Journal of Quality Measurement and Analysis *JQMA* 20(1) 2024, 159-169
e-ISSN: 2600-8602 *http://www.ukm.my/jqma*
https://doi.org/10.17576/jqma.2001.2024.12

PREDICTING HAZE PHENOMENON USING CHAOS THEORY IN INDUSTRIAL AREA IN MALAYSIA

(Peramalan Jerebu Menggunakan Teori Kalut di Kawasan Perindustrian Malaysia)

HAZLINA DARMAN* & NOR ZILA ABD HAMID

ABSTRACT

Predicting the occurrence of haze is of great importance due to its negative impact on human health, the environment, and the economy. This study aims to develop a model for predicting haze using chaos theory. The data were taken from an industrial area, Klang, Selangor Malaysia during Southwest Monsoon. The model is trained using historical data on haze occurrences and the accuracy of the prediction is evaluated using a testing dataset. A chaos model, namely local mean approximation method (LMAM) will be used to predict the haze phenomenon. Results show that the chaos-based approach is effective in forecasting the onset and duration of haze events. The predicting model can provide early warnings for policymakers and relevant authorities, enabling them to take proactive measures to mitigate the effects of haze on public health and the environment. The model also presents a promising alternative to traditional forecasting techniques and highlights the potential applications of chaos theory in atmospheric science.

Keywords: haze forecasting; PM10; chaos theory; Cao method; sustainability development goals

ABSTRAK

Peramalan jerebu adalah sangat penting kerana memberi kesan negatif kepada kesihatan manusia, alam sekitar, dan ekonomi negara. Kajian ini akan membangunkan model peramalan jerebu menggunakan teori kalut. Data siri masa di ambil di kawasan perindustrian Klang, Selangor Malaysia semasa berlakunya musim Monsun Barat Daya. Teori kalut membahagikan data kepada dua: data latihan dan data ujian. Data latihan digunakan untuk membina model peramalan dan data ujian digunakan untuk membandingkan ketepatan model peramalan. Kaedah Penghampiran Purata Setempat (KPPS) akan digunakan dalam proses peramalan jerebu. Hasil kajian ini menunjukkan kaedah kalut sesuai diaplikasi dalam peramalan jerebu. Model peramalan jerebu mampu memberikan amaran awal kepada pihak berkuasa dan penggubal dasar untuk mengambil tindakan proaktif dalam mengurangkan kesan buruk jerebu kepada kesihatan awam dan persekitaran.

Kata kunci: peramalan jerebu; PM10; teori kalut; kaedah Cao; matlamat pembangunan mampan

1. Introduction

The haze phenomenon is a type of air pollution that often occurs during the hot and dry weather, especially during the Southwest Monsoon season, from the end of May until mid-September every year. The dry periods caused by the El-Nino prolong the duration of poor air quality (Latif *et al.* 2018). Health Management Actions Due to Haze Plan released by Kementerian Kesihatan Malaysia (2020) define haze as a situation of deterioration of air quality due to the presence of fine suspended particles in the atmosphere. These particles have high concentration which absorb and scatter sunlight that impairs visibility. The primary sources of haze can be categorized into two main groups: natural sources and human activities. Natural sources of haze include volcanic eruptions, wildfires, and dust storms. Meanwhile, human activities include engine emissions from vehicles, fuel combustion, agricultural activities, and open burning (Latif *et al.* 2017).

Haze produces negative impacts to humans' health, flora, and fauna due to the fine suspended particles. Kementerian Kesihatan Malaysia (2020) has listed the impacts of haze to humans' health, such as infection to respirotary system, reducing lung function, eyes iritation, nose and throat irritation, severe headache, dizzy, nausea and vomit. The negative impact of haze on the economy can be significant. Haze can disrupt various sectors and have adverse effects on businesses and industries, such as loss in tourism and business, productivity loss, and cost of mitigation and adaption (Quah *et al.* 2021). The cost of hospital admissions due to haze has ranged from MYR1.8 million to MYR118.9 million, and the incremental cost of illness has increased from MYR21 million to MYR410 million (Latif *et al.* 2018).

Haze contains pollutants or particles of PM_{10}, $PM_{2.5}$, ozone, nitrogen oxide, sulfur dioxide and carbon monoxide. PM_{10} is fine particulate matter, with a diameter of 10 micrometeres or less. It is the largest component of haze compared to other particles and it can come from various sources such as dust, smoke, and vehicle emissions. When the concentration of PM_{10} in the air increases, it indicates a higher level of air pollution. Thus, PM_{10} will be used as an indicator for haze or air quality in this study. Studies by Hamid and Noorani (2014), Abdullah *et al.* (2017; 2020), Shang *et al.* (2021), and Ul-Saufie *et al.* (2012) had also used PM_{10} to predict the haze in their papers.

