New Environmental Prediction Model Using Fuzzy logic and Neural Networks

Abou-Bakr Ramadan [1], Ahmed El-Garhy [2], Fathy Zaky [3] and Mazhar Hefnawi [4]

[1] Department of National Network for Monitoring Radioactivity, Nuclear and Radiological Regulatory Authority Cairo, Egypt

[2] Department of Electronics, Communications and Computers, Faculty of Engineering, Helwan University Cairo, Egypt

[3] Department of Electronics, Communications and Computers, Faculty of Engineering, Helwan University Cairo, Egypt

[4] Department of National Network for Monitoring Radioactivity, Nuclear and Radiological Regulatory Authority Cairo, Egypt

Abstract

This work introduces a new prediction model. This prediction model is designed to accomplish its task by only one type of measurements while other prediction models need at least three types of measurements. This feature makes this model less expensive than other models. The user who works with other models such as Statistical model, Chemical model, Physical model and neural network model needs more than two types of measurements and if any type of these measurement is not available the user must buy the unavailable data to operate his or her model. This work uses this model for predicting the Gamma radiation levels measurements in ambient air. The results from this model are good enough to depend on it for environmental prediction, recognizing the artificial phenomena and covering lost or missing data and making a temporally monitoring system. This model can be used in any continuous environmental monitoring system.

Keywords: *Data Mining, Environmental Prediction model, Fuzzy Logic, Neural Networks, Radiation Prediction.*

1. Introduction

Deterministic models (i.e. theoretical or detailed atmospheric diffusion models) are based on a fundamental mathematical description of atmospheric processes in which effects are generated by causes [1]. Such models aim to resolve the underlying chemical and physical equations that control pollutant concentrations and therefore require detailed emission data and meteorological conditions for the region of interest. An excellent example is the urban air-shed model (UAM) [1], [2]. This model can be used to obtain an accurate picture of the factors involved in ozone production. However, the model is highly sophisticated because it requires a high level of human resources and intense data input [2], [3]. There are generally severe limitations in both spatial and temporal accuracy of the data. In addition, some input data are not easily acquired by environmental protection agencies or local industries. This means that if these inputs are unknown, then the application of the UAM is problematic. Therefore, it is much more practical to rely on statistical models. Statistical models are based on semi-empirical statistical relations among available data and measurements. They attempt to determine the underlying relationship between sets of input data (predictors) and targets. Examples of statistical models are correlation analysis [4] and time series analysis [5]. However, the complex and sometimes non-linear relationships of multiple variables can make statistical models awkward and complicated [6]. Other statistical approaches frequently used include several artificial neural network implementations [7], [8], [9]. The use of these artificial intelligence-based networks has been shown to give acceptable results for atmospheric pollution forecasting of pollutants such as SO2, Ozone and Benzopyrene. But the artificial networks model still needs detailed emission data from different types of measurements to operate in a suitable way. For this case our environmental prediction model is designed to operate with only one type of measurement. The results from this model have a well accepted result to depend on them for prediction, recognizing the artificial or strange phenomena, covering lost or missing data and making a temporally monitoring system.

2. New Environmental Prediction Model Construction

The new Environmental Prediction Model Construction model is a hybrid system, which consists of four parts.

Fuzzy part: Converting each measurement to fuzzy number or linguistic status according to allowed limit and converting the time for every measurement to fuzzy number [10] as shown in table 1.

Table 1: Fuzzy number for time segmentation

Fuzzy number	***Linguistic meaning***	***Fuzzy value for measurement time***
S1	Segment 1	From 0:00 to 1:00
S2	Segment 2	From 1:00 to 2:00
S3	Segment 3	From 2:00 to 3:00
S4	Segment 4	From 3:00 to 4:00
S5	Segment 5	From 4:00 to 5:00
S6	Segment 6	From 5:00 to 6:00
S7	Segment 7	From 6:00 to 7:00
S8	Segment 8	From 7:00 to 8:00
S9	Segment 9	From 8:00 to 9:00
S10	Segment 10	From 9:00 to 10:00
S11	Segment 11	From 10:00 to 11:00
S12	Segment 12	From 11:00 to 12:00
S13	Segment 13	From 12:00 to 13:00
S14	Segment 14	From 13:00 to 14:00
S15	Segment 15	From 14:00 to 15:00
S16	Segment 16	From 15:00 to 16:00
S17	Segment 17	From 16:00 to 17:00
S18	Segment 18	From 17:00 to 18:00
S19	Segment 19	From 18:00 to 19:00
S20	Segment 20	From 19:00 to 20:00
S21	Segment 21	From 20:00 to 21:00
S22	Segment 22	From 21:00 to 22:00
S23	Segment 23	From 22:00 to 23:00
S24	Segment 24	From 23:00 to 0:00

