Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 9, No. 1, 2017 Article ID IJIM-00857, 7 pages
Research Article

The fuzzy logic in air pollution forecasting model

F. Abbasi *†

Received Date: 2016-01-02 Revised Date: 2016-09-23 Accepted Date: 2016-10-21

Abstract

In the paper a model to predict the concentrations of particulate matter PM10, PM2.5, SO2, NO, CO and O3 for a chosen number of hours forward is proposed. The method requires historical data for a large number of points in time, particularly weather forecast data, actual weather data and pollution data. The idea is that by matching forecast data with similar forecast data in the historical data set it is possible then to obtain actual weather data and through this pollution data. To aggregate time points with similar forecast data determined by a distance function, fuzzy numbers are generated from the forecast data, covering forecast data and actual data. Again using a distance function, actual data is compared with the fuzzy number to determine how the grade of membership. The model was prepared in such a way that all the data which is usually imprecise, chaotic, uncertain can be used.

Keywords : Suspended particles; Fuzzy number; Mathematical model; Distance function.

1 Introduction

THe first test predict every phenomena, especially weather started a bout 650 Bc By the Babylonians. They hare been According to the clouds, short term fore casting of climate changes. Forecast methods climates completed in after centuries increasingly. As a result the development of mathematics and physical models using equations differential are formulated about trivial derivatives. The equations that describe the climate can be solved numerically. Thus in 1961 E-LORENZ the limitations of this model is expressed as follows:
A: chaotic para meters,
B: The effectiveness of a few days, at most a week.
In recent years many such statistical forecasting methods, Fuzzy logic, neural networks, fuzzy neural emerged forecasting. Using short-term numerical weather prediction, forecast pollution levels began [7, 8]. It is very difficult, apart from in formation about the weather because the air pollution emission depends first of all on emission. Measuring low emission, social and urban is nearly impossible. In addition to the field of emission calculation models D3 with two measuring temperature in a few hundred meters above the ground that such measurements is in limited areas of world with the help of sodar. Fuzzy set theory is useful in this situation [10]. The use of this method in many mathematical models to predict known. Fuzzy theory is used when the Data transferred to the model is ambiguous or incomplete. Many everyday phenomena of nature are vague. Based on this theory, fuzzy is helpful. The problem is that our knowledge do not have complete information about future climate and we have only numerical forecast. Forecast pm10 can termination due to significant impact on human life have been selected and at each stage, meteorological data with a mathematical device

*Corresponding author. k.9121946081@gmail.com
†Department of Mathematics, Ayatollah Amoli Branch, Islamic Azad University, Amol, Iran.

use.
The model presented in this paper predict the concentrations of pollution. Especially PM10, PM2.5, SO, 2, NO, CO, O3. This model of most models excels because of forecast pollution levels very depending stage time between data for select the number of hours in front. Because, in most cases, the time is 1 hour. But this model can also be used with small steps or larger time. For example predicted 10 minutes, 1 hour, 1 day have. In, addition , this model can predict pollution contingency is a special day to a other day forecast weather conditions we have that day. The model presented in this paper is used based on the data for the desired regions under the circumstances (having a proper database of data).
The paper is organized in the following way. In Section 2 we describe basic terms used in forecast pollution concentrations model. In Section 3 we present three terms: fuzzy numbers, grouping and fuzzy grouping. In Section 4 we talk over the algorithm of changing sequence into a fuzzy number. In Section 5 we present air pollution forecasting model and we describe in details its steps. In Section 6 we give some conclusions.

2 Preliminaries

In this section, some notations and background about the concept are brought.

Definition 2.1 *(Forecast) Forecasts, is the estimation process unknown situations. A forecast, provides about future events and can turn past experiences to predict future events. In recent years ,forecasts, to forecast demand in daily business practices has become vendors.[1]*

Definition 2.2 *(Pollutants, air pollution and particulate matter) The definition of pollution, it is said that the spread of contaminants to a material amount too much hurt in the environment to human health and other harmful organisms.*
Contaminants: in general , air pollutants are divided tow.

