



## Prediction of Real SO<sub>2</sub> Time Series through Chaotic Approach in Kota Kinabalu, Sabah

AHMAD BASRI RUSLAN<sup>1</sup>, AHMAD HABIL AHMAD ZUBIR<sup>1</sup>, FRANKLIN CHAI JUN<sup>1</sup>, MOHAMAD DARIUS ELMAN BIN MOHAMAD ADNAN<sup>1</sup>, MOHAMAD LUQMANUL HAKEEM BIN MOHD FAISAL<sup>1</sup>, ANGELA RINA FRANCIS<sup>1</sup>, NOR ZILA ABD HAMID<sup>2</sup>, KHAIRUNNISA CHE JUSOH<sup>3</sup>

<sup>1</sup>*Akademi Sains Pendang, Kedah, Malaysia,*

<sup>2</sup>*Department of Mathematics, Faculty of Science and Mathematics, Sultan Idris Education University, Perak, Malaysia,*

<sup>3</sup>*SMK Seksyen 10 Kota Damansara, Selangor, Malaysia*

Email: ahmadbasiruslan@gmail.com | Tel: +60193325687

Received: November 28, 2023

Accepted: December 03, 2023

Online Published: December 04, 2023

### Abstract

When there is an excess of carbon, sulphur and any dangerous gaseous emission, then the air is polluted. Breathing through polluted air can cause harm and may lead to several diseases such as respiratory failure and cancer. This study is about the prediction of real SO<sub>2</sub> time series through the chaotic approach in Kota Kinabalu, Sabah. Before the prediction model can be built, the existence of chaotic behavior needs to be confirmed. Chaotic behavior can be detected through the phase space plot and Cao method. The observed time series is confirmed to be chaotic. Prediction through an improved local linear approximation method (ILLAM) is done. Three parameters are determined which are delay time  $\tau$ , embedded dimension  $m$  and the number of nearest neighbor  $k$ .  $\tau$  is set to one and  $m$  is determined through the Cao method. Previously, parameter  $k$  is chosen through trial and error. In this study, the improved method is applied to determine the value for  $k$ . The improved method is done by plotting the graph of  $k$  against  $cc$ . The  $k$  value that gives the maximum  $cc$  is recorded. The prediction result shows that at  $k = 68$ , the performance of the prediction model is  $cc = 0.9304$ . With the combination of these three parameters, the prediction shows 93.04% well correlated with the original time series. The improved method in determining the value of  $k$  can be continued in predicting another time series in the future.

**Keywords:** Sulphur dioxide, linear approximation method, chaotic approach, number of the nearest neighbor.

### 1.0 Introduction

Sulfur dioxide (SO<sub>2</sub>) exist in the atmosphere are from the combustion of coal, vehicle exhaust and industrial area. SO<sub>2</sub> is colorless and gives an unpleasant feeling if smells it. SO<sub>2</sub> dissolves in water and turns the water into an acidic state which is sulfuric acid (H<sub>2</sub>SO<sub>4</sub>). It is corrosive to metals and tissue. The emission of SO<sub>2</sub> in the hotspot area is needed to be lessened. The hotspot area is defined as where the pollutions occur surely will expose danger to the citizen and may lead to the degradation of health quality (CARB, 2019). It is unavoidable in a highly-populated area and industrial area due to the economic factors and for living. These communities are probably exposed to the risk of getting cancer and respiratory failure (Dury et al., 1999). Therefore, an SO<sub>2</sub> prediction model is needed to help local authorities to create a proper plan in the future accordingly about the air quality. The development of a prediction model on air pollutant time series through the chaotic approach is increasing. These are the examples of air pollutant time series through the chaotic approach that has been successfully run such as ozone (O<sub>3</sub>) (Hamid & Noorani, 2013) (Hamid, Noorani & Adenan, 2017), particulate matter below 10 micrometers or less in diameter (PM<sub>10</sub>) (Hamid & Noorani, 2014) and carbon monoxide (CO) (Ruslan & Hamid, 2019). However, it is not only limited to air pollutants where (Adenan & Noorani, 2015) also successfully apply the chaotic approach to river flow time series and (Hamid, 2018) (Bahari & Hamid, 2019) applied to temperature time series.



