CHAOTIC ANALYSIS FOR THE ATMOSPHERIC AIR POLLUTANTS
USING R STATISTICAL SOFTWARE

Tamil Selvi. S ,
Research Scholar, Bharathiar University, Coimbatore
KCG College of Technology, Chennai, Tamil Nadu, India.
E-mail: tssekar9773@gmail.com

Abstract
Exposure to air pollution is an unavoidable part of our urban life. In this study, the interaction patterns of air pollutants, SO2, NO2, CO and NO are investigated for the study area in Manali, near Chennai. This is surrounded with high industrial corridors and residential area, which operates with a large number of heavy vehicles. It is necessary to investigate this area, since the chances for the sources of these pollutants are from vehicles and petrochemical industries. We have calculated the Lyapunov exponent for the gaseous pollutants and the values obtained for the largest Lyapunov exponent turned out to be positive and ranging between 0.52 and 1.44 which according to the theory of chaos, is a condition for the presence of deterministic chaos and random behavior in time series. The values of correlation dimension and Hurst exponent, shows that random behavior can be discarded. We therefore conclude that the time series for the air pollutants are chaotic and very sensitive to initial conditions. The study presented here can be replicated in other cities that represent similar situations where a reliable air quality forecasting model may be important for environmental management.

Keywords: air pollutants, Lyapunov exponent, deterministic chaos, Hurst exponent, random behavior

INTRODUCTION

Environmental pollution is a serious and growing issue in both industrialized and developing country in the world. The primary cause of air pollution in the study area of Manali is industrial operations. Air pollution has been becoming a necessary evil with rapid industrialization and urbanization around the world, after it results in kinds of human health problems, such as ophthalmic, respiratory and cardiovascular diseases [Brunekreef and Holgate 2002, Giles et.al; 2011, Gudmundsson 2011, Jamrozik and Musk 2011, Miller 2007, Nandasena 2010]. Air pollutants can be divided into anthropogenic and natural pollutants according to their sources, or primary and secondary pollutants, which stem from reactions of primary pollutants, when taking production process into account [UNEP 2004].

This study focuses only on air pollution or the air quality in Manali, Chennai so that an effective policy can be chalked out for a holistic growth of the country without damaging the environment. The different sources of air pollution are classified under following categories transport, industries, residential and others. Manali, which is in northern part of Chennai mainly an industrial area comprising of petro chemical industries. Manali area is a home to petroleum refinery (17 petro chemical industries) and other industries.
The air pollutant comes out in the form of smoke, particulate matter, hydrocarbons, oxides of sulphur and nitrogen, carbon monoxide and VOC. These are the predominant sources of air pollution as shown in Table 1. Petroleum industry is the industry, which unites anthropogenic activities to explore for, produce, transport worldwide, and process around 3.5 billion tons of crude oil and 2.5 Giga m$^3$ of natural gas and their derivatives every year. One of the most valid reasons for the air pollution is that, more than 2500 refined products are extracted including liquefied petroleum gas, gasoline, kerosene, aviation fuel, diesel fuel, fuel oils, lubricants and feed socks for the petrochemical industry are the direct results of these activities. Bigger refineries integrated with petrochemical plants may produce additionally different kinds of synthetic derivatives from pure chemicals to additives for fuels and lubricants, synthetic polymers, elastomers etc.

Chemicals in refinery atmosphere act on human in three ways such as 1) local action as dermatitis or absorption through skin 2) direct inhalation and 3) indigestion or absorption into the stomach. Toxic effects are produced by prolonged contact with airborne, solid or liquid chemical compounds even in small quantities because of their properties such as carcinogenicity, mutagenecity and corrosiveness. [Vijayan G. Iyer 2007]

<table>
<thead>
<tr>
<th>POLLUTANT</th>
<th>SOURCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO$_2$</td>
<td>The progressing area of the petroleum products and desulphurization plant and from the captive power plan</td>
</tr>
<tr>
<td>NO$_x$</td>
<td>Boilers using coal, biomass and husks as fuel. Refining process of petroleum</td>
</tr>
<tr>
<td>CO</td>
<td>Boilers</td>
</tr>
</tbody>
</table>