Categories: first and second type.
Contaminants first: these are pollutants the have the same basic shape, the air can be entered, for example, carbon monoxide emissions from forest fires naturally and by man is produced by Pollution: If the chemical reaction between air pollutants and natural components of the air takes place, which has produced other pollutants to the second type is called.

Definition 2.3 *(Air pollution) Air pollution in the presence of undesirable substances in the air an amount that would not cause harmful effects, Air pollution subs tances the are related to the following general forms found:*
-Aerosol,
-Sand and gravel,
-Foam,
-Smoke,
-Soot,
-Toxic gases and toxic,
-Steam.

Definition 2.4 *(Particular matter) The definition of particulate matter refers to particles that are dispersed (solid or liquid) placed in a gaseous medium it is also pm a term that describes particles in the air. In other words, the so-called particulate matter(PM) particles and droplets that are air borne particles can noise sources a highly variable sizes.*

Definition 2.5 *(Time series , fuzzy logic and fuzzy time series) Definition of time series:*
The time series analysis largely from branches of engineering, physical sciences and economics is located and it can be said that most branches of science led to the study data, which occur in the form of time series.

The purpose of a time series of statistical data that are collected at intervals of every regulatory equal and such statistical data used statistical methods that favors called time series analysis. In the other words: the time series is of observations xt which is registered: characterized by $\{x_t, t \in T\}$ Timer serie.What is continuous and continuous time in the practice of the time such as day, week, month and year are used. Classic components of a time series time can be a series of different components. The components are as follows:

$$X_t = T_t + S_t + C_t + E_t$$

Where in
T_t: Rondo component term changes in average,
S_t : seasonal changes in the calendar component cyclical,
C_t: cyclical changes in the components of a very

changes and period,
E_t: The remaining components of the random or unexplained changes.
Definition of fuzzy logic:
Any expression of reality is not quite right or wrong the fact that something wrong is complete the full truth. Something Between zero and one, the concept of multi-valued or gray. However, the gray between black and white fuzzy things. Fuzzy logic [7] versus Aristotelian logic that only ten against two black and white yes and no zero and one see is lo cated. This logic is located between zero and one and the absolute absence of accountability (only zero and one) member of accrued amount collection. Fuzzy logic is a mechanism by which the complex systems that are they use mathematical and classic modeling techniques impossible or very difficult easily modeled with much more flexibility.
The definition of fuzzy time series:
Y(t) (variable values at time t) consider the following set of real number s (t=o,1,2,) and Also, changes in the range by the fj (t) be separate.. Then, the fj(t) is nota fuzzy time series y(t).[9]

3 Fuzzy numbers, grouping and fuzzy grouping

Definition 3.1 *(Fuzzy numbers) The conception a lot of fuzzy numbers is.[6, 10] Before we give the definition that we use in the paper let us denote the family of fuzzy sets on $\Re$ by $F_s(\Re)$.*

The fuzzy set $A \in F_s(\Re)$, whose membership function:

$$\mu_A : \Re \to [0,1] \tag{3.1}$$

satisfies the following conditions:
1. $\exists x \in \Re,\ \mu_A(x) = 1$,
2. $\forall x_1, x_2,\ \lambda \in [0,],\ \mu_A(\lambda x_1 + (1-\lambda x_2)) \geq min(\mu_A(x_1), \mu_A(x_2))$,
3. μ_A is an interval continuous function,
will be called a fuzzy number. In the paper we use fuzzy numbers with membership function given by

$$\mu_G(x) = \begin{cases} exp(\frac{-(x-m_1)^2}{2\delta_1^2}), & x \leq m_1, \\ 1, & x \in (m_1, m_2), \\ exp(\frac{-(x-m_2)^2}{2\delta_2^2}), & x \geq m_2, \end{cases} \tag{3.2}$$

where,
$m_1, m_2, \delta_1, \delta_2 \in \Re$,
$m_1 \leq m_2,\ \delta_1, \delta_2 > 0$.
The membership function is approximated by the Gaussian function. Let us denote the fuzzy numbers set by FN.
Discrete fuzzy number:
In the algorithm of changing a sequence into a fuzzy number we use a discrete form of the fuzzy number which is defined in the following way. Let us assume that $A \in F_s(\Re)$ and $\mu_A(x_i)$ for $x_i \in \Re,\ i = 1, ..., n$. We will call a set of pairs $A = \{(x_i, \mu_A(x_i)|i = 1, ..., n\}$ with the membership functions $\mu_A(x_i)$ a discrete fuzzy number. Let us denote by $d - FN$ the set of all discrete fuzzy numbers.

