fuzzy system models, fuzzy numbers, fuzzy matrix, fuzzy weather forecast, air pollution, forecasting

Diana DOMAŃSKA*, Marek WOJTYLAK**

AIR POLLUTION FORECASTING MODEL CONTROL

In the paper we discuss the analysis of multidimensional data. We consider the relationship between them using a special fuzzy number form. Calculations are kept on set of actual and historical meteorological data. Our model using to forecast pollution concentrations is important in today because pollutions have very big influence on our life in particular pollutions PM10 (particulate matter less than 10 μm in diameter). The effects of inhaling particulate matter have been widely studied in humans and animals and include asthma, lung cancer, cardiovascular issues, and premature death. Because of the size of the particle, they can penetrate the deepest part of the lungs. In Air Pollution Forecasting Model for the chosen weather forecast we find similar weather forecasts. Next, we find real meteorological situations from the historical data which correspond to them and we create fuzzy numbers, that is, the fuzzy weather forecasts. Then we estimate the validity of the weather forecast on the basis of the historical data and its accuracy. We investigate it with the help of a set of indicators, which corresponds to the parameters of the weather forecast, using the similarities rule of the weather forecast to the meteorological situation, a proper distance and data analysis. This comprehensive analysis allows us to investigate the effectiveness of forecasting pollution concentrations, putting the dependence between particular attributes describing the weather forecast in order and proving the legitimacy of the applicable fuzzy numbers in air pollution forecasting.
Models are created for data, which are measured and forecasting in Poland. By reason of this data our models are testing in real sets of data and effects are received in active system.

1. INTRODUCTION

The first trials of forecasting everyday phenomena, particularly meteorological, began around 650 B.C. [1] by the Babylonians. They tried to predict short-term weather changes based on the appearance of clouds. Methods of weather forecasting were increasingly perfected in subsequent centuries. In the XX century, as a result of the development of mathematics and physics, models which used partial differential equations were formulated. These equations which describe the state of the atmosphere, could be solved numerically. However, in 1961 E.Lorenz showed the limitation of possibilities of these models — first of all their chaotic character. These models are only effective for a few days — maximum a week. However, for a 3-day term their effectiveness is high.
In recent years many prediction approaches, such as statistical [2], fuzzy [3], [4], neural networks [5], [6], neuro-fuzzy predictor [7] have emerged. Using numerical short-term weather prediction, research into the forecasting of air pollution concentrations began [8], [9]. This task is very difficult because apart from the information about meteorological conditions, the emission of air pollution depends first of all on the immission. At this moment, emission is quite accurately measured from a single, high pointer emitter (e.g. carbon power stations). Measurement of low emission, communal and municipal, is almost impossible. Moreover, 3D models of immission calculating (e.g. Gaussian puff modelling system) require a field of wind and a field of temperature measure from several hundred metres above ground level. Such measurements are only conducted in a very few places in the world with the help of a sodar. In this

* University of Silesia, Institute of Informatics, ul. Będzińska 39, 41-200 Sosnowiec, Poland, e-mail: *ddomanska@poczta.onet.pl*
** Institute of Meteorology and Water Management (IMGW) ul. Bratków 10, 40-045 Katowice, Poland, e-mail: *monitoring.katowice@imgw.pl*

situation Fuzzy sets theory is helpful [10], [11], [12]. Use of this method is known in many mathematical forecasting models. It is usually used when the information transferred to the model is imprecise or incomplete [13], [14]. Many everyday phenomena of an ambiguous, continuous and imprecise nature may be effectively described using this theory.

The problem is with knowledge. We do not have precise knowledge about the weather in the future. We only have numerical forecasting, i.e. conditions which may announce many similar meteorological situations. Because of the essential influence of emission, historical data must come from areas similar to the place for which we are calculating a forecast. Similarity of the area includes the following parameters: surface roughness (topography), number of habitants, percentage of industry, rural, heavy traffic, low buildings and green regions. These are the only accessible factors that determine emission. The result of a working APFM (Air Pollution Forecasting Model) is a forecast of air pollution concentration, among others PM10 for the next day. It is a specially chosen pollution because PM10 has a huge influence on human life [15], [16].

In each stage we use meteorological data with a mathematical apparatus [17], [18]. In particular in APFM we use the weather forecasts derived from the Consortium for Small Scale Modelling (COSMO) model based on the Local-Model (LM) of Deutscher Wetter Dienst (DWD).

In paper we assume that objects are similar when their distance are small. Objects availing in paper are vectors and matrices. In vector space R^d for vectors we use (1) for terms distance between objects.

$$d_k^d(x,y) = \left(\sum_{i=1}^{d} \left| x_i - y_i \right|^k \right)^{\frac{1}{k}}, x, y \in R^d, k > 0. \tag{1}$$

For $k \geq 0,\ k \in N$ the distance (1) is metric. The distance between matrix objects is composition of vector objects. For terms distance between matrices we use (2).

$$d_{k_1 k_2}^{n \times m}(A,B) = d_{k_1}^{n}\left(\left[d_{k_2}^{m}\left(a_{i.}, b_{i.}\right)\right], \mathbf{0}_n\right),\ for A = \left[a_{ij}\right], B = \left[b_{ij}\right], A, B \in R^{n \times m}, \tag{2}$$

where $\mathbf{0}_n$ is a zero vector and $a_{i.}, b_{i.}$ are i-th rows in matrices A and B.

In the first instance we introduce a term, time horizon set T, in which the forecast will be calculated. $T = \{t = i \cdot \Delta t : i = 0, \ldots, n_T\}, \Delta t > 0$, where Δt means a time step (usually $\Delta t = 1$ hour). We will identify the term from set T with 0 hour UTC. We assume that for each term from T we have values of d_f parameters of a numerical weather forecast (e.g. temperature, sea level pressure, wind direction and speed, cloud cover — high, medium, low). For the term weather forecast we will understand a matrix $F \in R^{(n_T+1) \times d_f}$. Moreover, we assume that we possess the data from days for many years in every term $t \in T$. Every term t describes the state of the atmosphere with the aid of the d_s parameters measured near the surface (e.g. temperature, wind direction) and the value of concentrations whose size we are forecasting. The set of meteorological data for each subsequent term $t \in T$ defines the meteorological situation. The meteorological situation will be represented by a matrix $S \in R^{(n_T+1) \times d_s}$. The aerosanitary situation is the number of sequences of concentrations in $t \in T$ terms, so it is a time series belonging to R^{n_T+1}. In order for the model to function properly it is essential to have all the historical data.

Let us denote the set of weather forecasts as WF, the set of meteorological situations as MS, the set of pollution concentrations as AS.

1.1 FIRST STAGE OF THE MODEL

In the first stage, because of the huge data range, we start from *min-max* normalisation for every weather forecast in every column separately. Let us define:

$f^* \in WF$ — a chosen weather forecast for which we are calculating the forecast of pollution concentrations.

