Chapter 18
Nonlinear Time Series Prediction of Atmospheric Visibility in Shanghai

Jian Yao and Wei Liu

Abstract. Atmospheric visibility has recently become more essential to both the aviation safety and environmental pollution studies. Due to the characteristic of nonlinear time series, the visibility is difficult to predict by traditional statistical method. In this study, fuzzy time series models are used to predict the atmospheric visibility in Shanghai. The irregular dynamic of visibility was firstly investigated by the histogram as well as the autocorrelation analysis to identify the long-term memory of its behavior. Observed single-variable time series data were used to construct the fuzzy forecasting model. Parameters needed to construct the model were chosen to extract the rule of visibility variation. The results revealed that fuzzy time series could well predict the variation of visibility. The relative error between model outputs and observations was within the practically acceptable limits, which points out that atmospheric visibility could be explained and well predicted by the fuzzy time series.

1 Introduction

1.1 Application of Time Series Analysis to Environmental Research

Prediction of air quality takes place in the environmental management. Among the various statistical prediction method, time series analysis is one or the most useful tool for simulation of pollutant concentration. Time series analysis applies the previous data to find the regularities and estimate future values in consecutive time moments. It has been widely used in the field of economics, finance, and

Jian Yao · Wei Liu
Key Laboratory of Nuclear Analysis Techniques, Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China
e-mail: `yaojian@sinap.ac.cn`

W. Pedrycz & S.-M. Chen (Eds.): Time Series Analysis, Model. & Applications, ISRL 47, pp. 385–399.
DOI: 10.1007/ 978-3-642-33439-9_18

market forecasting, especially for the phenomenon with periodical cycle. However, the way to predict the non-linear data or those without distinct relationship remains unclear.

A time series is a sequence of statistical data arranged over time. It is obtained at determined time moments from environmental system of interest. This analysis is fundamental to scientific, engineering, and business research. A time series can be divided into two categories, single variable and multiple variables. The major model of single variable time series is auto-regressions model (ARs), and the model for multivariable time series is vector auto-regressions model (VARs). In mathematical expression, the stochastic process is defined as the behavior of a random variable { z, t € T}, where T is the range of the variable of t. If T = $(-\infty, \infty)$, then the stochastic process can be represented by $\{Z_t, -\infty < t < \infty\}$.

A time series can be represented as follows:

$$X = \{x_t, t = 1,\ldots, N\} \tag{1}$$

where t is the time index and N is the total number of observations.

The stationary time series is an important random series with the following properties:

$$E(Y_t) = E(T_{t-s}) = \mu_y \tag{2}$$

$$V(Y_t) = Var(Y_{t-s}) = \sigma_y^2 \tag{3}$$

$$Cov(Y_t, Y_{t-s}) = Cov(Y_{t-j}, Y_{t-j-s}) = \gamma_s, \text{ for all t, t-s, t-j-s} \tag{4}$$

where μ_y, σ_y^2, and γ_s are some constants. If time series is not stationary, then the process exhibits changes over time.

The disadvantage present in traditional time series analysis is that the time series should be converted to stationary and periodic series prior to its analysis (Kantz and Schreiber 1997). To alleviate this problem, we explored the fuzzy time series and established its inference engine. The prediction results of fuzzy time series were compared with those formed by the traditional time series method to determine the feasibility of this fuzzy time series. The visibility concentration was analyzed by the fuzzy time series with data mining technique, which was useful in extracting the hidden knowledge and characteristic pattern from existing database.

Mining Fuzzy time series forms the focus of this paper. The main purpose of time series data mining is to abstract the phase space of time delay and represent it by mathematical equation. Fuzzy theory is broadly used to forecast the phenomenon with uncertainty. This study presents the inference engine of fuzzy time series deduced from previous information or other important predictors. The fuzzy time series use the fuzzy interval to cope with the random perturbation of observed data. An important information extraction method, data mining, is the technique to analyze data with hidden patterns. The trend of visibility can be captured, as being revealed by the experimental analysis.

1.2 Definition of Atmospheric Visibility and Its Importance

Visibility is defined as the greatest distance in a given direction at which an object can be visually identified with unaided eyes. The object could be a dark object positioned prominently against the sky on the horizon in the daytime, or a known, preferably unfocused moderately intense light source at nighttime (Wark et al. 1998). Visibility impairment is a basic form of air pollution that people can see and recognize without special instruments.