## 2.0 Problem Statement

EPA's national ambient air quality standards for SO<sub>2</sub> are designed to guard against exposure to the entire group of sulfur oxides (SO<sub>x</sub>). SO<sub>2</sub> is the component of greatest concern and is used as the display for the larger group of gaseous sulfur oxides (SO<sub>x</sub>). Other gaseous SO<sub>x</sub> (such as SO<sub>3</sub>) are found in the atmosphere at concentrations much lower than SO<sub>2</sub>. Control measures that reduce SO<sub>2</sub> can generally be expected to reduce people's exposures to all gaseous SO<sub>x</sub>. This may have the important co-benefit of lowering the formation of particulate sulfur pollutants, such as fine sulfate particles. Emissions that lead to high concentrations of SO<sub>2</sub> generally also lead to the structure of other SO<sub>x</sub>. The largest sources of SO<sub>2</sub> emissions are from fossil fuel combustion at power plants and other industrial services. Therefore, the prediction of SO<sub>2</sub> substances really needs to be done as it corresponds to the harmful pollution index.

## 3.0 Research Objective

The objectives of this study are to i) identify the presence of chaotic dynamics and ii) build a prediction model of SO<sub>2</sub> time series in Kota Kinabalu, Sabah. The prediction will be done by using a prediction method which is the improved local linear approximation method (ILLAM) introduced by (Hamid & Noorani, 2013). In this study, the modeling through a chaotic approach is divided into two stages which are i) reconstruction of phase space and ii) prediction process. In the first stage, it is important to determine the existence of chaotic behavior in the SO<sub>2</sub> time series. The study can be continued if and only if the time series is chaotic. If it is not, this study cannot be continued onto the prediction stage. MATLAB R2010 software and TSTOOL toolbox are involved in solving every related mathematical equation in this study.

## 4.0 Literature Review

The free radical of SO<sub>2</sub> gas in the atmosphere is toxic (USEPA, 2016). Excessive inhalation of SO<sub>2</sub> can be dangerous to health and can cause death if inhaled over a long period of time (Kobayashi et al., 2020). According to (Kumar, 2017), SO<sub>2</sub> gas is free into the air and mixed with rainwater will be oxidized in turn undergo a chemical reaction that eventually into sulfuric acid. This sulfuric acid is harmful to health as well impacting the environment. Free SO<sub>2</sub> in the atmosphere is toxic, contributing to air quality which is bad in cities and damages human health and quality of life. It is invisible to the naked eye and has a rancid odor (De et al., 2017). It easily reacts with other materials to form harmful compounds such as sulfuric acid (Wang et al., 2019). SO<sub>2</sub> affects human health when humans breathe and sneeze. It makes discomfort exist in the nose, throat, ducts breathing causes coughing, shortness of breath and chest pain. The government of Australia in environmental conservation clusters expressed the effects of inhaling this SO<sub>2</sub> can be felt very quickly and most will feel severe symptoms within 10 or 15 minutes after breathing (Australian Department of Environment and Heritage, 2015). High-risk groups such as those with asthma or problems breathing will make worse if exposed to SO<sub>2</sub>. Therefore, a study of the SO<sub>2</sub> time series needs to be done.

These environmental problems are influenced by urbanization factors, in addition, to economic and service factors (Shafie et al., 2018). A study by Mabahwi et al. (2015) defines air pollution as a state of deterioration of physical or chemical changes through natural processes or human activities such as smoke from motor vehicles, domestic activities as well as industrial activities. Shafie et al. (2018) stated three sources of air quality deterioration problems in Malaysia namely through moving sources, stationary sources and open combustion. Source moving is from a motor vehicle and immobile source is from industrial activity. When an area has a population distribution high, daily movements necessarily use a motor vehicle emit high amounts of pollutants. High population distribution area such as Shah Alam also has many factories. According to statistics current from Google Maps (2020), there are more than 200 factories built in Shah Alam. Smoke from factories and motor vehicles is also present contributes to air pollution in areas where population distribution is high. However, there are other factors that also contribute to air pollution such as urbanization factors (Kalhoor and Mahdisoltani, 2015; Rai, 2017). Urbanization is rapid due to demographic factors i.e., population growth natural and inward migration of the population to the city (Shafie and Mahmud, 2015). Urbanization is defined as an increase in the maximum population living in urban areas thus changing the physical environment i.e., from natural habitats surrounding to human habitat (Malik et al., 2018).