$k = k_1 = k_2$ — first parameter used to control APFM system, it decides about dispersion between elements from set WF. In determining the parameter k we follow the data diversification. For parameter $k > 1$ and matrices which size is meaningful i.e. the number of rows is greater than 24 and the number of columns is greater than 9, the distances between the matrices are small so the grouping of data is not good. Experiments have shown it in paper [19].

Θ — a real number set representing distances between every normalised weather forecast from WF and f^* so $\Theta = \{\omega_1, \ldots, \omega_q\}, q = |WF| - 1$ where $\omega_i = d_k^{(n_T+1) \times d_f}\left(f_i, f^*\right)$ and $f_i \in WF$. $|A|$ means cardinality of set A.

We prefer fractional distance — distance (1) for $k_1, k_2 \in (0,1)$ because of better element diversification in a multidimensional space [19].

In the next step we define second parameter ε. Parameter ε decides about cardinality of similar elements from set WF. ε is decided in (3).

$$\forall_{i=1,\ldots,q}\, \omega_i < \varepsilon, \tag{3}$$

where $\omega_i \in \Theta$ for $i \in \{1, \ldots, q\}$, i is the number of an element. The result of the first stage is set $\varepsilon - WF\left(f^*\right)$.

1.2 SECOND STAGE OF THE MODEL

In the second stage, in connection with results from the first stage, we create subset $\varepsilon - MS^F \subset MS$. Every weather forecast is related with meteorological situation with date. Therefore for $\varepsilon - MS^F$ we consider pairs $(f, s), f \in WF, s \in MS$. Then, we set parameters describing the meteorological situations and the time horizon. After review of the chosen meteorological situations, we get a sequence of values:

$$\forall_{t \in T} \forall_{i \in 1,\ldots,d_s} \left(\xi_{t,i}^{(1)}, \ldots, \xi_{t,i}^{(m)}\right),$$

where

$$m = \left|\varepsilon - MS^F\right|. \tag{4}$$

We modify this sequence into a fuzzy number using a special form of the fuzzy number given by (5). For each attribute i and in each hour $t \in T$ we have individual fuzzy number. This fuzzy number is approximate to the Gaussian function. The fuzzy number (5) was chosen based on our own calculations and based on paper [20].

$$\mu_{i,t}(x)=\begin{cases}\exp\left(\dfrac{-\left(x-m_{1_{i,t}}\right)^2}{2\cdot\sigma^2_{1_{i,t}}}\right) & if\ x\le m_{1_{i,t}},\\ 1\ if\ x\in\left(m_{1_{i,t}},m_{2_{i,t}}\right), & \\ \exp\left(\dfrac{-\left(x-m_{2_{i,t}}\right)^2}{2\cdot\sigma^2_{2_{i,t}}}\right) & if\ x\ge m_{2_{i,t}},\end{cases} \tag{5}$$

where $m_{1_{i,t}} \le m_{2_{i,t}}, \sigma_{1_{i,t}} > 0, \sigma_{2_{i,t}} > 0$ for $m_{1_{i,t}}, m_{2_{i,t}}, \sigma_{1_{i,t}}, \sigma_{2_{i,t}} \in R$, for each attribute i in each hour $t \in T$. An individual fuzzy weather forecast consists of a time series (5). We receive fuzzy weather forecast (6) — equivalent to the real weather.

$$\phi^* = \left[\mu_{i,t}\right]\ for\ 1 \le i \le d_s, t \in T \tag{6}$$

where i is a number of attribute and t is an hour. ϕ^* is a function for which we determine membership matrix composed from fuzzy numbers. $\phi^*_{it}: R^{d_s} \to [0,1]$. In ϕ^* we can take property values. We receive $\phi^*(s) = \left[\mu_{it}(s_{it})\right]$ which is membership matrix, $s \in R^{(n_T+1)\times d_s}$. A time series is a sequence of the regularly sampled quantities of an observed system. In natural phenomena we have a different types of time series. The time series controls the chaotic behaviour [21] independent phenomena in the world.

Fuzzy weather forecast is need to be able to clearly and precisely define the quality of a weather forecast and assign meteorological situation relative to a fuzzy weather forecast [14]. The assignment of a coefficient quality of a weather forecast is the point of entrance for the exact determination of the individual influence of an attribute on forecasting pollution concentrations in the future.

1.3 THIRD STAGE OF THE MODEL

In the third stage we review all meteorological situation $s \in MS \subset R^{(n_T+1)\times d_s}$. Then, for every meteorological situation s we calculate $\phi^*(s)$ and number $\rho(s) = \left|\phi^*(s)\right|$ using formula (7).

$$\left|\phi^*(s)\right| = d_k^{n_T+1}\left(\left[d_k^{d_s}\left(\phi^*_{.t}(s_{.t}), \mathbf{1}_{d_s}\right)\right]_{t\in T}, \mathbf{0}_{n_T+1}\right), s \in MS, k \in [0,1], \tag{7}$$

Let us fix $\eta \in [0,1]$ and determine a set $\eta - MS \subset MS$ that $|\eta - MS| = r$, where $\rho(s) < \eta$ for $s \in \eta - MS$. For subset $\eta - MS$ we consider pairs $(s,p), p \in AS$. Then we fix the weight of the meteorological situations using following formula $w(s) = 1 - \rho(s), s \in \eta - MS$.

1.4 FOURTH STAGE OF THE MODEL

In the fourth stage we choose r time series from set AS, where $r = |\eta - MS| > 1$. Afterwards, for every chosen time series we get a function $p^{(j)}: T \to R_0^+, j = 1,\ldots,r$ with weight $w^{(j)} \in R$, representing pollution concentrations. For each $t \in T$ we create a sequence $\left(p^{(1)}(t),\ldots,p^{(r)}(t)\right)$.

Then we take these sequences and we carry out an aggregation process to obtain one time series.

We have used two methods α -aggregation (10) and $\alpha\beta$ -aggregation (12). We base these methods on the well-known methods: (8) and (9).

$$\forall_{t\in T}\mu_{a,t}=\frac{\sum_{i=1}^{r}w^{(i)}p^{(i)}(t)}{\sum_{i=1}^{r}w^{(i)}}, \tag{8}$$

$$\forall_{t\in T}\mu_{m,t}=\max_{i=1,\dots,r}\left\{p^{(i)}(t)\right\}. \tag{9}$$

where a means average aggregation, m means maximum aggregation.
Let us denote for each $t\in T$ the following time series $\mu_{a,t}$ as a time series received from method (8), $\mu_{m,t}$ as a time series received from method (9) and $\mu_{r,t}$ as a time series received from the actual researched data. For $l\approx\frac{n_T}{4}$ we determine two numbers based on knowledge about actual aerosanitary situation. We forecast pollution concentrations having partial knowledge, that is real number of pollution concentrations. We execute some calculations on time series and we get parameters α from the first function (10) and α,β from the second function (12).

