Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090104

Alpha Model: A Mathematical Modeling Approach Applied to an Air Quality Monitoring Network

C. El Aoun[1], *N. Boutabba*[2,3] *and H. Eleuch*[4,*]

[1] Institut des Hautes Etudes Comerciales de Sfax, B.P 43 - 3061 Sfax, Tunisia
[2] Physics Department, Cegep St-Laurent, Montreal, Qc, H4L 3X7, Canada
[3] University of Carthage, Tunis 1040, Tunisia
[4] Department of Physics, McGill University, Montreal, Qc, H3A 2T8, Canada

Received: 12 Jun. 2014, Revised: 30 Jun. 2014, Accepted: 1 Aug. 2014
Published online: 1 Jan. 2015

Abstract: In this paper we apply the alpha Model to air quality prediction purposes. This model is used to predict the ground level ozone concentration which is influenced by multiple pollutants such as NO2, CO, Pb and particulate matter. The model was previously elaborated and successfully tested on other complex systems. Our main object here is to show that the alpha model is a powerful technique for analyzing performance complexities of implementing input-output relations. Air quality is taken as an example.

Keywords: Air quality prediction, model, alpha model.

1 Introduction

Our modern society is data oriented. The major problems in our modern life are related to predictions: prediction of financial stock markets, volcanic eruptions, tsunamis, epidemiological widespread diseases, political stability, economical growth and financial crisis etc. These phenomena have extreme complicated dependence on big number of observable and non observable parameters. Concrete scientific models based on fundamental laws even combined with heavy computations fail to describe such phenomena. Several empirical methods based on historical data and time series have been developed for example neuronal networks [1,2,3] and genetic algorithms[4]. These techniques are often considered as black-box models where explicit relations between the input and the output data can not be seen. Recently, a new method, the alpha model has been developed by two of us [5]. The alpha model gives explicit relations between the inputs and the outputs data based on historical or time series data. It was applied to forecast of several financial stock markets and diffuse solar irradiance [5,6]. Here we extend the application sets of the alpha model to study and predict the air pollution. Our main goal is to show the prediction power of this technique.

Air pollution occurs when unwanted chemicals or other contaminants are released into the air in large enough amounts to harm the health of people, plants, animals, and our environment. Air quality degradation is one of the important environmental hazards and a lot of research is therefore conducted in this field. It is the result of our modern life, in fact one of the main causes of air pollution is manufacturing which produces gases or vapors, dust, smoke or soot. However, Some other pollutants are emitted from natural processes such as forest fires, decaying vegetation, dust storms, and volcanic eruptions. The major pollutants are particles of diameter 10 μm or less (PM10), very small particles of diameter 2.5μm or less (PM2.5), nitrogen dioxide (NO2), carbon monoxide (CO) and ozone (O3). Moreover, solar radiation, wind speed, opaque cloud cover and ozone itself contribute considerably in the ozone formation[7]. Several authors have developed models based on neural networks to forecast air quality and air pollutant concentration [8,9,10]. They are important tools and they play a crucial role in protecting air quality. The various models used to predict air quality and to analyze travel road strategies attempt to minimize possible negative impacts of pollution on the environment[11,12,13]. In this work, we focus on the analysis of two data sets

* Corresponding author e-mail: heleuch@physics.mcgill.ca

described below using the alpha model [5,6] to predict the ozone formation.

The analysis part of the modeling did lead to clear results regarding air quality prediction in the area where the data have been collected.

2 Method and materials

2.1 Data sets

Two data sets were used. The first one is (856 points) describing the daily ozone concentration in Houston summer (April-October) along the Gulf coast during the years 1982-1985 (in ppm). The meteorologic variables considered as predictors are the solar radiation(in W/m^2), the wind speed (in m/s), the opaque cloud cover (in percentage) and the ozone of a previous day. More information about this data can be found in the Geophysical Statistics Project, through NCAR.

The second data set are daily max 8-hours ozone formation, hourly pollutant and meteorological averages (for the first 173 days of 2012) in Los Angeles, California. These data have been taken from the US environmental protection agency web site. Pollutants are carbon monoxide (CO) (in parts per million unit (ppm)), Ozone (in parts per million unit (ppm)), Nitrogen dioxide (in parts per billion ppb), PM10 and PM2.5 (in $\mu g/m^3$) and Pb (in $\mu g/m^3$).

