes of determining the influence of atmospheric factors on the concentration of pollutants in the form of PM in the atmosphere is the possibility of predicting pollutant immissions. One of the computational techniques used in this approach is advanced statistical analysis based on neural networks that allow for the representation of complex non-linear time series functions. Until now, many authors have used artificial neural networks with different architectures to predict PM immissions depending on atmospheric parameters. In order to assess and forecast air quality in the city of Amman, Jordan, Hamdan et al. (2021) used a Nonlinear Autoregressive Exogenous Model (NARX) of the neural network with the Marquardt-Levenberg learning algorithm. The forecasting concerned the PM_{10}, SO_2, O_3, CO and NO_2 pollution for meteorological data from 2015-2017, which were used as predictors to train the artificial neural network, while data from 2018 was used to test it. Cortina-Januchs et al. (2015) proposed a model based on an artificial neural network for predicting PM_{10} concentrations for the city of Salamanca, Mexico. The forecasting model was based on the relationship between PM_{10} concentration and meteorological variables in the form of wind speed and direction, temperature and relative humidity. Studies to forecast PM_{10} concentrations one day in advance based on observations from 2007-2014 for the Düzce province in Turkey were conducted by Taspinar (2015). Forecasting was carried out using artificial neural networks based on variables in the form of averaged daily values of air temperature, wind direction, wind speed, relative humidity and PM_{10} concentration. Perez and Reyes (2006) developed an integrated model of an artificial neural network for forecasting maxima of average daily PM_{10} concentrations one day in advance, based on data obtained from five monitoring stations in the city of Santiago, Chile. Nejadkoorki and Baroutian (2012) proposed a model based on a feedback neural network with back-propagation and a Marquardt-Levenberg learning algorithm, predicting average daily PM_{10} concentrations in the Tehran metropolitan area. Modeling was based on a set of data in the form of average solar radiation, wind direction and speed, NO and CO concentrations, as well as average and maximum PM_{10} concentrations and temperature, from 9 air monitoring stations in the years 2001-2009. Studies in the field of forecasting the concentration of particulate matter (PM) using artificial neural networks were also conducted by others (Rumaling et al., 2022; Chellali et al., 2016).

To date, there have not been many works devoted to forecasting the concentration of PM particles in the atmosphere in Poland. A comparison of models for predicting the level of PM_{10} in the air in the city of Lublin was presented by Kujawska et al. (2022). In addition to artificial neural network (ANN), linear regression (LR), K-nearest neighbor regression (KNNR), support vector machines (SVM) and others were used in the work. Data from January 2017 to December 2019 was used to develop the models, covering meteorological variables in the form of temperature, relative humidity, wind speed and direction, and concentrations of chemical compounds SO_2, PM_{10}, NO_2, NO_x, CO, O_3 and C_6H_6. In their research, Pawul and Śliwka (2016) used artificial neural networks to predict PM_{10} concentrations as factors determining the occurrence of smog phenomena. A set of input data consisting of meteorological data (maximum, minimum and average temperature, average wind speed, average temperature from the previous day) and average daily PM_{10} concentrations was used to create the neural networks. The data was recorded in 2014 and 2015 at three measuring stations operating in Krakow.

The aim of the work is to examine the possibility of using artificial neural networks to forecast the concentration of PM_{10} from road transport in one of the largest cities in Poland, Szczecin. Forecasting was carried out using measurement data for the concentration of particulate matter PM_{10} and meteorological parameters in the form of average daily air temperature, average daily atmospheric pressure, average daily relative humidity, average daily wind speed and daily precipitation (rainfall and snowfall). The data covered the period from January 1, 2020 to December 31, 2023. NARX neural networks with four different architectures were used for predictive analysis.

2. Materials and methods

2.1. Research area

Szczecin is the largest city in north-western Poland, one of the countries bordering the Baltic Sea. It is inhabited by approximately 401,990 people (Urząd Statystyczny w Szczecinie, 2020). It is located right next to the Polish-German border at the mouth of the Oder River with the following coordinates: 14°33′10″E east longitude and 53°25′44″N north latitude (Instytut Meteorologii i Gospodarki Wodnej-Państwowy Instytut Badawczy, 2020-2023). The area of the city is 300.55 km^2, of which 1202.8 km^2 is covered by land under water. The city of Szczecin is located in the Szczecin Coastland macroregion, which consists of eleven mesoregions. Szczecin is located in the area of four of them: Lower Odra River Valley, Goleniów Plain, Szczecin Heights (consisting of the Warszewskie Hills and the Stobno Embankment) and Bukowe Hills. Szczecin's natural hydrographic systems are formed by two arms of the Oder River: the Eastern Oder (Regalica) and the Western Oder. Standing water reservoirs are also an important complement. The largest of them are Dąbie Lake and Głębokie Lake. The described location and the presence of large industrial and production plants in the city and its vicinity have an impact on the climate of Szczecin. Mild winters and quite cool, humid summers are also due to the influence of the oceanic and continental climate. South-westerly winds prevail in the city, occurring mainly in June, July, September, November and December. In April and May, easterly and north-easterly winds prevail (Instytut Meteorologii i Gospodarki Wodnej-Państwowy Instytut Badawczy, 2020-2023).

Szczecin is a city with a well-developed port and industrial function, where, due to its geographical location, transit routes from the north to the south of Europe and from the west to the east intersect. As an object of research, it is a representative of a medium-sized city, through which run the communication routes important for international transit, including land and sea transport. Urban transport in Szczecin is based mainly on the use of buses, trams and taxis. It is complemented by ecological means of transport such as scooters, bicycles and electric scooters. Currently, there are 81 bus lines and 14 tram lines in Szczecin with a total length of 118.7 km, on a network of 54.5 km length. A significant number of residents still travel by their own means of transport, by car. In 2020, 236,707 passenger cars and 32,905 trucks were registered in Szczecin. There were 594.4 passenger vehicles and 95 trucks per 1000 inhabitants (Główny Urząd Statystyczny, 2022).

2.2. Data used in the analysis

In order to carry out the analysis, annual measurement data for the concentration of PM_{10} particles and meteorological parameters were used. The data covered the period from January 1, 2020 to December 31, 2023. The concentration of PM_{10} particles was obtained for the measurement station no. 1 in Szczecin from the Internet database belonging to the Chief Inspectorate for Environmental Protection, which is an office of the Polish government administration (Główny Inspektorat Ochrony Środowiska, 2022). The measuring station is located in the city center at the crossroads of Piłsudskiego St. and Wyzwolenia Ave. (Figure 1) and performs continuous measurements, with 1 and 24 hour averaging times, for pollutants from road transport. It is a station classified as "urban traffic", with the following geographical coordinates: latitude 53.432169 and longitude 14.553943 (international code: PL0249A). The concentration of PM_{10} particles is presented in the μg/m^3 unit as an average daily value.