$$g(\alpha)=\sum_{t=0}^{l}\left(\alpha\mu_{a,t}+(1-\alpha)\mu_{m,t}-\mu_{r,t}\right)^2, \tag{10}$$

where

$$\alpha=\frac{\sum_{t=0}^{l}\left(\mu_{m,t}^2-\mu_{m,t}\mu_{a,t}-\mu_{m,t}\mu_{r,t}+\mu_{a,t}\mu_{r,t}\right)}{\sum_{t=0}^{l}\left(\mu_{a,t}-\mu_{m,t}\right)^2}, \tag{11}$$

if $\sum_{t=0}^{l}\left(\mu_{a,t}-\mu_{m,t}\right)^2\neq 0$. When we determine the optimal value of the parameter α for (10) we receive formula (11). We proceed analogically with value of α,β for (12) and we receive formulas (13), (14).

$$h(\alpha,\beta)=\sum_{t=0}^{l}\left(\alpha\mu_{a,t}+\beta\mu_{m,t}-\mu_{r,t}\right)^2, \tag{12}$$

where

$$\alpha=\frac{\sum_{t=0}^{l}\mu_{a,t}\mu_{r,t}\sum_{t=0}^{l}\mu_{m,t}^2-\sum_{t=0}^{l}\mu_{m,t}\mu_{r,t}\sum_{t=0}^{l}\mu_{a,t}\mu_{m,t}}{\sum_{t=0}^{l}\mu_{m,t}^2\sum_{t=0}^{l}\mu_{a,t}^2-\left(\sum_{t=0}^{l}\mu_{m,t}\mu_{a,t}\right)^2}, \tag{13}$$

$$\text{if } \sum_{t=0}^{l}\mu_{m,t}^2\sum_{t=0}^{l}\mu_{a,t}^2-\left(\sum_{t=0}^{l}\mu_{m,t}\mu_{a,t}\right)^2\neq 0$$

$$\beta = \frac{\sum_{t=0}^{l}\mu_{m,t}\mu_{r,t}\sum_{t=0}^{l}\mu_{a,t}^{2} - \sum_{t=0}^{l}\mu_{m,t}\mu_{a,t}\sum_{t=0}^{l}\mu_{a,t}\mu_{r,t}}{\sum_{t=0}^{l}\mu_{m,t}^{2}\sum_{t=0}^{l}\mu_{a,t}^{2} - \left(\sum_{t=0}^{l}\mu_{m,t}\mu_{a,t}\right)^{2}}, \tag{14}$$

$$\text{if } \sum_{t=0}^{l}\mu_{m,t}^{2}\sum_{t=0}^{l}\mu_{a,t}^{2} - \left(\sum_{t=0}^{l}\mu_{m,t}\mu_{a,t}\right)^{2} \neq 0$$

From (10) we receive parameter α given by (11) and from (12) we receive parameters α, β given by (13) and (14).
Then using methods (10), (12) and having knowledge about collateral information i.e. first ten hours real pollution concentrations in the day we being forecast, we can calculate for each $t \in T$ the final time series $\mu_{f,t}$ [18]. Fig. 1 shows the respective stages of the model at work.

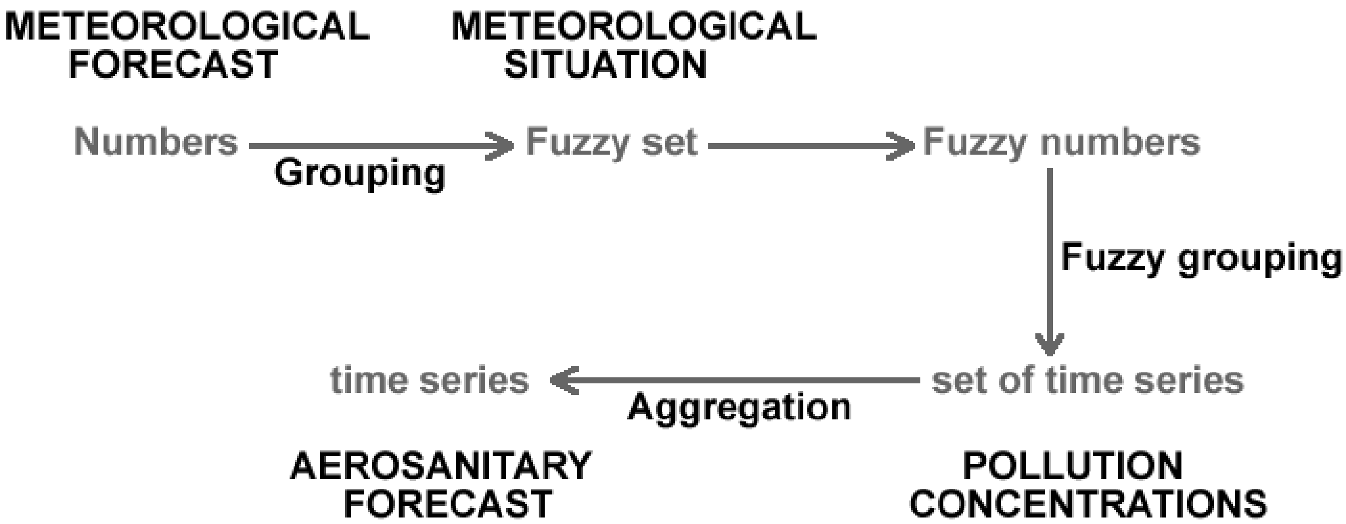

Fig. 1. The stages of Air Pollution Forecast Model.

2. CHARACTERISTICS OF FUZZY WEATHER FORECAST

A fuzzy weather forecast ϕ^{*} is determined for each attribute i individually and is evenly distributed on each hour $t \in T$. It is valued on the basis of data similarity and proper weights of classification. We researched the behaviour of fuzzy weather forecasts using different sets of forecast data. This is necessary because we have weather forecasts from a short period of time (only five years). Therefore, continuous work in a COSMO LM model weather forecast [22], [23] is not heterogeneous to finding the period of a weather forecast which is the best estimate of real meteorological situations. In Figs 2, 3, 4 fuzzy weather forecasts are shown along with real meteorological situations. The fuzziness is a good measure with which to mark the quality of a weather forecast both its elements and the whole weather forecast because fuzziness characterises the scattering of real data around the prognosis.

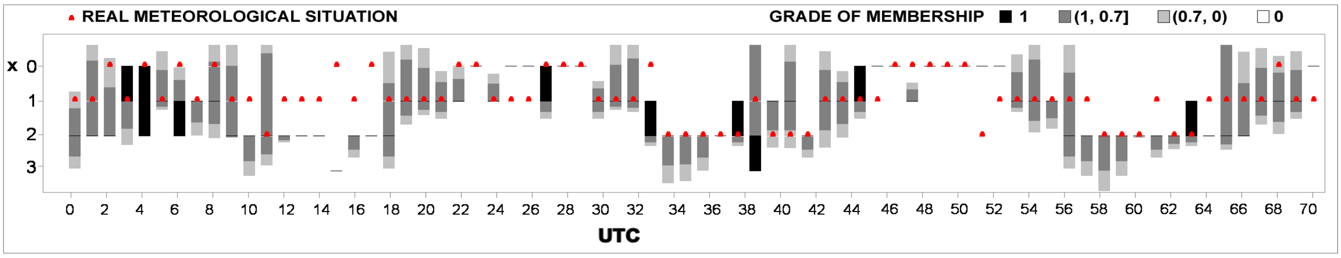

Fig. 2. Fuzzy weather forecast for wind speed attribute on 10 January 2006.