The haze prediction can be employed using various methodologies, namely chaos (Hamid 2020; Hamid & Noorani 2014), multiple linear regression (Abdullah *et al.* 2017; Chen & Wang 2019), neural network (Pinghua 2018; Zhang *et al.* 2021), long-short term memories (Liu *et al.* 2019; Wu *et al.* 2021), and deep learning (Hasmarullzakim & Abdullah 2018; Idris & Yassin 2021). From study of Shahizam *et al.* (2021), Hamid (2020), Hamid and Noorani (2017) and several other papers showed that prediction using chaos method produced small error and high correlation coefficient. Therefore, since the chaos approach can be used to produce accurate predictions and only PM_{10} time series is required in prediction, this method will be used in this study.

Dynamics of time series can be either deterministic or random, where deterministic time series can be predicted and random time series cannot be predicted. Chaos is a form of dynamics that exists between deterministic and random dynamics. Any time series that has chaos dynamics can be predicted, however only short-term predictions are permitted due to a number of reasons (Abarbanel 1996). There are several methods to determine if the dynamics of a time series is chaos, such as Cao method and phase space plot (Hamid 2018; 2020; Adenan *et al.* 2021; Zhao *et al.* 2023), Lyapunov exponents (Hu *et al.* 2023; Pino-Vallejo *et al.* 2018; Wang *et al.* 2023), and correlation dimension (Saeed *et al.* 2017; Yu *et al.* 2013). Study by Shahizam *et al.* (2021), Hamid (2020), Hamid and Noorani (2017) used Cao method and phase space plot to do in chaos method and they had produced good prediction model. Thus, based on these previous successful outcomes of Cao method and phase space plot, the same methodology will be employed in this study.

There are two main methods for prediction in chaos method namely, local mean approximation method (LMAM) (Bahari & Hamid 2019; Shahizam *et al.* 2021) and local linear approximation method (LLAM) (Hamid *et al.* 2013; 2021). These two methods are known as chaotic models because they are involved in reconstructing the phase space at *m*-dimension. In this paper, LMAM will be applied to predict PM_{10} time series.

1.1 *Haze and Sustainable Development Goal (SDG)*

Haze poses a significant challenge to achieving several Sustainable Development Goals (SDGs). Firstly, haze directly impacts SDG3, which focuses on ensuring good health and well-being for all. The fine particles in haze can worsen respiratory conditions and lead to various health issues.

Haze also affects SDG7, which aims to ensure access to affordable and clean energy. The particles in haze can reduce solar radiation, impacting the efficiency of solar power generation. This hinders the transition to clean and renewable energy sources, which are crucial for mitigating climate change and achieving sustainable energy access for all.

Furthermore, haze undermines SDG11, which focuses on creating sustainable cities and communities. Poor air quality caused by haze can lead to reduced livability and hinder efforts to create inclusive, safe, and resilient cities. It also affects SDG13, which addresses climate action, as haze is often associated with the burning of fossil fuels and deforestation, contributing to greenhouse gas emissions and exacerbating climate change.

