We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Urban air pollution forecasting using artificial intelligence-based tools

Min Li[1,2] and Md. Rafiul Hassan[1]
[1]Department of Computer Science and Software Engineering,
Melbourne School of Engineering,
[2]Centre for Women's Health, Gender and Society, Melbourne School of Population Health,
The University of Melbourne, Carlton 3010,
Victoria, Australia
racli@unimelb.edu.au
mrhassan@csse.unimelb.edu.au

1. Introduction

The detrimental effects of urban air pollution (UAP) have been represented as growing problems in recent years (Per Nafstad et al, 2004; World Health Organization). The harm represented by air pollution has been largely demonstrated from the impacts on human health and well being, such as asthma, eye irritation and even cancer (Nyberg F et al., 2000; J. Sunyer et al., 1997). Thus, the studies on specifying the pollution sources and analyzing concentrations of airborne pollutant variables are addressed sedulous attention by environmentalists and computer scientists. In order to prevent any further decline in air quality, to develop tools for air pollution control by introducing alternatives to existing practices is necessary.

Over the last decade, artificial intelligence (AI) based techniques have been proposed as alternatives to traditional statistical ones on forecasting UAP (Mikko Kolehmainen et al., 2000). Air pollution phenomena have been measured by using physical reality as the start point. And then, for example, these data traditionally has been coded into differential equations. However, these kinds of techniques have limited accuracy due to their inability to predict extreme events (Mikko Kolehmainen et al., 2000; Yilmaz Yildirim & Mahmut Bayramoglu, 2006). Comparing the traditional approaches, the models which are constructed in AI can be entirely based on these traditional measure data to forecast UAP. AI is a branch of scientific research enabling a structure to simulate intelligent behavior in computers. It is able to make a system deal with cognitive uncertainties in a manner more like human beings (Nils J. Nilsson., 1998). Thus, using AI techniques for modeling and forecasting can promote the development on UAP research.

There are several AI techniques which have been proposed as feasible and reliable ways for UAP forecasting, such as artificial neural networks (ANNs) (Harri Niska et al., 2004),

support vector machines (SVMs) (Wei-Zen Lu & Wen-Jian Wang, 2005) and fuzzy logic (FL) (Francesco Carlo Morabito & Mario Versaci, 2003). ANNs are as simplified mathematical models of brain-like systems (Dahe Jiang et al., 2004). This kind of techniques can learn the associations, functional dependencies and patterns by generalizing training data (Yilmaz Yildirim & Mahmut Bayramoglu, 2006; W. Z. Lu et al., 2002). ANNs have been used on detecting pollution sources, such as carbon monoxide (CO) (A.B.Chelani & S.Devotta, 2007; Ming Cai et al., 2009; Patricio Perez et al., 2004), particles measuring 10μm or less (PM$_{10}$) (Jef Hooyberghs et al., 2005; Patricio Perez & Jorge Reyes, 2006) and sulfur monoxide (SO) (U. Brunelli et al., 2007). Although ANN is regarded as one of the most popular AI methods on environmental researches, their inherent drawbacks (Wei-Zen Lu & Wen-Jian Wang, 2005), e.g., getting over-fitted into the training rules, can stuck in a local minima during training, poor generalization performance, determination of the appropriate network architecture, etc, impede the practical applications. The kernel-based hyper plane separation technique as SVM is another reliable and cost-effective AI technique (Wei-Zen Lu & Wen-Jian Wang, 2005) for classification and regression. For instance, it is built for predicting whether a new example belongs to one category or the other within a two categories' dataset. Although SVM has many potential problems as a new forecasting tool, there are only a few studies where SVMs has been reported to perform well by some promising results, e.g., SVM is superior to the conventional radial basis function (RBF) network in predicting air quality parameters with different time series.

Compared to other AI techniques, FL can offer a clear insight into the model for forecasting (Giorgio Corani, 2005). FL is a form of multi-valued logic to deal with reasoning that is approximate rather than precise. For example, it can be used on description of metrological impacts on UAP species (Oleg M. Pokrovsky et al., 2002; Md. Rafiul Hassan 2009; Md. Rafiul Hassan et al., 2007). However, FL suffers from the computational complexity associated with handling a large number of initially generated inappropriate rules, and thereby its interpretability is reduced (Md. Rafiul Hassan 2009).

A Hidden Markov model (HMM) is a classic approach for time series phenomena analysis and prediction. It has been widely used in the fields like DNA sequencing and speech recognition (Behzad Zamani et al. 2010). A significant hypothesis on HMM is based on the relationships between the attributes of particular data items in the dataset considered. Recently, Rafiul Hassan has developed a hybrid tool of HMM with fuzzy logic for time series forecasting (Md. Rafiul Hassan 2009; M. Maruf Hossan et al. 2008).

Contribution to the book chapter: The aim of this book chapter is to analysis the existing AI methodologies for UAP forecasting. In order to achieve this, the following approaches have been developed:

- We represented and summarized previous research on AI-based tools for UAP forecasting. This research is based on the analysis of current reliable AI methodologies which have already been used on predicting UAP.
- Based on Md. R. Hassan's previous research, we describe a HMM-FL model which combines the HMM's data pattern identification method to the generation of fuzzy logic for the prediction of UAP time series data. The dataset of testing PM$_{10}$ was introduced for experiment and results analysis.

- We compared the AI based tools which we described on this book chapter, and analysis their results on UAP forecasting.

Organization of the book chapter: This book chapter is organized as five sections. The introduction is provided in Section 1. We have introduced our topic, 'UAP forecasting Using AI-Based Tools', and described why using AI-based tools for UAP forecasting is important in this section. The contributions to this book chapter are presented from Section 2 to Section 4. Research on AI-based tools for UAP forecasting is described in Section 2. Then, Section 3 is designed for representing HMM-FL model. We briefly introduced some related principles and algorithms firstly, such as HMM and Fuzzy rules. Then, we construct HMM-FL model for predicting UAP time series. Section 4 is on experiment and comparison analysis. Furthermore, Section 5 is the discussion and conclusion of the whole book chapter.

2. Previous AI-based Methodologies for UAP Forecasting

In this section, we review some of the significant AI based methodologies which has been designed for forecasting UAP. Some of them combined AI methods, such as ANN, SVM and FL, with other methods. A chronological list of the major developments is preset in Table 1.

As one of the most compromising AI methods in estimation of environmental complex air pollution problems, ANN has been used by many scientists, such as (Ulku Sahin et al., 2005) and (P. Viotti et al., 2002). In the study of Ulku Sahin et al., ANN approach was used for predicting SO_2 concentration in Bahcelievler region. In this paper, the results were used to compare to nonlinear regression for actual measured values (Ulku Sahin et al., 2005). By comparing maximum and minimum values of observed SO_2 which were predicted by ANN model and nonlinear regression respectively, the results which are from ANN are quite realistic. P. Viotti et al's paper is another good example of using KNN for forecasting air pollution time series. ANN is the main technique to predict short and middle long-term concentration levels for some of the well-known pollutants in the city of Perugia. P. Viotti et al. reported in their study that the ANN has given great results in the middle and long-term forecasting of almost all the pollutants, although the ANN forecasts appear to be worse than the 1-hour ones.

Authors	Forecasting Models and techniques	Challenges	Datasets
Ulku Sahin, et al., 2005	ANN	Focus on modelling of SO_2 distribution and predicting its future concentration	Meteorological variable and SO_2 concentrations from Istanbul-Florya meteorogical station and Istanbul-Yenibosna air pollution station
P. Viotti, et al., 2002	Various ANN models	To forecast short and middle long-term concentration levels for some of the well-known pollutants	Variables monitored in Perugia, particularly for the area of Fontivegge
Wei-Zhen Lu & Wen-Jian Wang, 2005	A support vector machine (SVM) model	To examine the feasibility of applying SVM to predict air pollutant levels in advancing time series	An air pollutant database in Hong Kong downtown area
Oleg M. Pokrovsky, et al., 2002	FL based model	To study the impact of metrological factors on the evolution of air pollutant levels and to describe them quantitatively	The developed model is based on simulation of diurnal cycles of principal meteorological variables and the corresponding diurnal patterns of various air pollutants.
Luis A. Diaz-Robles, et al., 2008	A hybrid model combining Box-Jenkins (ARIMA) method and ANN	To improve forecast accuracy for an area with limited air quality and meteorological data.	Hourly and daily time series of PM_{10} and meteorological data during 2000-2006 at the Las Encinas monitoring station in Temuco.