As we all known, impairment of atmospheric visibility constitutes many common and vexing problems for different public authorities in multiple countries throughout the world. First, low visibility is obviously a problem for traffic safety. Secondly, reduced visibility is a cause of delays and disruption in air, sea and ground transportation for passengers and freight. Of cause, impaired visibility is also a symptom of environmental problems because it is evidence of air pollution (Hyslop 2009); in addition, it has been shown that impaired visibility in urban environment and mortality are correlated (Thach et al. 2010). Therefore, visibility degradation is a major problem in atmospheric pollution in many mega cities around the world. Impairment of visibility is not just an aesthetic problem, but could also be used as a visual indicator of ambient air quality in urban areas (Watson 2002). Improvement of visibility requires an understanding of what constituents in the atmosphere impair visibility as well as the origins of those constituents.

From 1973 to 2007, visibility had decreased substantially over the globe except for Europe (Wang et al. 2009). In the Asian region, dozens of studies have reported a severe decline in visibility (Vingarzan and Li 2006; Chang et al. 2009; Tsai et al. 2003). Many analysts have been conducted worldwide (Dzubay et al. 1982; Larson et al. 1988; Johnson et al. 1990; Wilson and Suh 1997; Kim et al. 2001; Clancy et al. 2002). Furthermore, visibility impairment due to urban aerosol has been the subject of numerous air pollution studies around the world over the past several decades (Chan et al. 1999; Lee and Sequeira 2002; Zhang et al. 2004). It is known that the impairment of visibility is attributed primarily to the scattering and absorption of visible light by suspended particles, as well as by gaseous pollutants (e.g. NO_2) in the atmosphere (Appel et al. 1985; Hodkinson 1966; Groblicki et al. 1981; Latha and Badarinath 2003). Among them, fine particulates, which include sulfates, nitrates, organic and elemental carbon, and soil, effectively scatter or absorb visible light and thus reduce visibility (Malm et al. 1994, 1996; Sisler and Malm1994; Latha and Badarinath 2003; Kim et al. 2006; Tan et al. 2009a, 2009b).

Previous studies revealed that the size, chemical composition, and mass concentration of airborne particles substantially affect visibility (Conner et al. 1991; Malm and Pitchford 1997). Although the extinction of visible light from gaseous species can also impair visibility, such species have a much weaker influence (Chan et al. 1999; Dzubay et al. 1982). Also the PM Science Assessment Report (North American Research Strategy for Tropospheric Ozone (NARSTO) 2004) published recently by the NARSTO suggested that the chemical and physical properties of PM with an aerodynamic diameter less than 2.5 μm have to be better characterized, as they are responsible for adverse health effects linked to chronic respiratory diseases (Dockery and Pope 1994) and visibility impairment (Malm 1999).

In addition to air pollutants, many meteorological elements such as relative humidity (RH), pressure, wind, and temperature may directly or indirectly contribute to the degradation of visibility (Lee 1990; Green et al. 1992; Malm et al. 1994; Raunemaa et al. 1994; Tsai and Cheng 1998, 1999). Relative humidity in and of itself does not reduce visibility, but as RH increases, hygroscopic particles progressively absorb more water, thus increasing their scattering cross section and proportionately reducing visibility. Therefore RH directly affects the particles that contribute to visibility reduction.

However, other meteorological variables, such as wind speed, temperature, and barometric pressure, have little to no direct effect on visibility but may have an effect on the concentration of atmospheric particles because of atmospheric dispersive characteristics. According to studies by Chang (1999) and Tsai and Cheng (1999), lower wind speeds cause particulates to gather and subsequently prevent them from spreading, which in turn indirectly affects air quality.

Nevertheless, either fine particulates, or meteorological parameters are hardly to be controlled and forecasted. It becomes necessary to explore a new method to predict atmospheric visibility.

1.3 Stochastic Property for Environmental Phenomenon and Visibility

The analysis of visibility time series starts from the following two definition.

Definition 1: Visibility is a chaotic occurrence in the environmental system.

Definition 2: Since we get the measured visibility data in the monitoring station, we call the observational results in time domain as a "visibility time series".

The chaotic nature exists in many fields such as air pollution concentration, stock price index, rainfall, and earthquake. The nonlinear of visibility comes from many reasons such as the scale-invariant and clustering characteristics. Because the times of visibility are a scale-dependent process, it is not easy to extract the information by the traditional way. The irregular dynamic behavior, or chaos, could be explained by the influence of some non-linear interdependent parameters in the system.

The chaotic behavior could be investigated by many tools such as histograms or spectral analysis. The chaotic indicator, the correlation dimension was the tool for the evaluation of pollution concentration to know the possible chaotic characteristic. The autocorrelation could identify the long-term memory and the possibility of scale invariance.