Library Part: Library part consists of 24-sub library. Every sub library contains a history measurements in fuzzy value formed in patterns for every hour in the day [11].
Neural Network part: Consists of one neural network, which measures the similarity degree between the input pattern and each pattern in one sub library in the library [12].
The pattern consists of 14 elements .The first part which is called "Tail" part of the pattern contains 12 elements that expresses one type of measurements at sequence time. The second part of the pattern is called "Head" contains two elements. One element expresses one measurement. This element is called "Result". The other element expresses the time for that radiation measurement as shown in figure 1.

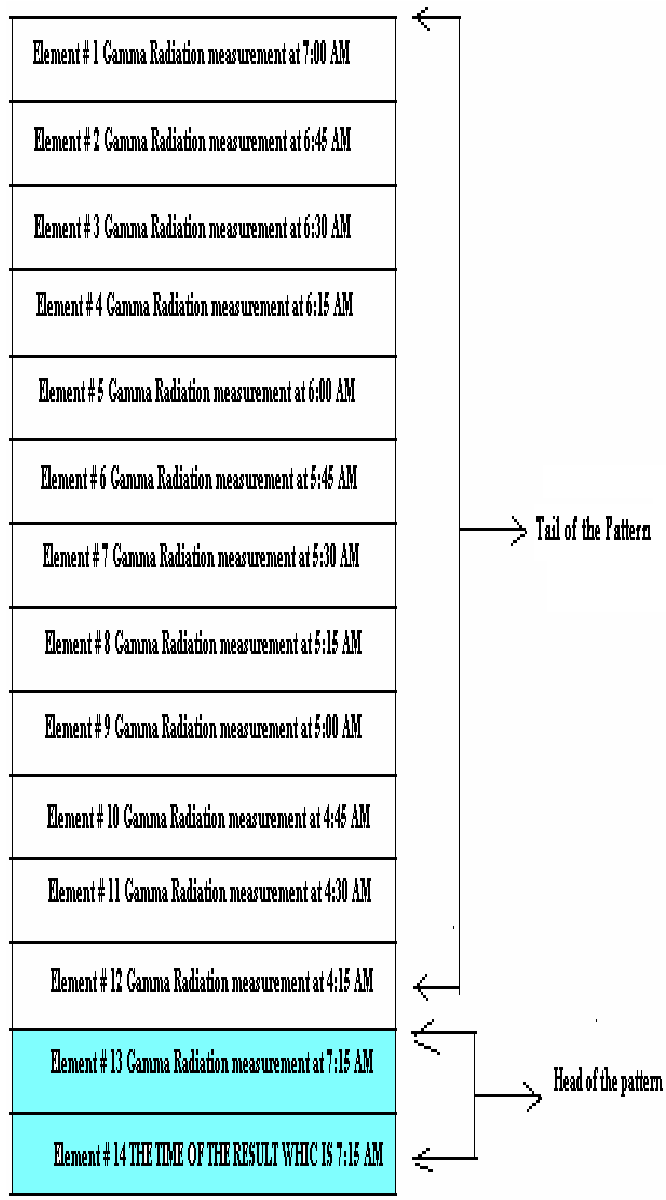

Fig. 1 Pattern construction.

Recognizer part: Determines the predicated Gamma measurement part by selecting the head of pattern from the library, which has the highest similarity degree.
These four parts are working together as a environmental model for prediction [13].
Figure 2 shows that The input layer for the neural network has 24 neurons which are twelve neurons from the tail of input pattern and twelve neurons from the tail of library pattern. The output layer for neural network has one neuron. The output value from this neuron is the similarity degree between the input pattern and library pattern [14].

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 3, March 2012
ISSN (Online): 1694-0814
www.IJCSI.org

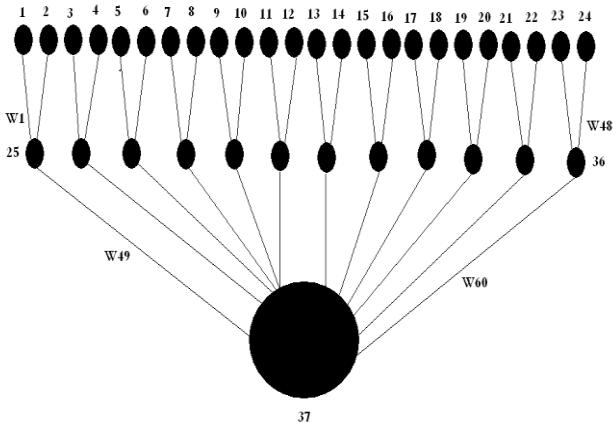

Fig. 2 Neural Network construction.