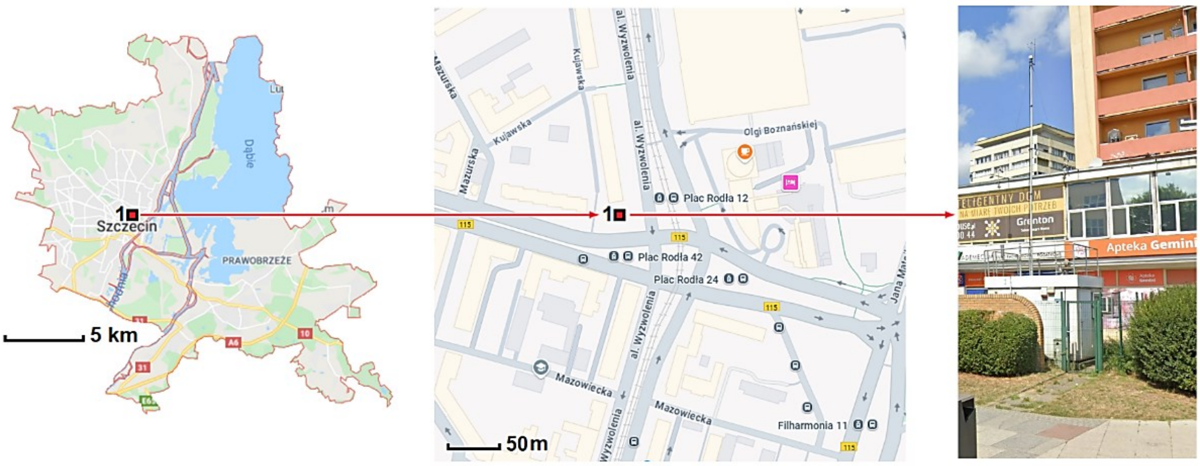

Figure 1. Map of the study area with measurement point.

Source: Own creation based on Google Maps.

Meteorological parameters in the tests were used in the form of average daily air temperature (T, °C), average daily atmospheric pressure (P, hPa), average daily relative humidity (H, %), average daily wind speed (Ws, m/s) and daily precipitation (rainfall and snowfall) (P_{RS}, mm). All meteorological data were obtained on the basis of the Meteorological Yearbook 2020-2023, published by the Institute of Meteorology and Water Management - National Research Institute in Poland (Instytut Meteorologii i Gospodarki Wodnej-Państwowy Instytut Badawczy, 2020-2023).

2.3. Methods

2.3.1. Data preprocessing

In our research the collected data has some missing values in the area of PM_{10} concentration. The remaining parameters in the form of air temperature, atmospheric pressure, relative humidity, wind speed and precipitation were complete. Taking into account, that the imputation of missing values will not lead to accurate results since air pollution is not predictable, to fully retain the representative character of the models the missing values were excluded from the study.

Since meteorological parameters and concentration of PM_{10} have different units of measurement and the orders of magnitude of the measured values, so it is necessary to normalize the data. Therefore, the scaling of the original data using the Min-Max method was performed, in which the final values are in the range [0,1]. Normalization is carried out according to equation (Ibrahim, Wahab, 2022):

$$X' = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where: X' is the normalized value, X_i is the original value, while X_{min} and X_{max} are the smallest and largest values from the input data set, respectively.

2.3.2. Prediction model

In the work, ANN artificial neural networks of NARX type were used to forecast the concentration of air pollution in the form of PM_{10} from road transport. Forecasting was carried out by taking into account the impact of complex relationships between input data, such as meteorological conditions (air temperature, atmospheric pressure, relative humidity, wind speed, precipitation), on the state of atmospheric pollution by PM_{10} particles.

The concentration of PM_{10} particles in the atmosphere in the time domain is a non-linear dynamic system. Therefore, the study used a neural network model for solving problems in the area of non-linear dynamic systems of time series, i.e. a non-linear autoregressive model with an external NARX (Nonlinear Autoregressive Exogenous Model) input. Due to its features, such as good dynamic properties and strong anti-interference capabilities, NARX is one of the most frequently used models in many areas to describe non-linear dynamic systems (Boussaada et al., 2018; Haris et al., 2022; Magallanes-Quintanar et al., 2022; Riverol, Harrilal, 2018; Zheng et al., 2022).

The NARX Neural Network is a Recurrent Neural Network (RNN) that performs one-step ahead predictions of a discrete non-linear system based on the system's current and previous inputs and outputs. The network is recursive because the current network output is fed back as network input in the next time step so that the network can make subsequent predictions one step ahead. In this model, the variables affecting the value of the time series to be predicted are exogenous variables.

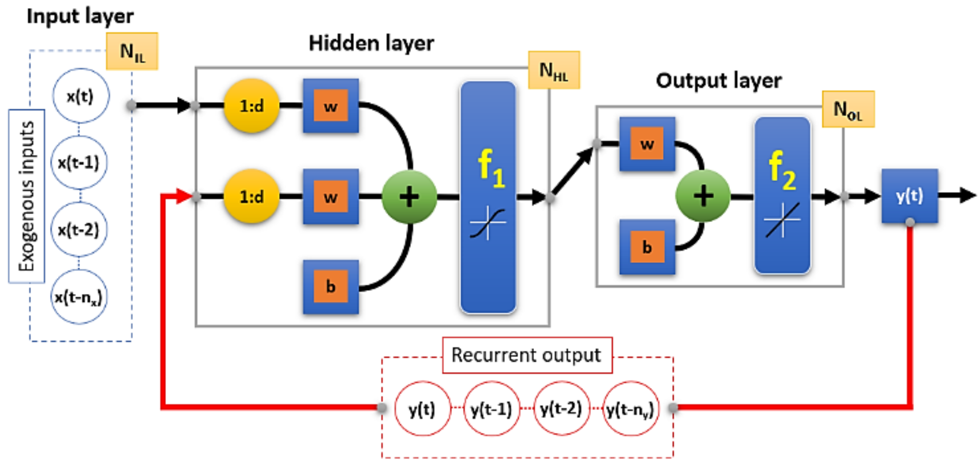

Figure 2. General structure of the NARX ANN.

Source: Own creation.

Figure 2 shows the general diagram of architecture for the NARX ANN. The NARX architecture consists of three layers: the input layer, the hidden layer, and the output layer. Each layer has an appropriate number of neurons, depending on the system under study. The input layer consists of current and previous inputs and outputs, with a number of N_{IL} neurons. The exogenous input $x(t\text{-}n_x)$ represents the time series of data with time step (t), used to estimate the output parameter $y(t\text{-}n_y)$, whose past values are also entered into the hidden layer. In the hidden layer, consisting of one or more neurons (N_{HL}), the exogenous data stream and output parameters are weighted (w), and modified with appropriate deviations (b). The parameter (d) defines the degree of delay. These results are then, in the standard NARX network, filtered by the sigmoid neuron activation function f_1, which is used to create a newly trained stream model, and a linear transfer function f_2 in the output layer with a number of N_{OL} neurons. The output layer leads to the obtaining the predicted value of y(t), which is then passed back to the hidden layer.