Table 1. Some Previous AI-based Methodologies for UAP Forecasting at a Glance

Authors	Forecasting Models and techniques	Challenges	Datasets
Giorgio Corani., 2005	Feed-forward neural networks (FFNNs), pruned neural networks (PNNs) and lazy learning (LL)	Prediction of ozone and PM_{10}.	A dataset related to ozone and PM_{10} which are the major concern for air quality of Milan.
Yilmaz Yildirim & Mahmut Bayramoglu, 2006	An adaptive neuro-fuzzy logic model	To estimate the impact of meteorological factors on SO_2 and total suspended particular matter (TSP) pollution levels.	Datasets based on SO_2 and TSP detection in Zonguldak city (Turkey).
Mikko Kolehmainen, et al., 2000	A model using the Self-Organizing Map (SOM) algorithm, Sammon's mapping and fuzzy distance metrics. Overlapping Multi-Layer Perceptron (MLP) models were applied to the clustered data.	By using airborne pollutant, meteorological and timing variable to develop a form of air quality modelling which can forecast urban air quality.	The data applied to the city of Kuopio during the years 1995-1997.
M. Maruf Hossain, et al., 2008	A hybrid approach of Hidden Markov Model (HMM) with fuzzy logic	To model hourly air pollution at a location related to its traffic volume and meteorological variable.	A dataset that was originally put together as part of a study on air pollution related to traffic volume and meteorological variables on a road, conducted by the Norwegian Public Roads Administration.

Table 1. Some Previous AI-based Methodologies for UAP Forecasting at a Glance (Continued)

Among the fewer models which are based on SVMs, Wei-Zhen Lu, et al. (Wei-Zen Lu & Wen-Jian Wang, 2005; Wei-Zen Lu et al., 2004) introduced an SVM methodology for UAP forecasting. This study examined the feasibility of applying SVM to predict air pollutant level in advancing time series based on the monitored air pollutant database in Hong Kong downtown region. In this methodology, the SVM was firstly trained by data sets selection from the original dataset. Then, the SVM were used again for forecasting the pollutant levels in different time series. Results of the comparisons in forecasting between the SVM model and classical radial basis function (RBF) network show that SVM has a better generalization performance and superior to the conventional RBF network in predicting air quality parameters.

Besides ANN and SVM, FL approach for UAP forecasting has been developed recently. For example, in Oleg M. Pokrovsky et al.'s study (Oleg M. Pokrovsky et al., 2002), a FL based method has been used to model the impact of meteorological factors on the evolution of air pollutant levels and to describe them quantitatively. The model is based on simulation of diurnal cycles of principal meteorological categories, such as wind speed and direction, and the corresponding diurnal patterns of air pollutants, such as O_3. Another found from the research is that UAP phenomena can be simulated by sequences of its conservation inside some fuzzy sets and the transition from one fuzzy set to another. Thus, the development of the transition rules should be important in these kinds of cases.

Compared to above AI-based methodologies which are all used as single AI tools for UAP detection, AI-based methodologies are always combined with some other methods. Luis A. Diaz-Robles, et al. (Luis A. Diaz-Robles, et al., 2008) constructed a hybrid Box-Jenkins Time Series (ARIMA) and ANNs model to forecast particulate matter in urban areas which is the case of Temuco, Chile. Due to the inability of ARIMA to predict extreme events, the systems which based single ARIMA have limited accuracy. An improved forecasting accuracy was achieved by using the ARIMA and ANNs combined model. There is another model that predicts hourly NO_x and NO_2 concentrations (Gardner and Dorling, 1999) and neural models for ozone concentrations (Comrie, 1997; Yi and Prytok, 1996) were constructed for UAP predicting. Most of these works have focused on comparing feed-forward neural networks with the traditional methodologies, such as the ARIMA model and linear regression.

Combination of several AI-based methodologies is another idea on UAP forecasting research. From the research of Giorgio Corani (Giorgio Corani, 2005), there are three models which have been combined for air quality prediction in Milan. They are feed-forward neural networks (FFNNs), pruned neural networks (PNNs) and lazy learning (LL). FFNN is currently recognized as state-of-the-art approach for statistical prediction of air quality, while PNNs and LL are two alternative approach derived from machine learning. They are all constructed for forecasting ozone and PM_{10} which are the two major concerns for air pollution of Milan. From the results, it shows LL provides the best performances on indicators associated to average goodness of the prediction, such as correction, mean absolute error, etc. In addition, PNNs are superior to the other approaches in detecting the exceedances of alarm and attention thresholds.

Neuro-fuzzy methodology (S. Chiu, 1997) has been tested by many researchers for UAP prediction. Yilmaz Yildirim, et al. (Yilmaz Yildirim & Mahmut Bayramoglu, 2006) introduced an adaptive neuro-fuzzy logic method in their study. The adaptive neuro-fuzzy logic method is a hybrid of fuzzy logic and Neural-like architecture methodology. It is used to estimate the impact of meteorological factors on SO_2 and total suspended particular matter (TSP) pollution levels over an urban area. The model forecasts satisfactorily the trends in SO_2 and TSP concentration levels, and their performance are between 75-90% and 69-80%, respectively (Yilmaz Yildirim & Mahmut Bayramoglu, 2006). Francesco Carlo Morabito et al. proposed a hybrid fuzzy neural model for predicting time series of pollutant concentration levels in urban air (Francesco Carlo Morabito, et al., 2003). Through the use of the fuzzy surface concept, the manageable model has been carried out for the reduction of the model. In order to manage the multidimensional state problem, the use of ellipsoidal rules has been tested by designing and compiling a software code.

The AI-based model which developed by Mikko Kolehmainen, et al. (Mikko Kolehmainen, et al., 2000), is a typical model which can forecast UAP for the next day using airborne pollutant, meteorological and timing variables. This model combines Self-Organising Map (SOM) algorithim, Sammon's mapping and fuzzy distance metrics. Firstly, the clusters of data were characterized by statistics. Then, several overlapping Multi-Layer Perceptron (MLP) models were used on these cluster data. After this, by using a combination of the MLP model, the actual levels for individual pollutants could be calculated.

Recently, Md. R. Hassan introduced a novel hybrid of HMM and Fuzzy Logic model to analysis time series data for UAP forecasting. This hybrid HMM-FL model has the potential to achieve high levels of performance on hourly air pollution forecasting system. This model is able to reduce complexity and simultaneously improved forecasting accuracy. Compared to other techniques, the efficiency of the HMM-FL model is higher than well-performed fuzzy rule finding methods and KNN. In order to introduce the HMM-FL model, some principles are described in the Section 3 (M. Maruf Hossain, et al., 2008).

3. The HMM-Fuzzy Combination Model

3.1 Preliminaries
HMM-Fuzzy Model is combined HMM with Fuzzy Logic and Fuzzy Rule. In this section, we briefly introduce Hidden Markov Model (HMM), Fuzzy Logic (FL) and Fuzzy Rule.

3.1.1 Hidden Markov Model
A Hidden Markov Model (HMM) is a statistical model for modelling a wide range of time series data (Phil Blunsom, 2004). It is based on Markov process which is a time-varying random phenomenon for specific property holds. HMMs have been widely used in areas like speech, handwriting and gesture recognition (Lawrence R. Rabiner, 1989).

Figure 1 shows an example of a Markov process. It is a simple model for predicting air pollution. 'Clear', 'Mist' and 'Dirty' are used to represent the quality of air. 'High', 'Medium' and 'Low' are the percentage of pollutants in the air. In Markov process, 'Clear', 'Mist' and 'Dirty' are represented as states, while 'High', 'Medium' and 'Low' are index observations.