2.2 Alpha Model

Several Models in finance and economics are inspired from physical concepts such as kinetic exchange models of markets [14], quantum finance [15,16], correlation function approach [17] and alpha model [6]. Kinetic exchange models of markets apply the kinetic theory of gases to the exchange markets. Quantum finance uses the quantum mechanical formalism to analyze the option stock markets. We have recently adapted the correlation analysis [17] (widely used in electronics [18], plasma physics[19], astrophysics[20], atomic physics [21,22], statistical mechanics [23], condensed matter physics [24, 25,26,27] and telecommunication[28]) to identify the business cycle turning points.

The alpha model was previously used for finance applications in the context of e-commerce negotiation and stock exchange forecast [5]. It was also proposed as a model for prediction of hourly solar irradiance values [6]. In the present paper, we investigate the alpha model for air quality prediction. This model was developed based on analogy between the physics (Coulomb's) law and buyer-seller relation in finance. It describes the attraction between buyer and seller analogous to the attraction between two opposite charges. The case of seller/seller or buyer/buyer is considered as two similar charges pushing each other. The alpha model involves an explicit relation between the observable input and output variables with set of parameters adjusted from historical or time series data. As described in our previous work in finance applications[5], the parameter α represents the ratio between the final price and the starting price. It depends on the mean observable variables that influence the final price of the auction. The alpha model is a method aided-decision for selling or buying auctions. In fact, the prediction of the final price is possible if we have have enough information on α. In this work, we apply the alpha model for the two sets of data described above. In this section we are focusing on two goals. The first goal is to identify which predictors in this complex system have more impact and influence on the ozone concentration. We need then to study the relation between the input parameters (predictors) and the output variable (ozone concentration) at the same day. The second goal is concentrated on predicting the output variable in day $j+1$ from the input parameters at previous day j. Let us first present the alpha model for exploring the dominant pollution factors of the ozone concentration for the two data sets described in the previous section. The variable pollutants taken into account in the area of Houston (first data set) are the opaque cloud cover ($opcov$), the wind speed ($WSPD$) and the solar radiation (SR) at the same day $j+1$ and the ozone (O_3) of the previous day j. In this case the expression of the alpha model can be written as [5,6]

$$\alpha_{j+1}^{Houston} = \frac{(O3_j)^{2\beta_1}}{[r_1 + r_{01}]^2}, \tag{1}$$

with

$$r_1 = \left(opcov_{j+1}\right)^{\beta_2} \left(WSPD_{j+1}\right)^{\beta_3} \left(SR_{j+1}\right)^{\beta_4}. \tag{2}$$

For the second set the input variables of the alpha model are the pollutants at time $j+1$ (CO, PM10, NO_2, PM2.5, Pb) and O_3 at time j while the output is the ground level ozone concentration at time $j+1$. In this case the expression of the alpha model can be written as [5,6]

$$\alpha_{j+1}^{Los-Angeles} = \frac{(O3_j)^{2\beta_5}}{[r_2 + r_{02}]^2}, \tag{3}$$

with

$$\begin{aligned} r_2 = {} & \left(CO_{j+1}\right)^{\beta_6} \left(PM10_{j+1}\right)^{\beta_7} \left(NO2_{j+1}\right)^{\beta_8} \times \\ & \left(PM2.5_{j+1}\right)^{\beta_9} \left(Pb_{j+1}\right)^{\beta_{10}}. \end{aligned} \tag{4}$$

Where β_i and r_{0j} are the parameters to be calculated from the historical data. r_{0j} takes into account the residual weak influence of all other variables that are not considered here. α_{j+1} represents the predicted level of the

Table 1: The parameter values of the simulation for Houston

β_1	β_2	β_3
-0.24606	-0.0903078	-0.301688
β_4	MSE	
-0.00202147	0.00061273	

Table 2: The parameter values of the simulation for Los Anglos

β_5	β_6	β_7	β_8
-0.0335725	-0.194444	-0.0623862	0.0433624
β_9	β_{10}	MSE	
-0.0177553	0.14281	$(8.17)10^{-5}$	

ground ozone concentration $O3_{j+1}$ divided by the value of the ozone concentration a day before $O3_j$

$$\alpha_{j+1} = \frac{O3_{j+1}}{O3_j}. \tag{5}$$

If $\alpha > 1$, then the ozone concentration will increase. For $\alpha < 1$ the ozone concentration will decrease.