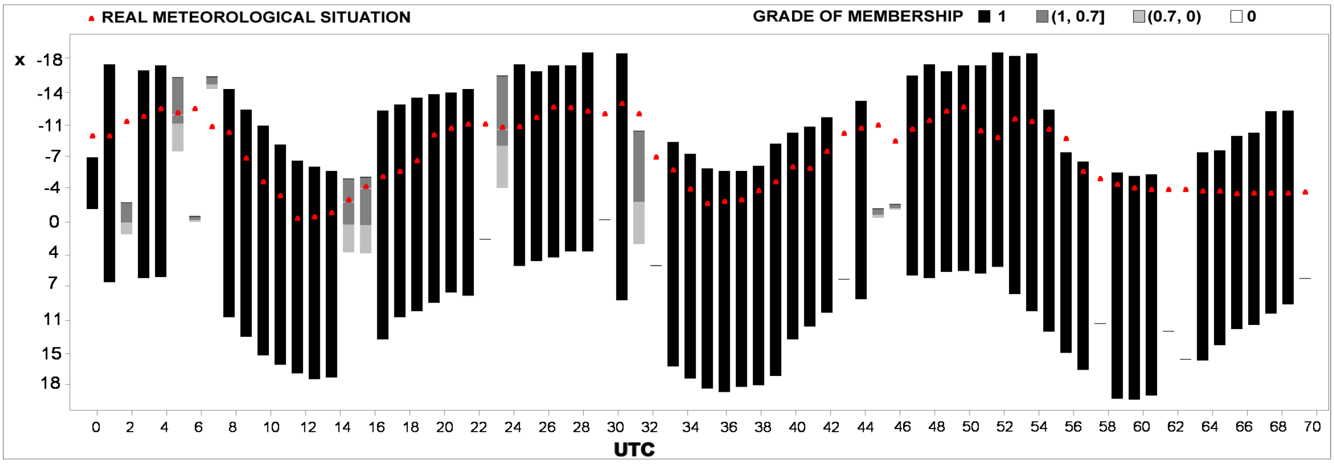

Fig. 3. Fuzzy weather forecast for temperature attribute on 10 January 2006.

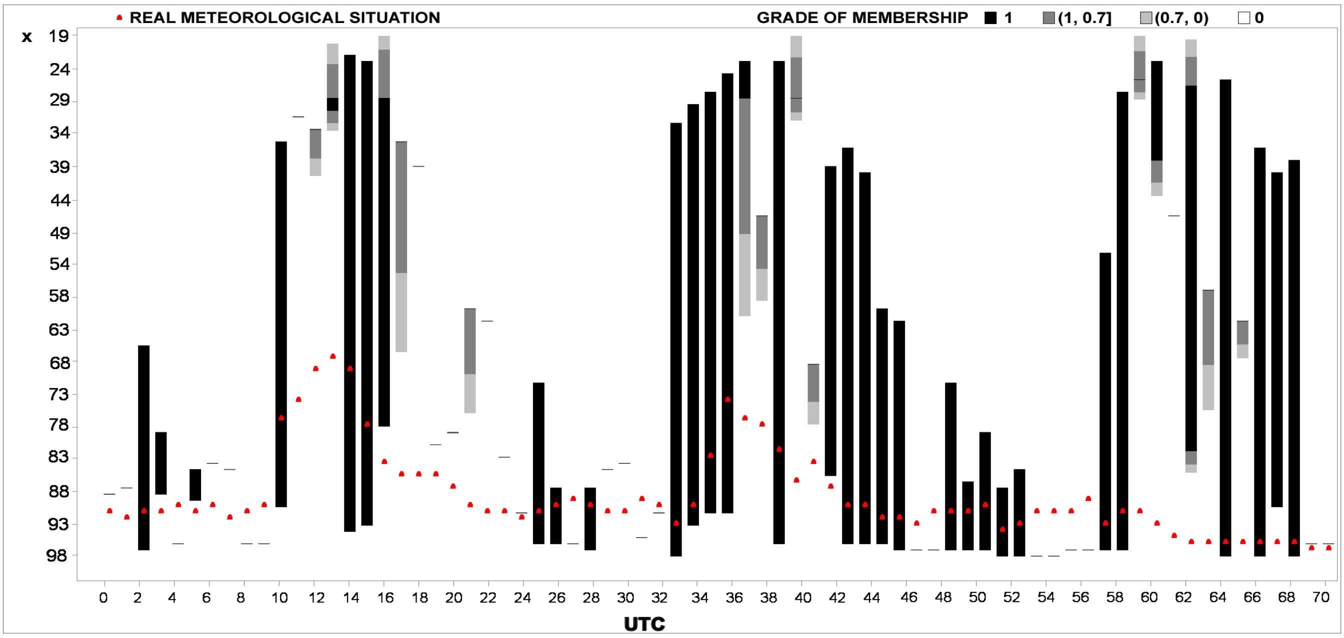

Fig. 4. Fuzzy weather forecast for humidity attribute on 10 January 2006.

3. PROPOSED FEATURES TO ESTIMATE THE QUALITY OF A FUZZY WEATHER FORECAST

The first feature is research volume for all attributes i in each hour $t \in T$. We receive the first number $F^{i,t}$ characterised by numbers.

$$F^{i,t} = \int_{-\infty}^{\infty} \mu_{i,t}(x)dx = \frac{\sqrt{2\pi}}{2}\left(\sigma_{1_{i,t}} + \sigma_{2_{i,t}}\right) + m_{2_{i,t}} - m_{1_{i,t}}, \tag{15}$$

where $i \in \{1,\ldots,d_s\}, t \in \{0,\ldots,n_T\}, \sigma_{1_{i,t}}, \sigma_{2_{i,t}}, m_{1_{i,t}}, m_{2_{i,t}} \in R^{(n_T+1)\times d_s}$.

Let us define $f_i(t) = F^{i,t}$ for each $t \in T$. In Figs 5, 6, 7 the integral of functions $\mu_{i,t}$ for all $t \in T$ and for the chosen attributes i are shown using real meteorological situations: wind speed, temperature and humidity. In Fig. 8 fuzzy weather forecast is shown for all attributes.