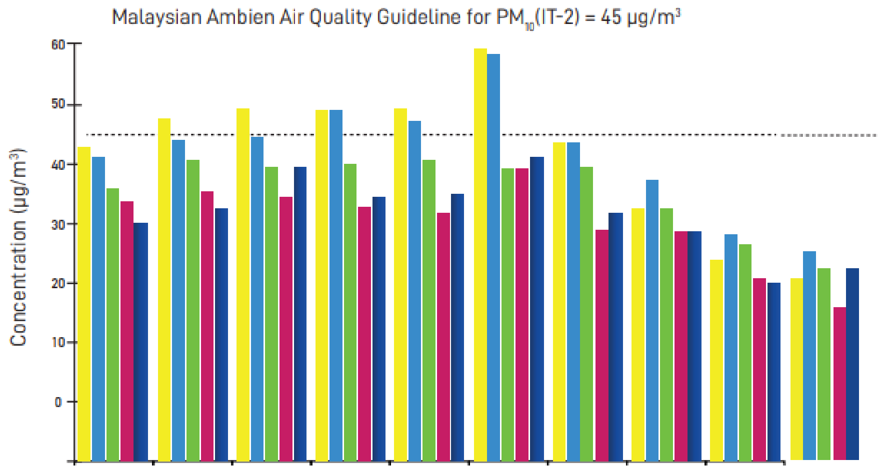
1.2 *Dataset*

For this study, the PM_{10} time series will be used to predict the haze in Klang, an industrial city located in Selangor, Malaysia. This dataset had been provided by Department of Environment, Malaysia. According to the official portal of the Selangor State Council, there are 29 areas in Klang that are gazetted as industrial areas (Dewan Negeri Selangor 2016). Since 1980s, Klang was reported to experience high concentration of particulates where Soleiman *et al.* (2003) had reviewed the severe haze episodes experienced in Klang Valley. Their paper stated that Klang has high potential for pollution because the rapid urbanization and industrial expansion. Thus, Klang area will be an excellent location to study the haze prediction. The trend of the annual average PM_{10} concentration levels in ambient air from 2010 to 2019 are as shown in Figure 1 (Department of Environment 2019). Based on land use categories, PM_{10} concentration is at highest level in industrial areas, followed by urban, suburban, and rural areas. Given that industrial sites frequently contribute to the release of pollutants into the atmosphere, it is critical to study haze prediction in this area. The industries, power plants, and other industrial facilities that produce pollutants including particulate matter, sulphur dioxide, and nitrogen oxides are often concentrated in these locations. These pollutants can interact with other substances in the atmosphere to create haze, which can be harmful to human health and the quality of the air.

The data used in this study is obtained from Department of Environment, Malaysia where 3671 hourly observations of PM_{10} time series were made over 153 days from 1st May 2017 to 30th September 2017. This specific time period is chosen due to the occurrence of the Southwest Monsoon and El-Nino phenomena where the haze concentration will be high during this period of time. Statistical analysis for PM_{10} time series is given in Table 1 and the trend of PM_{10} time series during Southwest Monsoon in Klang, Selangor is represented in Figure 2.

Table 1: Statistical analysis of hourly PM_{10} concentration

Total data	Number of train data	Number of test data	Minimum value	Maximum value	Mean value	Standard deviation	Median	Mode	Skewness
3672	3336	336	0.58	353.547	38.329	0.364	35.36	38.329	2.658



	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Industrial	44	48	49	48	49	58	44	33	23	26
Urban	42	45	46	48	46	57	44	35	28	32
Suburban	36	41	39	40	41	49	39	31	26	29
Background	34	36	34	34	33	49	29	25	21	22
Rural	30	34	39	35	36	43	32	25	22	28

Figure 1: Annual average concentration of particulate matter (PM_{10}) by land use, 2010 – 2019 (Department of Environment 2019)

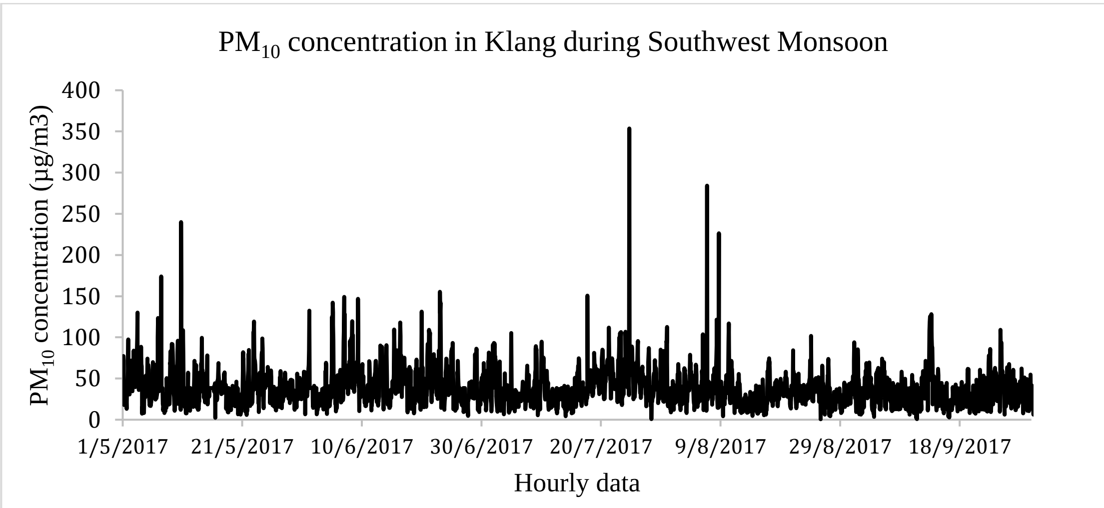


Figure 2: PM_{10} time series during Southwest Monsoon in Klang, Malaysia

2. Methodology

In the beginning of this study, the raw data needed to do the pre-processing to check for the missing data. There were 189 missing data from the total of 3672 data in PM_{10} time series which were imputed using series mean method. Series mean is a simple and straightforward method that enables to fill in the missing values with a single value derived from the mean of the series. This helps to preserve the overall distribution and statistical properties of the data.