2 Model Developments

2.1 Theory of Fuzzy Time Series Analysis

A visibility time series can be defined as the following.

Definition 3: Visibility variation is a sequence of numerical data represented as follows:

$$X_{vis} = \{x_{vis\,t},\ t = 1,\ldots\ldots,N\} \tag{5}$$

Definition 4: A fuzzy time series of visibility can be estimated by the event characterization function, ECF.

The event of time series is defined by the event characterization function, ECF, as follows

$$ECF = g(t) = g(X_i, X_{i-1}, \ldots\ldots X_1) \tag{6}$$

The event characterization function is defined in such a way that its value at t time index correlates highly with the occurrence of an event at some specified time in the future (Povinelli 1999).

In analyzing the event in time series data mining, g (t) $=X_{i+1}$, the ECF can be chosen as:

$$g(t) = \frac{X_{i+1} - X_i}{X_i} \tag{7}$$

(7) offers the clear relationships of X_{i+1} and X_i in predicting the visibility.

Event characterization functions can be defined by different ways for event predictions at different time series. The event characterization function varies according to the objective of prediction. For example, if x_t represents today's monitoring results and the target is to predict the change of tomorrow's visibility, then the event characterization function can be defined in the form given above.

The membership function of the fuzzy time series was specified in the form *f*(*x, a, c*) with two parameters *a* and *c,* and it is a mapping on a vector X. Depending on the sign of the parameter *a*, it is appropriate for representing concepts such as "very large" or "very small".

$$f(x,a,c) = \frac{1}{1+e^{-a(x-c)}} \tag{8}$$

The value of X_{vis} at time step (n+1) is determined by the membership function *f(x , a ,c)* and the value of its previous step, as shown below.

$$X_{vis}(t_{n+1}) = X_{vis}(t_n) + f(x,a,c) \cdot [\, X_{vis}(t_{n+1}) - X_{vis}(t_n) \,] \tag{9}$$

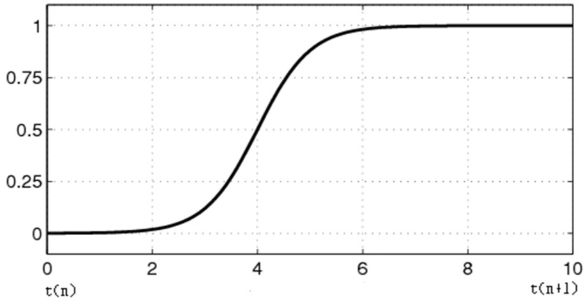

Fig. 1 The membership function of fuzzy time series in the interval [t_n, t_{n+1}]

2.2 Data Mining Technique of Visibility Time Series

How to extract the useful knowledge from time series data? The data mining could be a good answer. Data mining technique can extract hidden and useful information using various effective ways using pattern recognition, machine learning, artificial intelligence, and statistical methods. (Han and Kambe 2005).

Data mining is the analysis of data with the goal of uncovering hidden patterns. It is defined as extracting useful and meaningful information using statistic, machine learning, artificial intelligence and pattern recognition techniques from large data sets. (Han and Kamber 2005) Povinelli defines it as "combining of data mining, time series analysis and genetic (Povinelli 1999). Weiss and Indurkhya defined it as "the search for valuable information in large volumes of data". (Weiss and Indurkhya 1998).

Data mining is the process of discovering hidden and useful information from huge data. The data mining technique and nonlinear time series analysis to analyze a time series were combined in time series data mining. The event is considered as an interesting pattern when data mining is applied to time series data. The prediction algorithm based on data mining of fuzzy time series is shown in figure 2.

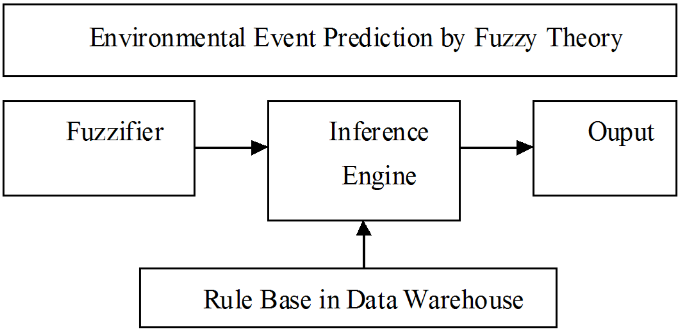

Fig. 2 The prediction algorithm of fuzzy time series with data mining technique

2.3 Prediction of Visibility by ARIMA Model

The model proposed by Box-Jenkins, with the method of Autoregressive Integrated Moving Average (ARIMA), is the most frequent used traditional time series analysis methods. A seasonal univariate $\mathrm{ARIMA}(p,d,q)(P,D,Q)\mathrm{s}$ model is given by

$$\Phi(B)[\Delta y_t - \mu] = \Theta(B)a_t \qquad \mathrm{t} = 1, \cdots\cdots, N \tag{10}$$