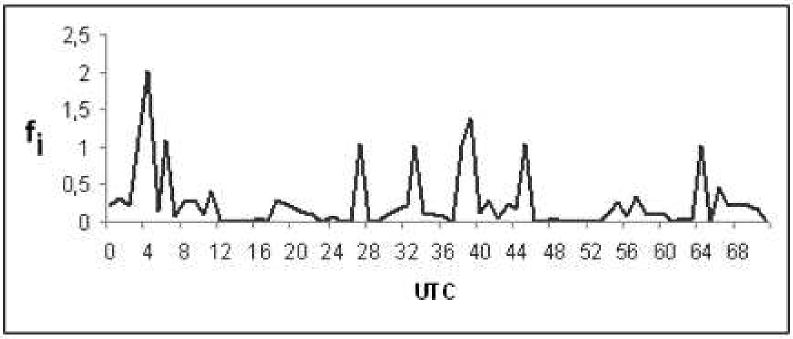

Fig. 5. f_i for wind speed attribute on 10 January 2006.

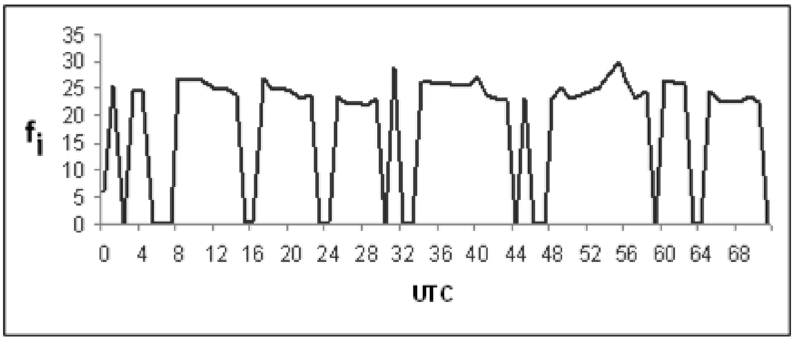

Fig. 6. f_i for temperature attribute on 10 January 2006.

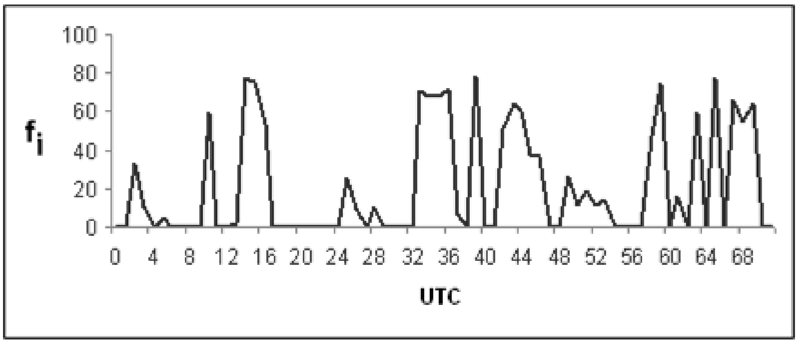

Fig. 7. f_i for humidity attribute on 10 January 2006.

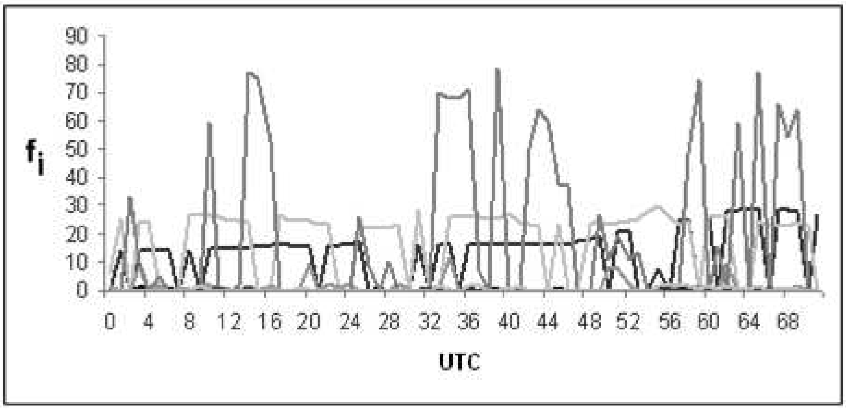

Fig. 8. f_i for all attributes on 10 January 2006.

For each $\mu_{i,t}$ it is the grade of membership of x.

The second feature is researching the quality of a fuzzy weather forecast by comparing it to a real meteorological situation. In this way we keep an attribute characterised by a grade of membership for each hour. In Figs 9, 10, 11 grades of membership for wind speed, temperature and humidity are shown.

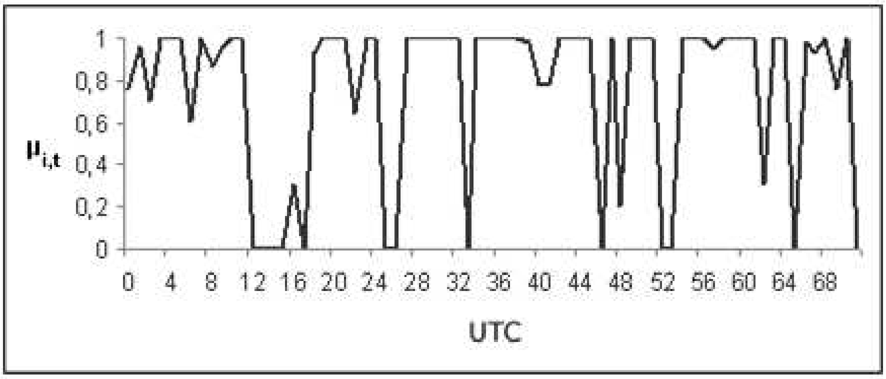

Fig. 9. The grade of membership $\mu_{i,t}$ for wind speed attribute and $T = 72$ on 10 January 2006.

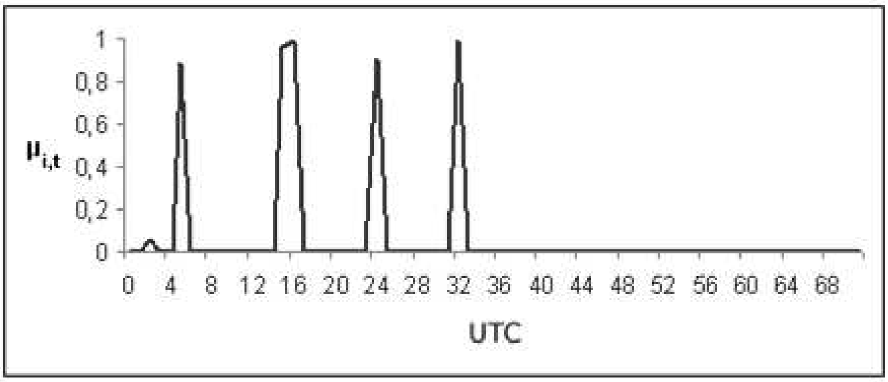

Fig. 10. The grade of membership $\mu_{i,t}$ for temperature attribute and $T = 72$ on 10 January 2006.