Chaos method involved two steps; (i) reconstruction of phase space and (ii) prediction process. Two methods will be applied to reconstruct phase space, namely phase space approach and Cao method. Phase space approach is a simple approach to apply, while Cao method able to differentiate the nature of time series. For that reasons, these two methods will be applied in this study. Meanwhile, Local Mean Average Method (LMAM) will be applied in prediction process. This method is chosen because it is one of the chaotic model since it involved with the reconstruction of phase space. More details on these two steps can be found in next sections.

2.1 *Reconstruction of phase space*

The time series in one-dimensional scalar is given below:

$$X = \{x_1, x_2, \dots, x_N\} \tag{1}$$

where X is the PM10 time series for *N* observations. *X* is later divided into two parts: training dataset, X_{train} and testing dataset, X_{test} as shown in equation below:

$$X_{train} = \{x_1, x_2, x_3, \dots, x_l\} \tag{2}$$

and

$$X_{test} = \{x_{l+1}, x_{l+2}, x_{l+3}, \dots, x_N\} \tag{3}$$

The training dataset will be used to determine all the chaos parameters and the testing dataset will be used during prediction process. The first 3336 hourly data (139 days) will be used to train the model while the next 336 hourly data (14 days) will be used to test the performance of the model.

To reconstruct the phase space, two parameters will be determined, time delay τ and embedding dimension, *m*. The dataset X_{train} will be reconstructed into m-dimensional phase space:

$$Y_j^m = \left\{x_j, x_{j+\tau}, x_{j+2\tau}, \dots, x_{j+(m-1)\tau}\right\} \tag{4}$$

Fraser & Swinney (1986) had introduced average mutual information method to calculate time delay τ as shown below. *I*(*T*) represents the mutual information using various value of τ and *T* is various value of time delay τ.

$$I(T) = \frac{1}{N}\sum_{t=1}^{N} p(x_t, x_{t+\tau})\, log\left[\frac{p(x_t, x_{t+\tau})}{p(x_t)p(x_{t+\tau})}\right] \tag{5}$$

where x_t is the original time series, $x_{t+\tau}$ is time series shifted by τ, $p(x_t)$ and $p(x_{t+\tau})$ are the probability of x_t and $x_{t+\tau}$, accordingly and $p(x_t, x_{t+\tau})$ is their joint probability. By varying τ, *I*(*T*) is obtained. Later, graph *T* against *I*(*T*) is plotted and parameter τ is chosen based on the first minimum value of *T*. Next, the phase space graph is plotted using the obtained τ value. This graph can determine if the dynamic of the time series is random or chaos. If an attractor exists, the time series is then classified as chaos (Sivakumar 2002).

Cao (1997) had introduced Cao method that has the following advantages: (i) does not depend on other parameters, (ii) use only parameter τ, and (iii) does not depend on the size of data. Thus, this paper will apply Cao method to determine the embedding dimension, m using the following formula:

$$E_1(m) = \frac{E(m+1)}{E(m)} \tag{6}$$

and

$$E(m) = \frac{1}{N-m\tau}\sum_{j=1}^{N-m\tau}\frac{\left\|Y_j^{m+1}-Y_N^{m+1}\right\|}{\left\|Y_j^{m}-Y_N^{m}\right\|} \tag{7}$$

where N is the number of observations, $\|\bullet\|$ is the maximum norm and $\left\|Y_j^m\right\|$ is the nearest neighbour to Y_j^m. Parameter m is determined when $E_1(m)$ saturated. $E_1(m)$ can also be used to determine the dynamic of the time series, whether is random or chaos. Meanwhile, Cao (1997) also introduced term $E_2(m)$ where if exist $E_2(m) \neq 1$, then the dynamic of the time series is chaos. $E_2(m)$ is given as:

$$E_2(m) = \frac{E^*(m+1)}{E^*(m)} \tag{8}$$

where

$$E^*(m) = \frac{1}{N-m\tau}\sum_{j=1}^{N-m\tau}\left|x_{j+m\tau}^m - x_{N+m\tau}^m\right| \tag{9}$$

2.2 *Prediction: local mean approximation method*

For prediction process, this study will use local mean approximation method (LMAM). Phase space as shown in Eq. (4) will be applied and prediction process via chaos method is interpreted through equation using the obtained value of τ.