Where

$$\Phi(B) = \varphi p(B)\Phi P(B) \tag{11}$$

$$\Theta(B) = \theta q(B)\Theta Q(B) \tag{12}$$

and μ is an optional model constant. It is also called the stationary series mean, assuming that, after differencing, the series is stationary. An optional log scale transformation can be applied to y_t before the model is fitted. In this section, the same symbol, y_t, is used to denote the series either before or after log scale transformation.

Independent variables $x_1, x_2, ..., x_m$ can also be included in the model. The model with independent variables is given by

$$\Phi(B)\left[\Delta\left(y_t - \sum_{i=1}^{m} c_i x_{it}\right) - \mu\right] = \Theta(B)a_t \tag{13}$$

where $c_i, i = 1, 2, ..., m$, are the regression coefficients for the independent variables.

Basically, two different estimation algorithms are used to compute maximum likelihood (ML) estimates for the parameters in an ARIMA model. Melard's algorithm is used for the estimation when there is no missing data in the time series. The algorithm computes the maximum likelihood estimates of the model parameters. The details of the algorithm are described in (Melard 1984), (Pearlman 1980), and (Morf et al. 1974). A Kalman filtering algorithm is used for the estimation when some observations in the time series are missing. The algorithm efficiently computes the marginal likelihood of an ARIMA model with missing observations. The details of the algorithm are described in the following literature: (Kohn and Ansley 1986) and (Kohn and Ansley 1985).

The Conditional least square, CLS, IS used as the forecasting method in the ARIMA model. Define $\hat{y}_t(l)$, the *l*-step-ahead forecast of y_{t+1} at the time *t*, can be described as:

$$\hat{y}_t(l) = D(B)\hat{y}_{t+l} + \Phi(B)\mu + \Theta(B)\hat{a}_{t+l} + \sum_{i=1}^{m} c_i \Phi(B)\Delta x_{i,t+1} \tag{14}$$

Note that

$$\hat{y}_{t+l-i} = \begin{cases} y_{t+l-i} & \mathit{if}\ l \le i \\ \hat{y}_t(l-i) & \mathit{if}\ l > i \end{cases}$$

$$\hat{a}_{t+l-j} = \begin{cases} y_{t+l-i} - \hat{y}_{t+l-i-1}(1) & \mathit{if}\ l \le i \\ 0 & \mathit{if}\ l > i \end{cases} \tag{15}$$

3 Results and Discussion of Numerical Experiment

3.1 Experimentals

Three experiments were performed in this study. The first is the statistical analysis. The second was tradition forecasting method by ARIMA model. And the third was the fuzzy forecasting. The statistical analyses calculate the basic attributes of this time series which helps us to clarify its feature. The fuzzy forecasting applied the method in section 2.1 and 2.2 to find the rules in the data mine. The results were compared with the results from ARIMA model. Meanwhile, the trend of the time series was represents by the dimensionless time series together with the autocorrelation analysis for the long term memory in the data mining.

Model evaluation performance was compared by the observed data with the forecasted data. The scattering diagram was shown and the statistical value such as the root mean square error (RMSE); mean absolute percentage error (MAPE); maximum absolute percentage error (MaxAPE); mean absolute error (MAE); and maximum absolute error (MaxAE) were calculated as well.

3.2 Study Area and Data Collection

A time series of hourly average visibility observations, which was obtained from the monitoring station at Shanghai, was analyzed by descriptive statistics and statistical methods and fuzzy model to examine the temporal structures of visibility. The length of time for analysis was one year, which is enough to discriminate the most important feature of this time series.

3.3 The Correlation between Model and Observed Values

The model performance evaluation was accomplished by the comparison of forecasted value with the observed value. The correlation coefficient, which represents the relationship between the two quantities, was calculated as follows:

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{16}$$

Where $\bar{X}$ is the sample mean.

The values of the correlation coefficient were shown in Table 1.