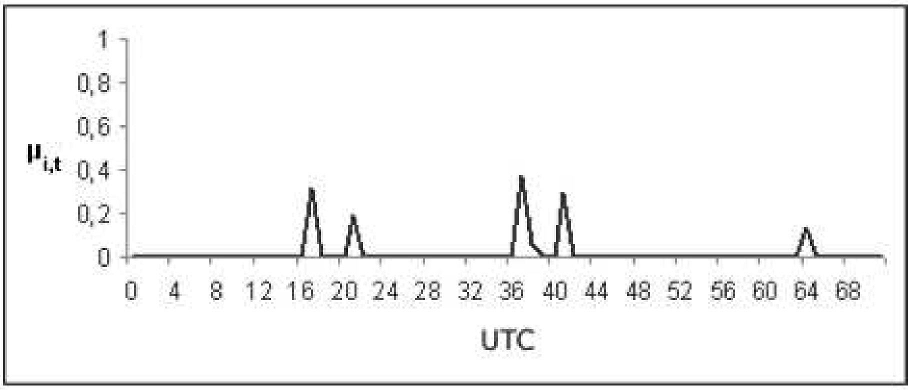

Fig. 11. The grade of membership $\mu_{i,t}$ for humidity attribute and $T = 72$ on 10 January 2006.

On the basis of the received sets of features, in section 4 we propose several indices which describe the qualities of the weather forecast. These will represent distinctness, precision and its credibility.

4. MODEL CONTROL

The APFM is controlled by three parameters which have a fundamental influence on forecasting. These parameters are: k for distance (2), ε and η. Distance $d_{kk}^{(n_T+1)\times d_f}$ depends on the coefficient k which has influence on which weather forecasts will be chosen. The ε parameter determines how many meteorological situations will be chosen in the first stage and the η parameter indicates how many aerosanitary situations will be chosen in the third stage. All the parameters can be changed independently from each other, but each of them has a considerable impact on the model.

In particular parameters k and ε have an influence on the quality of fuzzy weather forecast ϕ^*.

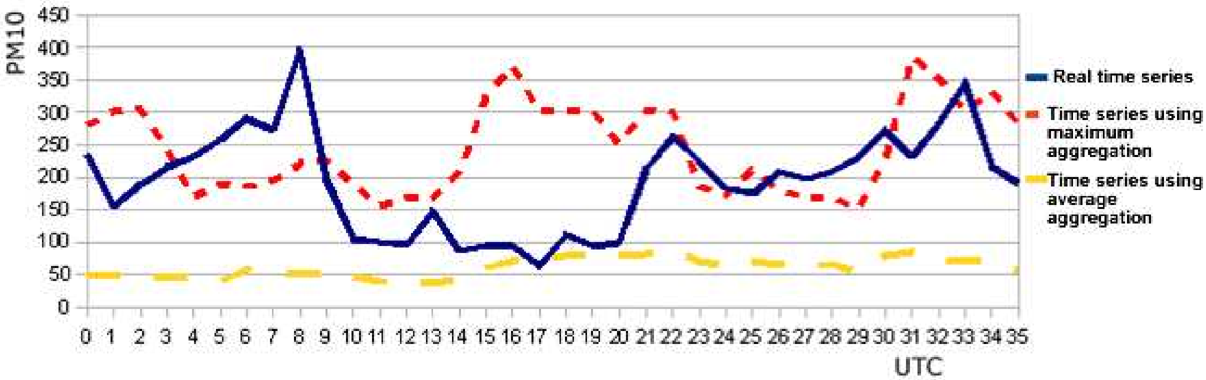

Fig. 12. Pollution concentration runs: time series using methods (8), (9) using distance (2) with $k = 0.1$ and real time series on 9 January 2006.

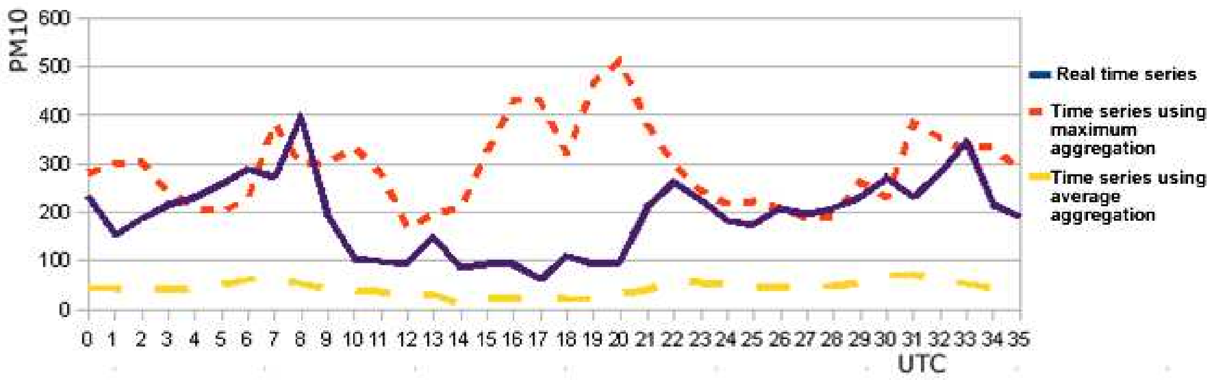

Fig. 13. Pollution concentration runs: time series using methods (8), (9) using distance (2) with $k = 2$ and real time series on 9 January 2006.

In Figs 12, 13 we show representatives graphs which have been chosen from dozens experimental results.

Experiments showed that the results for distance (2) with $k = 0.1$ and $k = 2$ are different. On the selection of the set of similar weather forecasts their symmetrical difference is not equal to 0. In Fig. 12 we can observe that for a low value of k we have better results than for the big ones.

Thanks to the parameter ε we receive a set $\varepsilon - WF(f^*)$. In Fig. 14 we can see that the set of similar weather forecasts is too small when its cardinality is 5. The set of similar weather forecasts is too big when its cardinality is 20. This has an important influence on forecasting. The optimal number of elements in the set $\varepsilon - MS^F$ is about 10. Too many variable samples have the effect that for attribute i in each hour $t \in T$, $m_{1_{i,t}}$ is more distant from $m_{2_{i,t}}$, which causes greater fuzziness. And this has an influence on the inaccuracy of the pollution concentrations forecast. This can be seen in Fig. 15.

