

On Modeling Smoke-Haze Incidence with Cluster and Regression Analyses

Kacha Chansilp, Paradee Chuaybamroong, Kittisak Kerdprasop, and Nittaya Kerdprasop

Abstract—Thailand and many countries in the Southeast Asia have long been suffered from the regional smoke-haze incidences. Smoke-haze is a kind of air pollution event frequently occurred from forest fires that had been intentionally set for vegetation purpose. The smoke-haze can cause serious health problem from high concentrations of small particulate matters that can retain in the lung or even spread through the whole body to cause obstruction in major organs. In the northern part of Thailand, smoke-haze normally occurs during the dry season from late January to early April with the peak polluted air around March. Controlling burning activity is an obvious solution but impractical when burning areas are in remotely high mountains that are hard to reach by ground officers. Monitoring incidences as well as estimating pollution level are more or less efficient and practical ways to handle the smoke-haze situation. We thus propose the application of machine learning technology to learn smoke-haze patterns from historical events. The specific approach used in our work is the cluster analysis with the k-means and Kohonen self-organizing map algorithms. The cluster with serious pollution effect is then further analysed to induce the predictive regression model using the meteorological factors. The built model can serve as a predictive pattern useful for invoking an early warning sign for air pollution awareness.

Index Terms—Air pollution predictor, cluster analysis, multivariate regression, smoke-haze incidence modeling.

I. INTRODUCTION

Important air pollution incidences in Asia relates to haze. The natural haze phenomenon has been defined by the World Meteorological Organization as extremely small and dry particles suspending in the atmosphere with tremendous amount to cause dusky sky [1]. Winter haze events in many areas of China cause health problems from the toxic air pollutants [2]-[5]. Many forecasting models and simulation methods have been applied to study the transport characteristics of haze events [6]-[11].

Unlike haze that can occur naturally, smoke-haze is caused by humans to set large area of fires mainly for

agricultural purpose. Countries in the Southeast Asia, including Thailand, have long been suffered from the smoke-haze incidences [12]-[16]. Smoke-haze has also recently been a problem in Russia [17]. High concentrations of hazardous heavy metals and small particulate matters (PM) are dangerous constituents of haze and smoke-haze. PM10 and PM2.5 are particulate matters with the aerodynamic diameters less than or equal to 10 and 2.5 micrometers, respectively. The smaller PM, the more harmful because these extremely tiny particles can retain in the human lung and spread through the whole body. PM also causes airway damages, cardiovascular impairments, and adverse effects in infants [18].

In the Southeast Asian countries, smoke-haze is a periodic and transboundary incidence occurred during the dry season [19]-[22]. The man-made forest fire season in Thailand can start from the late January and extend possibly to the last week of April or even early May, depending on the coming of rain. In northern Thailand, the extent and intensity of fire-related air pollution have been increasing recently. Controlling burning is obviously an effective solution, but it is quite impractical because there are so many large areas in valleys and high mountains almost unreachable by officers. Monitoring burning incidences as a warning sign is more or less a feasible way for handling in advance the smoke-haze situation [23].

We thus propose in this paper a data-driven modeling, also known as data mining, as an alternative approach to traditional simulation method. The advantage of data mining method is its automatic approach on deriving, or learning, discriminative patterns from the provided historical events as training data. The derived model of our approach is a pattern of smoke-haze occurrence that can transport among mountain regions. Our study area and details of data records are presented in Section II. Our model building strategy is explained in Section III. Model evaluation results are presented in Section IV. We finally conclude the paper in Section V.

II. STUDY AREA AND DATA CHARACTERISTICS

A. Study Area

Northern Thailand is the focus area of our study because of its intensity in smoke-haze from the burning of land. The terrain in the north is forested and mountainous, as shown in Fig. 1. The smoke-haze problem normally starts from the border provinces such as Chiang Rai in the up north and Mae Hong Son on the west side. These two provinces have many

Manuscript received July 27, 2019; revised June 1, 2020. This work was supported in part by grants from the National Research Council of Thailand (NRCT) and Suranaree University of Technology (SUT) through the funding of Data and Knowledge Engineering Research Unit in which the third and fourth authors are principal researchers.

K. Chansilp and K. Kerdprasop are with the School of Computer Engineering, Suranaree University of Technology (SUT), Nakhon Ratchasima 30000, Thailand (e-mail: kacha@sut.ac.th, kerdpras@sut.ac.th).

P. Chuaybamroong is with the Department of Environmental Science, Thammasat University, Rangsit Campus, Thailand (e-mail: paradee@tu.ac.th).