The model of the NARX neural network in mathematical form can be written as follows (Liu et al., 2020):

$$y(t) = f\left[y(t-1), y(t-2), \dots, y\left(t-n_y\right), x(t-1), x(t-2), \dots, x(t-n_x)\right] + e(t) \quad (2)$$

where: x(t) and y(t) are the input and output values of the network in time t (time step), respectively, n_x and n_y are the input and output layers of the network, respectively, e(t) represents noise signals, and f(.) is an unknown non-linear function.

Choosing for our study a non-linear autoregressive neural network with exogenous input (NARX) was dictated by its wide use and effectiveness for forecasting of time-series data in different areas of knowledge (Al-Sbou, Alawasa, 2017; Khan et al., 2014; Ghahari et al., 2012; Pisoni et al., 2009; Buevich et al., 2021). It was found, that the NARX neural network model outperforms other models in forecasting and is desired for time series with such characteristics as unpredictable and chaotic nature (Paul, Sinha, 2016; Yu et al., 2019; Peña et al., 2020). Due to its properties, the model has also been used in the field of environmental pollution for forecasting particulate matter concentration in the air (Kaur, Mandal, 2020; Rumaling et al., 2022; Adnane et al., 2022; Gündoğdu, 2020; Wiktorzak, Sawicki, 2023; Gündoğdu, Elbir, 2024).

3. Results and discussion - forecasting PM_{10} concentration using NARX Neural Network Model with Levenberg-Marquardt Algorithm

The study was conducted on the basis of the data described in section 2.2. All data were initially subjected to the normalization process. Modelling using artificial neural networks of the NARX type was carried out using the MATLAB R2021a software.

Figure 3 shows the architecture of the NARX network used in the research, which consists of three layers: input layer with 5 neurons, a hidden layer, and output layer with 1 neuron. The performance of the NARX network and the accuracy of the results is primarily determined by three hyperparameters: input delay, output delay, and the neuron number in the hidden layer (Zeng et al., 2024). This study assumes, that the number of input delays and number of feedback delays is the same and equal to 2. One of the fundamental issue in supervised machine learning is overfitting, which prevents from perfectly generalizing the models to well fit observed data on training data. Overfitting can occur when there are too many neurons in the network's hidden layer, leading the model to become highly trained in fitting the training data and hence performing poorly when tested with unknown data (Basheer, Hajmeer, 2000). Therefore, the study of the influence of the hyperparameter in the form of the number of neurons in the hidden layer on the accuracy of the model was conducted for 5, 10, 20 and 50 neurons.

Another important hyperparameter is the number of epochs, where an epoch is the number of iterations until the MSE (mean squared error) reaches minimum value. The epochs in this study were set to 1000, which allowed for visualization of the results and identify the point at which over-fitting formation occurred.

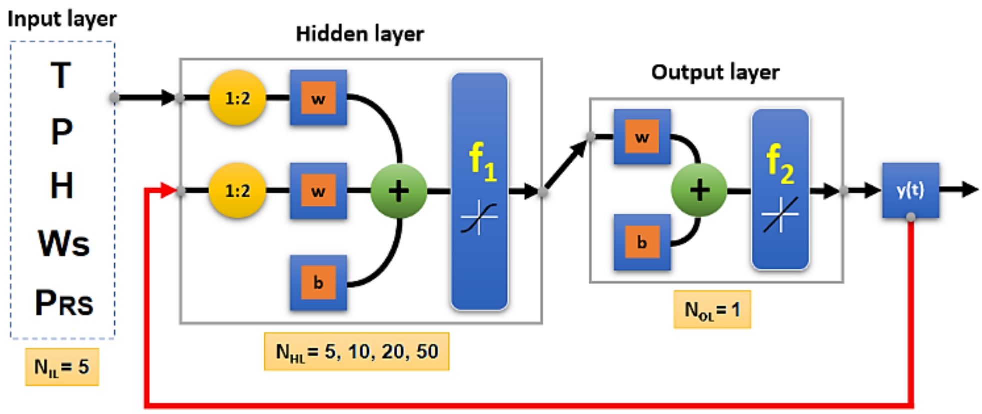

Figure 3. Scheme of the architecture of the NARX model used in the research.

Source: Own creation.

The Levenberg-Marquardt algorithm was used in the training process. This algorithm was chosen due to its undoubted advantages, which include: giving a solution even though its start-point is far from the final minimum, its processing time is one of the lowest compared to other algorithms, the training algorithm stops when it finds the maximum epoch and the best performance value is achieved, or when it finds that the gradient value is lower than its minimum (Viejo et al., 2019). The total number of input variables as independent parameters x(t) = [T, P, H, Ws, P_{RS}] and the output variable as dependent parameter y(t) = [PM_{10}] is 1319, of which 70% was used for network training, 15% for model validation and 15% for testing. The quality of the model was tested using a standard error evaluation procedure and the results are tabulated for comparison. For this purpose, the accuracy of the tested models was assessed on the basis of the Mean Squared Error coefficient (MSE) and the correlation coefficient (R). Mean Squared Error represents the average of the squared difference between the original and predicted values over the data set. It is a measure of how close a fitted line is to actual data points. When the MSE is closer to zero then the performance of the model is more favorable. The correlation coefficient R, also known as the Pearson's correlation, is a measure of the strength of a linear association between two variables and can take on values between -1 and 1. A perfect linear relationship is for R = -1 or R = 1. An R value of 0 denotes no linear correlation.