3. New Environmental Prediction Model Algorithm.

The New Environmental Prediction Model operates as following steps

1. The input to the model is one pattern. The tail and element number 14 of this pattern are known but element number 13 is unknown.
2. Create a Conversion of tail to fuzzy value according to allowed limits.
3. Create a Conversion of the time (element #13 of the input pattern) to fuzzy value according to table 1.
4. The library part receives the time fuzzy value and selects the matching sub library according to this value to the neural networks part.
5. The neural network compares between the tail of the input pattern and every tail of patterns stored in the selected sub library. Therefore, the neural network operates number of times equal to the number of patterns stored in the selected sub library.
6. For every neural network operation, it sends the resulted similarity degree and head of its current pattern from selected sub library to recognizer part.
7. For every neural network operation, the recognizer part receives and stores the highest similarity degree and its head.
8. At the last operation of neural network, the recognizer part will have the highest similarity degree and its head.
9. This head is the element number 14 in the input pattern, which was an unknown and predicted gamma level.
10. Update the input pattern by replacing the first element in the input pattern with resulted predicated gamma level from step number 9.
11. Shifts up the contents of the reminder elements to one level (e.g. contents of element number one is stored in contents element number two, contents of element number two is stored in contents of element number three and so force). At the last, the content of element number 12 is deleted.
12. Increment the time value by amount of the time for the next measurement and store it in element number 14 in the input pattern. If time value does not equal the desired time value then go to step number three.

Figure 3 shows the block diagram for the algorithm operation

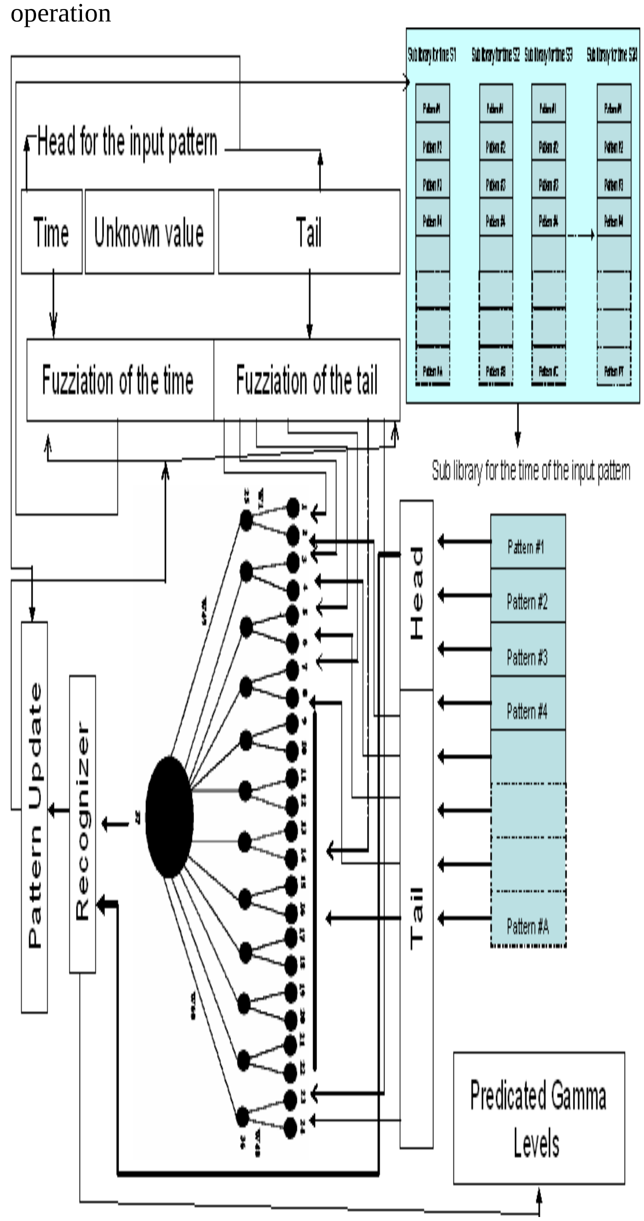

Fig. 3 Radiation Prediction Model Operation.

4. Results and Discussion.

This model is implemented on prediction of Gamma radiation levels measurements in ambient air. The conversion process to fuzzy values [15] of Gamma radiation levels measurements is done according to allowed limits [14] as shown in table 2.