Table 1 The values of the correlation coefficient for different variables

Parameter	Description	R value
T	Temperature	0.486[a]
V	Wind velocity	0.214[a]
RH	Relative Humidity	0.162[a]
$PM_{2.5}$	Particulate matter with diameter less than 2.5 μ m	-0.125[a]

Note: Confidence level: a: 0.01

3.4 The Trend of Time Series Data

The non-dimensional time series were used to compare the trend of time series data. The dimensionless time series were represented by

$$X = [\hat{x}_i, i = t, \cdots, n], \hat{x}_i = \frac{x_i}{\overline{x}}. \tag{17}$$

Where, X is the domain of time series, x_i is the individual value, $\overline{x}$ is the average value of the time series.

Figure 3 shows the trend of visibility and its related variables, the occurrence sequence of the hourly averaged visibility data in this study. This figure reveals that the characteristic of stochastic perturbation is obvious, and this is also the reason why a simple linear regression cannot be used in the prediction of atmospheric visibility.

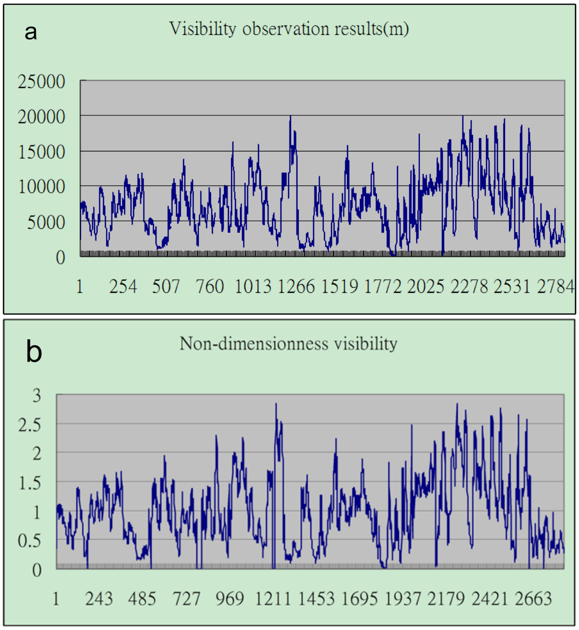

Fig. 3 The trend of visibility (a) is the visibility observation results and (b) is the non-dimensionless plot of the results

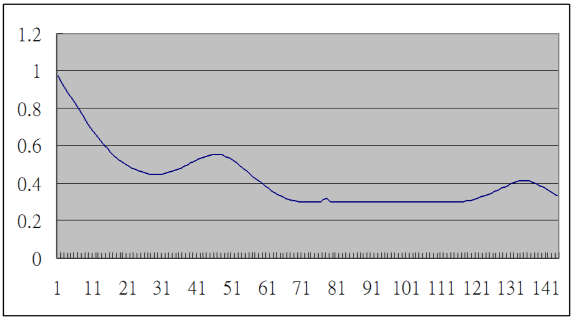

Fig. 4 Autocorrelation coefficients of visibility time series

The autocorrelation of this study is shown in figure 4. As shown here, the value of autocorrelation in decreasing. The results reveal that the long term memory effect is not obvious. There are probably many factors which will influence the variation of atmospheric visibility. The possible influence factors were listed in Table 1.

3.5 Analysis of Data by Fuzzy Time Series

The comparison of observed value and prediction results were shown in figure 5. The observed value is shown along the horizontal axis and the forecasted value is presented on the vertical axis. The results reveal that the relationship was fine. The fuzzy model could explain the tendency of the variation of atmospheric visibility time series. The forecasting results of fuzzy model are shown in Figure 6.

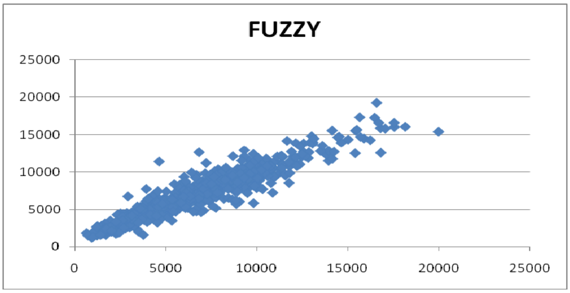

Fig. 5 Observed value and predicted value by fuzzy time series

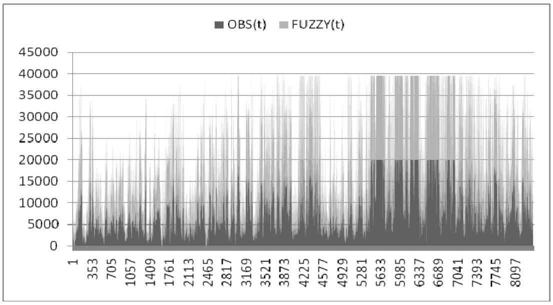

Fig. 6 The forecasting results of fuzzy model

3.6 Comparison of ARIMA Model

Table 2 is the summary statistics of the results by ARIMA (1, 1, 1) model. Figure 7 is the prediction results of ARIMA model. This figure also shows that the ARIMA model could also predict the atmospheric visibility time series.