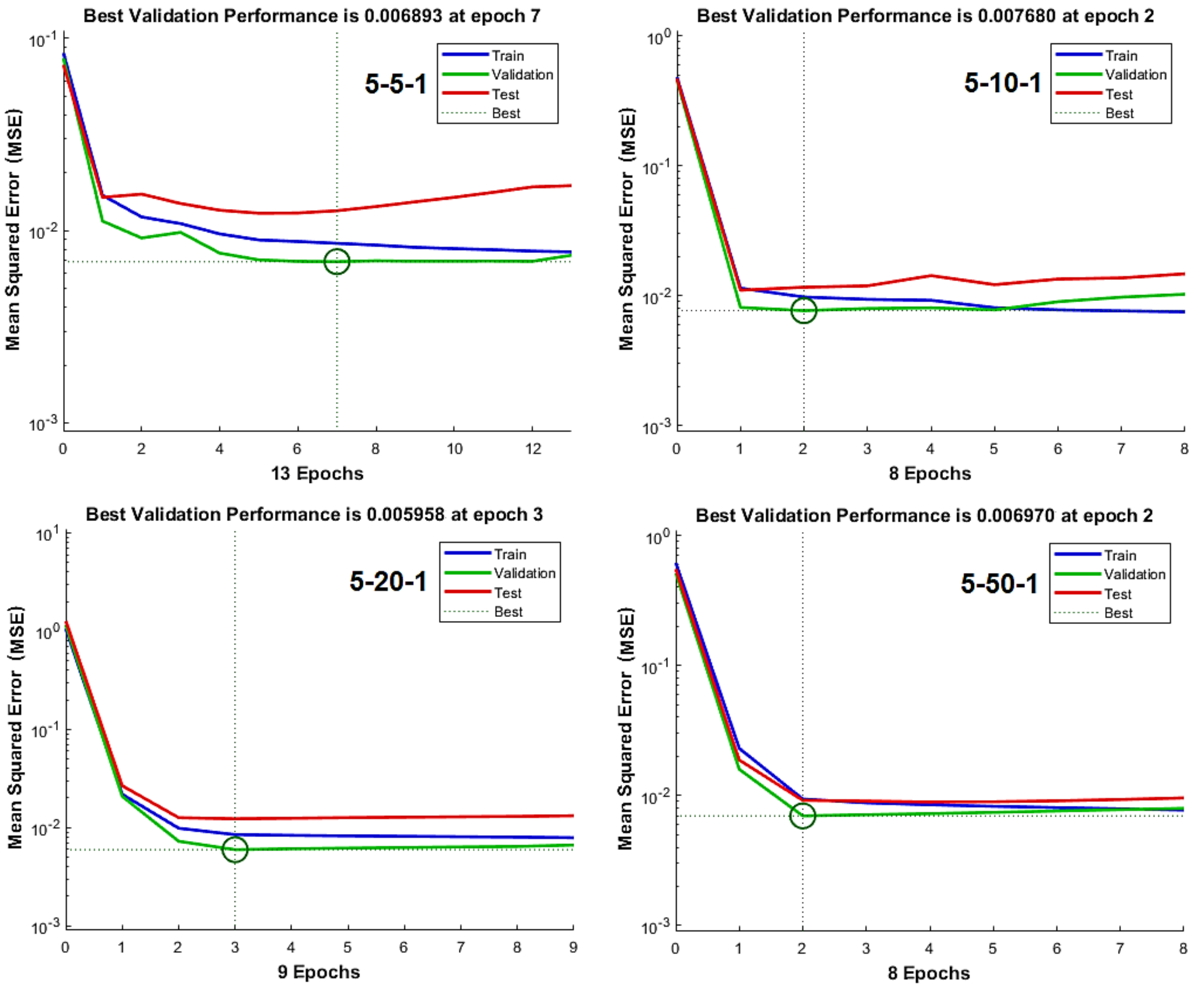

Figure 4. Learning performance for models 5-5-1, 5-10-1, 5-20-1 and 5-50-1.

Source: Own study.

Figure 4 shows a graph presenting the learning performance of the tested models, trained with the Levenberg-Marquardt method, with the 5-5-1, 5-10-1, 5-20-1 and 5-50-1 architecture. The performance function is shown as the mean squared error on the ordinate axis, while the respective learning epochs of the network are recorded on the abscissa axis. Lines on the graphs in blue, green and red indicating training, validation and testing, respectively. The figure shows that the best validation of the neural network performance was achieved for the 5-20-1 system in epoch 3. At the beginning the model started training with a high MSE value, which decreased until the 9 epoch, but the testing and validation errors were increasing from 2 and 3 epoch, respectively. Therefore, after 9 epoch, the model training was stopped, and an optimized model was produced with minimum value of MSE for validation, equal to 0.00596. In the case of the others architectures (5-5-1, 5-10-1, 5-50-1), higher MSE values of validation performance were obtained.

Figure 5 shows the correlation of training, validation, testing, and model's overall accuracy between the input (target) and output values of the model. The ordinate axis presents the results provided by the neural network for the given input parameters (Output). The abscissa shows the values of the actual result (Target), to which the results provided by the trained neural network should aim. The solid straight lines represent the best fit obtained for the linear regression between the results and objectives of training (blue), validation (green), testing (red), and overall accuracy (black), while the dashed line for each plot represents a perfect fit.

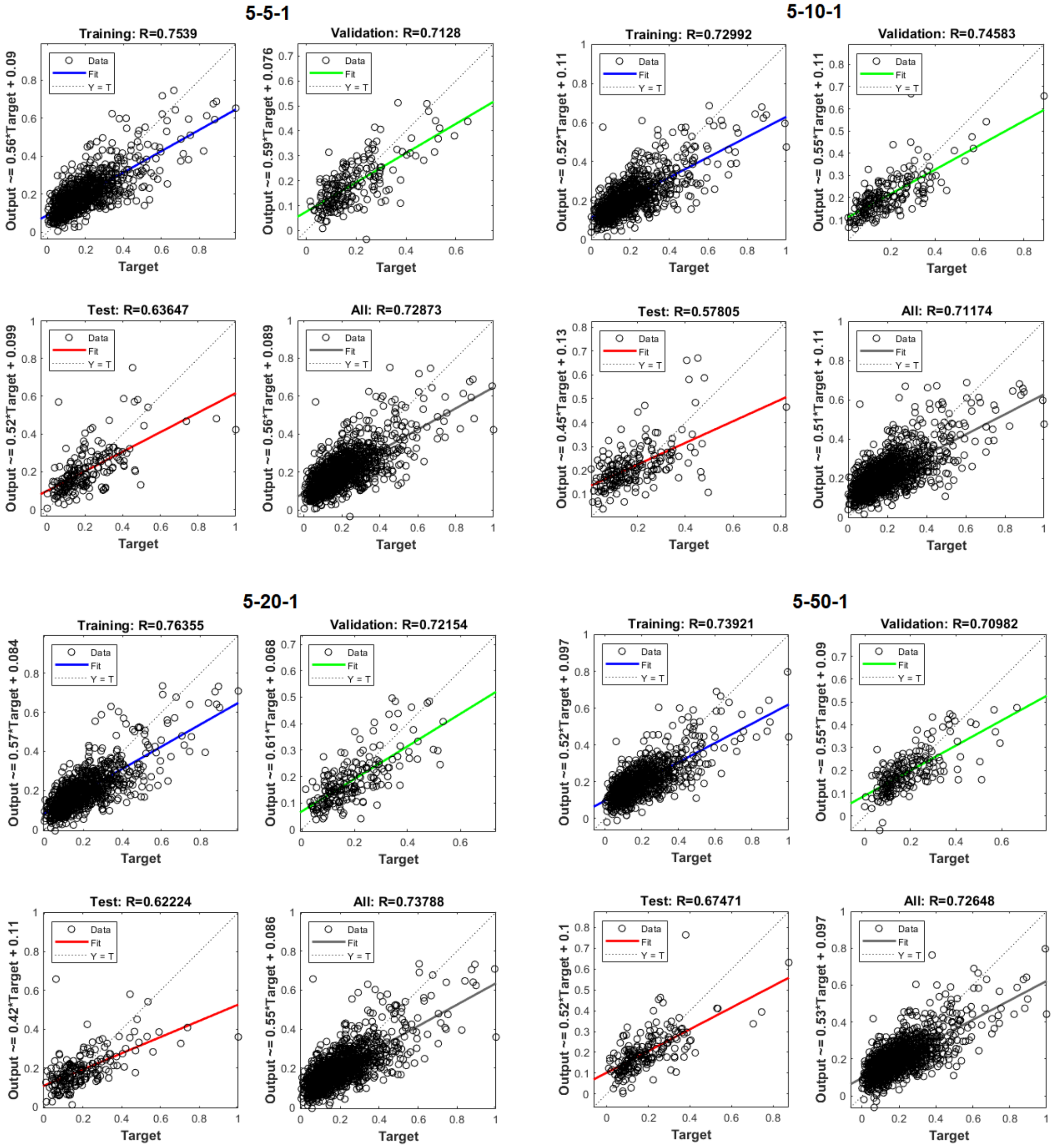

Figure 5. Regression results for training, validation, test and model's overall accuracy for the models 5-5-1, 5-10-1, 5-20-1 and 5-50-1.