For Malaysia, urbanization is said to be getting more serious so far contributing to the improvement of the living environment of the community through the provision of infrastructure and services such as housing, transportation, user-friendly communication and services (Hassan et. al, 2013). According to Embong (2011), the history of urbanization in Malaysia shows an increase rapid from 55.1% in 1995 and increased sharply to 70% in 2008 with a growth rate of 2.2%. Therefore, the acceleration and the progress of the urbanization process in Malaysia are clearly increasing the increasing population and affecting the deterioration of air quality. Statistics from the Department of



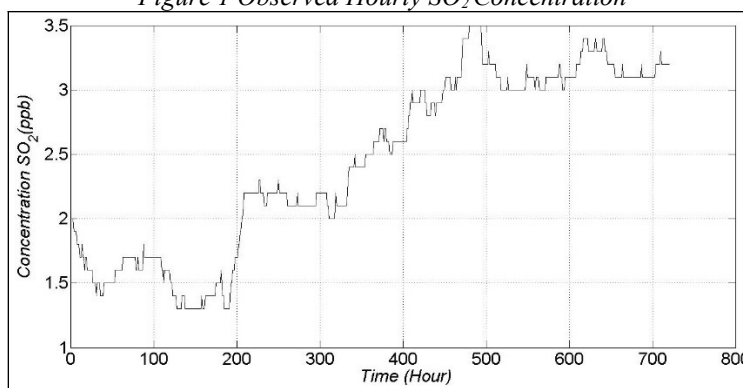
Statistics Malaysia (JPM) which conducted the census in 2010 showed that the total population in Malaysia was led by three main states are Selangor (5.45 million) followed by Johor (3.35 million) and Sabah (3.21 million). A high population will directly increase and intensify daily activities in each of their respective places. Shafie et al. (2018) in their study concluded that urban air pollution that occurs in The Klang Valley was detected due to population factors and population density high apart from economic and service factors. This situation is proven with annual PM10 readings in the Klang Valley higher than in the Klang has a small population. Hence, the speed and activity of the process of urbanization in Malaysia significantly affect the decline of air quality in the city. Pollution in the city especially in the distribution area's high population remains an issue that should be emphasized. Effects of air pollution this is not only on human health but also capable of having a negative impact on surrounding buildings and plants (Azam et al., 2016). In most urban areas with a high population, densities are vulnerable to pollutants due to emissions from stationary and moving sources (Sulaiman, 2015). Studies in 2019 showed that pollution killed 8.3 million humans every year (McCarthy, 2019). India and China each recorded deaths due to pollution 2, 326 711 and 1, 865 566. These deaths are the cause from air, water, soil and chemical pollution (Fuller, Sandilaya & Hanrahan, 2019). Deaths due to air pollution are as many as 1, 242 987 (China) and 1, 240 529 (India). Factors causing China and India to be the top two countries that contribute to air pollution are because of economic activity and population size (Myllyvirta, 2020). Therefore, the study of highly populated areas with high populations is done.

**5.0 Data And Methods**

**5.1 Data**

SO<sub>2</sub> time series is observed in Kota Kinabalu, Sabah. It is the largest city in Sabah. Kota Kinabalu is the state capital city, the economic center of the state and the seat of the Sabah state government. SO<sub>2</sub> exists in the air composition. It is a natural component of air. It is formed from the combustion of trees, brush and other organic matters. It is also released in large amounts by human activities (Press, 2020). Small amounts of sulfur as impurities exist in all fossil fuels such as coal, oil and natural gas. When those fuels burn, sulfur is converted to sulfur dioxide through chemical reactions. This finally leads as a component of air pollution when being released in high amounts. Figure 1 shows the observed time series in Kota Kinabalu on Jun 2014 recorded hourly. In this study, the observed data for a total is 720 hours for a month. The time series is separated into two parts which are the training part and the test part. The training part is 552 hours (3 weeks) and the test part is 168 (1 week). The training part is being used for analyzing the chaotic behavior and to determine the parameter. While the test part, is used to see the prediction performance. Table 1 shows the descriptive statistics of the data.