$$Y_{j+1}^m = f\left(Y_j^m\right) \tag{10}$$

The prediction of Y_{j+1}^m is done based on the neighbours of Y_j^m. Y_j^m is the last phase space and Y_{j+1}^m is the next phase space that we would like to predice. The neighbours are labelled as $Y_{j'}^m$ where $1 < j < j'$. Not all neighbours are used to predict Y_{j+1}^m in LMAM. Only neighbours with the minimum value of Euclidean distance $\left\|Y_{j'}^m - Y_j^m\right\|$ will be used, which is also known as k nearest neighbours. In this study, parameter k = 50 is determined using trial and error method. After $Y_{j'}^m$ is determined, the one-hour ahead $Y_{j'+1}^m$ will be listed. The prediction of Y_{j+1}^m is taken as the average of $Y_{j'+1}^m$ values:

$$Y_{j+1}^m = \frac{\sum_{q=1}^{k} Y_{j_q'+1}^m}{k} \tag{11}$$

2.3 *Performance measure*

The correlation coefficient (*CC*) was used to explain the relationship between the observed data and predicted data. *CC* value ranges between -1 and +1. If the *CC* value is closer to +1, the relationship is stronger, which indicate the prediction is accurate and close to the observed data. Contradictly, if the *CC* is closer to 0, the relationship is weaker, indicate the prediction is not that accurate. The correlation is given by:

$$CC = \frac{\Sigma(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\Sigma(O_i - \overline{O})^2}\sqrt{\Sigma(P_i - \overline{P})^2}} \tag{12}$$

where O is observed data, P is predicted data, $\overline{O}$ and $\overline{P}$ is their mean, while i is the length of prediction. R-squared is another method to measure the performance of the model in this paper.

R-squared is one measure of how well a model can predict the data, and falls between 0 and 1. The higher the value of R-squared, the better the model is at predicting the data. The formula is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{13}$$

where N is total observations, y_i is observed value at time i, $\hat{y}_i$ is the predicted value at time i, and $\overline{y}$ is the mean of variable y.

3. Results and Discussion

To find the value of τ, the average mutual information method was applied. Figure 3 concluded that $\tau = 8$ since the first minimum T was eight. Using $\tau = 8$, the graph of phase space plot is constructed in two dimensions and three dimensions, as shown in Figure 4 and Figure 5. Theses figures showed that there exists an attractor inside the region except the presence of some outliers in the time series. According to Sivakumar (2002), the presence of reasonably defined attractor in a low dimensional phase space suggests the existence of chaotic behavior in time series where this can be seen in these two graphs. Thus, it can be concluded that PM_{10} time series are chaos and can be predicted using non-linear prediction method. Figure 6 displays the result of *$E_1(m)$* and *$E_2(m)$* obtained using Cao method. *$E_1(m)$* starts to saturate when the value of m is greater than $m_0 = 5$. Therefore, $m = m_0 + 1 = 6$ is the minimum embedding dimension. In addition, if $E_1(m)$ saturates with increasing m, then chaotic behaviors exists in the time series (Cao 1997). Also, if exists at least a point where $E_2(m) \neq 1$, it assures the presence of chaotic behavior in the PM_{10} time series.