Definition 3.2 *(Grouping) Let us introduce the distance*

$$d^d : \Re^d \times \Re^d \longrightarrow [0, \infty)$$

$$d^d(x, y) = (\sum_i |x_i - y_i|^k)^{\frac{1}{k}} \tag{3.3}$$

where $x, y \in \Re^d,\ k > 0$.
For $k \in N,\ k \geq 1$ the distance (3.3) is a metric. Let us denote the set of matrices $\Re^{n \times m}$ by X. In the grouping we find matrices similar to the chosen $f^* \in X$ (k-nearest neighbours method). From X we choose a subset $\epsilon - x(f^*) \subseteq X$ that:

$$f \in \epsilon - x(f^*) \Leftrightarrow d^x(f^*, f) < \epsilon \tag{3.4}$$

where $\epsilon > 0,\ f \in X$ and

$$d^x(f^*, f) = d^m([d^n(f_j, f_j^*)]_{j=1,...,n},\ 0) \tag{3.5}$$

where f_j means jth column of the matrix f. We need e to be as small and near to 0 as possible and $|\epsilon - x(f^*| > 1$, because the matrices from $\epsilon - x(f^*)$ will be close to each other. 0 - vector in which all the elements are equal to 0.

Definition 3.3 *(Fuzzy grouping) Let $y \in Y$ and $\phi \in F$. Membership of a matrix y to the fuzzy numbers matrix ϕ is a matrix given by*

$$\phi(y) = [\mu_{ij}(y_{ij})]_{i=1,...,n,\ j=1,...,m} \tag{3.6}$$

where μ_{ij} is a membership function of the fuzzy number ϕ_{ij} for i = 1,. . . ,n and j = 1,. . . ,m.
Module of a membership of a matrix y to the fuzzy numbers matrix ϕ is given by

$$|\ \phi(y)\ | = d^{n,m}([\mu_{11}(y_{11}), ..., \mu_{1m}(y_{1m})$$

$$, ..., \mu_{n1}(y_{n1}), ..., \mu_{nm}(y_{nm})]^T,\ 0) \tag{3.7}$$

where μ_{ij} is a membership function of the fuzzy number ϕ_{ij} for i = 1,. . . ,n and j = 1,. . . ,m. In the fuzzy grouping at first we have a chosen $\phi^* \in F$ and we fix $\eta > 0$. From Y we choose subset $\eta - Y$ that:

$$s \in \eta - Y \Leftrightarrow |\phi^*| > \eta \tag{3.8}$$

We want the membership of a matrix s to the fuzzy numbers matrix ϕ^* to be the nearest to the matrix in which all elements are equal to 1, so $|\phi^*(s)|$ is the smallest. In this way $|\phi^*(s)| = 0$ membership of a matrix s to the fuzzy numbers matrix ϕ^* consist of same ones so every real number will be fully included in the fuzzy number.

4 Changing a sequence into a fuzzy number algorithm (SFN)

The way of changing a sequence into a fuzzy number is described in [5]. In Section 4.1 we change the number sequence X into a discrete fuzzy number B, in Section 4.2 we change B into a fuzzy number which membership function is of the form defined by (3.2).

4.1. Algorithm 1

Input: number sequence $X = \{\xi_1, \xi_2, ..., \xi_m\}$ where $\xi^i \in \Re$, $m \in N$, and for i = 1,. . . ,m and a parameter $n \in N(n > 1)$.