Table 1. The sources of gaseous air pollutants in a Petro chemical industry

Carbon monoxide (CO), one of tasteless, colorless and non-odorous gases, is produced primarily by car exhaust and incomplete combustion of organic materials. Hence, CO can be easily inhaled and readily combined with the hemoglobin in the blood, which prevents oxygen (O$_2$) transportation in the blood and causes intoxication effects such as headache, giddiness, tiredness and buzzing in the ears etc. Sulfur dioxide (SO$_2$) is one of the precursors for acid rain because not only a sulfurous acid (H$_2$SO$_3$) reaction is formed when SO$_2$ is compounded as mist in the atmosphere, but also SO$_2$ has a chemical reaction with O$_2$ and forms sulfur trioxide (SO$_3$). SO$_3$ then becomes liquid SO$_3$ because its’ dew point is 22 °C. At the dew point a sulfuric acid (H$_2$SO$_4$) reaction is formed. When inhaling SO$_2$ at about 30 to 40 ppm, the body will feel giddy and experience breathing difficulty [Aplin et.al. 2000, Lin 1992]. Those suffering from respiratory diseases, such as asthma are particularly at high risk of being affected by oxides of nitrogen [WHO 2000]. Anthropogenic SO$_2$ and NO$_x$ emissions are chemically converted to sulphuric and nitric acids in the atmosphere in the gaseous (primarily via reaction with OH) and aqueous phases. These acids are responsible for acid rain, stone leprosy and decrease in the soil alkalinity [Hewitt 2000, Wright, Schindler 1995, Munn et.al. 2000]. As per 1995, Chennai has the highest per unit area NO$_x$ emissions (170t/km$^2$) [Garg 2001]. Nitrogen oxide (NO$_x$) is so-called “thermal induced compound” because it has two reactions during combustion. The first reaction is from oxidation in which a nitric (N) substance after combustion is formed with the numerator of O$_2$. The second reaction involves the reaction of O$_2$ and N$_2$ from the air during the combustion process [Li et.al. 1993, Takabashi et.al. 1995]. Time-series analysis helps in obtaining the air quality predictions that are generally used for issuing the warnings to public.
Recent studies have shown the nonlinear and irregular behavior of air pollutant concentrations [Raga and LeMoyne 1996, Li et. al. 1994, Chen et. al. 1998, Kocak et. al. 2000] Because of the complex and nonlinear nature of air pollutants, an approach is required that considers the nonlinearity among the data. The approach based on nonlinear dynamical systems theory or chaos theory captures the complex and nonlinear behavior of the time-series that, in turn, can be used for prediction purposes. The technique investigates the deterministic, chaotic, or random behavior of the apparently irregular time-series [Abarbanel et. al. 1993, Grassberger et. al. 1991]. The gaseous air pollutant time series for $\text{SO}_2$, $\text{NO}_2$, CO and NO are shown in Figure 1.
Figure 1. The gaseous air pollutant time series for SO$_2$, NO$_2$, CO and NO

In the decade of the 80’s, the analysis of time series through dynamical systems theory [Takens and Floris 1981] has established that relevant information may be obtained by the calculation of a set of parameters as time lag, mutual information [Fraser 1986], embedding dimension, correlation dimension, Hurst exponent and Lyapunov exponents, among others are important [Grassberger 1983]. The largest Lyapunov exponent is one of the most widely used indicator to detect chaotic behavior in a dynamical system consisting of a time series [Eckmann et.al. 1986]. The significantly positive Lyapunov exponents obtained are an indication that the evolution of concentrations is very sensitive to initial conditions and the estimation of correlation dimension and Hurst exponent indicate that they are not random but chaotic.

In the present study, the problem of air pollution was, therefore, approached by using the concepts from nonlinear time-series analysis techniques. The objective is to examine the behavior (deterministic, chaotic, or stochastic) of the NO$_2$, SO$_2$, NO, CO concentration and its predictability of Manali industrial area in Chennai, Tamil Nadu, India.

**AREA DESCRIPTION**

Chennai is one of the metropolitan cities of India. The study area, Manali (13°09N, 80°15E) is an industrial complex and is the most air pollution-sensitive area, situated about 20 km North of Chennai, connected by road. It comprises an area of 16 km$^2$, intersected by villages and is inside the inhabited area. The Manali industrial area of Chennai has witnessed high levels of air pollution in the past few decades. Efforts must be taken to control the air pollution in the industrial corridor’s of Manali since this area is not only covered by industries but also by the residential. The mean maximum temperature during summer is 45°C and the mean minimum temperature during winter is 20°C. The relative humidity is around 70-80%. Manali has an average rainfall of 6 cm. The total population of 28,597 consists of 15,080 males and 13,517 females (census, 2001). There was an increase of 20% decadal variation of population from 1951 to 2001.
The area consists of various industries like oil refineries, fertilizer plants, chemicals, fabric yarn and steel, etc. which have existed for more than two decades. There are 28 categories of industries (major–20, minor–3 and small–5) located in Manali, which include CPCL, MFL, TNPP, Manali Petro Chemical, etc.