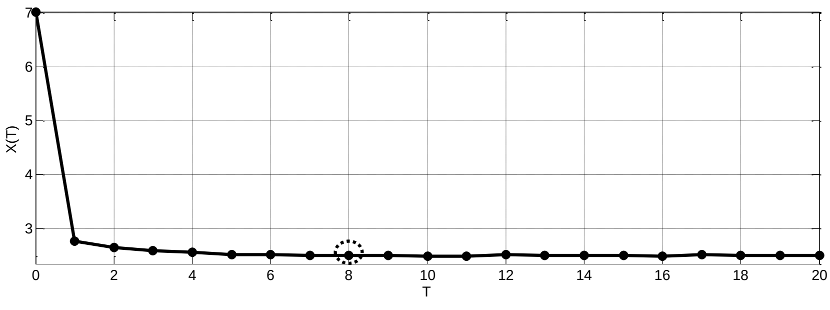


Figure 3: Average mutual information method to obtain $\boldsymbol{\tau}$

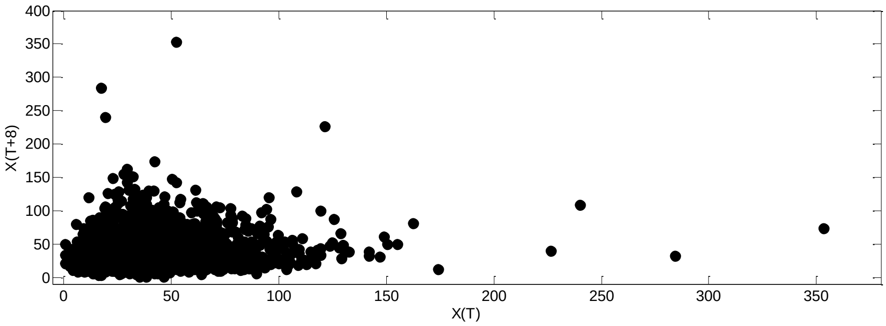


Figure 4: Attractor plot in two dimensions

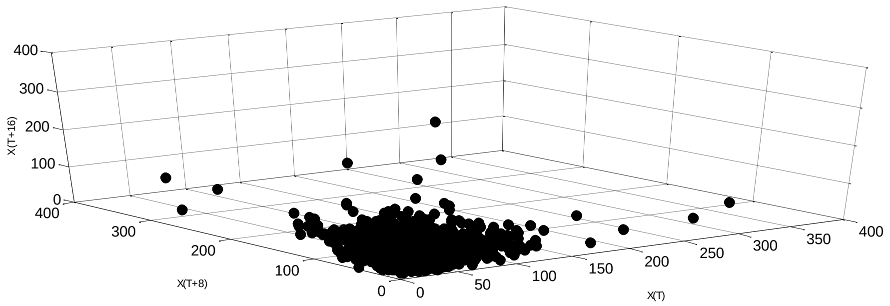


Figure 5: Attractor plot in two dimensions

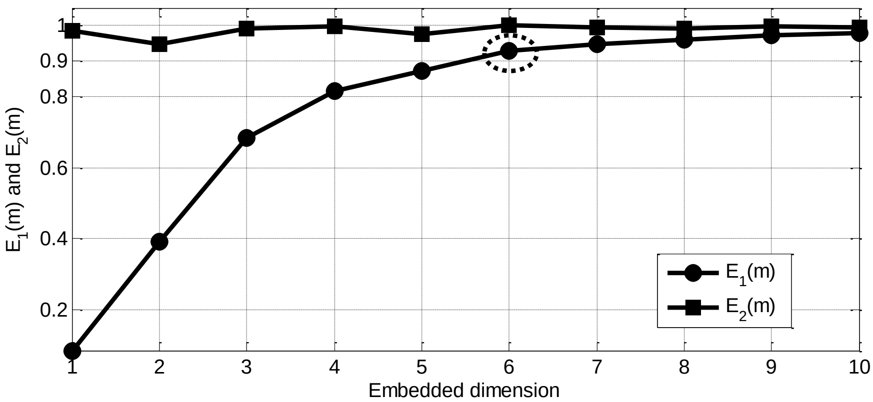


Figure 6: Embedding dimension using Cao method

With $\tau = 8$ and $m = 6$, the phase space of Eq. (4) is reconstructed and the prediction model is developed. Performance of the model will be measured using correlation coefficient (CC) where the correlation value obtained is 0.6420, which indicated that the relationship between the predicted PM_{10} and observed PM_{10} is moderately positive. This reflects that the predicted values are in good agreement with the observed values. Also, the obtained value of R-square is 41.21% which denoted 41.21% variation of PM_{10} is around its mean. Thus, the prediction performance of PM_{10} time series was moderately good. Figure 7 shows comparison of the predicted PM_{10} values fitted the trend of the observed data moderately. The observed data in the graph is the testing data that has been shown in Eq. (3).