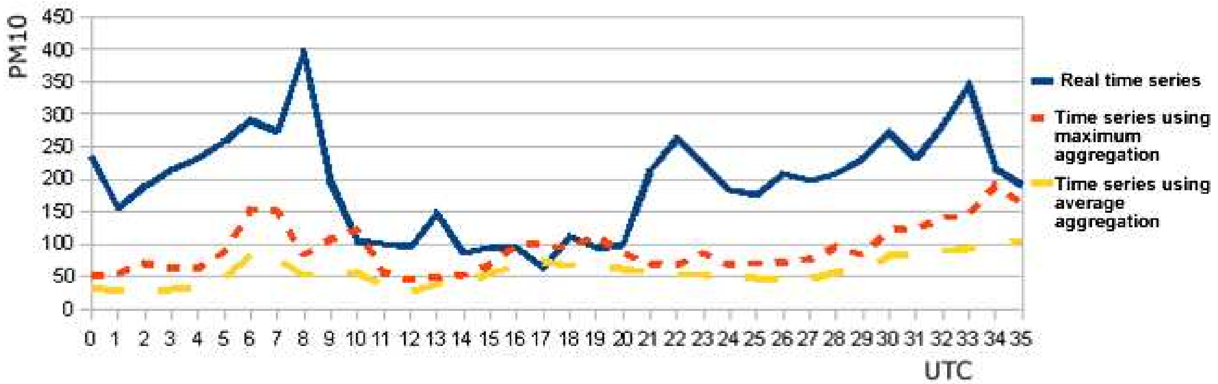

Fig. 14. Pollution concentrations runs: time series using methods (8), (9) using $\varepsilon = 5$ and real time series on 9 January 2006.

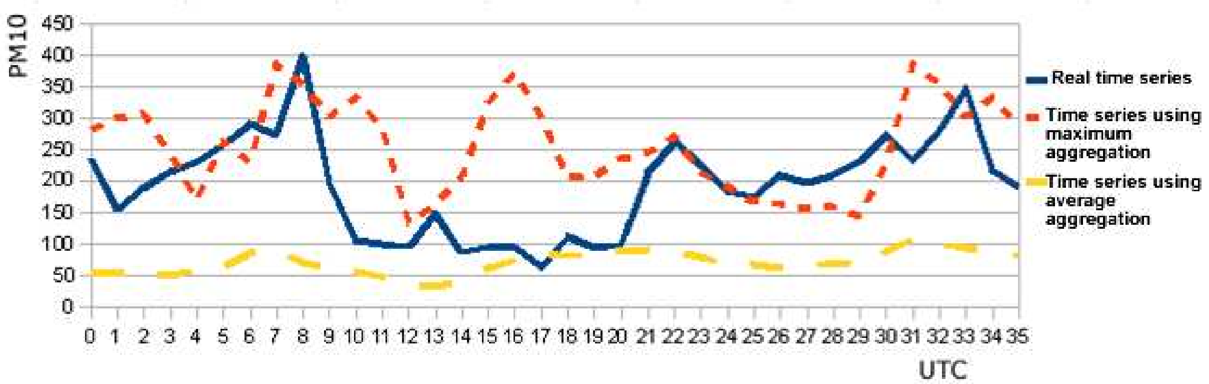

Fig. 15. Concentration pollution runs: time series using methods (8), (9) using $\varepsilon = 10$ and real time series on 9 January 2006.

Through parameter η we receive $\eta - MS$ which we use to forecast pollution concentrations in the fourth stage. In Fig. 16 we can see that $\eta = 10$ have too small influence on final time series and notably wander away from this final time series. The bigger η the better, the optimal cardinality of the set $\eta - MS$ is about 20. This can be seen in Fig. 17.

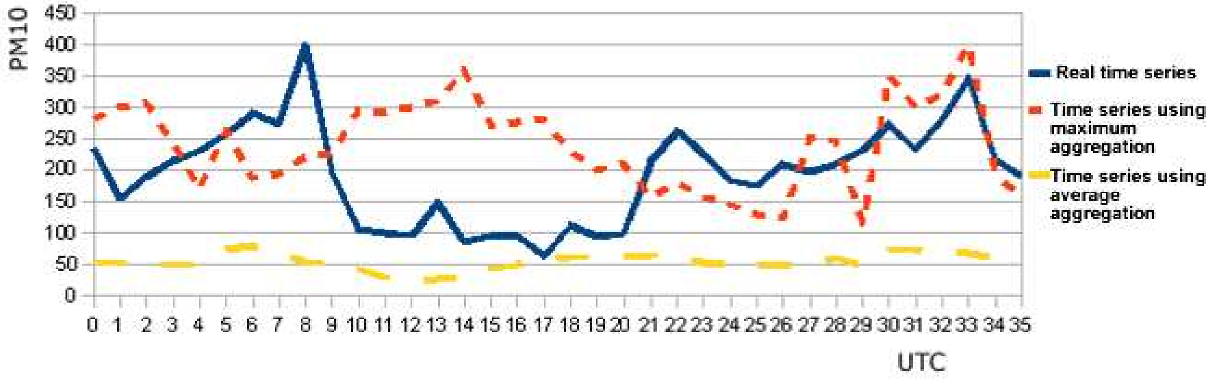

Fig. 16. Pollution concentration runs: time series using methods (8), (9) using $\eta = 10$ and real time series on 9 January 2006.

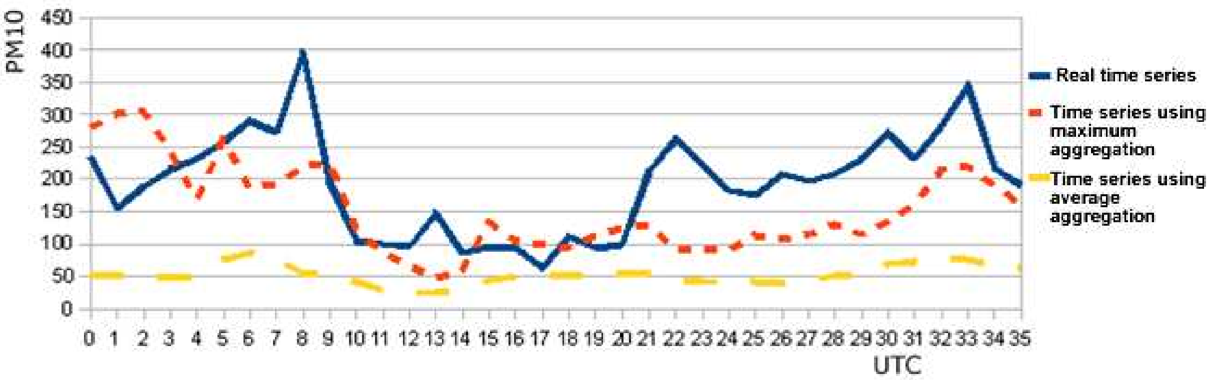

Fig. 17. Pollution concentration runs: time series using methods (8), (9) using $\eta = 20$ and real time series on 9 January 2006.

5. CONCLUSIONS

Computations were performed for weather forecasts in 2003-2007, meteorological situations in 1997-2007 and pollution concentrations in 1998-2007 with $\Delta t = 1$. We have $d_f = 28$ attributes describing weather forecasts, while the number of meteorological situations was equal to $d_s = 9$. Attributes describing meteorological situations were chosen based on investigations by [24]. The effect of the suggested method for the prediction of a weather forecast was introduced for data from COSMO LM model, but the same method can be used for different weather forecasts based on numerical models. The condition which has to be met is to have real meteorological data.