Source: Own study.

Table 1.
Comparison of results evaluation parameters for the NARX model

Model (architecture)	Training		Validation		Test		All
	MSE	R	MSE	R	MSE	R	R
5-5-1	0.00858	0.7539	0.00689	0.7128	0.01268	0.6365	0.7287
5-10-1	0.00977	0.7299	0.00768	0.7458	0.01162	0.5781	0.7117
5-20-1	0.00848	0.7636	0.00596	0.7215	0.01230	0.6222	0.7379
5-50-1	0.00939	0.7392	0.00697	0.7098	0.00917	0.6747	0.7265

Source: Own elaboration based on conducted research.

Table 1 presents a comparison of the MSE and R coefficients obtained for the NARX model with Levenberg-Marquardt algorithm for the tested architectures. The selected measure for assessing the quality of PM_{10} concentration forecasts in time series is the MSE coefficient, which, tending to the lowest value, indicates a better representation of the model quality. Whereas in case of testing the strength of the linear relationship between the input (target) and output values of the model, the correlation coefficient R was used. As can be seen, the highest R-value for the model's overall accuracy, equal to 0.7379, was obtained for the model with the 5-20-1 architecture. However, with the increase in the number of neurons in the hidden layer to 50 (5-50-1 architecture), the value of R decreased to 0.7265. At the same time, Pearson's linear correlation coefficient R also reaches the highest value for the 5-20-1 model in the training process, equal to 0.7636. The quality of the model is reflected by the value of the MSE coefficient, which for the 5-20-1 model obtained the lowest values equal to 0.00848 and 0.00596 in the training and validation process, respectively. Considering the above data, it can be concluded that the best forecast results for PM_{10} concentration in the atmosphere were obtained using the 5-20-1 model configuration, which is by far the most stable among those tested.

Comparison of the obtained parameters also allows detecting any overfitting. An indicator of a good model with no overfitting is when the validation correlation coefficient is close to the value from the training stage (Viejo et al., 2019). If there is a large gap between the training and the validation parameters, it is a sign of overfitting. As shown in the table 1, the training and validation correlation coefficients values R, for all tested architectures, are very close. Also similar values of MSE coefficients, for the training and validation stage, were obtained. Simultaneously, the easiest way to detect overfitting is monitoring the gap between training and validation performance using learning curves (plots of training/validation performance over epochs). Generally, overfitting occurs, when training performance continues to improve, while the validation performance worsens. Figure 4 shows a decrease of training MSE value over the entire range for all architectures, what means that the quality of training increases. Meanwhile, after a sharp initial decline, the validation MSE for all architectures slowly begins to increase from a certain point. In order to avoid overfitting, learning performance was determined at the optimal point, where the model achieved the minimum validation MSE (the point marked by a circle in figure 4).

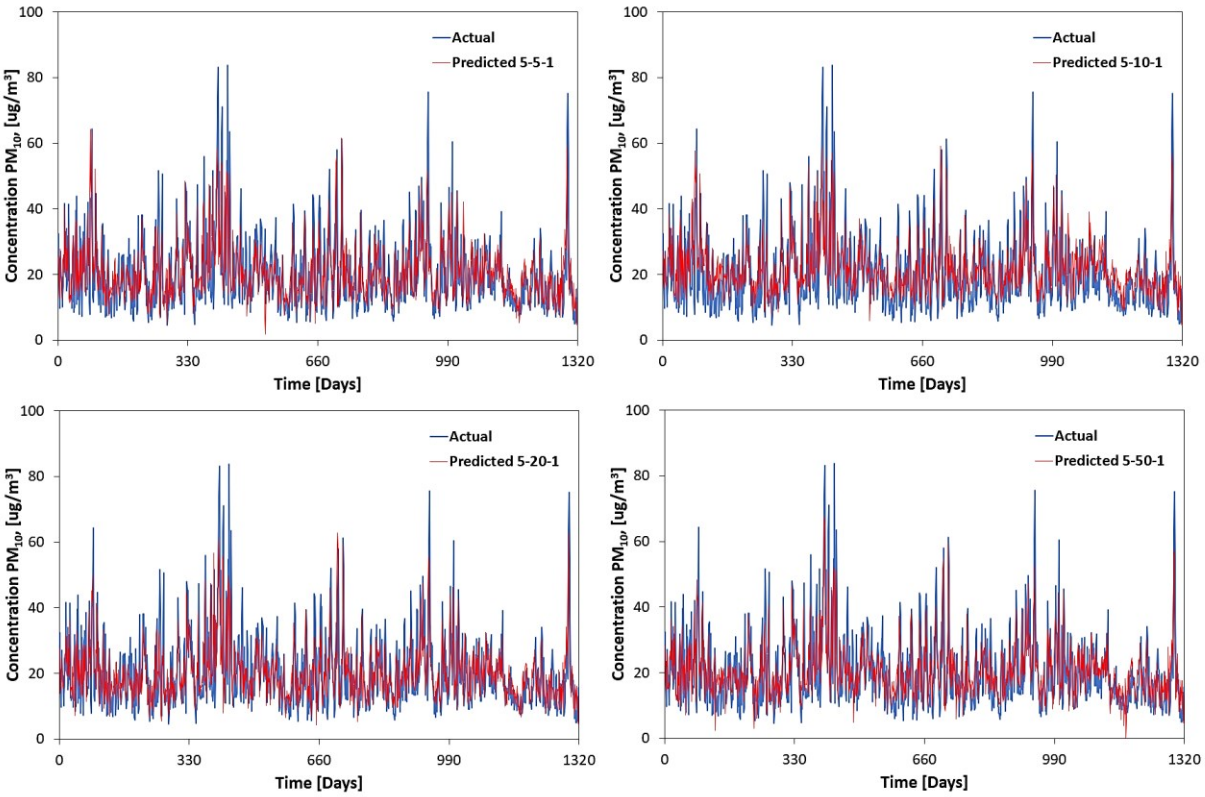

Figure 6. Comparison of PM_{10} concentration values observed and obtained from the 5-5-1, 5-10-1, 5-20-1 and 5-50-1 models.

Source: Own study.