N. Kerdprasop is with the School of Computer Engineering and the Data and Knowledge Engineering Research Unit, SUT, Thailand (e-mail: nittaya@sut.ac.th).

high mountains along their borders to neighbor countries: Lao on the north and Myanmar on the west. Therefore, they are vulnerable to smoke-haze transport from the neighborhoods. Forest fires in Chiang Rai and Mae Hong Son also affect other valley provinces such as Lamphun and Lampang. The main areas of our study are thus mountainous regions along the Lao and Myanmar borders and provinces in lower river valleys.

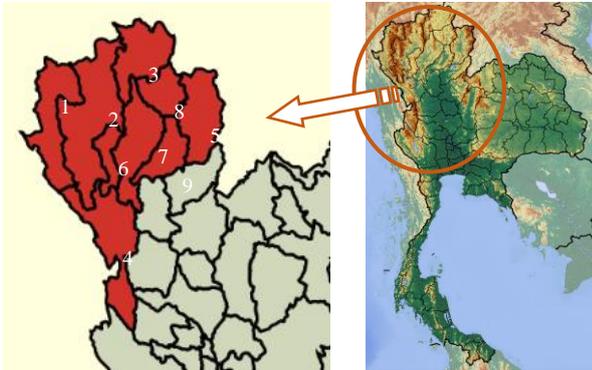


Fig. 1. The study area comprising of nine northern provinces (left) and terrain characteristics (right, source: <https://maps-for-free.com>).

TABLE I: DATA ATTRIBUTES AND THEIR EXPLANATION

Att. #	Meaning
1	Day 1 to 121 (1 January - 30 April 2016)
2	Average PM10 (ug/m ³) of <i>Chiang Mai</i> province on each day
3	Average PM10 (ug/m ³) of <i>Lampang</i> province on each day
4	Average PM10 (ug/m ³) of <i>Lamphun</i> province on each day
5	Average PM10 (ug/m ³) of <i>Chiang Rai</i> province on each day
6	Average PM10 (ug/m ³) of <i>Mae Hong Son</i> province on each day
7	Average PM10 (ug/m ³) of <i>Nan</i> province on each day
8	Average PM10 (ug/m ³) of <i>Phrae</i> province on each day
9	Average PM10 (ug/m ³) of <i>Phayao</i> province on each day
10	Average PM10 (ug/m ³) of <i>Tak</i> province on each day
11-13	Temperature (°C) in high, low, and average categories each day
14-16	Dew point (°C) in high, low, and average categories each day
17-19	Humidity (%) in high, low, and average categories each day
20-22	Sea level pressure (hPa) in high, low, and average scales each day
23-25	Visibility (km) in high, low, and average categories each day
26-27	Wind speed (km/h) in high and average categories each day

B. Data Attributes

The start of the dry season around mid to late February is the onset of slash and burn farming. The lingering smoke from the fires causes serious pollution problem to people living along the valley and lowland areas during March to April every year. We thus collect air pollution data, with a specific attention to PM10, of nine provinces in the mountainous and valley areas.

The five mountainous provinces (Fig. 1) are:

- (1) *Mae Hong Son* on the west,
- (2) *Chiang Mai* on the central north,
- (3) *Chiang Rai* on the up north,
- (4) *Tak* on the west of lower northern area, and
- (5) *Nan* on the farthest to the east.

The river valley provinces (Fig. 1) are:

- (6) *Lamphun*,
- (7) *Lampang*,
- (8) *Phayao*, and

(9) *Phrae*.

Pollution data are collected from the Air Quality and Noise Management Bureau of Thailand [24]. The PM10 data from these nine provinces are collected in 2016 from January to April, totaling of 121 days. Each data record contains 24-hour average PM10 values from ground stations in each province. For some missing PM10 records, we impute with the nearest neighbor technique.

We also include data collected from the local airport stations [25] as meteorological factors. The data attributes used in our modeling process are summarized and shown in Table I.

III. MODELING METHODOLOGY

The steps in our method for geographical polluted area clustering and pollution predictive modeling are illustrated in Fig. 2. The first step is data collection. The PM10 data [24] from the ground stations in nine provinces are collected from the first of January to the last day of April 2016. We average PM10 values from all stations in each province and use this mean as the pollution value for that province. To gain better understanding regarding the PM10 episodes of the nine provinces, we have to explore the characteristics of PM10 occurrences and levels using a graph plot.