Assume the initial probability of getting 'Dirty' is 0.2. If given a sequence of observations: 'High-Low-Low', the state sequence is able to be identified as: 'Dirty-Clear-Clear'. Thus, the probability of the sequence in this case is 0.2x0.2x0.3.
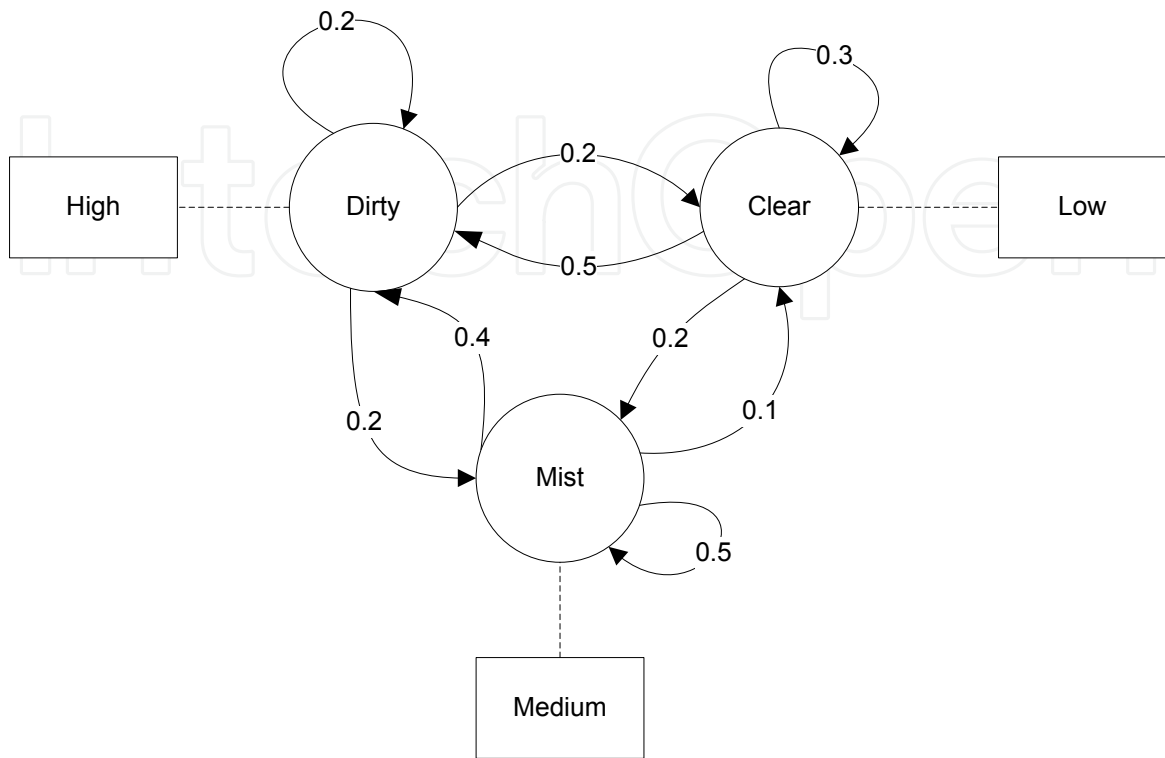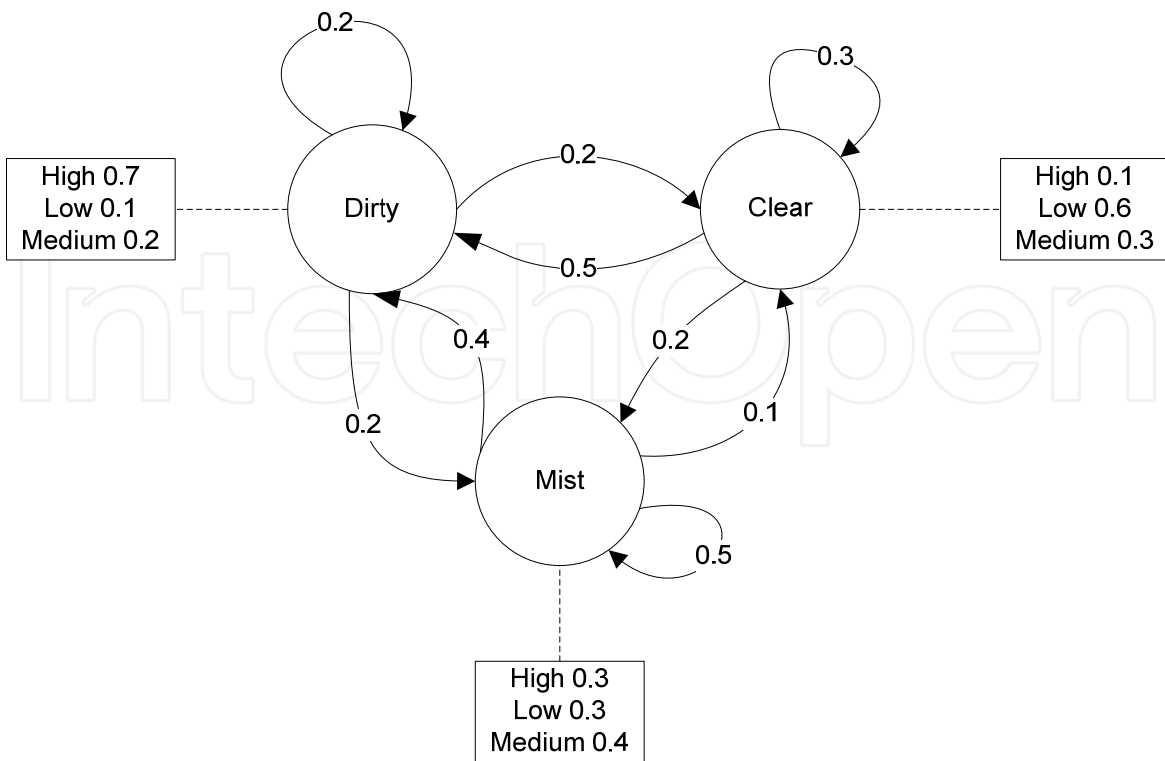

Fig. 1. Markov process example

Fig. 2. Hidden Markov model example

Figure 2 depicts an example of how the previous model is able to be extended into a HMM. In this example, we could not detect exactly what state sequence ('High', 'Low' and 'Medium') is able to produce the observations ('Dirty', 'Clear' and 'Mist'). Because the state sequences are 'hidden', the state sequence that was most likely to have produced the observation could be calculated.

HMM can be described as the following equation:

$$\lambda = (A, B, \pi) \tag{1}$$

where λ represents HMM in equation (1).

A is a transition array, storing the probability of state j following state i. Note the state transition probabilities are independent of time:

$$A = [a_{ij}], \ a_{ij} = P(q_t = S_j \mid q_{t-1} = S_i) \tag{2}$$

B is the observation array, storing the probability of observation k being produced from the state j, independent of t:

$$B = [b_i(k)], \ b_i(k) = P(x_t = v_k \mid q_t = S_i) \tag{3}$$

π is the initial probability array:

$$\pi = [\pi_i], \ \pi_i = P(q_1 = S_i). \tag{4}$$

S is our state alphabet set, and V is the observation alphabet set:

$$S = (S_1, S_2, \ldots, S_N) \tag{5}$$

$$V = (V_1, V_2, \ldots, V_M) \tag{6}$$

Q is defined to be a fixed state sequence of the length T, and O is the corresponding observations:

$$Q = q_1, q_2, \ldots, q_T \tag{7}$$

$$O = o_1, o_2, \ldots, o_T \tag{8}$$

Two assumptions are made by the model. The first, called the Markov assumption, states that the current state is dependent only on the previous state, which represents the memory of the model:

$$P(q_t \mid q_1^{t-1}) = P(q_t \mid q_{t-1}) \tag{9}$$

The independence assumption states that the output observation at time t is dependent only on the current state; it is independent of previous observations and states:

$$P(o_t \mid o_1^{t-1}, q_1^t) = P(o_t \mid q_t) \tag{10}$$

3.1.2 Fuzzy Logic and Fuzzy Rule

Fuzzy logic usually processes non-linear datasets by mapping input data (features) vectors into scalar output well. It is because that the fuzzy rules can be used to map the non-linear relationship between inputs and outputs. A fuzzy IF-THEN rule consists of an IF part (antecedent) and a THEN part (consequent) which can be shown as follows: (Sudhir Agarwal & Pascal Hitzler, 2005; Rouzbeh Shad et al, 2009)

If antecedent proposition **Then** consequent proposition

The antecedent is a combination of terms, while the consequent is exactly one term. In this standard syntax, a term is an expression of the form X=T, where X is a linguistic variable and T is one of its linguistic terms.

For example, a simple air pollution prediction that used the detection of percentage of PM_{10} in air looks like this:

IF the percentage of PM_{10} is High THEN the air is Dirty.

In this example, the linguistic variable is 'the percentage of PM_{10} is High', and its linguistic term is 'the air is Dirty'. In Hassan's paper, Takagi-Sugeno Fuzzy Model (TS) was used on predicting UAP.

A dynamic TS fuzzy model is described by a set of fuzzy "IF…THEN" rules with fuzzy sets in the antecedents and dynamic linear time-invariant systems in the consequents. A generic TS fuzzy rule can be written as follows:

$$i^{th} \text{ Rule: If U is } M_j \text{ Then output is } (D_{j0} + D_{j1}u_1 + D_{j2}u_2 + \ldots + D_{jk}u_k) \tag{11}$$

where U is the input data vector $(u_1, u_2, \ldots, u_k)$, i.e. $u_i \in U$, M_j is the set of membership functions M_{ji} for j^{th} rule, i.e. $M_{ji} \in M_j$; M_{ji} is the membership function for i^{th} feature of jth rule and D_{ji}s represent linear parameters.

While $(D_{j0} + D_{j1}u_1 + D_{j2}u_2 + \ldots + D_{jk}u_k)$ is the output from an individual rule j, the output y from the all rules (assume c is the total number of fuzzy rules) is computed as follows:

$$\frac{\sum_{j=1}^{c} W_j y_j}{\sum_{j=1}^{c} W_j} \tag{12}$$

where,

$$W_j = \min_i M_{ji}(u) \tag{13}$$

Wj represents the weight or firing strength of jth rule for a data vector U, and $M_{ji}(u)$ is the degree of membership for j^{th} rule and i^{th} feature of an attribute rule u. y_i is the output from j^{th} rule for data vector.