Figure 6 presents a comparison of PM_{10} concentration values recorded in the period from January 1, 2020 to December 31, 2023 with the predicted values obtained from the 5-5-1, 5-10-1, 5-20-1 and 5-50-1 models. According to the values of the MSE (validation) and R coefficients (overall accuracy), the least accurate approximation of the model data with the real ones is reflected by the PM_{10} time series for the 5-10-1 architecture, which is clearly visible through the lack of coverage between the dependencies, mainly in the low PM_{10} concentrations area. The best fit between the model data and the real data is achieved using the 5-20-1 architecture, accurately reflecting the time course of PM_{10} concentration. Building a model that gives a smaller error in the forecasting process in non-linear dynamic systems of time series for atmospheric pollutants is related to the complexity of the system. Achieving a higher correlation between model data and real data is related to the possible use of additional input data in the form of atmospheric factors and chemical factors directly or indirectly affecting the instantaneous concentration of PM_{10} particles (Hamdan et al., 2021).

Considering that public safety and air quality are among the main parameters of the Smart City Index, the possibility of using the ANN NARX model to implement air quality policy in urban agglomerations becomes a very attractive proposition for local governments. This is all the more justified because the idea of a Smart City assumes the use of modern technologies and innovative solutions to improve the quality of residents life. Currently, the rapid expansion of urban agglomerations leads to increased traffic density and road congestion, which in turn leads to an increase in air pollution levels and deterioration in the quality of life in urban areas. Despite the desirability of using the model to forecast PM_{10} particulate matter pollution from road transport, its implementation is associated with certain difficulties. Overcoming these difficulties involves the following activities:

1. To improve the quality of the model, it is necessary to take into account the complexity of the system by introducing additional independent variables.
2. In order to obtain reliable data for over a larger area (to adequately represent the air quality in the city), it is necessary to build a network of stations monitoring PM_{10} concentrations within the city. Simultaneously, the location of measurement points must be justified by the complex topography of the location.
3. Traditional continuous air quality monitoring systems are often very costly. Therefore, it is possible to use of low-cost wireless PM sensors, in the form of a distributed sensor network linked to a cloud system. The system in this form is cost-effective, small-sized, and consumes little power.
4. Use of geostatistical techniques or machine Deep Learning algorithms to investigate the spatial relationships between measurement points (stations).
5. Introduction of the mechanism periodically updating the ANN NARX parameters, in order to better match to the latest air quality data, including integration with real-time data sources (recalibration of the model based on the data available from network stations located at the place of deployment).
6. Should be avoided using external weather stations or weather forecasts. Therefore, each individual measurement point (station) should be equipped with an appropriate number of sensors to allow the measurement of all independent variables. This will allow direct access to data and reduce the data transmission time to the Master Control Unit.
7. Real-time different parameters monitoring data is valuable for constructing innovative studies that respond to dynamic temporal variations. However, missing values are prevalent due to network miscommunication, maintenance, device replacement or failure. The absence of data could cause bias in the statistical analysis, leading to invalid conclusions. Moreover, the lost data makes many data modelling techniques ineffective because they presume complete information for all the variables included. Therefore, it is important to ensure stable operation of the system (network communication-data transmission, power management) with imputation of missing values.

4. Conclusions and future research direction

The possibility of using artificial neural networks gives an accessible tool for obtaining a forecast in the area of information about the state of air quality. It enables further development towards the creation of an integrated air quality supervision system for metropolitan areas such as Szczecin.

In the research, the process of forecasting pollution in the form of PM_{10} particles from road transport was carried out by using artificial neural networks of the NARX (Nonlinear Autoregressive Exogenous Model) type with the Levenberg-Marquardt gradient algorithm. Four models with the 5-5-1, 5-10-1, 5-20-1 and 5-50-1 architecture were tested, which differed in the number of neurons in the hidden layer, by using input parameters from the years 2020-2023. The forecasting process was based on 5 independent variables: air temperature, atmospheric pressure, relative humidity, wind speed and precipitation (rainfall and snowfall). The study of the model quality was based on the analysis of statistical indicators in the form of the mean squared error coefficient (MSE) and the Pearson's linear correlation coefficient (R).

It was found that among the analyzed architectures, the best performance results in the learning proces were achieved by the 5-20-1 architecture, obtaining the lowest MSE values equal to 0.00848 and 0.00596 in the training and validation process, respectively. All correlation coefficients R values for the model's overall accuracy were very similar, however the highest value, equal to 0.7379, was obtained for the model with the 5-20-1 architecture. Experimental results demonstrated that the NARX neural network with the Levenberg-Marquardt algorithm required a very small number of epochs (from 2 to 7), during the learning stage in the tested architectures, which significantly shortened the training time. Furthermore, it was found that the all tested models did not show signs of overfitting.

The results obtained for the PM_{10} pollution forecasting by using NARX artificial neural networks models are promising. The proposed model can be useful not only to predict the concentration of particulate matter, but also to simulate PM_{10} imissions by changing the values of input variables. Since it is not possible to construct one universal neural network model that will allow for the prediction of PM_{10} in different areas and therefore the network must be built and trained individually for each case, the obtained results may be particularly important for local governments responsible for air quality policy. It is possible to use multiple measurement stations to increase the universality of the model, but this requires the comparability and compatibility of data obtained from different stations, as well as the use of geostatistical techniques or machine learning algorithms to investigate the spatial relationships between measurement stations.

Nevertheless, in order to improve the quality of the model in subsequent studies, it is necessary to take into account the complexity of the system by introducing additional independent variables in the form of atmospheric and chemical factors. The NARX model for the PM_{10} pollution forecasting presented in this research was designed to operate offline, and its parameters were built based on the available training data from a certain period. A future research on model improvement may focus on mechanism adapted to periodically updating the NARX's parameters, in order to better match to the latest air quality data, including integration with real-time data sources.

References

1. Adnane, A., Leghrib, R., Chaoufi, J., Chirmata, A. (2022). Prediction of PM_{10} concentrations in the city of Agadir (Morocco) using non-linear autoregressive artificial neural networks with exogenous inputs (NARX). *Materials Today Proceedings, 52, 1*, pp. 146-151, doi:10.1016/j.matpr.2021.11.340
2. Akhbarizadeh, R., Dobaradaran, S., Torkmahalleh, M.A., Saeedi, R., Aibaghi, R., Ghasemi, F.F. (2021). Suspended fine particulate matter (PM2.5), microplastics (MPs), and polycyclic aromatic hydrocarbons (PAHs) in air: their possible relationships and health implications. *Environmental Research, 192, 110339*, doi: 10.1016/j.envres.2020.110339
3. Al-Sbou, Y.A., Alawasa, K.M. (2017). Nonlinear Autoregressive Recurrent Neural Network Model For Solar Radiation Prediction. *International Journal of Applied Engineering Research, 12, 14*, pp. 4518-4527.
4. Bae, S., Pan, X.C., Kim, S.Y., Park, K., Kim, Y.H., Kim, H., Hong, Y.C. (2010). Exposures to particulate matter and polycyclic aromatic hydrocarbons and oxidative stress in schoolchildren. *Environmental health perspectives, 118, 4*, pp. 579-583, doi:10.1289/ehp.0901077
5. Basheer, I.A., Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods, 43, 1*, pp. 3-31, doi: 10.1016/s0167-7012(00)00201-3
6. Boussaada, Z., Curea, O., Remaci, A., Camblong, H., Bellaaj, M.N. (2018). A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation. *Energies, 11, 620*, pp. 1-21, doi:10.3390/en11030620
7. Buevich, A., Sergeev, A., Shichkin, A., Baglaeva, E. (2021). A two-step combined algorithm based on NARX neural network and the subsequent prediction of the residues improves prediction accuracy of the greenhouse gases concentrations. *Neural Computing and Applications, 33, 1*, pp. 1547-1557, doi:10.1007/s00521-020-04995-4
8. Canagaratna, M.R., Jayne, J.T., Ghertner, D.A., Herndon, S., Shi, Q., Jimenez, J.L., Silva, P.J., Williams, P., Lanni, T., Drewnick, F., Demerjian, K.L., Kolb, C.E., Worsnop, D.R. (2004). Chase studies of particulate emissions from in-use New York City vehicles. *Aerosol Science and Technology, 38, 6*, pp. 555-573, doi:10.1080/02786820490465504