KEY AIRBORNE POLLUTANTS
Airborne pollutants mainly consist of natural contaminants, aerosols, gases and vapors. They can be broadly classified into two categories, primary and secondary. Primary air pollutants are emitted from direct sources. They consist of gaseous pollutants, sulphur dioxide (SO2), nitrogen oxides (NOx) and particles, suspended particulate matter (SPM) less than 100 mm in aerometric diameter, respirable suspended particulate matter (RSPM/PM10) less than 10 mm in aerometric diameter. This study covers classical pollutants, viz. SO2, NO2, CO and NO and these are investigated based on the measured database of Central Pollution Control Board (CPCB) from June 1, 2013 to December 17, 2014.

THEORY AND METHODOLOGY
Considerable methods are available in the literature to identify the existence of chaos in a time series. Among these methods, the correlation dimension, the Lyapunov exponent [Wolf et.al. 1988] have been widely employed. To identify the chaotic characteristics of time series, main causes of atmospheric pollution should be analyzed, associated with a timely prediction of atmospheric pollution. Thus following the combination of air quality in Manali, the characteristics of chaotic time series data were analyzed based on the daily air pollution information from the phase space reconstruction, correlation dimension and the maximal Lyapunov exponent. The following steps were performed for nonlinear time series analysis: (i) Hurst exponent (ii) phase space reconstruction; (iii) delay time (s) determination after averaged mutual information estimation; (iv) embedding dimension (m) determination; (v) maximum Lyapunov exponent estimation; and (vi) correlation dimension.

Chaotic Behavior in Air Pollution Time Series
Linear methods for analysis and prediction of time series may be an initial point in the study of behavior of air pollution and meteorological data. However, the most relevant results have been obtained only with the introduction of nonlinear tools. The reason for this is that in this type of data, small causes may have large effects in future values, which is the essence of nonlinearity. According to Lorenz (1963) nonlinearity is inherent to atmospheric systems. Sivakumar et al (2007) points out the existence of nonlinear deterministic behavior in the air quality index in time series such as ozone. Raga and Le Moine (1996) have applied nonlinear dynamic tool on air quality data (NO, CO, SO2 and O3).

Lyapunov exponent
The largest Lyapunov exponents for the gaseous pollutant time series are calculated using the algorithm proposed by Rosenstein et.al. 1993. The Lyapunov exponent illustrates the system's Sensitivity on Initial Conditions (SIC). [Feder 1988]. Positive Lyapunov exponents are considered evidence of chaos, while negative exponents suggest a mean reverting behavior. When the positive exponent is larger, the more chaotic is the system and, conversely, the shorter is the time scale of the system's predictability.
Since there are as many exponents as there are dynamical equations, only the most positive exponent is calculated here. The sum of the positive Lyapunov exponents (base e) is called entropy \( S \); its reciprocal is roughly the time over which meaningful prediction is possible [Sprott and Rowlands 1995]. Lyapunov exponents may be positive, zero or negative. If the maximal Lyapunov exponent is negative, we are in the presence of a time series indicating a dissipative system with a stable fixed point. If the motion settles down onto a limit cycle, the maximal Lyapunov exponent is zero. If a predominantly deterministic system is affected by random noise, the maximal Lyapunov exponent is infinite [Kantz and Schreiber 2005]. A positive finite maximal Lyapunov exponent would be an indication of chaos.

**Maximum predictable time scale**

According to [Liu 2004] the maximum predictable time scale \( T_f \) has a relationship with maximum Lyapunov exponent as [Peitgen 2004]. Much chaotic behavior in the system implies the larger Lyapunov exponent (\( \lambda \)) which causes in the smaller predictable time scale (\( T_p \)).

\[
T_p = \frac{1}{\lambda}
\]

**Correlation Dimension**

In order to use Rosenstein’s algorithm, an estimation of the embedding dimension of the system is necessary. One way to calculate embedding dimension is by analyzing the correlation dimension as a function of embedding dimension. In this study, the presence of chaos has been investigated using correlation dimension method. The correlation dimension is the representation of the variability of the process and furnishes information about the number of dominant variables present in the evolution of the corresponding dynamical system. It can indicate, not only the existence of chaos in the air pollution variability process, but also reveals whether the process is deterministic or stochastic, if not chaotic.