*Figure 1 Observed Hourly SO<sub>2</sub> Concentration*



**Table 1** Statistics Descriptions of SO<sub>2</sub> Time Series

Statistics	Average	Median	Mode	Std. Deviation	Maximum	Minimum
Value	2.442	2.5	3.1	0.6867	3.5	1.3

**5.2 Chaotic Approach**

Prediction through chaotic approach is done by two steps were i) reconstruction of phase space and ii) prediction. At the first step, the presence of chaotic behavior needs to be confirmed exists in the selected time series. Chaotic



behavior detection is done through the phase space plot and Cao method. After the confirmation of the existence of chaotic behavior, research can be continued with the prediction step. Meanwhile, for the prediction model, the improved local linear approximation method (ILLAM) will be used. ILLAM gives the best prediction result compared to the autoregressive linear and local linear approximation method (LLAM) in predicting the O<sub>3</sub> time series (Hamid & Noorani, 2013). Recent studies by (Ruslan & Hamid, 2019) on CO time series show that ILLAM gives the best prediction result compared to the local mean approximation method (LMAM) and LLAM. Hence, ILLAM is chosen for predicting SO<sub>2</sub> time series.

### 5.3 Phase Space Plot

The recorded data is written in the scalar form of one-dimensional vector  $X$ :

$$X = \{x_1, x_2, x_3, \dots, x_N\} \quad (1)$$

with  $N$  is the total data. As explained earlier, the data is divided into two parts which are  $X_{train}$  and  $X_{test}$ . Both are arranged in the form of:

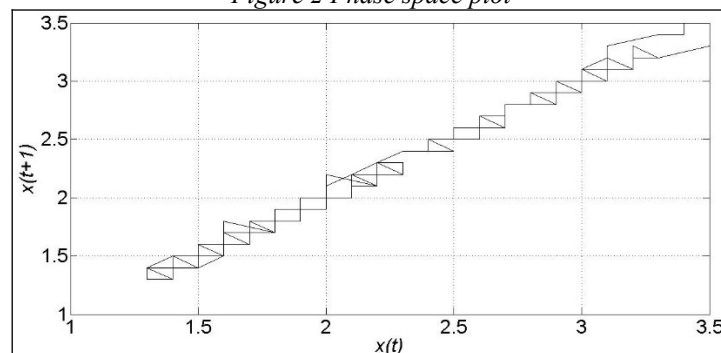
$$X_{train} = \{x_1, x_2, x_3, \dots, x_{552}\} \quad (2)$$

$$X_{test} = \{x_{553}, x_{554}, x_{555}, \dots, x_{720}\} \quad (3)$$

$X_{train}$  is used to find the parameters in reconstructing the phase space.  $X_{test}$  is saved for calculating the prediction performance. With  $x_t$  is the SO<sub>2</sub> concentration time series, a graph in two-dimensions are plotted in the plane of  $\{x_t, x_{t+\tau}\}$ . Before proceeding with plotting the graph, the  $\tau$  value need to be determined. In this study, the value for  $\tau$  is set to one (1). Previous studies show that prediction with  $\tau = 1$  gives an excellent prediction performance (Sivakumar, 2002), (Sivakumar, 2003), (Jayawerdana, 1997).

Figure 2 shows the phase space plot for the SO<sub>2</sub> time series. The accumulated form is called as a strange attractor. The existence of a strange attractor shows that the observed time series is chaotic (Ruelle & Takens, 1971). To support the existence of chaotic behavior, Cao method is added to strengthen this study.

Figure 2 Phase space plot



### 5.4 Cao Method

There are several methods to determine the value of  $m$ . False Nearest Neighbor (FNN) and embedded correlation were used to determine the value for  $m$ . However, these methods need more parameters and need to be set for another threshold value. Cao method is not only limited to compute the parameter for  $m$ . It is also used to detect the presence of chaotic behavior of the observed time series. Cao method was introduced by (Cao, 1997) is a practical way to find the value of  $m$ . The Cao method is chosen because it does not contain any subjective parameter except  $\tau$  and it does not depend on the number of data in the observed time series. Hence, the Cao method is chosen in this study. The  $m$  from Cao method is computed through:

$$E1(m) = \frac{E(m+1)}{E(m)} \quad (4)$$

and



$$E(m) = \frac{1}{N - m\tau} \sum_{n=1}^{N - m\tau} \frac{\|Y_n^{m+1} - Y_{jj}^{m+1}\|}{\|Y_n^m - Y_{jj}^m\|}$$