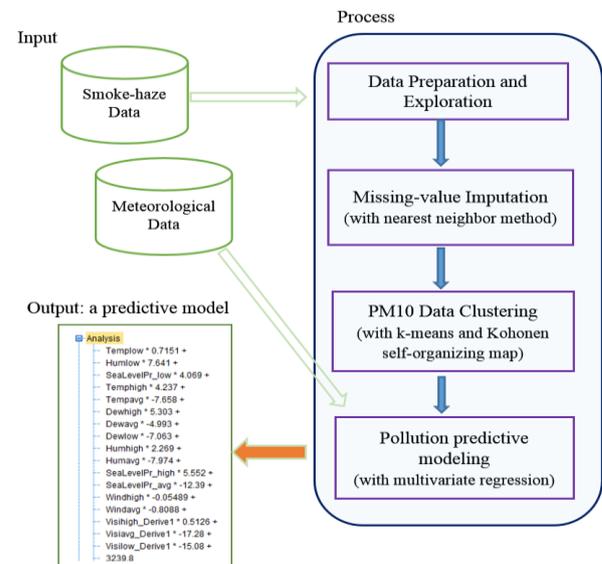


Fig. 2. Major steps in clustering and modeling haze occurrences.

The next step is checking for data completeness. There exist some missing values in the PM10 records from some stations. We therefore impute the missing values with the nearest neighbor method that estimates potential PM10 value from the closest station.

After completing the missing PM10 values, all the 121 data records are clustered to form groups of provinces experiencing the same level of air pollution due to smoke-haze. We apply two automatic data clustering algorithms for this step including the k-means clustering [26] and the Kohonen self-organizing map [27]. The main reason for applying two clustering methods is to select the best approach that yields reasonable result suitable for the

subsequent step of predictive model building.

Our criteria for selecting clustering results are number of groups that should much less than nine, which is the number of provinces to be clustered, and the silhouette coefficient that should higher than 0.5 to reflect a good formation of clusters.

We then consider from the clustering results to extract group of provinces that experience the strongest impact from air pollution due to smoke-haze. The meteorological data (attributes 11-27 in Table I) for that specific group are then used to build a predictive model for estimation PM10 level from the meteorological proxy.

IV. EXPERIMENTATION AND RESULTS

Based on the design of our analysis methodology, there are three main phases of our work: preliminary data exploration, automatic data clustering based on the PM10 concentrations in each province, and the creation of predictive model to estimate amount of PM10 using the meteorological factors such as humidity, wind speed, and so on as pollution predictors. The results from each phase are demonstrated and explained in the following subsections.

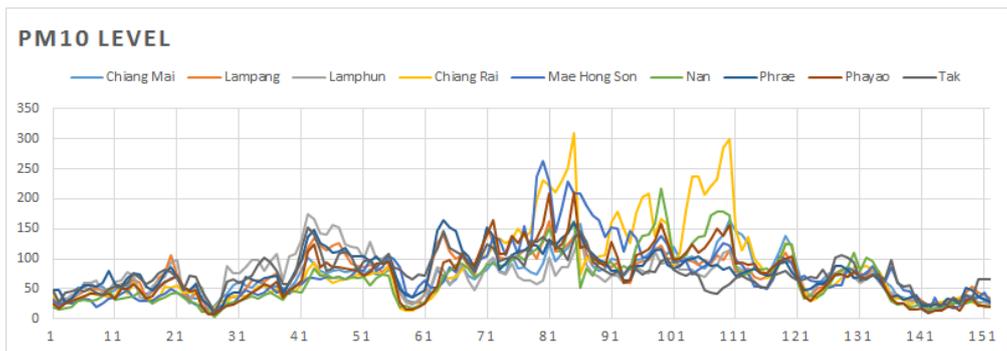


Fig. 3. PM10 concentrations (ug/m³) in the nine northern provinces during Jan-May (or day 1-152) of the year 2016

TABLE II: AIR POLLUTION INDEX FROM THE CONCENTRATION OF PM10 POLLUTANT

Category	Index Level	PM10 Particles	Health Impact
Low	1	0-16	Minimal
	2	17-33	
	3	34-50	
Moderate	4	51-58	Minor breathing discomfort to sensitive people
	5	59-66	
	6	67-75	
High	7	76-83	Breathing discomfort to asthma patients, elderly, and children
	8	84-91	
	9	92-100	
Very high	10	101 or more	Breathing discomfort to all and serious impact on people with heart or lung disease

B. Cluster Analysis Results

We perform cluster analysis twice with two different data subsets: PM10 concentrations from nine provinces during

A. Data Exploration

To correctly planning the analysis experimentation, we have to firstly explore the pollution situation based on the observed PM10 concentrations from the first day of January extending to the last day of May in year 2016. The levels of average PM10 from the ground stations in each province are displayed in Fig. 3.

