ISSN: 0971-1260
Vol-22-Issue-16-August-2019

A Review on Application of Artificial Neural Networks (ANN) For Retrieval of Atmospheric and Air Quality Data

Prasenjit Adak

School of Chemical Engineering and Physical Sciences, Lovely Professional University, Phagwara, Punjab-144411

Email: prasenjit.23966@lpu.co.in

Abstract: Artificial Neural Network (ANN) is an efficient approach to model complex inter-dependency of numerous environmental factors. In many cases it has been proved to be more efficient and accurate than the conventional statistical models. Present literature review attempt to collate some notable research works that involved the development or use of various ANN algorithms to improve the atmospheric and air quality data retrieval process.

Introduction

Air pollution is one of the most discussed environmental issues now a days. Especially, the recent increment in concentration of significantly harmful pollutants such as PM_{10} (particulate matters with aerodynamic diameter less than 10 μm) and $PM_{2.5}$ (particulate matters with aerodynamic diameter less than 2.5 μm) in the atmosphere of highly populated cities like Delhi, Kanpur, Faridabad and Patna is a matter of serious concern. Decreased ambient air quality is reportedly associated with various health issues [1], [2], [3], [4], [5]. In order to formulate any management policy to reduce the ambient concentration of air pollutants, proper quantification of the same is of utmost importance. There are several approaches are practised for quantification of air pollutant concentration in local atmosphere. Direct measurement using the air quality monitoring devices is the most reliable approach in this context but the collected data can not provide any meaningful insight if the underlying pattern and the inter-relationship between various meteorological factors are not recognized. This is particularly very important for forecasting and trend analysis study. Therefore, despite the fact that, the modelling approach to predict air quality models is not very accurate, the air quality models are used in numerous studies, especially in the field of policy making and planning [6], [7]. Previously many researchers have studied and recognized the data of ambient air quality are stochastic time series. Therefore, the historical data can be used to forecast the air quality to some extent in short-term [8], [9], [10].

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

The conventional statistical methods for air quality forecasting generally involves time-series methods or regression methods. In time-series methods, the meteorological parameters are not included as the input. As a result, the prediction may exhibit serious uncertainties if proper measures are not taken [11]. On the other hand, the regression method is mostly assumes simple multivariate linear relationship between the meteorological parameters and the pollutant concentration. Therefore, the complex interaction of various meteorological parameters are mostly overlooked. Moreover, the conventional empirical and semi-empirical models using the huge dataset generated by the meteorological sensors requires enormous amount of time. In order to overcome these limitations, artificial neural networks (ANN) can be used for air quality forecasting. Previously, studies have been conducted using neural networks to model chaotic time series [12]. This approach uses less assumption and more adaptive to the changes which makes ANN capable of being deployed in various aspects of environmental studies including runoff modeling, river stage forecasting, watershed management, rainfall forecasting, ambient air quality prediction and adsorption studies [13], [14], [15].

Neural networks: a brief overview

An ANN is a computer program that mimic the biological nervous system in terms of its basic working principle. The program can efficiently identify the underlying patterns in a dataset and learn from the variations of dependent variable and independent input parameters. It not only can effectively identify the complex inter-relationship between the input variables and the output data but also can make prediction of output based on the values of numerous input parameters.

An ANN system consists of several interconnected nodes each of which acts like a biological neuron. A typical ANN consists of an input layer, one or more hidden layers and an output layer. When input data is fed into the input layer, it passes through an activation function which, depending on the weightage of the input parameters, generates the value of the output parameter. This process is called feed-forward process. Then the generated output value is compared with actual measured output value and the weightage of the input parameters are adjusted accordingly. This process is called back-propagation. This entire process of feeding the pair of input and output data into the ANN is called the training process. After rigorous train of the ANN, the weightage of the input parameters are adjusted. After this, a set of input variables feed into the system to obtain the value of output. The predicted output value is then compared with measured value of the output variable. This process is called testing of ANN. The available set of known input and output data is generally segregated into training dataset and testing dataset. Some widely used programming

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

languages and tools for the construction of ANN are Matlab, Python, R and Java. There are several approaches and algorithms presently in use for development of ANN. Some of the commonly used approaches are as follows.