For using the alpha model to predict the ozone concentration, the two formulas (1) and (2) will be slightly modified. We use in this case the values of the pollutants of the day j in order to predict the ozone concentration in the next day $j+1$, so in this case the alpha model for Houston is then [5,6]

$$\alpha_{j+1}^{Houston} = \frac{(O3_j)^{2\beta_{11}}}{[r_3 + r_{03}]^2}, \tag{6}$$

with

$$r_3 = (opcov_j)^{\beta_{12}} (WSPD_j)^{\beta_{13}} (SR_j)^{\beta_{14}}, \tag{7}$$

and for Los-Angeles is given by [5,6]

$$\alpha_{j+1}^{Los-Angeles} = \frac{(O3_j)^{2\beta_{15}}}{[r_4 + r_{04}]^2}, \tag{8}$$

where

$$r_4 = (CO_j)^{\beta_{16}} (PM10_j)^{\beta_{17}} (NO2_j)^{\beta_{18}} \times (PM2.5_j)^{\beta_{19}} (Pb_j)^{\beta_{20}}. \tag{9}$$

3 Results and discussion

For the first simulation, we consider the first set of variables described above. 756 points are used to estimate the fitting parameters of the alpha model while the other set of 100 points is used to validate the model and to predict the ozone formation. The result of this simulation is plotted in Fig.1 where the daily ozone concentration is plotted during the last 100 days of summer 1985 (July24-October31) in Houston. The values of the fitting parameters of the alpha model are given in table 1. Where MSE represents the mean squared error. These parameters show that the ozone concentration in Houston during 1986 was less sensitive to opaque cloud and very sensitive to the wind speed and highly sensitive to the previous day ozone concentration. We can deduce that the ozone concentration in Houston is approximatively proportional to wind speed in power of 0.6 ($WSPD^{0.6}$) and to the square root of the ozone concentration of the previous day. The second simulation is calculated for the concentration of ozone in Los Angeles during the first 173 days of 2012. 142 points are used to estimate the fitting parameters of the alpha-model. 31 daily measures of the ozone concentration (May 22- June 21, 2012) are used to validate the alpha-model prediction. The calculated values of the fitting parameters for this simulation using the second set of the observable variables are given in table 2.

As β_9 has the smallest value and β_6 has the highest value, the ozone concentration in Los Angeles is less sensitive to the $PM2.5$ and NO_2 (nitrogen dioxide) and highly sensitive to CO and Pb as well as the previous day ozone concentration. As a good approximation, the ozone concentration is proportional to the previous day ozone concentration, to $CO^{1/3}$ and to $Pb^{-1/4}$.

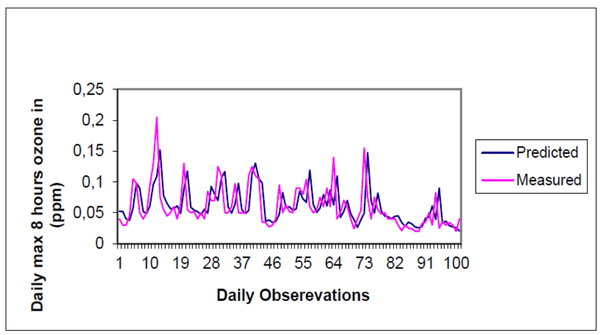

Fig. 1: Ozone daily prediction during 100 days of 1985 (July24-October31) in Houston.

Let us now analyze the predictability of the ozone concentration using the alpha model. Fig.1 and Fig.2 show good fits between the predicted values with alpha-model and the real values. So, the alpha-model has small error for predicting the ozone concentration. For these simulations we obtain the following values of the β parameters (see table 3 and 4).

4 Conclusion

In this paper, we have shown that the alpha-model can be applied to air quality prediction which enlarges the fields

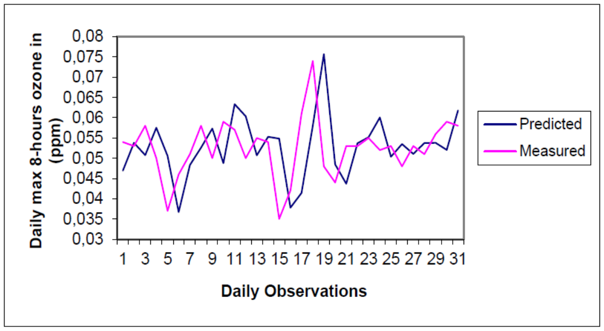

Fig. 2: Daily ozone prediction during one month (May 22- June 21, 2012) in Los Angeles.