We can observe from the pollution level distribution that the peak period of air pollution mainly caused by smoke-haze is between days 61 to 121, which are March to April. During the peak period the average level of PM10 concentration among the nine provinces is around 105, whereas before this peak period the pollution level is at 57.

The PM10 concentration level higher than 101 is considered harmful due to the UK standard [28] (as shown in Table II). We therefore partition dataset into two subsets: data during days 1-60 and those during days 61-121. Pollution data in May are excluded from further experimentation because our focus is on the peak period of smoke-haze events. The prepared two data subsets of average PM10 for each of the nine province are used in the next step of automatic clustering.

days 1-60 and PM10 during days 61-121. The two clustering algorithms adopted in our experimentation is k-means and Kohonen self-organizing map (SOM). The result (shown in Fig. 4) turns out that the k-means clustering yields a smaller number of clusters than the Kohonen SOM. Even though the silhouette coefficient value of k-means is lower than the value obtained from the Kohonen SOM, we choose the k-means method from its reasonable grouping results.

Graph	Model	Silhouette	Number of Clusters	Smallest/Largest
	Kohonen...	0.883	8	0.5
	K-mean...	0.572	5	0.333

Fig. 4. Comparative performance of k-means and Kohonen SOM.

From both experiments performed on the two data subsets, we obtain the same number of clusters, that is, five. But provinces in each group are different. Details of clustering results are summarized in Table III and geographically shown in Fig. 5.

From the clustering results of provincial locations based

on the PM10 concentrations, we can notice (in Table III) that at the non-peak period during January-February the air quality of the four provinces (Chiang Rai, Mae Hong Son, Phayao, Nan) in cluster 5 is good with the low level of PM10 pollutant. The air quality in these areas drastically changes in an unsatisfactory way during the peak period of burning (March-April). The pollution level in the four provinces is worsen rapidly from the low level (index = 3) to the harmful level (index = 10) in less than a month.

TABLE III: GEOGRAPHICAL-BASED CLUSTERING RESULTS FROM THE PM10 CONCENTRATIONS DURING PEAK AND NON-PEAK PERIODS

Period	Cluster no.	Provinces in a cluster	Average PM10 of a cluster	Pollution level
Non-peak (1 January - 29 February)	1	Lamphun	77.33	7
	2	Tak	72.05	6
	3	Phrae, Lampang	64.98	5
	4	Chiang Mai	56.29	4
	5	Chiang Rai, Mae Hong Son, Phayao, Nan	46.11	3
Peak (1 March - 30 April)	1	Chiang Rai	143.68	10
	2	Mae Hong Son	119.06	10
	3	Phayao, Nan	108.68	10
	4	Tak, Phrae, Lampang	97.39	9
	5	Chiang Mai, Lamphun	88.24	8

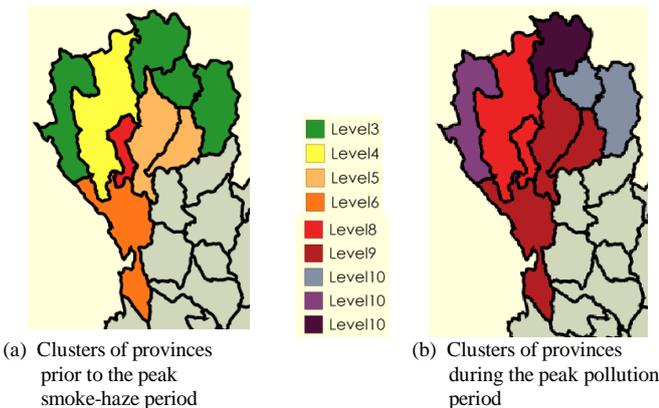


Fig. 5. Clustering results of the nine provinces based on the PM10 concentrations during (a) Jan-February, 2016 and (b) March-April, 2016.

Among the most affected four provinces, Chiang Rai is the worst one with the highest PM10 concentration at 309 $\mu\text{g}/\text{m}^3$ in late March. We are therefore interested in deriving a predictive model to estimate the pollution level of Chiang Rai province using meteorological factors.