In fuzzy model, both the Mamdani and the TS model (Jun Young Bae et al., 2009; F. Khaber et al., 2006) can be used, because they are depending on the desired proposition and implication of the rule (Rouzbeh Shad et al., 2009). Compared to the Mamdani model which the consequent part is a fuzzy proposition, the TS model is a crisp function of the antecedent variables. Thus, TS model was used in the hybrid HMM and Fuzzy Logic model for it produces numerical output.

3.2 Hybrid HMM and Fuzzy Logic Model
In order to improve fuzzy rule generation, Hassan et al. have introduced a hybrid HMM and fuzzy rule generation tool (M. Maruf Hossain et al., 2008). The HMM in this model is trained using the Baum-Welch algorithm (David J.C. MacKay, 2007) and available training data vectors. There are four phrases in this model: (M. Maruf Hossain et al., 2008).

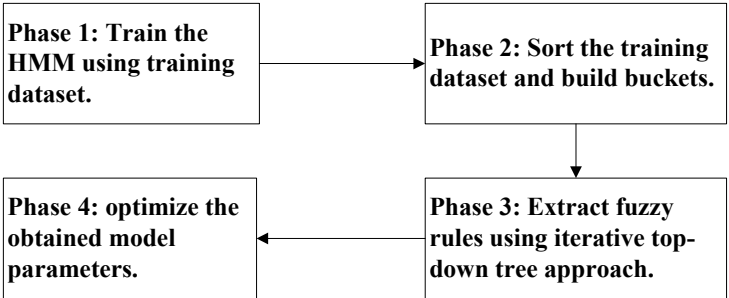| Phase 1: Train the HMM using training dataset.	→	Phase 2: Sort the training dataset and build buckets.
Phase 4: optimize the obtained model parameters.	←	Phase 3: Extract fuzzy rules using iterative top-down tree approach.

Fig. 3. The process phrases of the hmm-FL model

Firstly, an HMM is trained using the training dataset (Phrase 1) and then the training datasets are sorted and put into a number of buckets by using the HMM-log-likelihood values which are calculated in the training stage (Phrase 2). Then, a recursive divide and conquer algorithm (top-down tree approach) is used to generate a set of fuzzy rules (Phrase 3). Finally, a gradient descent method is used for further optimization of the fuzzy rule parameters (Phrase 4). The following four subsections describe more details of these four phrases respectively.

3.2.1 Generating HMM-log-likelihood values
Initially, an HMM structure is built for re-estimating the parameter values of a given dataset. Each data vector is able to form a pattern. In HMM-Fuzzy model, the HMM-log-likelihood (Behzad Zamani et al., 2010) is generated from a single HMM as the first phrase. The following equations show how HMM-log-likelihood values are computed.

For a given HMM and a sequence of observation, it is common to compute $P(O|\lambda)$, the probability of the observation sequence. This value can be used to evaluate how well a model predicts a given observation sequence. The probability of the observations O for a specific state sequence Q is: (M. Maruf Hossain et al., 2008; Phil Blunsom, 2004)

$$P(O|Q,\lambda) = \prod_{t=1}^{T} P(o_t|q_t,\lambda) = b_{q1}(o_1) \times b_{q2}(o_2) \dots b_{qT}(o_T) \qquad (14)$$

In addition, the probability of the state sequence is:

$$P(Q|\lambda) = \pi_{q1}a_{q1q2}a_{q2q3} \cdots a_{qT-1qT} \tag{15}$$

By using these two equations, we can calculate the probability of the observations as:

$$P(O|\lambda) = \sum_Q P(O|Q,\lambda)P(Q|\lambda) = \sum_{q1\ldots qT} \pi_{q1}b_{q1}(o_1)a_{q1q2}b_{q2}(o_2)a_{q2q3} \cdots a_{qT-1qT}b_{qT}(o_T) \tag{16}$$

The probability of the observations is known as the log-likelihood value, and it can be called the generated scalar value as well. We can accord Rabiner (Lawrence R. Rabiner, 1989)'s method to proof why the log-likelihood value can determine the similarity between two data patterns of k-dimensional vectors for sorting the data patterns: the log-likelihood value can show the probability that the vector was produced by the model, and the probability acts as an indicator for how well a given model matches a given vector. Thus, the entire vector can be transformed into related scalar log-likelihood values. For example, there are four data vectors which the log-likelihood values are M_1, M_2, M_3 and M_4 respectively. We can assume that the value of M_2 and M_3 are within the same tolerance level, so the data vectors that associating to M_2 and M_3 are similar. If the values of M_1 and M_4 are not close to the values of M_2 and M_3, we can find out that the data vectors which corresponding to M_1 and M_4 are not similar to those of M_2 and M_3. By using this method, we can detect that data values with similar log-likelihood values are belong to the same group.

One thing we have to mentioned is that, if we want to calculate $P(O|\lambda)$, the evaluation of the probability of O is allowed. However, to evaluate O directly would be exponential. A better approach is to use caching calculations which can lead to reduced complexity. The cache can be implemented as a trellis of states at each time step. The cached value (called θ) for each state can be calculated as a sum over all states at the previous time step. We define the forward probability variable: (Phil Blunsom, 2004)

$$\theta_t(i) = P(o_1 o_2 \ldots o_t, q_t = s_i|\lambda) \tag{17}$$

The algorithm for this process is called forward algorithm which is used in HMM-FL model. The following example can explain well how HHM-log-likelihood is generated in this model.

If we want to predict the concentration of CO in the air during a certain time period for air pollution prediction, we should measure the number of cars per hour (A), wind speed (B), temperature 2 meters above the ground (C), wind direction (D) and etc. In this example, the set of predictor variable is A, B, C and D. The cached value for each state can be visualized as in Figure 4. We can use the values of these four variables to create a data vector for the particular time at each time unit. The patterns of these variables in every data vectors are assumed to appear consecutively and differently. In Hassan et al (M. Maruf Hossain et al., 2008)'s previous work, the HMM fed into these data patterns to re-estimate the parameter values, and the HMM was used as a pattern matching tool only. Once the HMM is trained well, this HMM is used to generate a log-likelihood value for every data vector in the dataset by using the forward algorithm in our project. Every data vector or pattern is able to generate one corresponding log-likelihood value. In this case, the table of the indices of data vectors and their corresponding log-likelihood values can be generated as shown in Table 2.
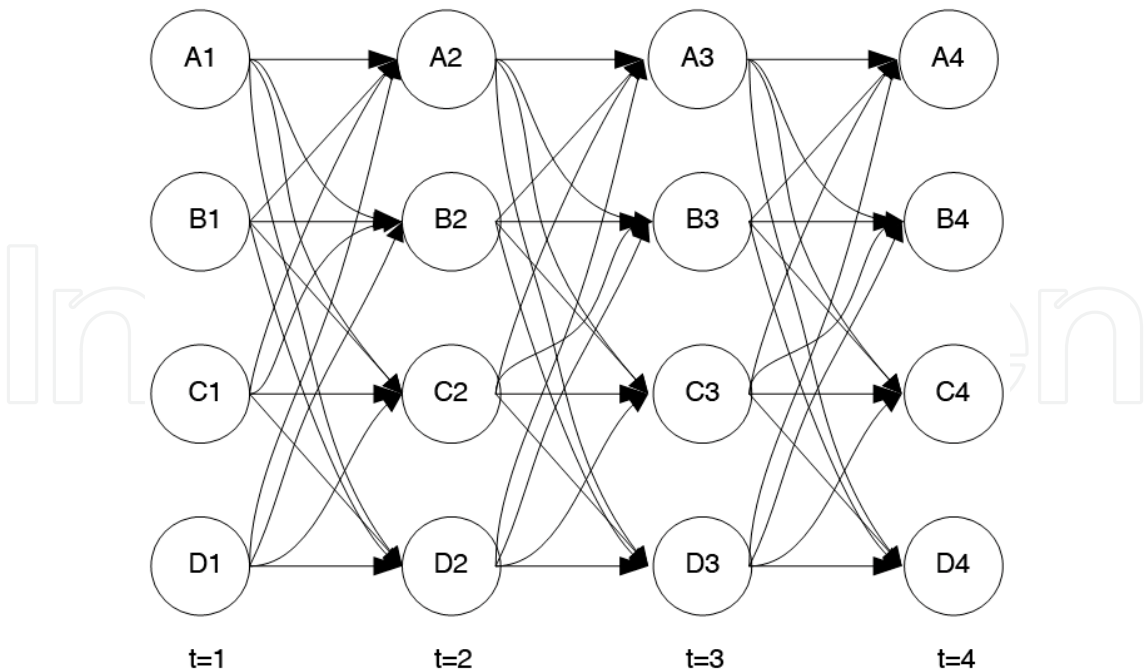
Fig. 4. A trellis algorithm example

Data vector index i	A	B	C	D		Data vector index i	Log-likelihood
1	8.5	9.25	8.25	8.25	→	1	-3.0090
2	10.5	11.25	10	10.75		2	-0.0010
3	8	9.5	8	8.5		3	-3.0080
… …						… …	
… …						… …	
i-1	10.75	11.5	10.5	11		i-1	-0.0009
I	10.5	11.25	9.5	11	→	I	-0.0008