Output: $B = d - FN$. Having the sequence X, we define:

$$\alpha = min\, X,\ \beta = max\, X,\ h = \frac{\beta - \alpha}{n - 1} \tag{4.9}$$

Let us define:

$$V_i = \alpha + h(i - \frac{1}{2}),\ i = 1, ..., n \tag{4.10}$$

$$\alpha_i = \{\xi \in X \mid V_{i-1} \leq \xi \leq V_i\},\ i = 1, ..., n \tag{4.11}$$

$$x_i = \alpha + (i - 1)h \tag{4.12}$$

Finally we define B as the discrete fuzzy number:

$$B = \{(x, \mu_B(x_i) \mid \mu_B(x_i) = \frac{|\alpha|}{max(|\alpha_1|, ..., |\alpha_n|)},\ \mu_B(x_i) \neq 0\} \tag{4.13}$$

4.2. Algorithm 2

Input: $x_1, x_2, ..., x_n \in \Re$, $y_1, y_2, ..., y_n \in \Re$ and weights $w_1, w_2, ..., w_n \in \Re$. In particular $w_i = 1$ or $w_i = \frac{1}{y_i}$ for i = 1,. . . ,n.

Output: Fuzzy number which membership function is of the form (3.2). Let us assume that the set of pairs $B = \{(x_i, y_i) \mid i = 1, ..., n\}$, where $y_i = \mu_B(x_i)$ for i = 1,. . . ,n is a discrete fuzzy number. We will determine the membership function (3.2) in the mean-square sense from the discrete fuzzy number B. Function (3.2) contains four parameters. For simplicity, let us denote them by $p_1 = m_1, p_2 = \delta_1, p_3 = m_2, p_4 = \delta_1$. To compute parameters p_1, p_2, p_3, p_4 we use the mean-square approximation. Practically, we can write that function (3.2) is dependent on 5 parameters: x and p_1, p_2, p_3, p_4.

$$\mu(x) = \mu(x, p_1, p_2, p_3, p_4) = \mu(x; p) \tag{4.14}$$

Afterwards we calculate the minimum value of this function. For this purpose we use the gradient method. Therefore, we define function:

$$\chi^2(p) = \frac{1}{n} \sum_i w_i(y_i - \mu(x_i; p))^2. \tag{4.15}$$

Let us define:

$$M = \{x_i \mid y_i = 1,\ i = 1, ..., n\} \tag{4.16}$$

$$p_1 = min_{x \in M}(x), p_2 = \frac{p_1 - x_1}{2}, \tag{4.17}$$

$$p_3 = max_{x \in M}(x), p_4 = \frac{x_n - p_3}{2} \tag{4.18}$$

To find the minimum of the function χ^2 we calculate partial derivatives and then we equal them to 0. Additional difficulties of finding the minimum of χ^2 are the dependencies between the parameters which come from the form of the function μ : $p_1 \leq p_3,\ p_2 \geq 0,\ p_4 \geq 0$. The following equalities arise:

$$\frac{\partial \chi^2}{\partial p_i} = \sum_i \mu((x_i; p) - y_i) \frac{\partial \mu}{\partial p_i} |_{x = x_i},\ \ i = 1, ..., 4 \tag{4.19}$$

The derivatives are used to compute the minimum of χ^2 in an iterated process. Thereby, we get a function of the form (3.2) with the proper parameters p_1, p_2, p_3, p_4.

5 Air pollution forecasting model (APFM)

The model forecasts pollution concentrations, e.g. particulate matters, for any chosen day (usually next day) or hours under condition of possessing the weather forecasts for that day. Let us