Table 2: Fuzzy numbers and their linguistic meaning.

Fuzzy number	***Linguistic meaning***	***Fuzzy value for measurement time***
UL_ST1	Under allowed limit stage one	0<= UL_ST1 <= (0.044* Allowed limit)
UL_ST2	Under allowed limit stage two	UL_ST1< UL_ST2 <= (0.087* Allowed limit)
UL_ST3	Under allowed limit stage three	UL_ST2< UL_ST3 <= (0.13* Allowed limit)
UL_ST4	Under allowed limit stage four	UL_ST3< UL_ST4 <= (0.174* Allowed limit)
UL_ST5	Under allowed limit stage five	UL_ST4< UL_ST5 <= (0.217* Allowed limit)
UL_ST6	Under allowed limit stage six	UL_ST5< UL_ST6 <= (0.261* Allowed limit)
NL_ST1	Near from allowed limit stage one	UL_ST6< NL_ST1 <= (0.304* Allowed limit)
NL_ST2	Near from allowed limit stage two	NL_ST1< NL_ST2 <= (0.348* Allowed limit)
NL_ST3	Near from allowed limit stage three	NL_ST2< NL_ST3 <= (0.393* Allowed limit)
AL_ST1	At allowed limit stage one	NL_ST3< AL_ST1<= (0.435* Allowed limit)
AL_ST2	At allowed limit stage two	AL_ST1< AL_ST2<= (1.304* Allowed limit)
AL_ST3	At allowed limit stage three	AL_ST2< AL_ST3<= (2.174* Allowed limit)
AbL_ST1	Above allowed limit stage one	AL_ST3< AbL_ST1<= (3.04* Allowed limit)
AbL_ST2	Above allowed limit stage two	AbL_ST1< AbL_ST2<= (3.91*
AbL_ST3	Above allowed limit stage three	AbL_ST2< AbL_ST3<= (4.34*
OL_ST1	Over allowed limit stage one	AbL_ST3< OL_ST1<= (13.04* Allowed limit)
OL_ST2	Over allowed limit stage two	OL_ST1< OL_ST2<= (21.73* Allowed limit)
OL_ST3	Over allowed limit stage three	OL_ST2< OL_ST3<= (30.43* Allowed limit)
OL_ST4	Over allowed limit stage four	OL_ST3< OL_ST4<= (34.18* Allowed limit)
VOL	Very over allowed limit	VOL > OL_ST4
NO_DATA	No data recorded at this date	Ø

This model was tested more than 100 times. The source for training data set is Cairo Gamma station. The results are as following: average error is about 6% and the ratio of the accepted predicted data is about 95% as shown in table 3.

Table 3: Classification of the results for the new prediction model

Pure true result or error = 0 %	32.3%
Excellent true result ***or error*** < 1 0 %	41.4%
Very good true result ***or error*** < ***15 %***	16.3%
Good true result ***or error*** <***2 0 %***	4.9%
Accepted true result ***or error*** < ***25 %***	1.2%
Unaccepted Prediction result ***or error*** > ***25 %***	4.2%
Average Error	5.995%
Total accepted results	95.8%

This model requires only one pattern to operate. Figure 4 compares between the predicted gamma level and actual gamma values. From this figures it is clear that the difference is very small. This means that we can depend on the output results from this model with the lowest cost in effort, time and money.