In order to have a more clear comparison of the forecasting, the errors of these two methods were shown in figure 8. The values of error for these two models were converting into a 100 percents scale. The results produced by the fuzzy forecasting model are shown in grey and the ARIMA results were shown in black. It is seen that these two models are similar. In some place the results of fuzzy is better while in some place the ARIMA model is better. The results reveal that both models are capable to predict the non-linear characteristics of the atmospheric visibilities.

Table 2 Summary statistics of the results by ARIMA model

Summary Statistics	Aver	Min	Max	Percentile 5	10	50	90	95
Stationary R square	.253	.253	.253	.253	.253	.253	.253	.253
R square	.959	.959	.959	.959	.959	.959	.959	.959
RMSE	719 .380	719 .380	719 .380	719 .380	719 .380	719 .380	719 .380	719 .380
MAPE	9 .079	9 .079	9 .079	9 .079	9 .079	9 .079	9 .079	9 .079
MaxAPE	81 .433	81 .433	81 .433	81 .433	81 .433	81 .433	81 .433	81 .433
MAE	517 .445	517 .445	517 .445	517 .445	517 .445	517 .445	517 .445	517 .445
MaxAE	3096 .390	3096 .390	3096 .390	3096 .390	3096 .390	3096 .390	3096 .390	3096 .390
Normalized BIC	13 .237	13 .237	13 .237	13 .237	13 .237	13 .237	13 .237	13 .237

Note: RMSE: root mean square error; MAPE: mean absolute percentage error; MaxAPE: maximum absolute percentage error; MAE: mean absolute error; MaxAE: maximum absolute error

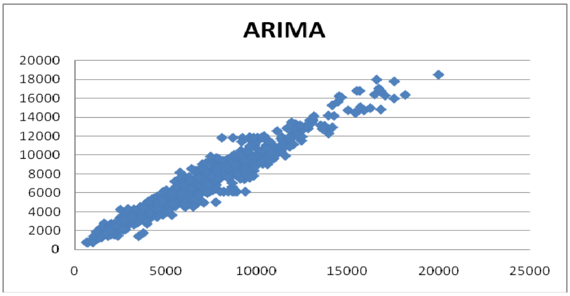

Fig. 7 Observed value and predicted value by ARIMA

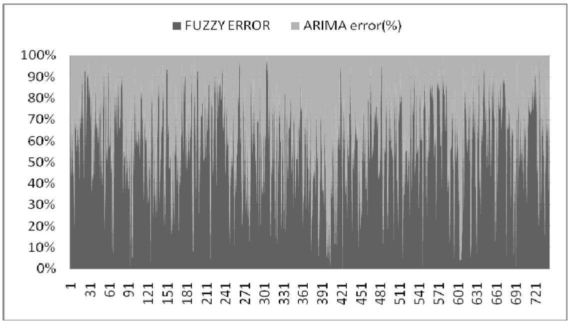

Fig. 8 The comparison between fuzzy time series and ARIMA

4 Conclusions

The study presents the results of atmospheric visibility forecasting by fuzzy time series. The proposed method, time series data mining, is based on the fuzzy logic to find the hidden rule in previous data. The fuzzy inference engine was obtained by the event characterization function, ECF, to abstract the previous information of variation in this study. In order to know the performance of this method, the forecasting was compared with the ARIMA model. The results reveal that fuzzy time series can handle the non-linear characteristics as effectively as ARIMA. However, the prediction of environmental event by fuzzy time series is simpler and does not include complex relationships. Comparing to the traditional forecasting method, the method can describe complicated nature of the environmental phenomenon, especially for non-periodical, non-cyclical data.