(5)

by referring to  $\|\cdot\|$  is the maximum norm.  $Y_{jj}^m$  is the nearest neighbor to  $Y_n^m$ . Graph  $m$  against  $E1(m)$  is plotted. If  $E1(m)$  stops changing when the value of  $m$  is higher than  $m_0$ , then  $m_0 + 1$  is the minimum embedded value needed which is  $m$ . In Figure 3, the value of  $E1(m)$  stops changing from 0.95. Therefore, the  $m_0 = 4$  and the value of  $m = 5$ . Cao method is not only limited to find the value of  $m$ . (Cao, 1997) also introduced parameter  $E2(m)$  in his study. The calculation for  $E2(m)$  as follow:

$$E2(m) = \frac{E*(m+1)}{E*(m)}$$

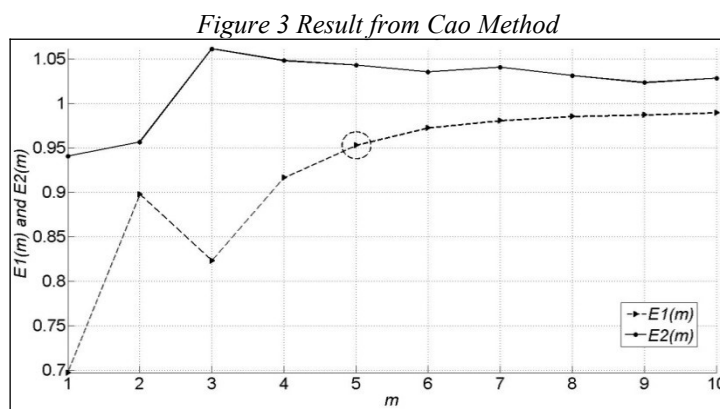
(6)

and

$$E*(m) = \frac{1}{N - m\tau} \sum_{n=1}^{N - m\tau} |x_{n+m\tau}^m - x_{jj+m\tau}^m|$$

(7)

Cao stated that if there is at least one value for  $E2(m) \neq 1$  then the observed time series is chaotic. By referring to Figure 3, all the value for  $E2(m) \neq 1$ . However, for different time series, the result for  $E2(m) = 1$  might be computed as well. For this time series, based on parameter  $E2(m)$  the observed time series is chaotic.



### 5.5 Prediction

Before the prediction can be done, three parameters need to be determined which are  $\tau$ ,  $m$  and  $k$ . The value of delay time is set as  $\tau = 1$ . While for  $m = 5$ , it is determined from Cao method. Parameter  $k$  previously is done by try and error method.  $k = 50$  (Hamid, Noorani & Adenan, 2017),  $k = 100$  (Hamid et al., 2013) and  $k = 200$  (Hamid & Noorani, 2013) were used from try-and-error method. Previous study by (Zaim, 2019) varies the value of  $k = 5, 10, 20, 30, 50, 100, 200, 1000$  and  $2000$ . After varying the  $k$  value, the study found that  $10 \leq k \leq 100$  is suggested for prediction through chaotic approach. However, (Hamid & Noorani, 2013) believes that there is an optimal value of  $k$  for making the prediction better. Latest studies by (Ruslan & Hamid, 2019) used an improved method to determine the value for  $k$ . Parameter  $k$  is determined through plotting the graph of  $k$  against  $CC$ . The value of  $k$  tested is ranged between  $1 \leq k \leq 200$ . The value of  $k$  that gives maximum value of  $CC$  is recorded. The result from (Ruslan & Hamid, 2019) shows that prediction through the improved method in determining the value of  $k$



shows a better prediction performance compared to trial-and-error method. Hence, this study will use the improved method in determining the value of  $k$  for predicting the  $SO_2$  time series.

Prediction through chaotic approach is computed through equation:

$$Y_{j+1}^m = f(Y_j^m) \tag{8}$$

$Y_j^m$  is the final reconstructed phase space and  $Y_{j+1}^m$  is a one-step ahead phase space. To forecast  $Y_{j+1}$ , the nearest neighbor(s) to  $Y_j$  are used. Euclidian distance between  $Y_t$  and the vectors before  $Y_j (j=1,2,3, \dots, t-1)$  is calculated. Take  $Y_p$  and  $Y_{p+1}$  as example to satisfy the linear equation  $Y_{p+1} = AY_p + B$ . Parameters  $A$  and  $B$  are calculated from the least square method. Prediction through the local linear approximation method (LLAM) is done through the equation in (9). The equation for LLAM is as follow:

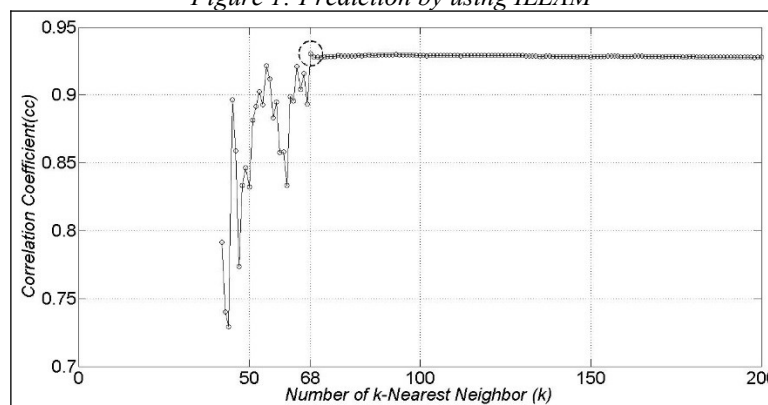
$$Y_{j+1}^m = AY_j^m + B \tag{9}$$

In this study, prediction is done through Improved Local Linear Approximation Method (ILLAM) introduced by (Hamid & Noorani, 2013). For LLAM model, parameter  $A$  and  $B$  are calculated using a series of  $X_{test}$ . Through ILLAM, for every new prediction,  $X_{test}$  is updated. In addition, LLAM only build one linear equation. But, for ILLAM if the number of  $X_{test}$  is  $n$ , then,  $n$  linear equation will be computed. The study by (Hamid & Noorani, 2013) was about prediction on the  $O_3$  concentration time series. The forecasted time series data is arranged in the form of  $X_{predict} = \{x_{553}, x_{554}, x_{555}, \dots, x_{720}\}$ . Then, the performance of the model is measured through correlation coefficient between  $X_{test}$  and  $X_{predict}$ . From the previous studies by (Hamid & Noorani, 2013) and (Ruslan & Hamid, 2019), the prediction through ILLAM shows a better performance compared to LLAM. Therefore, this study will proceed with predicting through ILLAM. The new equation for ILLAM is as follow:

$$Y_{j+1}^m = C_n Y_j^m + D_n \tag{10}$$

Figure 4 shows the result of prediction with ILLAM by plotting the graph of  $k$  against  $CC$ . The best value of nearest neighbor is at  $k=68$  that gives the maximum  $CC=0.9304$ . It is clearly can be seen for number of  $k$  -nearest neighbor that ranged between  $1 \leq k \leq 41$  cannot solve the equation (10). The result showed in command window on MATLAB software as Not a Number (*NaN*). Let  $k=2$ ,  $x_{jk} = \{5,5\}$  and the time series  $x_{jk^1} = \{5,9\}$  the value for the least square method cannot solve to find the value for  $C$  and  $D$ . Hence, the prediction cannot be done within that interval of  $k$ .

Figure 1: Prediction by using ILLAM



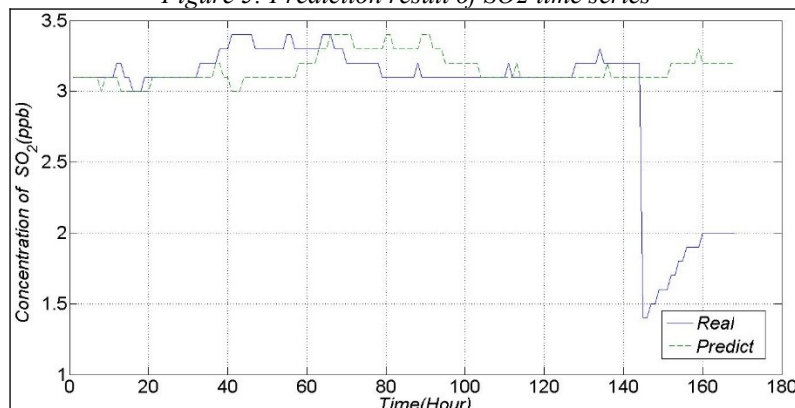
## 6.0 Results And Discussion

Based on phase space plot and Cao method ( $E1(m)$  and  $E2(m)$ ) the observed  $SO_2$  time series is chaotic. Three parameters are set which are  $i=1$ ,  $m=5$  and  $k=68$ . The result of prediction by using this combination is 0.9304. By



referring to Figure 4, the improve method in determining the number  $k$  shows a better prediction result. In addition, the graph plot of  $k$  against  $CC$  is not only shows the best number of  $k$  that gives maximum  $CC$ . It also helps to avoid from choosing a small number of  $k$  that maybe cannot solve the equation (10) and leads to  $NaN$ . Figure 5 shows the prediction result of  $SO_2$  time series.