Table 2. Generate log-likelihood values example

3.2.2 Grouping Similar Data Vectors

Grouping similar data vector is to split the range of log-likelihood values into equal sized buckets. Each bucket should contain the similar log-likelihood value of the data vectors. The fig shows there are five equal size buckets and the frequency values represents the number of similar data pattern.
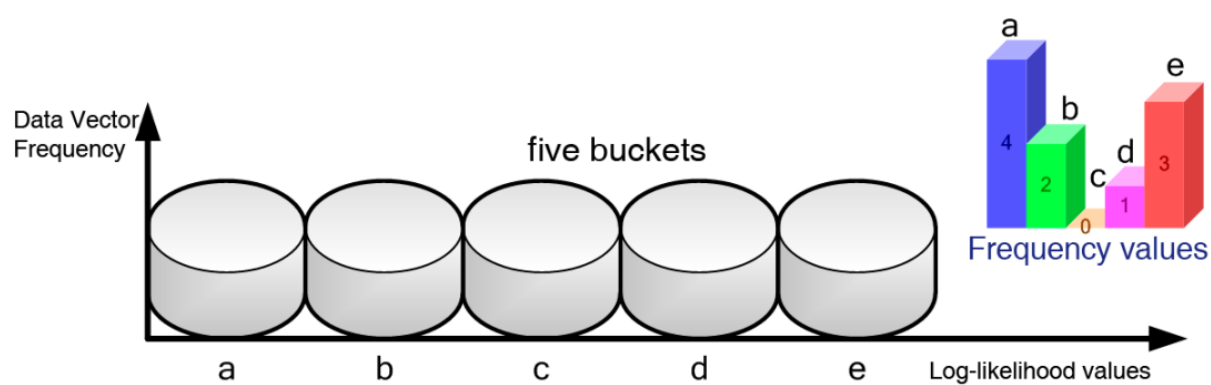
Fig. 5. Group data with similar log-likelihood values

For these buckets, each of them has a starting point and an ending point corresponding to the log-likelihood values (M. Maruf Hossain et al., 2008). The size of the bucket, W, can be used for guiding the rule extraction process. These data vectors are grouped for generating fuzzy rules and establishing the fuzzy model in the next phrase. The Figure 6 shows the pseudo-code to split the range of log-likelihood into buckets.

```
function split_values
  bucket_size = b;
  start_Range = minimum of the log-likelihood values;
  end_Range = maximum of the log-likelihood values;
  while (i<end_Range)
      bucket[j].start = i;
      bucket[j].end = i+bucket_size;
      bucket[j].data = find(log-likelihood.data >=i and log-
likelihood.data<i+bucket_size)
      i=i+bucket_size;
      j= j+1;
  end while
```

Fig. 6. The pseudo-code for split log-likelihood values into various buckets

3.2.3 The Fuzzy Model

The fuzzy rule extraction is a significant step in this model which after creating the buckets. (Md. Rafiul Hossan et al., 2009) In the fuzzy model, a divide and conquer (top-down tree) approach are used for the fuzzy rule generation. Initially, there is only one fuzzy rule which is generated for representing the entire input space of the training dataset. Under this circumstance, we use one global bucket to contain all the log-likelihood values of all the individual buckets. The process step is shown as Figure 7. In this process, mean squared error (MSE) is used to evaluate the performance of the developed model for the training dataset in this model.

The pseudo-code of the divide and conquer (top-down tree) approach (Joost Engelfriet, 1975) for rule extraction using buckets is shown below. Firstly, we set a threshold value T. If the prediction error for the training dataset is less than or equal to T, there should be no further

rules extracted and the algorithm is terminated. On the other hand, if the prediction error is greater than T, the input space is split into two parts with the help of the buckets produced in the second phrase. The method for splitting of the input space is to divide the total buckets into two equal parts. And then, we can create two individual rules for each of the parts. In this way, the total number of rules is increased by one. Then, we could use the extracted rule set to recalculate the training dataset. If the error threshold value is not greater than T, the buckets on the left side of the previous splitting are divided into two parts and the same process is iterated. This loop can be terminated only when the number of rules is equal to the number of buckets or the error threshold is less than or equal to T.

```
function rule_extraction

Threshold_Value = T ; (T is the desired error threshold value)
Extract only one rule using the entire training dataset
error = Calculate_Value(data,rules);
if(error>Threshold_Value)
  divide the total number of buckets into two parts and extract rules for each of
these parts;
  error = Calculate_Value(data,rules);
  left_flag = TRUE;
  right_flag = FALSE;
end if
  while(error>Threshold_Value)
      if(left_flag = = TRUE)
         divide the left part of buckets into two parts and extract rules for each of
         these parts;
         error = Calculate_Value(data,rules);
         left_flag = FALSE;
         right_flag = TRUE;
      else
         divide the right part of buckets into two parts and extract rules for each of
         these parts;
         error = Calculate_Value(data,rules);
         left_flag = TRUE;
         right_flag = FALSE;
      end if
   end while
return rules


function error = Calculate_Value(data,rules)

simulate results by using extracted rules
error = MSE(produced_output, actual_output);
return error;
```

Fig. 7. The pseudo-code for rule extraction

In this part, the Gaussian member function is chosen for fuzzy rule extraction. As mentioned at the beginning of this section, the inference in the TS model can be further applied in this phrase. In the step of fuzzy rule extraction, there are k membership functions existed for k variables in a data pattern. We can calculate the mean value μ and the standard deviation σ (Jun Young Bae et al., 2009; F. Khaber et al., 2006). Then, we could get the kth membership function which is:

$$M_{ji}(u) = e^{-\frac{1}{2}(\frac{u_i - \mu_{ji}}{\sigma_{ji}})^2}$$ (18)

In his equation, $M_{ji}(u)$ is the membership fuction for rule j and feature i.

3.2.4 Optimization of Extracted Fuzzy Rules

Gradient decent algorithm is used to optimize parameters for the extracted Fuzzy Rules in the last phrase (M. Maruf Hossain et al., 2008). In order to predict with better accuracy in the TS fuzzy model, the objective is to minimize the MSE for the training dataset. In the TS fuzzy model, every dataset has two parameters: one of them is the non-linear (premise) parameter, and the other is the linear (consequence) parameter. In our proposed model, the optimization technique ANFIS is used where a gradient decent method along with the least squared error (LSE) estimate is employed.

From the description of these four steps, we can understand how a hybrid AI-based tool is able to be used in UAP. This hybrid modeling is just like a "black box" (Mikko Kolehmainen et al., 2000) which combines HMM tools and the fuzzy model. The flowchart of the proposed model can show the main steps of this model clearly:
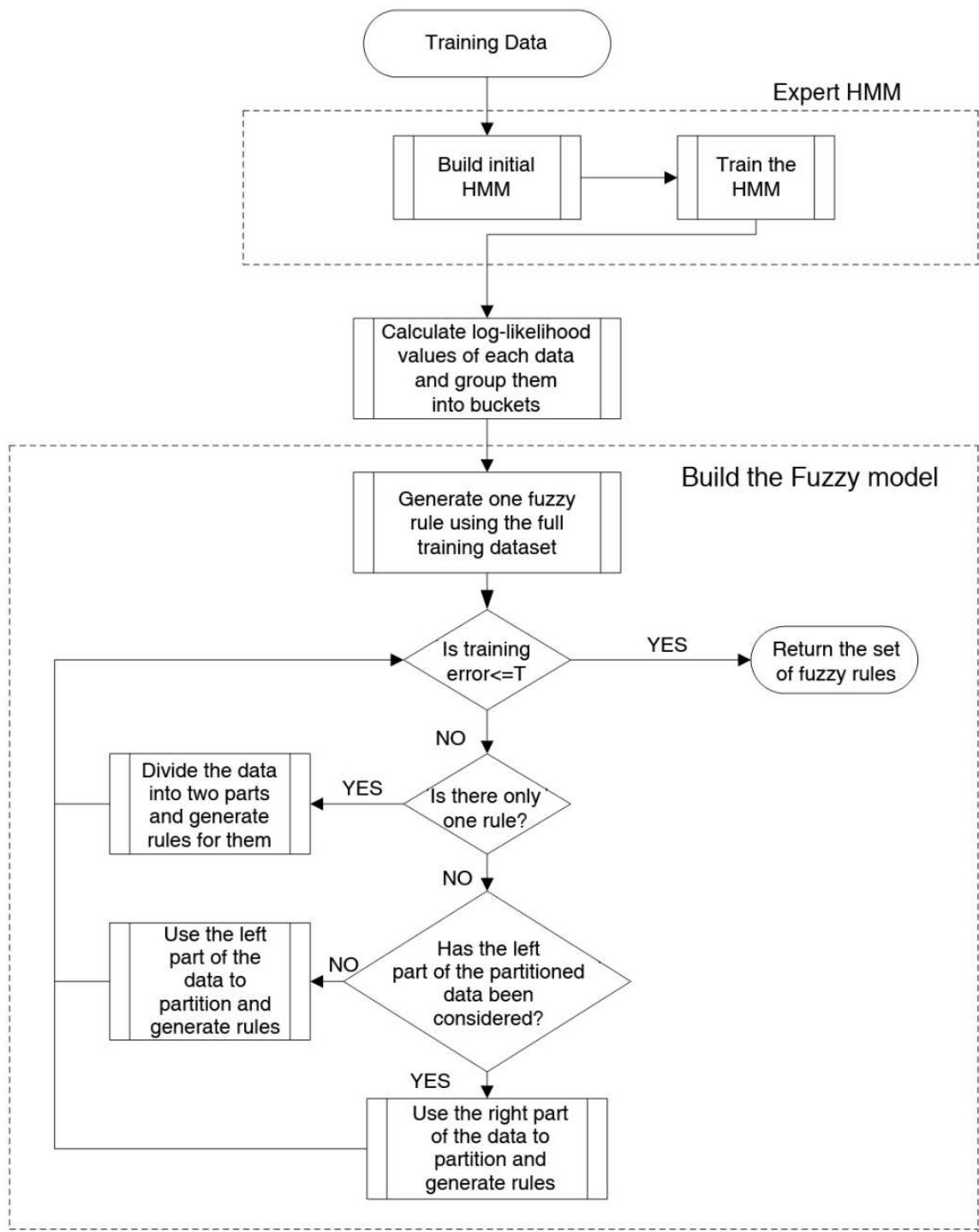
Fig. 8. The flowchart of main steps of HMM-FL model