References

Appel, B.R., Tokiwa, Y., Hsu, J., Kothny, E.I., Hahn, E.: Visibility as related to atmospheric aerosol constituents. Atmos. Environ. 19, 1525–1534 (1985)

Ansley, C.F., Kohn, R.: Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions. Ann. Stat. 13(4), 1286–1316 (1985)

Chan, Y.C., Simpson, R.W., Mctainsh, G.H., Vowles, P.D., Cohen, D.D., Bailey, G.M.: Source apportionment of visibility degradation problems in Brisbane (Australia)—using the multiple linear regression techniques. Atmos. Environ. 33, 3237–3250 (1999)

Chang, D., Song, Y., Liu, B.: Visibility trends in six megacities in China 1973–2007. Atmos. Res. 94(2), 161–167 (2009)

Chang, J.: The relationship of visibility with physical and chemical characteristics of suspended particles in Kaohsiung City. Master Thesis, National Sun Yat-Sen University, Kaohsiung, Taiwan (1999)

Clancy, L., Goodman, P., Sinclair, H., Dockery, D.W.: Effect of air-pollution control on death rates in Dublin, Ireland: intervention study. Lancet 360, 1210 (2002)

Conner, W.D., Bennett, R.L., Weathers, W.S., Wilson, W.E.: Particulate characteristics and visual effects of the atmosphere at Research Triangle Park. J. Air Waste Manage. Assoc. 41, 154–160 (1991)

Dockery, D.W., Pope, C.A.: Acute respiratory effects of particulate air pollution. Annu. Rev. Publ. Health 15, 107–132 (1994)

Dzubay, T.G., Stevens, R.K., Lewis, C.W., Hern, D.H., Courtney, W.J., Tesch, J.W., et al.: Visibility and aerosol composition in Houston, Texas. Environ. Sci. Technol. 16, 514–525 (1982)

Green, M.C., Flocchini, R.G., Myrup, L.O.: The relationship of the extinction coefficient distribution to wind field patterns in southern California. Atmos. Environ. 26, 827–840 (1992)

Groblicki, P.J., Wolff, G.T., Countess, R.J.: Visibility reducing species in the Denver Brown Cloud—1. Relationships Between Extinction and Chemical Composition. Atmos. Environ. 15, 2473–2484 (1981)

Han, J., Kamber, M.: Data Mining: Concepts and Techniques, p. 800. Academic Press, San Francisco (2005)

Hodkinson, J.R.: Calculations of color and visibility in urban atmospheres polluted by gaseous NO_2. Int. J. Air Water Pollut. 10, 137–144 (1966)

Hyslop, N.P.: Impaired visibility: the air pollution people see. Atmos. Environ. 43(1), 182–195 (2009)

Johnson, K.G., Gideon, R.A., Luftsgaarden, D.O.: Montana air pollution study: children's health effects. J. Off. Stat. 5, 391–408 (1990)

Kantz, H., Schreiber, T.: Nonlinear Time Series Analysis, p. 388. Cambridge University Press, Cambridge (1997)

Kim, K.W., Kim, Y.J., Oh, S.J.: Visibility impairment during yellow sand periods in the urban atmosphere of Kwangju, Korea. Atmos. Environ. 35, 5157–5167 (2001)

Kim, Y.J., Kim, K.W., Kim, S.D., Lee, B.K., Han, J.S.: Fine particulate matter characteristics and its impact on visibility impairment at two urban sites in Korea: Seoul and Incheon. Atmos. Environ. 40, 593–605 (2006)

Kohn, R., Ansley, C.F.: Estimation, prediction, and interpolation for ARIMA models with missing data. J. Am. Stat. Assoc. 81(395), 751–761 (1986)

Larson, S.M., Cass, G.R., Hussey, K.J., Luce, F.: Verification of image processing based visibility models. Environ. Sci. Technol. 22, 629–637 (1988)

Latha, K.M., Badarinath, K.V.S.: Black carbon aerosols over tropical urban environment—a case study. Atmos. Res. 69, 125–133 (2003)

Lee, D.O.: The influence of wind direction, circulation type and air pollution emissions on summer visibility trends in southern England. Atmos. Environ. 24A, 195–201 (1990)

Lee, Y.L., Sequeira, R.: Water-soluble aerosol and visibility degradation in Hong Kong during autumn and early winter, 1998. Environ. Pollut. 116, 225–233 (2002)

Malm, W.C.: Introduction to visibility, cooperative institute for research in the atmosphere (1999), http://vista.cira.colostate.edu/improve/Education/intro_to_visibility.pdf (accessed December 15, 2011)

Malm, W.C., Pitchford, M.L.: Comparison of calculated sulfate scattering efficiencies as estimated from size-resolved particle measurements at three national locations. Atmos. Environ. 31, 1315–1325 (1997)

Malm, W.C., Sisler, J.F., Huffman, D., Eldred, R.A., Cahill, T.A.: Spatial and seasonal trends in particle concentration and optical extinction in the United States. J. Geophys. Res. 99 (D1), 1347–1370 (1994)