In Figs 2, 3, 4 the fuzzy weather forecasts are clearly and explicitly shown and it can be seen that the fuzzy weather forecast has a little fuzziness when the grade of membership is large. It is described for single indicators: fuzziness average and grade of membership average in Tabs 1, 2. By analysing other examples, a fairly significant dependence between small fuzziness and membership can be observed a reverse correlation between them, because if the fuzziness is greater then the accuracy of the prognosis is minor.

Analyzing the influence of a fuzzy weather forecast for the fourth stage in APFM, it can be determined whether the fuzzy average is minor and the grade of membership average is minor in Tab. 2, forecasting pollution concentrations is not precise enough or fuzzy average is minor and the grade of membership average is large in Tab. 1, forecasting pollution concentrations is more precise.

Table 1. Results of the tests for attributes: wind speed, temperature and humidity on 9 January 2009.

Attribute	Fuzzy average	Grade of membership average
Wind speed	0.31	0.69
Temperature	15.10	0.16
Humidity	19.49	0.10

Table 2. Results of the tests for attributes: wind speed, temperature and humidity on 10 January 2009.

Attribute	Fuzzy average	Grade of membership average
Wind speed	0.25	0.75
Temperature	18.39	0.07
Humidity	21.01	0.02

BIBLIOGRAPHY

[1] Weather Forecasting Through the Ages, NASA Facts, http://aqua.nasa.gov, 2002.

[2] RIGATOSA G., ZHANGB Q., Fuzzy model validation using the local statistical approach, Fuzzy Sets and Systems 160(7) (2009) 882-904.

[3] LEE C., LIU A., CHEN W., Pattern Discovery of Fuzzy Time Series for Financial Prediction, IEEE Trans. Knowl. Data Eng 18(5) (2006) 613-625.

[4] KUNHUANG H., Heuristic Models of Fuzzy Time Series for Forecasting, Fuzzy Sets and Systems 123(3) (2001) 369-386.

[5] JIANG W., WANG P., Research on Interval Prediction of Nonlinear Chaotic Time Series Based on New Neural Networks, in: Proc. 6th World Congress Intell. Control and Automation, 2006, pp. 2835-2839.

[6] AJIT KUMAR GAUTAM, CHELANI A.B., JAIN V.K., DEVOTTA S., A new scheme to predict chaotic time series of air pollutant concentrations using artificial neural network and nearest neighbor searching, Atmospheric Environ. 42(18) (2008) 4409-4417.

[7] AZNARTE M.J.L., MANUEL BENÍTEZ SÁNCHEZ J., NIETO LUGILDE D., DE LINARES FERNÁNDEZ C, DÍAZ DE LA GUARDIA C., ALBA SÁNCHEZ F., Forecasting Airborne Pollen Concentration Time Series with Neural and Neuro-Fuzzy Models, Expert Systems with Applications 32(4) (2007) 1218-1225.

[8] OŚRÓDKA L., KRAJNY E., WOJTYLAK M., The use of numerical weather forecast for air pollution forecasting in an urban industrial agglomeration, in: Proc. 4th Annual Meeting EMS and 5th EC on Applied Climatology, 2004, EMS Vol. 1.

[9] OŚRÓDKA L., WOJTYLAK M., KRAJNY E., RORBEK K., Improving the methods of forecasting the high concentrations of pollutants in urban-industrial agglomerations, using the numerical weather forecast models, Wiad. MGW 27(1) (2004) 105-116.

[10] ZADEH L.A., Fuzzy Sets, Inf. and Control 8 (1965) 338-353.

[11] KLIR G., FOLGER T., Fuzzy Sets, Uncertainty and Information, Prentice Hall PTR, Englewood Cliffs, NJ, 1988.

[12] ZIMMERMAN H.J., Fuzzy Set Theory and its Applications, second ed., Kluwer, Dordrecht, 1991.

[13] SILVERT W., Ecological impact classification with fuzzy sets, Ecol. Model. 96 (1997) 1-10.

[14] HANSEN B.K., RIORDAN D., Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory, in: Proc. Workshop on Soft Computing in Case-Based Reasoning, International Conf. on Case-Based Reasoning, Canada, 2001, pp.175-178.

[15] KOWALSKA M., HUBICKI L., ZEJDA E.J., OŚRÓDKA L., KRAJNY E., WOJTYLAK M., Effect of ambient air pollution on daily mortality in Katowice Conurbation, Poland, Polish J. Environ. Stud. (2007) 227-232.

[16] KOWALSKA M., ZEJDA J.E., OŚRÓDKA L., KLEJNOWSKI K., KRAJNY E., WOJTYLAK M., Relationship between ambient air pollution and daily mortality in the Urban Area of Katowice-comparison on two periods 1994-1995 and 2001-2002, in: Proc. Public Conf. ”Particles and Health-State of the Research and Policy Implications”, (9), 2007.

[17] DOMAŃSKA D., WOJTYLAK M., Development data serving to forecasting pollution concentrations, 2009, pp. 351-359.

[18] DOMAŃSKA D., WOJTYLAK M., Selection criteria of forecast pollution concentrations using collateral informations, Computer Methods and Systems, Kraków, 2009, pp.213-218.

[19] BEYER K.S., GOLDSTEIN J., RAMAKRISHNAN R., SHAFT U., When is “Nearest Neighbour” meaningful?, in: Proc. of the 7th International Conf. on Database Theory, 1999, pp. 217-235.

[20] OŚRÓDKA L., WOJTYLAK M., KRAJNY E., DUNAL R., KLEJNOWSKI K., Application Data Mining for forecasting of high-level air pollution in urban-industrial area in southern Poland, in: Proc. of the 10th Int. Conf. on Harmonisation within Atmospheric Dispers. Modelling for Regulatory Purposes, 2005, pp. 664-668.

[21] CHEN L., CHEN G., Fuzzy Modeling, Prediction, and Control of Uncertain Chaotic Systems Based on Time Series, IEEE Trans. Circuits and Systems I, vol. 47, no. 10 (2000).

[22] BRUNET G., The first hundred years of numerical weather prediction, Proceedings of the 19th Int. Symposium on High Performance Computing Systems and Applications, Canada (2005).

[23] BANACOS P.C., SCHULTZ D.M., The use of moisture flux convergence in forecasting convective initiation: historical and operational perspectives, Weather and Forecasting, 20, 351-366 (2005).

[24] STAWIANY W., CABAN P., CIMANDER B., CISOWSKA E., HOŁDA I., KRAJNY E., KRUCZAŁA, OSTROWSKA E., OŚRÓDKA L., RORBEK K., SOCHA S., ŚWIĘCH-SKIBA J., WOJTYLAK M., Monograph of automatic measurements of air quality in Katowice Conurbation (1993-1999), Biuletyn Regionalnego Monitoringu Środowiska: Wojewoda Śląski (2000) 1-324.