**Phase space reconstruction**

The atmosphere as many other physical, geophysical, biological systems can be described as a mechanical dissipative multi-level system, which are fundamentally nonlinear. Nonlinear systems typically have a long term behavior, which is described by an attractor in phase space. It is well known that an attractor is strange attractor, if its dimension is non-integer which is fractal. The Grassberger - Procaccia algorithm used in the present study uses the concept of phase space reconstruction. For a scalar time series \( t \), where \( t = 1, 2, \ldots, N \), the phase space can be reconstructed using the method of delays.

\[
X(t) = \{x(t), x(t+\tau), \ldots, x[t+(m-1)\tau]\},
\]

where \( t = 1, 2, \ldots, N-(m-1)\tau/\Delta t \), \( m \) is the dimension of the vector \( Y_t \), also called the embedding and \( \tau \) is a delay time taken to be some suitable multiple of the sampling time \( \Delta t \) [Packard et.al. 1980]. The next step involves the calculation of the cumulative correlation function [Henderson and Wells 1988].

\[
C(l) = \frac{1}{N^2} \sum_{i,j=1}^{N} H(l - r_{ij})
\]
Where $r_{ij}$ is the Euclidean distance between the $i^{th}$ and $j^{th}$ points, $N$ is the total number of points, $l$ is the distance variable and $H$ is the Heaviside function with $H(x) = 0$ if $x \leq 0$ and $H(x) = 1$ if $x > 0$. Next we calculate the slope $d(m)$ of the linear part of the $\ln C(l)$ vs $\ln(l)$ curve by fitting a least square line. By repeating the process, for $m = 1, 2, 3, 4, 5…$ we obtain successive estimates of the attractor dimension. If the slope converges to a limiting value $d_\infty = d(M) = d(M+1) =….$, then $d_\infty$ is the true correlation dimension of the attractor and the corresponding embedding dimension $M$ is a measure of the number of variable sufficient to model the dynamics. A non-integer value of $d_\infty$ indicates the presence of a strange attractor, a term coined by Ruelle and Takens (1971) signifying deterministic but chaotic dynamics. It is interesting to note that there will be no saturation of slopes for a purely random time series.

The presence or absence of chaos can be understood using the correlation exponent versus the embedding dimension curve. If the correlation exponent saturates and the value is very low, the system consists of a low dimensional chaos. The saturation value of the correlation exponent is defined as the correlation dimension of the attractor. The nearest integer above the saturation value provides the minimum number of variables required to model the dynamics of the attractor. On the other hand, if the correlation exponent increases without bound with an increase in the embedding dimension, the system is generally considered as stochastic. The correlation functions and the exponents are computed for the four air pollution time series data. The correlation dimension for $SO_2$, $NO_2$, CO and NO are 3.05, 2.605, 4.326, 0.13 respectively.

![Figure 2 (a) Correlation dimension for time series NO$_2$](image)

![Figure 2 (b) Correlation dimension for time series SO$_2$](image)
Hurst exponent

The Hurst exponent is a measure of the degree to which the data can be represented by a random walk. Here, the root-mean-square displacement is calculated as a function of time, where each point in the time series is taken as an initial condition. The slope of the curve obtained is the Hurst exponent as shown in Fig.3. For ordinary Brownian motion the exponent is 0.5. Values between 0.5 and 1.0 indicate deterministic behavior and persistence. Hurst exponent significantly less than 0.5 are typical for deterministic behavior and anti-persistence [35].
Figure 3. Hurst exponent for the gaseous pollutants.
DISCUSSION
The evidence that the time series of atmospheric concentrations of SO2, NO2, CO and NO from Manali, Chennai, exhibit deterministic chaotic behavior imposes severe restrictions about the predictability of air quality in the city. Although a positive finite maximal Lyapunov exponent is not by itself an unambiguous indication of chaos. In our case, the results of calculation of additional statistical parameters confirm this behavior, which means that the series are deterministic and very sensitive to initial conditions as shown in Table 2. The Hurst exponent results prove the persistence nature of SO2 and CO and anti-persistence behavior for NO2 and NO. The maximal Lyapunov exponent ranges from 0.521 to 1.437 for the gaseous air pollutants, which shows the presence of deterministic chaos and random behavior. The maximum predictability period is only upto 1.9 days. The Correlation dimension values shows that minimum three or four dynamic variables are required to forecast the gaseous air pollutants. The embedding dimension and time lag are also calculated for the corresponding air pollutant concentrations. The Correlation dimension values confirms the presence of low dimensional chaos in the gaseous air pollutants. The Hurst exponent values clearly points out the presence of chaotic behavior and the random behavior is discarded.