9. Chellali, M.R., Abderrahim, H., Hamou, A., Nebatti, A., Janovec, J. (2016). Artificial neural network models for prediction of daily fine particulate matter concentrations in Algiers. *Environmental Science and Pollution Research, 23, 14*, pp. 14008-14017, doi: 10.1007/s11356-016-6565-9
10. Chrabołowska, M., Siemieniuk, A., Szatyłowicz, E. (2024). Smog in the city – the quality of suspended dust in Białystok. *Gaz, Woda i Technika Sanitarna, 7-8*, pp. 35-41, doi: 10.15199/17.2024.7-8.7
11. Cortina-Januchs, M.G., Quintanilla-Dominguez, J., Vega-Corona, A., Andina, D. (2015). Development of a model for forecasting of PM10 concentrations in Salamanca, Mexico. *Atmospheric Pollution Research, 6*, pp. 626-634, doi: 10.5094/APR.2015.071
12. European Environment Agency (2021). *Air quality in Europe 2021*. Retrieved from: https://www.eea.europa.eu/publications/air-quality-in-europe-2021, 12.05.2022.
13. Ganguly, R., Sharma, D., Kumar, P. (2019). Trend analysis of observational PM10 concentrations in Shimla city, India. *Sustainable Cities and Society, 51*, pp. 1-13, doi: 10.1016/j.scs.2019.101719
14. Ghahari, S.A., Queiroz, C., Labi, S., McNeil, S. (2012). Cluster Forecasting of Corruption Using Nonlinear Autoregressive Models with Exogenous Variables (NARX)-An Artificial Neural Network Analysis. *Sustainability*, *13, 20*, pp. 1-20, doi:10.3390/su132011366
15. Główny Inspektorat Ochrony Środowiska (2022). *Podstawowe informacje o stacji*. Retrieved from: https://powietrze.gios.gov.pl/pjp/current/station_details/info/987, 16.04.2022.
16. Główny Urząd Statystyczny (2021). *Ochrona środowiska 2021*. Retrieved from: https://stat.gov.pl/obszary-tematyczne/srodowisko-energia/srodowisko/ochrona-srodowiska-2021,1,22.html, 12.05.2022.
17. Główny Urząd Statystyczny (2022). *Bank Danych Lokalnych. Dane według dziedzin*. Retrieved from: https://bdl.stat.gov.pl/bdl/dane/podgrup/tablica, 24.05.2022.
18. Gündoğdu, S. (2020). Comparison of static MLP and dynamic NARX neural networks for forecasting of atmospheric PM_{10} and SO_2 concentrations in an industrial site of Turkey. *Environmental Forensics, 21, 3-4*, pp. 363-374, doi:10.1080/15275922.2020.1771637
19. Gündoğdu, S., Elbir, T. (2024). Elevating hourly $PM_{2.5}$ forecasting in Istanbul, Türkiye: Leveraging ERA5 reanalysis and genetic algorithms in a comparative machine learning model analysis. *Chemosphere*, *364, 143096*, doi: 10.1016/j.chemosphere.2024.143096
20. Hamdan, M.A., Bani Ata, M.F., Sakhrieh, A.H. (2021). Air Quality Assessment and Forecasting Using Neural Network Model. *Journal of Ecological Engineering, 22, 6*, pp. 1-11, doi: 10.12911/22998993/137444
21. Haris, M., Hasan, N.M., Qin, S. (2022). State of health prediction of supercapacitors using multi-trend learning of NARX neural network. *Materials Today Sustainability*, *20, 4*, p. 100201, doi: 10.1016/j.mtsust.2022.100201

22. Ibrahim, S., Wahab, A.N. (2022). Improved Artificial Neural Network Training Based on Response Surface Methodology for Membrane Flux Prediction. *Membranes, 12, 726*, pp. 1-25, doi:10.3390/membranes12080726
23. Instytut Meteorologii i Gospodarki Wodnej-Państwowy Instytut Badawczy (2020-2023). *Rocznik Meteorologiczny 2020-2023*. Retrieved from: https://danepubliczne.imgw.pl/data/dane_pomiarowo_obserwacyjne/Roczniki/Rocznik%20meteorologiczny/, 16.03.2025.
24. Jędruszkiewicz, J., Czernecki, B., Marosz, M. (2017). The variability of PM10 and PM2.5 concentrations in selected Polish agglomerations: the role of meteorological conditions, 2006-2016. *International Journal of Environmental Research and Public Health*, *27, 6*, pp. 441-462, doi: 10.1080/09603123.2017.1379055
25. Kaur, M., Mandal, A. (5-7.02.2020). *$PM_{2.5}$ Concentration Forecasting using Neural Networks for Hotspots of Delhi*. 2020 International Conference on Contemporary Computing and Applications (IC3A). Lucknow, India, pp. 40-43, doi: 10.1109/IC3A48958.2020
26. Khan, Z., Pathak, D.K., Pandey, A., Kumar, S. (08.06.2014). *Performance evaluation of nonlinear auto-regressive with exogenous input (NARX) in the foreign exchange market*. Proceedings of 10^{th} IRF International Conference. Chennai, India.
27. Kirešová, S., Guzan, M. (2022). Determining the Correlation between Particulate Matter PM10 and Meteorological Factors. *Eng., 3, 3,* pp. 343-363, doi: 10.3390/eng3030025
28. Kuchcik, M. (2020). Air pollution in Polish health resorts–a problem for both local authorities and patients. *Przegląd Geograficzny, 92, 1*, pp. 109-134, doi: 10.7163/PrzG.2020.1.7
29. Kujawska, J., Kulisz, M., Oleszczuk, P., Cel, W. (2022). Machine Learning Methods to Forecast the Concentration of PM10 in Lublin, Poland. *Energies*, *15, 6428*, pp. 1-23, doi: 10.3390/en15176428
30. Li, W., Lin, G., Xiao, Z., Zhang, Y., Li, B., Zhou, Y., Chai, E. (2022). A review of respirable fine particulate matter (PM2.5)-induced brain damage. *Frontiers in Molecular Neuroscience*, *15, 967174*, pp. 1-9, doi: 10.3389/fnmol.2022.967174
31. Liu, Q., Chen, W., Hu, H., Zhu, Q., Xie, Z. (2020). An Optimal NARX Neural Network Identification Model for a Magnetorheological Damper With Force-Distortion Behavior. *Frontiers in Materials, 7*, pp. 1-12, doi:10.3389/fmats.2020.00010
32. Liu, X., Ha, M., Xiaohui Lei, X., Zhang, Z. (2022). A Novel GRA-NARX Model for Water Level Prediction of Pumping Stations. *Water, 14, 19*, pp. 1-17, doi:10.3390/w14192954
33. Magallanes-Quintanar, R., Galván-Tejada, C.E., Galván-Tejada, J.I., Méndez-Gallegos, S.D.J., García-Domínguez, A., Gamboa-Rosales, H. (2022). Narx Neural Networks Models for Prediction of Standardized Precipitation Index in Central Mexico. *Atmosphere*, *13, 1254*, pp. 1-13, doi:10.3390/atmos13081254
34. Munir, S. (2017). Analysing Temporal Trends in the Ratios of $PM_{2.5}/PM_{10}$ in the UK. *Aerosol and Air Quality Research*, *17, 1*, pp. 34-48, doi: 10.4209/aaqr.2016.02.0081