Figure 5: Prediction result of  $SO_2$  time series



## 7.0 Conclusion

This is a pioneer study of the  $SO_2$  real-time series in Malaysia by using a chaotic approach. At the same time, this is also the first study that applied an improved method in determining the number of nearest neighbor. The prediction result shows that, chaotic behavior exists in this time series and the prediction results show a good performance. In the future, hoping that there will be an improvement in predicting another time series by using an optimal combination for all the three parameters for the best prediction performance. In addition, the chaotic approach can be extended beyond than air pollutions time series such as for electrocardiogram time series, nitrogen dioxide time series and many more.

## 8.0 Acknowledgement.

We extend our sincere gratitude to the Malaysian Department of Environment (DOE) for generously providing their data to facilitate this study. The research endeavors are made possible through the support of the FRGS grant, Project Code 2019-0005-102-02. Additionally, we would like to express our thanks to the Ministry of Education for granting us the opportunity to engage in this research as part of the Initiative #30 Program Pendidikan Pintar Cerdas (PPPC).

## References

- Adenan, N.H and Noorani, M.S.M (2013). River flow prediction using nonlinear prediction method. *International Journal of Mathematical and Computer Sciences*, Vol 7(11).
- Adenan, N.H and Noorani, M.S.M (2015). Peramalan data siri masa aliran sungai di dataran banjir dengan menggunakan pendekatan kalut. *Sains Malaysiana*, vol. 44, no 3.
- Azam, A. G., Zanjani, B. R. & Mood, M. B. (2016). Effects of Air Pollution on Human Health and Practical Measures for Prevention in Iran. *Journal of Research in Medical Sciences*, 21(5), 1-12.
- Bahari, M and Hamid, N.Z.A (2019). Analysis and prediction of temperature time series using chaotic approach. *IOP Conference Series: Earth and Environmental Sciences*, 286.
- Cadagli, M. (1991). Chaos and deterministic versus stochastic non-linear modelling. *Santa Fe Institute*, vol. 54, no. 2, pp. 303-328
- California Air Resources Board (2019). Air pollution hot spot. <http://www.greenfacts.org/glossary/abc/air-pollution-hot-spot.htm>. Retrieved on 31 December 2019.
- Cao, L. (1997) Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D*. 110 : 43-50
- Drury, R., Belliveau, M., Khun, J.S., Shipra, B. (1999). Pollution trading and environmental justice: los angeles failed experiment in air pollution policy. *Duke Environmental Law and Policy Forum*. 9(231)
- Embong, A. R. (2011). Pemandangan dan Kehidupan Bandar di Semenanjung Malaysia. *Akademika*, 81(2): 23-39.
- Fuller, R., Sandilaya, K. & Hanrahan. (2019). *Global Alliance on Health and Pollution*. Ministry of the Environment and Energy. Government Office of Sweden.