C. Multivariate Linear Regression Model to Predict PM10

There are six meteorological indexes used in the predictive model creation including the temperature ($^{\circ}\text{C}$), dew point ($^{\circ}\text{C}$), humidity (%), sea level pressure (hPa), visibility (km), and wind speed (km/h). These parameters, except wind speed, are further categorized into three groups: low, high, and average. The wind speed are categorized as high and average. Therefore, there are 17 factors used in the model building process.

We apply the multivariate linear regression technique to

derive a predictive model to estimate the level of PM10 to occur in Chiang Rai province. The obtained model is shown in Fig. 6. Among the 17 predictors, low humidity and average visibility distance are two most important factors to estimate pollution level. Importance of the best ten predictors are graphically shown in Fig. 7.

$$\begin{aligned}
 & \text{PM10 Concentration in Chiang Rai} \\
 & = 0.7151 * \text{Low Temperature} \\
 & + 7.641 * \text{Low Humidity} \\
 & + 4.069 * \text{Low Sea Level Pressure} \\
 & + 4.237 * \text{High Temperature} \\
 & + (-7.658) * \text{Average Temperature} \\
 & + 5.303 * \text{High Dew point} \\
 & + (-4.993) * \text{Average Dew point} \\
 & + (-7.063) * \text{Low Dew point} \\
 & + 2.269 * \text{High Humidity} \\
 & + (-7.974) * \text{Average Humidity} \\
 & + 5.552 * \text{High Sea Level Pressure} \\
 & + (-12.39) * \text{Average Sea Level Pressure} \\
 & + (-0.055) * \text{High Wind speed} \\
 & + (-0.081) * \text{Average Wind speed} \\
 & + 0.513 * \text{High Visibility} \\
 & + (-17.28) * \text{Average Visibility} \\
 & + (-15.08) * \text{Low Visibility} \\
 & + 3239.8
 \end{aligned}$$

Fig. 6. A predictive model to estimate amount of PM10 in Chiang Rai.

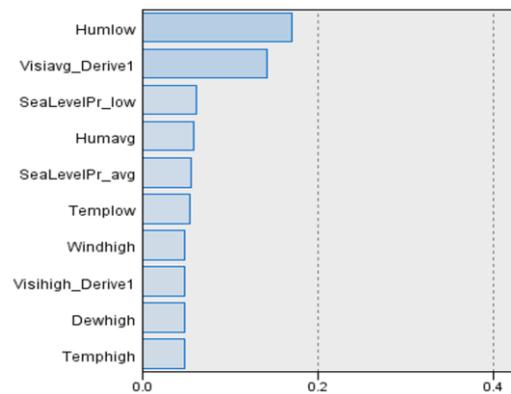


Fig. 7. Importance of predictors to estimate pollution level in Chiang Rai.

From the multivariate linear regression relationship and the importance of predictors, we can estimate the high pollution level when the humidity is low, the average visibility distance is short, and the sea level pressure is low. Wind speed has negative correlation to the PM10 concentration; that means wind can reduce level of air pollution.

V. CONCLUSION

Smoke-haze is a post-harvesting event constantly occurred in almost every region of Thailand during late January to early April. In some specific areas in the north that are surrounded by high mountains, the small particulate matters from farm burning are extremely severe air pollution incidence. This pollution event is preventive because it is man-made.

We thus try to study the occurrence patterns of smoke-haze with cluster analysis and to induce predictive model from the historical data. Our intention is to gain some insight regarding smoke-haze occurrence patterns and to apply the predictive model as a warning tool for officers and local people.

From our experimentation, we found that the technique of k-means clustering yielded a reasonable grouping result. The k-means clustering results are then analyzed to select group showing highest level of pollution for further analysis with predictive modeling.

The best modeling technique for this specific data is multivariate linear regression. The correlation coefficient of the linear model is as high as 0.837. The mean absolute error of model prediction is 30.513. We thus can conclude from this evaluation result that the predictive model is accurate enough to apply for the future situation.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

The first author is responsible for designing the research framework, organizing the experimentation steps and preparing the manuscript. The second author helps conducting experimentation and confirming the results. The third author takes part in the programming work. The last author helps editing the manuscript and discussing the results.