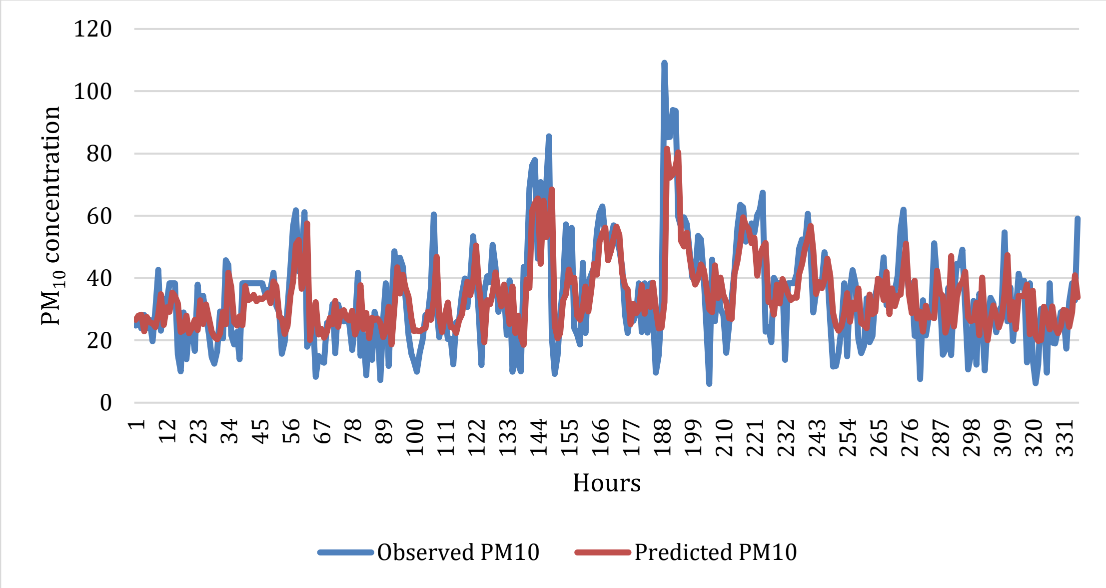


Figure 7: Comparison between the observed PM_{10} and predicted PM_{10}

4. Conclusion

In this study, the chaotic behavior of hourly PM_{10} time series at an industrial area, Klang Malaysia is detected using the phase space plot and Cao method. Later, the prediction is made using local mean approximation method which produced good haze prediction because the correlation coefficient is moderately positive. This result indicate chaos theory provides a good method to predict the haze. However, the limitation of this study is similar as studies by Hasmarullzakim and Abdullah (2018), Idris and Yassin (2021), and Shang *et al.* (2021) where the obtained model might not suitable to other areas due to different concentration of PM_{10}, ambient temperature, wind speed, and other factors.

For future research, the prediction using chaos method can also include other meteorological factors, such as air pressure, ambient temperature, windspeed, ozone, and others.

Acknowledgement

Thank you to the Department of Environment, Malaysia for providing necessary data. This study is under the sponsorship of University Pendidikan Sultan Idris (UPSI).

References

Abarbanel H.D.I. 1996. *Analysis of Observed Chaotic Data*. New York: Springer.

Abdullah S., Ismail M. & Fong S.Y. 2017. Multiple linear regression (MLR) models for long term PM10 concentration forecasting during different monsoon seasons. *Journal of Sustainability Science and Management* **12**(1): 60–69.

Abdullah S., Napi N.N.L.M., Ahmed A.N., Mansor W.N.W., Mansor A.A., Ismail M., Abdullah A.M. & Ramly Z.T.A. 2020. Development of multiple linear regression for particulate matter (PM10) forecasting during episodic transboundary haze event in Malaysia. *Atmosphere* **11**(3): 289.

Adenan N.H., Karim N.S.A., Mashuri A., Hamid N.Z.A., Adenan M.S., Armansyah & Siregar I. 2021. Traffic flow prediction in urban area using inverse approach of chaos theory. *Civil Engineering and Architecture* **9**(4): 1277–1282.