Table 3: The parameter values of the ozone concentration forecast for Houston

β_{11}	β_{12}	β_{13}
-0.166322	-0.05572	-0.198179
β_{14}	*MSE*	
0.01568	0.00101191	

Table 4: The parameter values of the ozone concentration forecast for Los Angeles

β_{15}	β_{16}	β_{17}	β_{18}
-0.0175861	-0.045441	-0.029776	-0.0615832
β_{19}	β_{20}	*MSE*	
0.0428351	-0.167781	$(9.34281)10^{-5}$	

of applications for this model. The advantage of this model compared to the input-output methods such as neuronal networks or genetic algorithms is the explicit analytical relation between the input and the output variables that can be established from historical data. This allows a better understanding of the dynamical behavior of the studied systems. It also identifies the most influent and the less influent factors on the dynamical system. Furthermore it opens the door for more general and explicit relations between the input and the output variables in particular systems. For example in this study we have deduced that, on one hand, that in Los Angeles the CO and Pb pollution factors are the most influent parameters, as well as the ozone concentration of the previous day, on the air quality. On the other hand, the ozone concentration in Houston is mostly determined by the wind speed and the ozone concentration of the previous day. Quantitatively speaking, the ozone concentration in Houston is approximatively proportional to the wind speed power 0.6 and to the square root of the ozone concentration of the previous day.

Acknowledgements

NB and HE would like to thank Prof. G. Brassard and Prof. E. Aimer (University of Montreal) for their fruitful comments.

References

[1] G. E. Hinton, R. R. Salakhutdinov Science **5786**, 313 (2006).
[2] D. Hansel and H. Sompolinsky Phys. Rev. Lett. **68**, 718 (1992).
[3] A.T.C. Goh, Artif. Intell. Eng. **9**, 3 (1995).
[4] Mitchell Melanie, An Introduction to Genetic Algorithms, A Bradford Book, MIT Press, Cambridge, 1999.
[5] Chokri El Aoun, Hichem Eleuch, Esma Aïmeur, Hella Ben Ayed, and Farouk Kamoun, Journal of Decision Systems **16**, 393 (2007).
[6] A. Mellit, H. Eleuch, M. Benghanem, C. Elaoun, A. Massi Pavan Energy Conversion and Management **51**, 771 (2010).
[7] M. Christian, Atmospheric Environment **35**, 1 (2001).
[8] K. Atakan , A. B. Oktay, Expert Systems with Applications **37**, 7986 (2010).
[9] W. G. Cobourn, Atmospheric Environment **44**, 3015 (2010).
[10] M. Kolehmainen, H. Martikainen and J. Ruuskanen, Atmospheric Environment **35**, 5 (2001).
[11] R. Beelena, M. Voogtb, J. Duyzerb, P. Zandveldb and G. Hoeka, Atmospheric Environment **44**, 4614 (2010).
[12] K. E. Kakosimos, O. Hertel, M. Ketzel and R. Berkowicz, Environmental Chemistry **7**, 485 (2010).
[13] T. Brown, P. Holmes, P. T. C. Harrison, Indoor Built Environ. **19**, 311 (2010).
[14] Bikas K Chakrabarti, Anirban Chakraborti, Satya R Chakravarty, Arnab Chatterjee, Econophysics of Income & Wealth Distributions. Cambridge University Press, Cambridge 2012.
[15] E. W. Piotrowski, M. Schroeder and A. Zambrzycka, Physica A **368**, 176 (2006).
[16] B. Baaquie, Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates, Cambridge University Press 2004.
[17] Markus Haas, Stefan Mittnik and Marc S. Paolella, A New Approach to Markov-Switching GARCH Models, Journal of Financial Econometrics, **2**, 493-530 (2004).
[18] J. Lindner, Binary sequences up to length 40 with best possible autocorrelation function, Electronics Letters **16**, 507 (1975).
[19] H. Eleuch, N. Ben Nessib and R. Bennaceur, Eur. Phys. J. D **29**, 391 (2004).
[20] N. Kylafis, D. Giannios and D. Psaltis, Advances in Space Research **38**, 2810 (2006).
[21] H. Jabri, H. Eleuch and T. Djerad, Laser Phys. Lett. **2**, 253 (2005).
[22] E. Sete et al. , J. Mod. Phys. **57**, 1311 (2010).
[23] H. Mori, Progress of Theoretical Physics **33**, 423 (1965).
[24] H. Eleuch, Eur. Phys. J. D **48**, 139 (2008).
[25] E. A. Sete and H. Eleuch, Phys. Rev. A **82**, 043810 (2010).
[26] H. Eleuch and N. Rachid, Eur. Phys. J. D **57**, 259 (2010).
[27] E. Giacobino et al., Quantum, C. R. Physique **3**, 41 (2002).
[28] L. Hosun , Sukyung Kim and Sungkwon Park, IEICE Transactions on Communications **E89-B**, 1423 (2006). (2004).