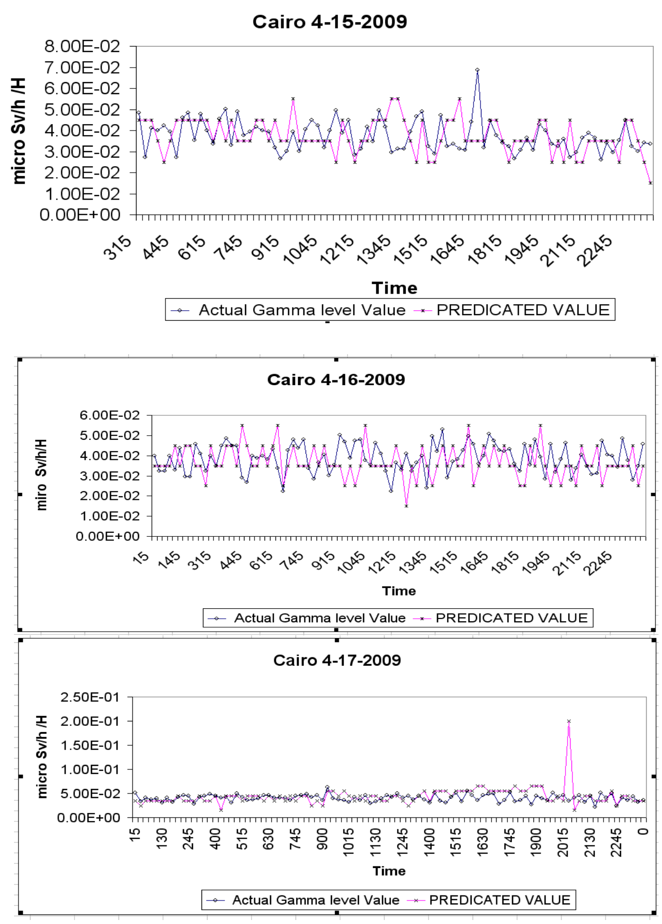

Fig. 4 Results of New Environmental Prediction Model for Gamma Radiation levels in Cairo city at 15,16 and 17 April 2009.

5. Conclusions

This prediction has a powerful feature in which this model can accomplish its task by only one type of measurements while other prediction models need at least three types of measurements.

This feature makes this model less expensive than other models [3], [4], [5] because any other model requires a large amount of data from different types of measurements. These types of measurements may be not available for the user. In this case, the user must buy the unavailable data to operate his or her model. On the other hand, the user can use only one type of date to operate his or her model.

This feature also saves a lot of effort and time because any other model requires a creation of a data preparation system to link different types of data with each other by the relationship between them. This will consume a lot of effort and time but in this model, we do not need to do this.

The results from this model are good enough to depend on it for prediction, recognizing the artificial or strange phenomena, covering lost or missing data and making a temporally monitoring system.

This model can be used in any continuous monitoring system, market forecast and decision-making.

References

[1] P.Zannetti, Computer modeling of air pollution, Boston : Southampton,science, art or fiction: Air pollution II. In: Computer Simulation, Computational Publications,1994.

[2] G.M.Johnson, An outline of air quality measurements and interpretations and tests of airshed models made possible by application of Airtrak instrumentation,USA: Explanatory notes and supplementary materials for seminar series held at EPA. Research Triangle Park,1991.

[3] M.Azzi,G.M.Johnson and R.Hyde,M.Young,"Prediction of NO_2 and O_2 and O_3 concentrations for NOx plumes photochemically reacting in urban air", Mathematical and Computational Modelling,1995, Vol. 21,pp. 39–48.

[4] S.A.Abdul-Wahab, W.Bouhamra,H.Ettouney, B.Sowerby and B.D.Crittenden, "Development of statistical model for prediction of ozone levels in Shuaiba Industrial Area in Kuwait", Environmental Science and Pollution Research (ESPR),1996,Vol 3,pp. 195–204.

[5] K.J.Hsu, "Time series analysis of the interdependence among air pollutants", Atmospheric Environment,1992, Vol 26,pp. 491–503.

[6] A.C.Comrie,"Comparing neural networks and regression models for ozone forecasting", Journal of the Air and Waste Management Association ,1997,Vol 47, pp. 653–663.

[7] M.Boznar, M.Lesjak. and P.Mlakar,"A neural network-based method for short-term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain", Atmospheric Environment,1993,Vol 27B, pp. 221–230.

[8] J.C.Ruiz-Sua´rez, O.A.Mayora-Ibarra,J.Torres-Jime´nez and L.G.Ruiz-Sua´-rez, "Short-term ozone forecasting by artificial neural networks",Advances in Engineering Software,1995, Vol 23,pp. 143–149.

[9] A. Elkamel, "Measurement and prediction of ozone levels around a heavily industrialized area: a neural network approach". Advances in Environmental Research,2001,Vol 5 ,pp. 47 -59.

[10] B.Walle , M. Turoff ,"Fuzzy relations for the analysis of traders' preferences in an information market game",European Journal of Operational Research,2009, Vol 195,pp. 905–913.

[11] E.Toyoda, "meteorological database for operational weather service",21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,2005, Vol 2, pp.50-55.

[12] T.Hill, L. Marquezb, M.O'Connor and W.Remusa , "Artificial neural network models for forecasting and decision making",International Journal of Forecasting,1994, Vol 10, pp.5-15.

[13] P.Zannetti, Computer modeling of water pollution, Boston : Southampton,science, art or fiction: Air pollution II. In: Computer Simulation, Computational Publications,1994.

[14] Law Four, Environmental law in Arab Republic of Egypt, 1995.

[15] J.J. Weinschenk, W.E. Combs and R.J. Marks ,"Forecasting Using fuzzy logic system", Proc IEEE Int'l Conference on Fuzzy Systems,2003, Vol 4, pp. 43–48 ,2003.