35. Nejadkoorki, F., Baroutian, S. (2012). Forecasting extreme PM10 concentrations using artificial neural networks. *International Journal of Environmental Research*, *6, 1*, pp. 277-284, doi: 10.22059/IJER.2011.493
36. Paul, R.K., Sinha, K. (2016). Forecasting Crop Yield: Arimax and Narx Model. *RASHI, 1*, pp. 77-85.
37. Pawul, M., Śliwka, M. (2016). Application of artificial neural networks for prediction of air pollution levels in environmental monitoring. *Journal of Ecological Engineering, 17, 4*, pp. 190-196, doi: 10.12911/22998993/64828
38. Peña, M., Vázquez-Patiño, A., Zhiña, D., Montenegro, M., Avilés, A. (2020). Improved Rainfall Prediction through Nonlinear Autoregressive Network with Exogenous Variables: A Case Study in Andes High Mountain Region. *Advances in Meteorology, 1828319*, pp. 1-17, doi:10.1155/2020/1828319
39. Perez, P., Reyes, J. (2006). An integrated neural network model for PM10 forecasting. *Atmospheric Environment*, *40, 16*, pp. 2845-2851, doi: 10.1016/j.atmosenv.2006.01.010
40. Pisoni, E., Farina, M., Carnevale, C., Piroddi, L. (2009). Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*, *22, 4-5*, pp. 593-602, doi:10.1016/j.engappai.2009.04.002
41. Polkowska, Ż., Dubiella-Jackowska, A., Zabiegała, B., Namieśnik, J. (2007). Composition of pollutants in the water samples collected along the highways with different traffic intensity. *Chemia i Inżynieria Ekologiczna, 14*, pp. 315-338.
42. Riverol, C., Harrilal, S. (2018). A Non-linear Autoregressive External Inputs (NARX) model for estimating the mixing volumes between batches in TRANSMIX. *International Journal of Heat and Mass Transfer*, *127, Part A*, pp. 161-163, doi:10.1016/j.ijheatmasstransfer.2018.07.011
43. Rumaling, M.I., Chee, F.P., Chang, H.W.J., Payus, C.M., Kong, S.K., Dayou, J., Sentian, J. (2022). Forecasting particulate matter concentration using nonlinear autoregression with exogenous input model. *Global Journal of Environmental Science and Management*, *8, 1*, pp. 27-44, doi: 10.22034/GJESM.2022.01.03
44. Srimuruganandam, B., Shiva Nagendra, S.M. (2012). Source characterization of PM_{10} and $PM_{2.5}$ mass using a chemical mass balance model at urban roadside. *Science of The Total Environment*, *433*, pp. 8-19, doi: 10.1016/j.scitotenv.2012.05.082.
45. Taspinar, F. (2015). Improving artificial neural network model predictions of daily average PM10 concentrations by applying principle component analysis and implementing seasonal models. *Journal of the Air & Waste Management Association, 65, 7*, pp. 800-809, doi: 10.1080/10962247.2015.1019652
46. Urząd Statystyczny w Szczecinie (2020). *Rocznik Szczecin 2020*. Retrieved from: https://szczecin.stat.gov.pl/files/gfx/szczecin/pl/defaultstronaopisowa/888/1/1/rocznik_szczecin_2020___tablice_przegladowe_2.pdf, 21.04.2022.

47. Viejo, C.G., Torrico, D.D., Dunshea, F.R., Fuentes, S. (2019). Development of Artificial Neural Network Models to Assess Beer Acceptability Based on Sensory Properties Using a Robotic Pourer: A Comparative Model Approach to Achieve an Artificial Intelligence System. *Beverages, 5, 33*, pp. 1-10, doi:10.3390/beverages5020033
48. Wiktorzak, A., Sawicki, A. (2023). Application of artificial neural networks in the development of the PM10 air pollution prediction system. *TASK Quarterly*, *27, 1*, pp. 1-9, doi: 10.34808/bamk-q919
49. World Health Organization (2021). *New WHO Global Air Quality Guidelines aim to save millions of lives from air pollution.* Retrieved from: https://www.who.int/news/item/22-09-2021-new-who-global-air-quality-guidelines-aim-to-save-millions-of-lives-from-air-pollution, 15.05.2022.
50. Wu, X., Zhu, B., Zhou, J., Bi, Y., Xu, S., Zhou, B. (2021). The epidemiological trends in the burden of lung cancer attributable to PM2.5 exposure in China. *BMC Public Health, 21, 737*, pp. 1-8, doi: 10.1186/s12889-021-10765-1
51. Yu, X., Chen, Z., Qi, L. (2019). Comparative Study of Sarima and Narx Models in Predicting the Incidence of Schistosomiasis in China. *Mathematical Biosciences and Engineering, 16, 4*, pp. 2266-2276, doi: 10.3934/mbe.2019112
52. Zeng, C., Liu, X., Chen, L., He, X., Kang, Z. (2024). Enhanced Short-Term Temperature Prediction of Seasonally Frozen Soil Subgrades Using the NARX Neural Network. *Applied Sciences*, *14, 10257*, pp. 1-16, doi: 10.3390/app142210257
53. Zhang, J., He, K., Ge, Y., Shi X. (2009). Influence of fuel sulfur on the characterization of PM_{10} from a diesel engine. *Fuel*, *88, 3*, pp. 504-510. doi: 10.1016/j.fuel.2008.09.001
54. Zheng, Y., Zhang, W., Xie, J., Liu, Q. (2022). A Water Consumption Forecasting Model by Using a Nonlinear Autoregressive Network with Exogenous Inputs Based on Rough Attributes. *Water*, *14, 329*, pp. 1-16, doi: 10.3390/w14030329