Recurrent Network Model (RNM): Recurrent Neural Network(RNN) are a type of ANN in which the output from previous step serves as input to the current step. Therefore, in order to make the ANN dynamic (especially if the objective of the study involves temporal variation of data) short-term internal state or memory is required to process the data sequence. Unlike the conventional ANN inputs and outputs are dependent on each other.

Changepoint Detection Model (CPDM): Regional ambient air quality is controlled by various factors such as, source strength, meteorological condition etc. The change in these factors over time leads to the alteration of air quality parameters. Therefore, it can be stated that the change of air quality parameters translate into a series of changepoints [16]. CPDM is used to detect whether any change in the observed data sequence has taken place.

Sequential Network Construction Model (SNCM): The basic idea behind this model is to optimize the network design for a specific task. A multilayer perceptron is introduced initially. Then the final weightage of the nodes are used in the next run. In this way, at each step a hidden layer is added to the network sequentially until desired performance is achieved.

Self Organizing Maps (SOM) Model: This ANN model is trained using unsupervised learning method. Unlike the error-correction approach such as back-propagation, the output nodes of the network compete with each other. The weightage of each node is calculated to find out the best representable node.

Implementation of ANN in the field of air quality and weather prediction: some noteworthy examples

Barai et. al. (2007) performed a comparative studies in which the researchers considered various types of neural network models and implemented on secondary data collected from the public database. The researchers concluded that the SOFM performed the best [17].

Previously, many researchers have implemented ANN for efficient prediction of various climatological parameters. Churnside et. al. (1993) used a neural network model to retrieve vertical temperature profile of atmosphere [18]. A multilayer perceptron based ANN model was used by the researchers in this study. Surface-level temperature, pressure, relative humidity and brightness

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

temperatures at six different frequencies as the input parameters. In certain cases, especially in strong inversion condition and near ground ANN performed better than it's conventional statistical counterpart. Loukachine and Loeb (2003) used an ANN featuring feed-forward error back-propagation (FFEB) approach to estimate top-of-atmosphere radiative fluxes [19]. Muller et. al. (2003) used a feed forward ANN for ozone profile retrieval [20]. The researchers developed a specialized neural network system called Neural Network Ozone Retrieval System (NNORSY). The input data of this system included spectral, geolocation (latitude), time and climatological data. It was demonstrated that NNORSY was faster and more efficient than the classical methods to solve the inverse retrieval problem for Global Ozone Monitoring Experiment (GOME) ozone profiles. Soares et. al. (2004) estimated hourly diffuse solar-radiation data in Sao Paulo city, Brazil using a perceptron neural network [21]. In this work the researchers used meteorological data of 4 years to train the neural network model. It was demonstrated that using atmospheric long-wave radiation data as an input improved the performance of the model. Therefore, it can be observed that ANN is effective in identifying the relevant factors that affect the dependent variable. The study proved that ANN based model performed better than correlation model. Chevallier et. al. (1998) developed an ANN based model named NeuroFlux to compute long-wave radiative budget [22]. It was observed that ANN based models can perform 22 times faster than band model and 10^6 times more than line-by-line model. Abhishek et. al. (2012) evaluated the possibility of applying ANN for weather forecasting [23]. The study revealed that using multiple hidden layers can increase the performance of the ANN model. Moreover, the researchers stated that proper selection of model can significantly reduce the data processing cost as the model uses raw data as input.

Besides the prediction of meteorological parameters, many works have been done to predict air quality and to examine the dependency of air quality on various environmental and anthropogenic factors. Melesse and Rodney (2005) used ANN to simulate carbon flux from forest, grassland and wheat field [24]. Field observations from 15 days of hourly heat flux data was used as input data. Training of the ANN model was performed using net radiation, sensible heat, latent heat and soil heat flux. The researchers confirmed that ANN model used in this study was efficient and reliable. Niska et. al. (2004) used multi-layer perceptron model and genetic algorithm to forecast nitrogen dioxide at traffic sites in Helsinki [25]. Reich et. al. (1999) used a three layered feed-forward back-propagation ANN to identify unknown sources of sulphur dioxide [26]. Measured concentration of sulphur dioxide was used to train the ANN model. Azid et. al. (2014) combined principal component analysis (PCA) and ANN to predict the air pollution index (API) at ten air quality monitoring stations in Malaysia [27]. Researchers used PCA to identify the most influential