define some terms used in the model:
• time horizon T - number of hours for which the forecast will be computed,
• time step , e.g. 1 h, 10 min, 1 day,
• weather forecast - forecast for a chosen day of the chosen attributes e.g. temperature, wind speed, pressure etc. computed for each i . , where i = 1,. . . ,T,
• meteorological situation - measured meteorological data; for a chosen day of the chosen attributes, e.g. temperature, wind speed, pressure etc. measured for each i . , where i = 1,. . . ,T,
• pollution concentration (aero sanitary situation) - measured pollution data; for a chosen day of the chosen pollution, e.g. PM10, PM2.5, SO2, NO, CO, O3 measured for each i . , where i = 1,. . . ,T.
The APFM model is divided into following steps:
Step 1. Data preparation.
Step 2. Defining the set of similar weather forecasts.
Step 3. Defining the subset of similar meteorological situations.
Step 4. Defining fuzzy numbers for subset of meteorological situations.
Step 5. Determining the grades of membership of a subset of meteorological situation to a fuzzy numbers and defining the set of similar meteorological situations.
Step 6. Defining the set of similar aero sanitary situations.
Step 7. Calculating the forecast outputs.
Steps of the APFM model are presented in Fig. 1. At first weather forecasts are grouped, then chosen meteorological situations are transformed into fuzzy sets. Next, fuzzy sets are described in form of the fuzzy numbers. Then using fuzzy grouping we obtain a set of pollution concentrations. Next, using standardisation methods we receive forecast aero sanitary situation. 5.1. Experts role
The individual steps of the model are controlled by three parameters:ϵ, k in the fractional distance and η. The expert decides about the values of the individual parameters.
5.2. Step 1
Input data:
time horizon T,
time step ΔT,
database of meteorological data,
chosen date l for which we will perform the pollution concentrations forecast.
Output data:
The set of weather forecasts $WF \subset \Re^{d_s \times T}$, where d_f a number of attributes.
The set of meteorological situations $WF \subset \Re^{d_s \times T}$, where d_s is a number of attributes.
The set of pollution concentrations $A \subset \Re^{T}$.
From the database of meteorological data we take weather forecasts from the past to the future, meteorological situations from the past to the previous day to l and pollution concentrations from the past to the previous day to l. Then we create sets WF, MS, AS.
5.3. Step 2
Input data:
WF from Step 1,
$f_L^* \in WF\ (weatherforecastfromdayl)$,
$k \in (0, 1)$,
$\epsilon > 0$.
Output data:
$\epsilon - WF(f_L^* \subset WF$.
Using grouping method for $X = WF$, $f^* = f_L^*$ we obtain $\epsilon - WF(f_L^*$.

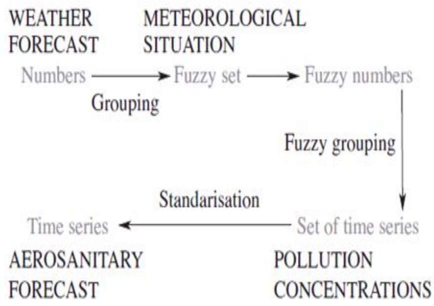

Fig. 1. Steps of the APFM model.

5.4. Step 3
Input data:
$\epsilon - WF(f_L^*)$ from Step 2,
MS from Step 1.
Output data:
$\epsilon - MS^{fL} \subset MS$.
We take set of dates D which corresponds to the elements of $\epsilon - WF(f_L^*)$. Next, from MS we take elements which corresponds to the dates from D. These elements create $\epsilon - WF(f_L^*$.
5.5. Step 4
Input data:
$\epsilon - MS(f_L^*)$ from Step 3.
Output data:
$\phi^* \in FN^{T\times d_s}$.
Let $m = |\epsilon - MS^{f_L^*}$. We create sequences:

$$\forall_t \in \{1, ..., T\},\ \forall_i \in \{1, ..., d_s\},\ (\xi_{ti}^{(1)}, ..., \xi_{ti}^{(m)} \tag{5.20}$$

where $[\xi_{ti}^{(j)}]_{t=1,...,T,\, i=1,...,d_s} \in \epsilon - MS(f_L^*)$.
Next, we change each of these sequences into a fuzzy number using SFN algorithm from Section 4 and arrange them into matrix:

$$\phi^* = [A_{t,i}]_{t=1,...,T,\, i=1,...,d_s} \tag{5.21}$$

where $A_{t,i} \in FN$.
5.6. Step 5
Input data:
ϕ^* from Step 4,
MS from Step 1,
$\eta > 0$.
Output data:
$\eta - MS \subset MS$,
w(s) for $S \in \eta - ms$ - weights of the similar meteorological situations.
Using fuzzy grouping method for $Y = MS$, $\phi = \phi^*$ we obtain $\eta - ms$.
Then we fix the weights of the meteorological situations using the formula:

$$w(s) = 1 - |\phi^*(s)| \tag{5.22}$$

where $S \in \eta - MS$.
5.7. Step 6
Input data:
$\eta - MS$ from Step 5,
AS from Step 1.
Output data:
$\eta - AS \subset AS$.
We take set of dates D which corresponds to the elements of $\eta - MS$. Next, from AS we take elements which corresponds to the dates from D. These elements create $\eta - AS$.
5.8. Step 7
Input data:
$\eta - AS$ from Step 6,
w(s) for $S \in \eta - MS$ from Step 5.
Output data:
Pollution concentrations forecasts using methods: average (u_a), maximum (u_m), α-standardisation, β-standardisation.
Let $v = |\eta - AS|$. For every time series from $\eta - AS$ we get a function

$$P^{(j)} : \{1, ..., T\} \longrightarrow \Re_0^t,\ j = 1, ..., v,$$

representing pollution concentrations.
For each $t \in \{1, ..., T\}$ we create a sequence:

$$(P_t^{(1)}, ..., P_t^{(v)}) \tag{5.23}$$

Then we take these sequences and we carry out a standardization process to obtain one time series. At first for this purpose we use two methods called: average and maximum.

$$\forall t \in \{1, ..., T\},\ u_\alpha(t) = \frac{\sum_i w^{(i)}.p^t}{\sum_i w^{(i)}} \tag{5.24}$$

where $w^{(i)}$ is the weight of the ith element of $\eta - MS$ for i = 1,. . . , r,

$$\forall t \in \{1, ..., T\},\ u_m(t) = max\{P^{(1)}(t), ..., P^{(v)}(t)\} \tag{5.25}$$

It should be noted details to the procedure the standardization refer to [3, 4].
The model presented in this paper, can be used for arbitrary regions having a proper database data.

Acknowledgement

This work was conducted with the support of the Ayatollah Amoli Branch, Islamic Azad University, Amol, Iran.

6 Conclusion

At the present days forecasting of pollution concentrations plays an important role both from the social point of view (including informing about exceed norms) and from the industrial point of view (including forecasting of consumption of the purchased units on the researched area). The models which are properly projected, implemented and use huge database make right forecast easier. The objective of the APFM model is to forecast any pollution concentrations. The APFM here can easily be adapted to regions in the world.

References

[1] J. S. Armstrong, *Principles of forcasting (hdbk)*, Kluwer Academic Publisher, (2001).

[2] K. S. Beyer, J. Goldstein, R. Ramakrishnan, U. Shaft, *When is "Nearest Neighbour" meaningful*, In Proceedings of the 7th international conference on database theory (1999) 217-235.

[3] D. Doman Ska, M. Wojtylak, *Selection criteria of forecast pollution concentrationsusing collateral informations* In: R. Tadeusiewicz, A. Ligeza, W. Mitkowski, M. Szymkat (Eds.) CMS'09 Computer Methods and Systems (2009a) 213-218.

[4] D. Doman ska, M. Wojtylak, *Air pollution forecasting model control*, Journal of Medical Informatics and Technology 14 (2010) 9-22.

[5] D.Doman ska, M. Wojtylak, *Change the sequence into fuzzy number*, LNCS (2010) 55-62.

[6] D. Dubois, H. Prade, *Ranking fuzzy numbers in the setting of possibility theory*, Information Sciences 30 (1983) 183-22.

[7] R. Noori, G. Hoshyaripour, K. Ashrafi, B.N. Araabi, *Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration*, Atmospheric Environment 44 (2010) 476-482.

[8] G. Nunnari, A. Nucifora and C. Randieri, *The application of neural techniques to the odeling of time series of atmospheric pollution data*, Ecological Modelling 111 (1998) 187-205.

[9] Q. Song, B.S. Chissom, *Fuzzy Time Series and its Models*, Fuzzy Sets and Systems 54 (1993) 269-277.

[10] L. A. Zadeh, *Fuzzy sets*, Information and Control 8 (1965) 338-353.

Fazlollah Abbasi is Ph.D of applied mathematics . He is member of Department of Mathematics, Ayatollah Amoli Branch, Islamic Azad University, Amol, Iran. Interested in fuzzy mathematics.