<table>
<thead>
<tr>
<th>POLLUTANT</th>
<th>SO2</th>
<th>NO2</th>
<th>CO</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORRELATION DIMENSION</td>
<td>3.05</td>
<td>2.605</td>
<td>4.326</td>
<td>0.13</td>
</tr>
<tr>
<td>EMBEDDING DIMENSION</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>TIME LAG</td>
<td>6</td>
<td>5</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>HURST EXPONENT</td>
<td>0.587</td>
<td>0.325</td>
<td>0.624</td>
<td>0.206</td>
</tr>
<tr>
<td>MAXIMAL LYAPUNOV EXPONENT</td>
<td>1.4377</td>
<td>0.6004</td>
<td>0.70264</td>
<td>0.52102</td>
</tr>
<tr>
<td>MAXIMUM PREDICTION</td>
<td>0.695</td>
<td>1.665</td>
<td>1.423</td>
<td>1.919</td>
</tr>
<tr>
<td>(Days)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Summary for the values of Correlation dimension, Lyapunov exponent, Hurst exponent and maximum prediction days for the gaseous air pollutants SO2, NO2, CO and NO
CONCLUSION

This is a study about higher air pollutant concentrations, which give rise to significant adverse health effects, and it is an issue of serious public health threat to inhabitants of the study area. The concentration of air pollutants, are increasing due to industrial activities and alarming increase in the number of vehicles may become a problem in the future, with greater health impact. The study concludes that the seriousness of the air pollution problem in Manali may aggravate further, if not brought under control. In this scenario, the city development and environmental enforcement authorities have to adopt future strategies to combat the menace of air pollution.

The maximum predicted time scale for the system is the reciprocal of maximum Lyapunov index, which means the maximum time scale that the system can forecast is up to 2 days. Because of the sensitivity of chaotic systems to the initial value, long-term prediction of chaotic system is impossible. The correlation dimension (attractor dimension) is calculated by G-P algorithm. The result shows the low correlation dimensions, which indicates that the air pollutant time series have chaotic characteristics. The dynamic system necessitates at least three or even four dynamic variables so as to explain properly and effectively the time series changes in Manali, Chennai. This work could fully explain the temporal evolution law of atmospheric pollutants and find out the main causes of air pollution in Manali. The situation of pollution concentration would change with significant variation in the weather. This shows that, for a certain period, the pollutant concentration is mainly affected by weather conditions. But as the atmospheric pollution system is a complex, it is not enough solely to prove the presence of strong nonlinear and chaotic characteristics in time series. More efforts are needed to reveal the inherent dynamic mechanism of air pollution. The criticality of the air pollutants and the temporal variation of air pollutant indexes should be further studied and improved. All these suggestions are the mainly technical procedures for developing a more appropriate air pollution control programs, which will benefit the facilitation of atmosphere and the mankind.

REFERENCES

- Aplin, G; Beggs, P; Brierley, G; Cleugh, H; Curson, P; Mitchell, P; Pitman, A; and Rich, D., 2000, Global Environmental crises, second edition, Oxford, ISBN 0-19-550827-0, Australia, 30-60
• Garg A., Shukla P.R., Bhattacharya S., Dadhwal V.K., 2001, “Sub-region (district) and sector level SO2 and NOx emissions for India: assessment of inventories and mitigation flexibility”, Atmos. Environ. 35 , 703–713.
• Giles, L.V; Barn, P.K.N; Romieu, I; Mittleman, M.A; Eeden, S; Allen, R; Carlsten, C; Stieb, D; Noonan, C; Smargiassi, A; Kaufman, J.D; Hajat, S.K.T & Brauer, M., 2011, “From Good Intentions to proven Interventions; Effectiveness of Actions to reduce the health impacts of Air pollution”, Environmental Health Perspectives, 119, 1, 29-36, ISSN 0091-6765.
• Jamrozik, E & Musk, A.W., 2011,” Respiratory Health Issues in the Asia- Pacific region: An overview”, Respirology, 16, 1, 3-12, ISSN 1440-1843.
• Kantz H and Schreiber T; 2005, Nonlinear time series analysis, 2nd ed. Cambridge, UK: Cambridge University Press.
• Li, Y.D; Yan, W.Y; and Juang, T.S, 1993, “Property of Polimer Substance”, KOU-LIH Book Corp. Taipei, 325-328.