Melard, G.: Algorithm AS 197: A fast algorithm for the exact likelihood of autoregressive-moving average models. Appl. Stat. 33(1), 104–114 (1984)

Morf, M., Sidhu, G.S., Kailath, T.: Some new algorithms for recursive estimation on constant, linear, discrete-time systems. IEEE Trans. Auto. Control AC-19, 315–323 (1974)

NARSTO.: Particulate Matter Assessment for Policy Makers: A NARSTO Assessment. In: McMurry, P., Shepherd, M., Vickery, J. (eds.). Cambridge University Press, Cambridge (2004)

Raunemaa, T., Kikas, U., Bernotas, T.: Observation of submicron aerosol, black carbon and visibility degradation in remote area at temperature range from −24 to 20 °C. Atmos. Environ. 28, 865–871 (1994)

Pearlman, J.G.: An algorithm for the exact likelihood of a high-order autoregressive-moving average process. Viometrika 67, 232–233 (1980)

Povinelli, R.J.: Time Series Data Mining: Identifying Temporal Patterns for Characterization and Prediction of Time Series Events. Ph.D. Dissertation, Marquette University, p.180 (1999)

Sisler, J.F., Malm, W.C.: The relative importance of soluble aerosols to spatial and seasonal trends of impaired visibility in the United States. Atmos. Environ. 28, 851–862 (1994)

Tan, J.H., Duan, J.C., Chen, D.H., Wang, X.H., Guo, S.J., Bi, X.H., et al.: Chemical characteristics of haze during summer and winter in Guangzhou. Atmos. Res. 94, 238–245 (2009a)

Tan, J.H., Duan, J.C., He, K.B., Ma, Y.L., Duan, F.K., Chen, Y., et al.: Chemical characteristics of $PM_{2.5}$ during a typical haze episode in Guangzhou. J. Environ. Sci. 21, 774–781 (2009b)

Thach, T.Q., Wong, C.M., Chan, K.P., Chau, Y.K., Chung, Y.N., Ou, C.Q., et al.: Daily visibility and mortality: assessment of health benefits from improved visibility in Hong Kong. Environ. Res. 110(6), 617–623 (2010)

Tsai, Y.I., Cheng, M.T.: Effects of sulfate and humidity on visibility in the Taichuang harbor area (Taiwan). J. Aero. Sci. 29, 1213–1214 (1998)

Tsai, Y.I., Cheng, M.T.: Visibility and aerosol chemical compositions near the coastal area in central Taiwan. Sci. Total Environ. 231, 37–51 (1999)

Tsai, Y.I., Lin, Y.H., Lee, S.Z.: Visibility variation with air qualities in the metropolitan area in southern Taiwan. Water Air Soil Poll. 144, 22 (2003)

Vingarzan, R., Li, S.M.: The Pacific 2001 Air Quality Study–synthesis of findings and policy implications. Atmos. Environ. 40(15), 2637–2649 (2006)

Wang, K., Dickinson, R.E., Liang, S.: Clear sky visibility has decreased over land globally from 1973 to 2007. Science 323(5920), 1468–1470 (2009)

Wark, K., Warner, C.F., Davis, W.T.: Air Pollution—Its Origin and Control, 3rd edn. Addison-Wesley Longman, Reading (1998)

Watson, J.G.: Visibility: science and regulation. J. Air Waste Manage. Assoc. 52, 628–713 (2002)

Weiss, S.M., Indurkhya, N.: Predictive Data Mining: A practical Guide, p. 228. Morgan Kaufmann, San Fransisco (1998)

Wilson, W.E., Suh, H.H.: Fine particles and coarse particles: concentration relationships relevant to epidemiologic studies. J. Air Waste Manage. Assoc. 47, 1238–1249 (1997)

Zhang, R., Wang, M., Sheng, L., Kanai, Y., Ohta, A.: Seasonal characterization of dust days, mass concentration and dry deposition of atmospheric aerosols over Qingdao, China. China Particuology 2(5), 196–199 (2004)

Abbreviations

ARs	auto-regressions model
VARs	vector auto-regressions model
NARSTO	North American Research Strategy for Tropospheric Ozone
RH	relative humidity
ECF	event characterization function
ARIMA	autoregressive integrated moving average
ML	maximum likelihood
CLS	Conditional least squares
FTS	fuzzy time series
RMSE	root mean square error
MAPE	mean absolute percentage error
MaxAPE	maximum absolute percentage error
MAE	mean absolute error
MaxAE	maximum absolute error