4. Experiment and Comparison

In this section, the experiment based on HMM-FL model for predicting UAP is described. And then, the comparison of all the AI-based tools for UAP forecasting which are introduced in Section 2 is analyzed as well.

4.1 Experiment of HMM-FL Model

On the previous study of Md. Rafiul Hassan et al, the dataset which contains 500 observations is a good example (M. Maruf Hossain et al., 2008) on the experiment of HMM-FL model. This dataset is related to traffic volume and meteorological variables on a road, which is conducted by the Norwegian Public Roads Administration as a part of research on air pollution. It is based on the concentration of PM_{10} which was measured at Alnabru in Oslo from October 2001 to August 2003. The predictor variables of this dataset are the logarithm of (A) the number of cars per hour, (B) wind speed (m/s), (C) temperature 2 meters above the ground (°C), (D) the temperature deference between 25 and 2 meters above the ground (°C), (E) wind direction (within the range of 0° to 360°), (F) hour of day and (G) day number as counted from 1st October, 2001. The response variable is hourly values of the logarithm of the concentration of $PM_{10.}$ Take (B) and (C) as examples, Figure 9 shows how the dataspace is being divided by the generated rules.
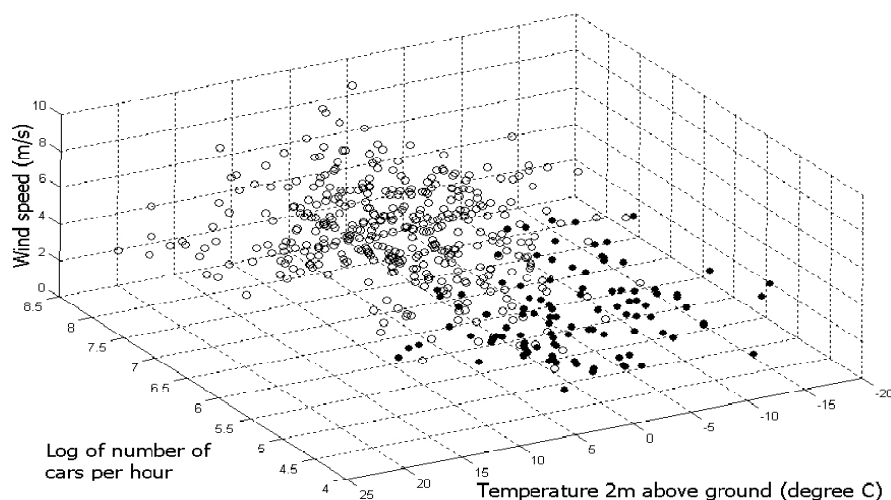

Fig. 9. Two groups in the dataspace after using HMM bucketing approach (M. Maruf Hossain et al., 2008)

In the HMM-FL model which is used on this dataset, the desired MSE was chosen to be 0.001 and the size of a bucket was 0.5. In addition, 500 epochs were chosen while executing the gradient descent algorithm for optimizing the extracted rules. In this experiment, HMM-FL model tool was executed in 10-fold cross validation. From the results, there are around 2.9 ± 1.3703 rules with confidence level of 95% or over which were generated in each fold. The fuzzy rule that actually divides the dataspace which shown in Fig. 11. and a membership function of the first attribute represented in Fig. 12 (M. Maruf Hossain et al., 2008).
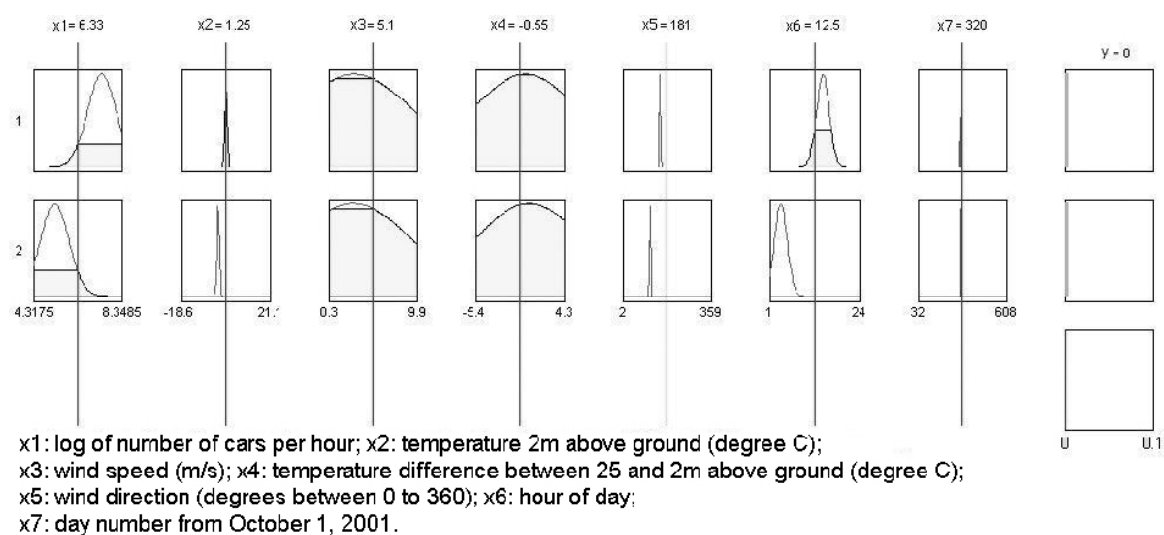
x1: log of number of cars per hour; x2: temperature 2m above ground (degree C);
x3: wind speed (m/s); x4: temperature difference between 25 and 2m above ground (degree C);
x5: wind direction (degrees between 0 to 360); x6: hour of day;
x7: day number from October 1, 2001.

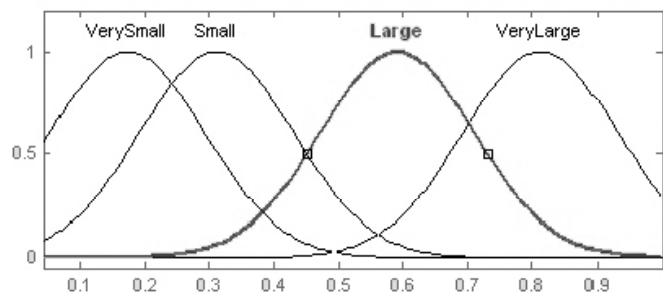Fig. 10. Two fuzzy rules for dividing the dataspace shown in Fig. 9.

Fig. 11. Membership function of the first attribute shown in Fig. 10.

4.2 Results Comparison

From the comparison of the existed AI-based methodologies, we can find out that ANNs is more popular than others for predicting UAP. Ulku Sahin et al (Ulku Sahin et al., 2005) and Viotti et al (Viotti et al., 2002) introduced the ANN-based tools. In Ulku Sahin et al's paper, they evaluated the performance by using ANN model or results to compare to other classical nonlinear methods. The correlation parameter is 0.999 and 0.528 for training and test data. P. Viotti et al also used ANNs on their UAP forecasting study. They tested various pollutants based on 48 hours and 500 hours respectively. Take Ozone as a example, they used a training set of 3500 patterns and a test set of 2300 patterns and two validation sets, 500 and 48 respectively. The number of neurons was 13 and about 10000 epochs were performed at constant learning rate of 0.3. The results for the two validations went to 0.126 (48 hours) of a relative MSE and 0.19 (500 hours) of a relative MSE. From these studies, we can see that ANNs' behavior has always been related to non-linear statistical regression. It seems that it is naturally suited for problems that show a large dimensionality of data, such as UAP prediction system which is the task of identification for systems with a large number of state variables.