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

parameters responsible for the degradation of local air quality. The PCA-ANN model was proven to be helpful in optimizing the sampling parameters. As a result, it helped in reducing the sampling time and cost. Adams et. al. (2016) used 13 years of air quality data (ambient concentration of fine particulate matter and nitrogen dioxide) from a mobile air quality monitoring station in Hamilton [28]. The air quality data and the land use information of the sampling sites were used as the input to train the ANN model. This work helped in identifying the most important parameters contributing in regional air quality. Kurt et. al. (2008) developed an online air pollution forecasting system in Istanbul using ANN based model [29]. The system was able to predict the ambient concentration of PM_{10}, carbon monoxide and sulphur dioxide for next three days. The researchers developed a public website to display these predictions. The system revealed that besides other factors, the day of week significantly affects the air quality. Kolehmainen (2000) constructed a hybrid ANN model to forecast air quality by combining SOM algorithm, Sammon's mapping and fuzzy distance metrics [30]. Numbers of multilayer perceptrons were used to construct the ANN model. The researchers reported efficient forecasting of gaseous and particulate pollutant levels in the study area.

Conclusion

The ANN based models has been proven to be an effective tool to predict the climatic and air quality parameters by many researchers. The applicability of ANN based prediction models is enormous especially where the theoretical relationship between various components are non-linear and too complex to be constructed in traditional methods. Although ANN can be beneficial to use when a large dataset is available to feed as training data, lack of the same may lead to significant uncertainty in output prediction.

References

[1] C.A. Pope III and D.W. Dockery, "Health effects of fine particulate air pollution: lines that connect", Journal of the air & waste management association, vol. 56, pp. 709-742, 2006.

[2] J.A. Bernstein, N. Alexis, C. Barnes, I.L. Bernstein, A. Nel, D. Peden, D. Diaz-Sanchez, S.M. Tarlo and P.B. Williams, "Health effects of air pollution" Journal of allergy and clinical immunology, vol. 114, pp. 1116-1123, 2004.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

[3] R. Bascom, P.A. Bromberg, D.L. Costa, R. Devlin, D.W. Dockery, M.W. Frampton, W. Lambert, J.M. Samet, F.E. Speizer and M. Utell. "Health effects of outdoor air pollution", American journal of respiratory and critical care medicine, vol. 153, pp. 477-498, 1996.

[4] JJ. Kim, "Ambient air pollution: health hazards to children", Pediatrics, vol. 114, pp. 1699-1707, 2004.

[5] M.N. Rao and H.V. Rao, “Air Pollution”, Tata McGraw-Hill, New Delhi, 2000.

[6] D. Melas, I. Kioutsioukis and I.C. Ziomas, (2000), “Neural network models for predicting peak photochemical pollutant levels”, J. Air and Waste Manage. Assoc., vol. 50, pp. 495-501, 2000.

[7] L. Hadjiiski and P. Hopke, "Application of artificial neural networks to modeling and prediction of ambient ozone concentrations”, J. Air and Waste Manage. Assoc., vol. 50, pp. 894-901, 2000.

[8] U. Kumar and V. K. Jain, "ARIMA forecasting of ambient air pollutants (O3, NO, NO2 and CO)", Stochastic Environmental Research and Risk Assessment, vol. 24, pp. 751-760, 2010.

[9] J.J. Kao and S.S. Huang, "Forecasts using neural network versus Box-Jenkins methodology for ambient air quality monitoring data", Journal of the Air & Waste Management Association, vol. 50, pp. 219-226, 2000.

[10] F. Giorgio and M. Piero, "Mathematical models for planning and controlling air quality”, Proceedings of IIASA Workshop, vol. 17. 1996.

[11] K. Chakraborty, K. Mehrotra, C.K. Mohan and S. Ranka, “Forecasting the behaviour of multivariate time series using neural networks”, Neural Networks, pp. 53-58, 1992.

[12] J. Principe and J.M. Kuo, “Dynamic modeling of chaotic time series with neural networks”, Proc. of Neural Infor. Proc. Sys., vol. 7, pp. 311-318, 1995.