Bahari M. & Hamid N.Z.A. 2019. Analisis dan peramalan siri masa suhu menggunakan pendekatan kalut. *Journal of Quality Measurement and Analysis* **15**(1): 43-52.
Cao L. 1997. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena* **110**(1–2): 43–50.
Chen J. & Wang J. 2019. Prediction of PM2.5 concentration based on multiple linear regression. *2019 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, pp. 457–460.
Department of Environment. 2019. *Annual Environmental Quality Report 2019*.
Fraser A.M. & Swinney H.L. 1986. Independent coordinates for strange attractors from mutual information. *Physical Review A* **33**(2): 1134–1140.
Hamid N.Z.A. 2018. Prediction of ozone pollution through chaotic approach. *International Journal of Engineering & Technology* **7**(4.38): 1635-1638.
Hamid N.Z.A. 2020. Prediction of PM10 time series at industrial area through an improved local linear approximation method. *ASM Science Journal* **13**: 1-7.
Hamid N.Z.A., Adenan N.H., Wahid N.B.A., Bidin B. & Saleh S.H.M. 2021. A pilot study using chaos theory to predict temperature time series in Malaysian semi urban area. *Turkish Journal of Computer and Mathematics Education* **12**(3): 997–1003.
Hamid N.Z.A. & Noorani M.S.M. 2014. Suatu kajian perintis menggunakan pendekatan kalut bagi pengesanan sifat dan peramalan siri masa kepekatan PM10. *Sains Malaysiana* **43**(3): 475-481.
Hamid N.Z.A. & Noorani M.S.M. 2017. Aplikasi model baharu penambahbaikan pendekatan kalut ke atas peramalan siri masa kepekatan ozon. *Sains Malaysiana* **46**(8): 1333–1339.
Hamid N.Z.A., Noorani M.S.M., Juneng L. & Latif M.T. 2013. Prediction of ozone concentrations using nonlinear prediction method. *AIP Conference Proceeding* **1552**(1): 125–131.
Hasmarullzakim A.M. & Abdullah A. 2018. Haze pattern prediction using deep learning. *UTM Computing Proceedings* **3**.
Hu Y., Liu Z., Yue H., Wang J., Gao C., Zhang Y., Tan C. & Chen R. 2023. Flame instability analysis of water hyacinth pellets in combustion based on the chaos theory. *Solid Fuel Chemistry* **57**: 228–241.
Idris A.C. & Yassin H. 2021. Deep learning method for haze prediction in Singapore. *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, pp. 1–6.
Dewan Negeri Selangor. 2016. Kawasan Perindustrian di Selangor. https://dewan.selangor.gov.my/question/kawasan-perindustrian-di-selangor (1 June 2023).
Kementerian Kesihatan Malaysia. 2020. *Pelan Tindakan Pengurusan Kesihatan Akibat Jerebu*.
Latif M.T., Othman M., Idris N., Juneng L., Abdullah A.M., Hamzah W.P., Khan M.F., Sulaiman N.M.N., Jewaratnam J., Aghamohammadi N., Sahani M., Xiang C.J., Ahamad F., Amil N., Darus M., Varkkey H., Tangang F. & Jaafar A.B. 2018. Impact of regional haze towards air quality in Malaysia: A review. *Atmospheric Environment* **177**: 28–44.
Latif M.T., Othman M. & Kamin K.H. 2017. Fenomena jerebu di Asia Tenggara: Punca dan penyelesaian. Salam LESTARI **37**: 1-4.
Liu X., Liu Q., Zou Y. & Liu Q. 2019. A LSTM-based approach to haze prediction using a self-organizing single hidden layer scheme. *Security with Intelligent Computing and Big-data Services*, pp. 701-706.
Pinghua Z. 2018. Empirical study on haze forecasting model based on neural network. *IOP Conference Series: Materials Science and Engineering* **439**: 032073.
Pino-Vallejo M., Tierra A., Haro A. & Perugachi N. 2018. Prediction of concentrations of PM2.5 in downtown Quito using the chaos theory. *AIP Conference Proceeding* **2003**: 020013.
Quah E., Chia W.-M. & Tan T.-S. 2021. Economic impact of 2015 transboundary haze on Singapore. *Journal of Asian Economics* **75**: 101329.
Saeed S., Aziz W., Rafique M., Ahmad I., Kearfott K.J. & Batoolb S. 2017. Quantification of non-linear dynamics and chaos of ambient particulate matter concentrations in Muzaffarabad city. *Aerosol and Air Quality Research* **17**(3): 849–856.
Shahizam S.I., Hamid N.Z.A. & Side S. 2021. Peramalan bahan pencemar karbon monoksida (CO) di Putrajaya menggunakan kaedah penghampiran purata setempat. *Journal of Science and Mathematics Letters* **9**: 63–71.
Shang K., Chen Z., Liu Z., Song L., Zheng W., Yang B., Liu S. & Yin L. 2021. Haze prediction model using deep recurrent neural network. *Atmosphere* **12**: 1625.
Sivakumar B. 2002. A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers. *Journal of Hydrology* **258**(1–4): 149–162.
Soleiman A., Othman M., Samah A.A., Sulaiman N.M. & Radojevic M. 2003. The occurrence of haze in Malaysia: A case study in an urban industrial area. In Rao G.V., Raman S. & Singh M.P. (eds). *Air Quality*: 221–238. Basel, Switzerland: Birkhäuser.
Ul-Saufie A.Z., Yahaya A.S., Ramli N. & Hamid H.A. 2012. Performance of multiple linear regression model for long-term PM10 concentration prediction based on gaseous and meteorological parameters. *Journal of Applied Sciences* **12**(14): 1488–1494.

Wang Y., Zhang H., Shi Z., Zhou J. & Liu W. 2023. Nonlinear time series analysis and prediction of general aviation accidents based on multi-timescales. *Aerospace* **10**(8): 714.
Wu X., Liu Z., Yin L., Zheng W., Song L., Tian J., Yang B. & Liu S. 2021. A haze prediction model in Chengdu based on LSTM. *Atmosphere* **12**(11): 1479.
Yu H.-L., Lin Y.-C., Sivakumar B., & Kuo Y.-M. 2013. A study of the temporal dynamics of ambient particulate matter using stochastic and chaotic techniques. *Atmospheric Environment* **69**: 37–45.
Zhang Z., Tian J., Huang W., Yin L., Zheng W. & Liu S. 2021. A haze prediction method based on one-dimensional convolutional neural network. *Atmosphere* **12**(10): 1327.
Zhao L., Jin H., Zou X. & Liu X. 2023. Short-term passenger flow forecasting for rail transit considering chaos theory and improved EMD-PSO-LSTM-combined optimization. *Journal of Advanced Transportation* **2023**: 1–16.

Department of Mathematics,
Faculty of Science and Mathematics,
Universiti Pendidikan Sultan Idris
35900, Tanjung Malim, Perak DR, MALAYSIA
E-mail: hazlinadarman@gmail.com[*], *nor.zila@fsmt.upsi.edu.my*

Received: 1 August 2023
Accepted: 10 January 2024

[*]Corresponding author