SVMs are not often used on UAP detection, but from Wei-Zhen Lu et al's view (Wei-Zhen Lu & Wen-Jian Wang, 2005), it can also be used for regression and time series prediction and have been reported to perform well by some promising results. By comparing SVM and

radial basis function (RBF) on different months, the mean absolute error (MAE) produced by the SVM method is smaller than the ones created by the conventional RBF network in both December and June. These experiments show that SVM is superior to RBF. It is because SVM can process robust predicting performance.

The results from FL model are also promising. In Yilmaz Yildirim's research (Yilmaz Yildirim & Mahmut Bayramoglu, 2006), adapitive neuro-fuzzy logic method has been proposed on testing SO$_2$ and total suspended particular matter (TSP) pollution levels over an urban area. It shows that for SO$_2$ and TSP the model indicating acceptable forecasting limits are between 75-90% and 69-80%. It is possible to predict the air quality levels with high accuracy with a better set of training patterns in this study.

The combination of AI-based tools which contain ANNs technologies have been also used by Luis A. Diaz-Robles et al (Luis A. Diaz-Robles, et al., 2008). In this experiment, they combined ARIMA and ANNs model to improve forecast accuracy for an area with limited air quality and meteorological data. By comparison, the hybrid model had better furcating performance than other models which were tested.

HMM-FL model is a novel hybrid AI-based model for UAP prediction. In the experiment which shows in Section 4.1, there are two other models which have been tested for comparing to HMM-FL model. They are an ANN model and a forecasting model using the subtractive clustering-based fuzzy model (S. Chiu, 1997). The ANN had 7 nodes in the input layer, 21 nodes in the hidden layer and I node in the output layer. The epochs and training goal were 500 and 0.001 respectively. MSE of HMM-FL model has the best results in this study, which is 0.0097 (M. Maruf Hossain et al., 2008). It shows that HMM-FL model has the potential to achieve high levels of performance on forecasting concentrations of UAP variables.

The results which come from these papers are collected and put in the Table 3. Besides MSE, in these papers, the mean absolute error (MAE) and the root mean square error (RMAE) are used as assessment indicators (Mikko Kolehmainen et al., 2000; Luis A. Diaz-Robles, et al., 2008). The MAE is used for measuring the average magnitude of the errors in a set of forecasts without considering their direction. It is usually on measuring accuracy for continuous variables. RMSE is the square root of MSE. It is a quadratic scoring rule which can measure the average magnitude. The MAE and RMSE are sometimes used together to diagnose the variation in the errors in a set of forecasts. Both of the MAE and RMSE can range from 0 to ∞. In addition, they are negatively-oriented scores which mean that lower values are better. They can be defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|o_i - p_i|, \tag{19}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2} \tag{20}$$

Where o_i is the actual values of pollutants' concentrations with {i=1,2,…,n} observations, n is the total observation number and p_i is the predicted pollutants value. The following table shows a part of results from statistic analysis.

AUTHORS	NUMBER OF OBSERVED	AI-BASED TOOLS	THE RESPONSE VARIABLE	STATISTICAL PARAMETERS	STATISTIC ANALYSIS
M. Maruf Hossain et al. 2008	500	HMM-FL model	PM_{10}	MSE	$0.0097(\mu g^2/m^6)$
S. Chiu et al. 1997	500	Fuzzy model following subtractive clustering	PM_{10}	MSE	$0.0102(\mu g^2/m^6)$
M. Maruf Hossain et al. 2008	500	a ANN model	PM_{10}	MSE	$0.0216(\mu g^2/m^6)$
Ulku Sahin et al. 2004	231	a ANN model	SO_2	Mean absolute error (MAE)	$0.103(\mu g/m^3)$
				Root mean square error (RMSE)	$0.448(\mu g/m^3)$
Wei-Zhen Lu et al. 2005	168	SVM method	Respirable suspended particulate (RSP)	MAE	$17.657(\mu g/m^3)$
			NO_2	MAE	$13.128(\mu g/m^3)$
			NO_x	MAE	$131.645(\mu g/m^3)$
Luis A. Diaz-Robles et al. 2008	2080	A novel hybrid model combining ARIMA and ANN	PM_{10}	MAE	$6.74(\mu g/m^3)$
				RMSE	$8.80(\mu g/m^3)$
Mikko Kolehmainen et al. 2000	--	A hybrid fuzzy model	NO_2	RMSE	$12.2 (\mu g/m^3)$
			CO	RMSE	$0.3 (mg/m^3)$
			PM_{10}	RMSE	$11.1(\mu g/m^3)$

Table 3. The statistic analysis of AI-based tools on UAP prediction

5. Discussion and Conclusion

The AI techniques which are used as UAP forecasting tool can give clear and intuitive results. It is because air quality time series contains complex linear and non-linear patterns, and most methodologies cannot be used on non-linear patterns except AI techniques methodologies (Harri Niska et al., 2004; Lovro Hrust et al., 2009). Thus, combining AI

techniques, such as ANNs, SVMs and FL, with some other methods can recognize different patterns and improve the performance of UAP prediction. This book chapter represents and summarizes the current reliable researches on which AI-based tool are implemented. Although single AI technique based tools are popular and efficient, they still cannot avoid their inherent drawbacks. For example, ANNs can get over-fitted into training rules and stuck in local minima during training, SVMs is more likely to be built as the kernel-based hyper plane separation techniques than as forecasting tools, and FL suffers from the computational complexity due to its interpretability reduces. Compared to single AI-based forecasting tool, there are many models based on hybrid AI-based tools, such as a hybrid ARIMA and ANNs tool from Luis A. Diaz-Robles et al and adaptive neuro-fuzzy based modeling from Yilmaz Yildirim et al.

Combination of the HMM and Fuzzy model is a novel hybrid AI based tool that can be used on UAP forecasting (M. Maruf Hossain et al., 2008). It can improve Fuzzy model by using the HMM's data partition approach which the relationship between data features. The Markov process can be used on detecting the current event according to the immediate past event in the data patterns. In addition, the top-down tree approach can generate optimized number of fuzzy rules for the non-linear data. All of these features can make the generated fuzzy model provide a better performance.

For the UAP prediction experiment, the datasets usually contain the response variables and predictor variables. In the testing of HMM-FL model, there are 7 predictor variables used for predicting the concentration of PM_{10}. In order to determine the efficiency of HMM-FL model, a fuzzy model which following subtractive clustering and another ANN model were tested for comparing results. By using MSE values for the evaluation, HMM-FL shows the better performance. The results represent that other techniques trained the input features as independent individuals which made complex systems. It further proves that HMM-Fuzzy model can reduce complexity and simultaneously improve the accuracy of predicting. However, a further performance can be achieved if a better weighting scheme which is used for generated fuzzy rules can be developed. In addition, larger size samples and various variables are required for further research.

6. References

Per Nafstad, Lise Lund Haheim, Torbjorn Wisloff, Frederisk Gram, Bente Oftedal, Ingar Holme, Ingvar Hjermann, and Paul Leren (2004). *Urban Air Pollution and Mortality in a Cohort of Norwegian Men*. Environmental Health Perspectives, Vol. 112, No. 5.

World Health Organization, *Urban Air Pollution with Particular Reference to Motor Vehicles*. World Health Orgnization Technical Report Series, No. 410.

Nyberg F, Gustavsson P, Järup L, Bellander T, Berglind N, Jakobsson R, Pershagen G (2000). *Urban air pollution and lung cancer in Stockholm*. Epidemiology. Sep;11(5):487-95.

J. Sunyer, C. Spix, P. Quenel, A. Ponce-de-Leon, A. Ponka, T. Barumandzadeh, G. Touloumi, L. Bacharova, B. Wojtyniak, J. Vonk, L. Bisanti, J. Schwartz, and K. Katsouyanni (1997). *Urban air pollution and emergency admissions for asthma in four European cities: the APHEA Project*. Thorax. 1997 September; 52(9): 760–765.

Mikko Kolehmainen, Hannu Martikainen, Teri Hiltunen, and Juhani Ruuskanen (2000). *Forecasting air quality parameters using hybrid neural network modeling.* Environmental Monitoring and Assessment 65:277-286.

Yilmaz Yildirim, Mahmut Bayramoglu (2006). *Adaptive neuro-fuzzy based modeling for prediction of air pollution daily levels in city of Zonguldak.* Chemosphere 63 1575-1582.

Nils J. Nilsson (1998). *Artificial Intelligence: A New Synthesis*, Chapter 1. Morgan Kaufmann Publishers, Inc.