- Hamid, N.Z.A (2018) Application of chaotic approach in forecasting highland's temperature time series. *IOP Conference Series: Earth and Environmental Sciences*, 169.
- Hamid, N.Z.A and Noorani, M.S.M (2013). An improved prediction model of ozone concentration time series based on chaotic approach. *International Journal of Mathematical and Computer Sciences*, Vol 7(11)
- Hamid, N.Z.A and Noorani, M.S.M (2014). A pilot study using chaotic approach to determine characteristic and forecasting of PM10 concentration time series. *Sains Malaysiana*, vol 43, no 3.
- Hamid, N.Z.A., Noorani, M.S.M. and Adenan, N.H (2017). Chaotic analysis and short-term prediction of ozone pollution in Malaysian urban area. *Journal of physics*, vol. 890, 2017
- Hamid, N.Z.A., Noorani, M.S.M., Juneng, L. and Latiff, M.T. (2013). Prediction of Ozone concentration using nonlinear prediction method. *Proceeding of the 20<sup>th</sup> National Symposium on Mathematical Sciences*, pp. 125-131.
- Hassan, N., Jaafar, N. I. M., Ariffin, R. N. R., Samah, A. A. & Jaafar, M. N. (2013). Planning Malaysia: *Journal of the Malaysian Institute of Planners*, XI, 21–40.
- Jayawardena A.W. 1997. Runoff forecasting using a local approximation method. *IAHS*, vol. 239, 167-171
- Kalhoor, K. dan Mahdisoltani, M. (2015). Urbanization and its Effects on the Environment and Society Along with Sustainable Development. *Third International Symposium on Environmental and Water Resources Engineering*, 1-12.
- Kobayashi, Y., Santos, J. M., Mill, J. G., Junior, N. C. R., Andreao, W. L., Albuquerque, T. T. A. & Stuetz, R. M. (2020). Mortality Risks Due to Long-Term Ambient Sulphur Dioxide Exposure: Large Variability of Relative Risk in the Literature. *Environmental Science and Pollution Research*. 1-10.
- Kumar, S. (2017). Acid Rain-The Major Cause of Pollution: Its Causes, Effects. *International Journal of Applied Chemistry*. 13(1). 53-58.
- Mabahwi, N. A., Ling, O., Leh, H., & Omar, D. (2015). Urban Air Quality and Human Health Effects in Selangor, Malaysia. *Procedia - Social and Behavioral Sciences*, 170, 282–291.
- Malik, N., Asmi, F., Ali, M. & Rahman, M. M. (2018). Major Factors Leading Rapid Urbanization in China and Pakistan: A Comparative Study. *Journal of Social Science Studies*.5(1), 148.
- McCarthy, N. (2019). Study: Pollution Kills 8.3 Million People. Anually. URL: <https://www.statista.com/chart/20360/premature-pollution-related-deaths/>.
- Myllyvirta, L. (2020). *Quantifying the Economic Cost of Air Pollution from Fossil Fuels*. Centre of Research on Energy and Clean Air.
- Rai, M. S. (2017). Impact of Urbanization on Environment. *International Journal on Emerging Technologies*, 8(1), 127-129.
- Robinson, M. S (2018). "Sulfur Dioxide." Chemical Compund. Retrieved January 07, 2020 from Encyclopedia.com: <https://www.encyclopedia.com/humanities/dictionaries-thesauruses-pictures-and-press-releases/sulfur-dioxide>
- Ruelle, D and F. Takens. (1971). On the nature of turbulence. *Communications in Mathematical Physics*, vol. 20, no. 3, pp 167-192
- Ruslan, A.B and Hamid, N.Z.A (2019). Application of improved chaotic method in determining number of k-nearest neighbor for CO data series. *International Journal of Engineering and Advanced Technology*. Vol 8, 6S3
- Shafie, S. H. M. & Mahmud, M. (2015). Pencemaran Habuk di Lembah Klang Melalui Analisis Statistik Boxplot, *Journal of Society and Space*, 11(11), 144–154.
- Shafie, S. H. M., Mahmud, M., Mohamad, S., Rameli, N.L.F., Ismail, K.M., Mohamed, A.F. & Abdullah, R. (2018). Pengaruh Penduduk Terhadap Pencemaran Udara Bandar di Lembah Klang dengan Mengaplikasi Sistem Maklumat Geografi (GIS), *The 2nd International Conference on Social and Economic Development*, 1-13.
- Sivakumar, B. (2002). A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers. *Journal of hydrology*, vol. 258, pp. 149-146.
- Sivakumar, B. (2003). Forecasting monthly streamflow dynamics in the western United States: a nonlinear dynamical approach. *Environmental Modelling & Software*, vol. 18, pp 721-728.
- Sulaiman, F. R. (2015). Atmosfera Bumi dan Tindakbalas Kimia Pencemaran Udara: *Meneroka Interaksi Dinamik dalam Kesihatan Persekitaran*. Fakulti Perubatan dan Sains Kesihatan, Universiti Putra Malaysia, Serdang, Selangor, 53-69.
- USEPA, (2016). Sulfur Dioxide Trends. U.S. Environmental Protection Agency. <https://www.epa.gov/air-trends/sulfur-dioxide-trends>.
- Wang, Z., Zhang, C., Lv, G., Sun, X., Wang, N. & Li, Z. (2019). Synergistic Reaction of SO<sub>2</sub> with NO<sub>2</sub> in Presence of H<sub>2</sub>O and NH<sub>3</sub>: A Potential Source of Sulfate Aerosol. *International Journal of Molecular Sciences*, 20(15), 1-14.
- Zaim, W.N.A.W.M. (2018). Forecasting seasonal ozone time series in Malaysia higher education areas through chaotic approach, *Sultan Idris Education University*.