[13] K. Thirumalaiah and M.C. Deo, “Hydrological forecasting using neural networks”, Journal of Hydrologic Engineering, ASCE, vol. 5, pp. 180-189, 2000.

[14] M. Boznar, M. Lesja, and P. Mlakar, “A neural network-based method for short term prediction of ambient SO2 concentrations in highly polluted industrial areas of complex terrain”, Atmos. Environ., vol. 27B, pp. 221-230, 1993.

[15] A. Basheer and M. Najjar, “Predicting dynamic response of adsorption columns with Neural Nets”, J. Computing in Civil Engineering, ASCE, vol. 10, pp. 211-220, 1996.

[16] J.O. Kyong and I. Han, “Using change-point detection to support artificial neural networks for interest rate forecasting”, Expert System with Applications, vol. 19, pp. 105-115, 2000.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

[17] S.V. Barai, A. K. Dikshit and S. Sharma, "Neural network models for air quality prediction: a comparative study", Soft computing in Industrial applications, pp. 290-305, 2007.

[18] J.H. Churnside, T.A. Stermitz and J.A. Schroeder. "Temperature profiling with neural network inversion of microwave radiometer data" Journal of Atmospheric and Oceanic Technology, vol. 11, pp. 105-109, 1994.

[19] K. Loukachine and N.G. Loeb, "Application of an artificial neural network simulation for top-of-atmosphere radiative flux estimation from CERES." Journal of Atmospheric and Oceanic Technology, vol. 20, pp. 1749-1757, 2003.

[20] M.D. Muller, A.K. Kaifel, M. Weber, S. Tellmann, J.P. Burrows and D. Loyola, "Ozone profile retrieval from Global Ozone Monitoring Experiment (GOME) data using a neural network approach (neural network ozone retrieval system (NNORSY))", Journal of Geophysical Research: Atmospheres, vol. 108, 2003.

[21] J. Soares, A.P. Oliveira, M.Z. Božnar, P. Mlakar, J.F. Escobedo and A.J. Machado. "Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique" Applied energy, vol. 79, pp. 201-214, 2004.

[22] F. Chevallier, F. Chéruy, N.A. Scott and A. Chédin, "A neural network approach for a fast and accurate computation of a longwave radiative budget", Journal of Applied Meteorology, vol. 37, pp. 1385-1397.

[23] K. Abhishek, M. P. Singh, S. Ghosh and A. Anand, "Weather forecasting model using artificial neural network", Procedia Technology, vol. 4, pp. 311-318, 2012.

[24] A.M. Melesse and R.S. Hanley, "Artificial neural network application for multi-ecosystem carbon flux simulation", Ecological Modelling, vol. 189, pp. 305-314, 2005.

[25] H. Niska, T. Hiltunen, A. Karppinen, J. Ruuskanen and M. Kolehmainen, "Evolving the neural network model for forecasting air pollution time series", Engineering Applications of Artificial Intelligence, vol. 17, pp. 159-167, 2004.

[26] S.L. Reich, D.R. Gomez and L.E. Dawidowski, "Artificial neural network for the identification of unknown air pollution sources", Atmospheric Environment, vol. 33, pp. 3045-3052, 1999.

[27] A. Azid, H. Juahir, M.E. Toriman, M.K.A. Kamarudin, A.S.M. Saudi, C.N.C. Hasnam, N.A.A. Aziz, F. Azaman, M.T. Latif, S.F.M. Zainuddin and M.R. Osman, "Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: A case study in Malaysia", Water, Air, & Soil Pollution, vol. 225, p. 2063, 2014.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

[28] M.D. Adams and P.S. Kanaroglou, "Mapping real-time air pollution health risk for environmental management: Combining mobile and stationary air pollution monitoring with neural network models", Journal of environmental management, vol. 168, pp. 133-141, 2016.

[29] A. Kurt, B. Gulbagci, F. Karaca and O. Alagha, "An online air pollution forecasting system using neural networks", Environment International, vol. 34, pp. 592-598, 2008.

[30] M. Kolehmainen, H. Martikainen, T. Hiltunen and J. Ruuskanen, "Forecasting air quality parameters using hybrid neural network modelling", Environmental Monitoring and Assessment, vol. 65, pp. 277-286, 2000.