Harri Niska, Teri Hiltunen, Ari Karppinen, Juhani Ruuskanen, and Mikko Kolehmainen (2004). *Evolving the neural network model for forecasting air pollution time series.* Engineering Applications of Artificial Intelligence 17, 159-167.

Wei-Zen Lu, Wen-Jian Wang (2005). *Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends.* Chemosphere 59: 693-701.

Francesco Carlo Morabito, Mario Versaci (2003). *Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data.* Neural Networks 16, 493-506.

A.B.Chelani, S.Devotta (2007). *Prediction of ambient carbon monoxide concentration using nonlinear time series analysis technique.* Transportation Research Part D 12, 596-600.

Jef Hooyberghs, Clemens Mensink, Gerwin Dumont, Frans Fierens, Olivier Brasseur (2005). *Aneural network forecast for daily average PM_{10} concentrations in Belgium.* Atmospheric Environment 39 (2005) 3279-3289.

U. Brunelli, V.Piazza, L.Pignato, F.Sorbello, S.Vitabile (2007). *Two-days ahead prediction of daily maximum concentrations of SO2, O3, PM_{10}, NO2,CO in the urban area of Palermo, Italy.* Atmospheric Environment 41 (2007) 2967-2995.

M. Maruf Hossain, Md. Rafiul Hassan, and Michael Kirley (2008). *Forecasting urban air pollution using HMM-Fuzzy Model.* PAKDD 2008, LNAI 5012, pp.572-581.

Luis A. Diaz-Robles, Juan C. Ortega, Joshua S. Fu, Gregory D. Reed, Judith C. Chow, John G. Watson, Juan A. Moncada-Herrera (2008). *A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile.* Atmospheric Environment 42, 8331-8340.

Giorgio Corani (2005), *Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning.* Ecological Modelling 185, 513-529.

Oleg M. Pokrovsky, Roger H.F.Kwok, C.N.Ng (2002). *Fuzzy logic approach for description of meteorological impacts on urban air pollution species: a Hong Kong case study.* Computers & Geosciences 28, 119-127.

Md. Rafiul Hassan (2009). *A combination of hidden Markov model and fuzzy model for stock market forecasting.* Neurocomputing 72, 3439-3446.

Md. Rafiul Hassan (2007). *Hybrid HMM and Soft Computing modeling with applications to time series analysis.* Department of Computer Science and Software Engineering, the University of Melbourne, Australia.

Joost Engelfriet (1975). Bottom-up and Top-down tree transformations – a comparison. Mathematical Systems Theory, Vol 9, No.3., 1975 by Springer-Verjag. New York Inc.

Dahe Jiang, Yang Zhang, Xiang Hu, Yu Zeng, Jiangguo Tan, Demin Shao (2004). *Process in developing an ANN model for air pollution index forecast.* Atmospheric Environment 38 (2004) 7055-7064. Elsevier Ltd.

Patricio Perez, Rodrigo Palacios and Alejandro Castillo (2004). *Carbon Monoxide Concentration Forecasting in Santiage, Chile.* Journal of the air and waste management association 54:908-913. ISSN 1047-3289.

Patricio Perez, Jorge Reyes (2006). *An integrated neural network model for PM₁₀ forecasting.* Atmospheric Environment 40 (2006) 2845-2851. Elsevier Ltd.

Harri Niska, Teri Hiltunen, Ari Karppinen, Juhani Ruuskanen, Mikko Kolehmainen (2004). *Evolving the neural network model for forecasting air pollution time series.* Engineering Applications of Artificial Intelligence 17 (2004) 159-167. Elsevier Ltd.

Comrie A. C. (1997). *Comparing neural networks and regression models for ozone forecasting.* Journal of Air and Waste Management Assiciation 47, 655-663.

Gardner M. W. Dorling S. R (1999). *Neural network modeling and prediction of hourly NO_x and NO_2 concentrations in urban air in London.* Atmospheric Environment 33, 709-719.

Yi J., Prybutok V. R. (1996). *A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area.* Environmental Pollution 92, 349-357.

Jun Young Bae, Youakim Badr, Ajith Abraham (2009). *A Takagi-Sugeno Fuzzy Model of a Rudimentary Angle Controller for Artillery Fire.* Institut National des Sciences Appliquees, INSA-Lyon, F-69621, France.

F. Khaber, K. Zehar, and A. Hamzaoui (2006). *State Feedback Controller Design via Takagi-Sugeno Fuzzy Model: LMI Approach.* International Journal of Computational Intelligence 2;3. The CReSTIC laboratory, I.U.T of Troyes , University of Reims, France.

Behzad Zamani, Ahmad Akbari, Babak Nasersharif, Mehdi Mohammadi and Azarakhsh Jalalvand (2010). *Discriminative transformation for speech features based on genetic algorithm and HMM likelihoods.* IEICE Electronic Express, Vol.7, No.4, 247-253.

Phil Blunsom (2004). *Hidden Markov Models.* Department of Computer Science and Software Engineering, The University of Melbourne.

Md. Rafiul Hassan, M. Maruf Hossain, Rezaul Karim Begg, Kotagiri Ramamohanarao, Yos Morsi (2009). *Breast-Cancer identification using HMM-Fuzzy approach.* Computers in Biology and Medicine. October 2009.

Ulku Sahin, Osman N. Ucan, Cuma Bayat and Namik Oztorun (2005). *Modeling of SO_2 distribution in Istanbul using artificial neural networks.* Environmental Modeling and Assessment (2005) 10: 135-142. Springer.

Lovro Hrust, Zvjezdana Bencetic Klaic, Josip Krizan, Oleg Antonic, Predrag Hercog (2009). *Neural network forecasting of air pollutants hourly concentrations using optimized temporal averages of meteorological variables and pollutant concentrations.* Atmospheric Environment 43 (2009) 5588-6696. Elsevier Ltd.

P. Viotti, G. Liuti, P. Di Genova (2002). *Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia.* Ecological Modelling 148 (2002) 27-46. Elsevier Science B.V.

Wei-Zhen Lu, Wen-Jian Wang, Xie-Kang Wang, Sui-Hang Yan, and Joseph C. Lam (2004). *Potential assessment of a neural network model with PCA/RBF approach for forcasting pollutant trends in Mong Kok urban air, Hong Kong.* Environmental Research 96 (2004) 79-87. 2003 Elsevier Inc.

Ming Cai, Yafeng Yin, Min Xie (2009). *Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach.* Transportation Research Part D 14 (2009) 32-41. 2008 Elsevier Ltd.

W. Z. Lu, W. J. Wang, X.K. Wang, Z. B. Xu and A. T. Leung (2002). *Using inproved neural network model to analyze RSP, NOx and NO2 levels in urban air in Mong Kok, Hong Kong.* Environmental Monitoring and assessment 87: 235-254, 2003. Kluwer Academic Publisher. Netherlands.

Rouzbeh Shad, Mohammad Saadi Mesgari, Aliakbar Abkar, Arefeh Shad (2009). *Predicting air pollution using fuzzy genetic linear membership kriging in GIS.* Computer ,Environment and Urban System 33(2009) 472-481. 2009 Elsevier Ltd.

Matthew J. Beal, Zoubin Ghahramani, Carl Edward Rasmussen (2002). *The Infinite Hidden Markov Model. Gatsby Computational Neuroscience Unit University College, London.* 17 Queen Square, London WC1N 3AR, England.

Lawrence R. Rabiner (1989). *A tutorial on Hidden Markov Models and selected applications in speech recognition.* Proceedings of the IEEE, Vol.77, No.2, Debruary 1989.

Sudhir Agarwal and Pascal Hitzler (2005). *Modeling Fuzzy Rules with Description Logics. Institute of Applied Informatics and Formal Description Methods.* (AIFB), University of Karlsruhe (TH), Germany.

S. Chiu (1997). *Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification.* Chapter 9 in Fuzzy Information Engineering: A guided Tour of Applications. Ed: D. Dubois, H. Prade, and R. Yager, John Wiley & Sons, 1997.

David J. C. MacKay (1997). *Ensemble Learning for Hidden Markov Models.* Cavendish Laboratory, Cambridge CB3, DHE, UK.

Air Pollution

Edited by Vanda Villanyi

Although the climate of the Earth is continually changing from the very beginning, anthropogenic effects, the pollution of the air by combustion and industrial activities make it change so quickly that the adaptation is very difficult for all living organisms. Researcher's role is to make this adaptation easier, to prepare humankind to the new circumstances and challenges, to trace and predict the effects and, if possible, even decrease the harmfulness of these changes. In this book we provide an interdisciplinary collection of new studies and findings on the score of air pollution.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Rafiul Hassan and Min LI (2010). Urban Air Pollution Forecasting Using Artificial Intelligence-Based Tools, Air Pollution, Vanda Villanyi (Ed.), ISBN: 978-953-307-143-5, InTech, Available from: http://www.intechopen.com/books/air-pollution/urban-air-pollution-forecasting-using-artificial-intelligence-based-tools

INTECH
open science | open minds