REFERENCES

- [1] World Meteorological Organization. International Meteorological Vocabulary, WMO - No. 182. (January 2018). [Online]. Available: <http://wmo.multitransms.com/MultiTransWeb/Web.mvc>
- [2] J. Li and Z. Han, "A modeling study of severe winter haze events in Beijing and its neighboring regions," *Atmospheric Research*, vol. 170, pp. 87–97, Mar. 2016.
- [3] T. Qiao, M. Zhao, G. Xiu, and J. Yu, "Simultaneous monitoring and compositions analysis of PM 1 and PM 2.5 in Shanghai: Implications for characterization of haze pollution and source apportionment," *Science of The Total Environment*, vol. 557–558, pp. 386–394, Jul. 2016.
- [4] Q. Zhang, "A heavy haze episode in Shanghai in December of 2013: Characteristics, Origins and implications," *Aerosol and Air Quality Research*, 2015.
- [5] F. Xu, N. Xiang, and Y. Higano, "How to reach haze control targets by air pollutants emission reduction in the Beijing-Tianjin-Hebei region of China?" *PLOS ONE*, vol. 12, no. 3, p. e0173612, Mar. 2017.
- [6] H. Che, X. Zhang, Y. Li, Z. Zhou, J. J. Qu, and X. Hao, "Haze trends over the capital cities of 31 provinces in China, 1981–2005," *Theoretical and Applied Climatology*, vol. 97, no. 3–4, pp. 235–242, Aug. 2009.
- [7] D. Liu and L. Li, "Application study of comprehensive forecasting model based on entropy weighting method on trend of PM2.5 concentration in Guangzhou, China," *International Journal of Environmental Research and Public Health*, vol. 12, no. 6, pp. 7085–7099, Jun. 2015.
- [8] Y.-Z. Tian *et al.*, "Seasonal and regional variations of source contributions for PM 10 and PM 2.5 in urban environment," *Science of The Total Environment*, vol. 557–558, pp. 697–704, Jul. 2016.
- [9] H.-X. Xu, G. Li, S.-L. Yang, and X. Xu, "Modeling and simulation of haze process based on Gaussian model," in *Proc. the 11th International Computer Conference on Wavelet Active Media Technology and Information Processing*, 2014, pp. 68–74.
- [10] X. Yan, W. Shi, N. Luo, and W. Zhao, "A new method of satellite-based haze aerosol monitoring over the North China Plain and a comparison with MODIS Collection 6 aerosol products," *Atmospheric Research*, vol. 171, pp. 31–40, May 2016.
- [11] G. Zhou, "Measuring and modeling aerosol: Relationship with haze events in Shanghai, China," *Aerosol and Air Quality Research*, 2014.

- [12] R. Field, "A drought-based predictor of recent haze events in western Indonesia," *Atmospheric Environment*, vol. 38, no. 13, pp. 1869–1878, Apr. 2004.
- [13] A. Heil and J. G. Goldammer, "Smoke-haze pollution: A review of the 1997 episode in Southeast Asia," *Regional Environmental Change*, vol. 2, no. 1, pp. 24–37, Aug. 2001.
- [14] L. C. Koe, A. F. Arellano, and J. L. McGregor, "Investigating the haze transport from 1997 biomass burning in Southeast Asia: its impact upon Singapore," *Atmospheric Environment*, vol. 35, no. 15, pp. 2723–2734, 2001.
- [15] J. Othman, M. Sahani, M. Mahmud, and M. K. Sheikh Ahmad, "Transboundary smoke haze pollution in Malaysia: Inpatient health impacts and economic valuation," *Environmental Pollution*, vol. 189, pp. 194–201, Jun. 2014.
- [16] E. Velasco and S. Rastan, "Air quality in Singapore during the 2013 smoke-haze episode over the Strait of Malacca: Lessons learned," *Sustainable Cities and Society*, vol. 17, pp. 122–131, Sep. 2015.
- [17] S. A. Sitnov, I. I. Mokhov, G. I. Gorchakov, and A. V. Dzhola, "Smoke haze over the European part of Russia in the summer of 2016: A link to wildfires in Siberia and atmospheric circulation anomalies," *Russian Meteorology and Hydrology*, vol. 42, no. 8, pp. 518–528, Aug. 2017.
- [18] S. Feng, D. Gao, F. Liao, F. Zhou, and X. Wang, "The health effects of ambient PM2.5 and potential mechanisms," *Ecotoxicology and Environmental Safety*, vol. 128, pp. 67–74, Jun. 2016.
- [19] E. Quah and D. Johnston, "Forest fires and environmental haze in Southeast Asia: Using the 'stakeholder' approach to assign costs and responsibilities," *Journal of Environmental Management*, vol. 63, no. 2, pp. 181–191, Oct. 2001.
- [20] E. Quah, "Transboundary pollution in Southeast Asia: the Indonesian fires," *World Development*, vol. 30, no. 3, pp. 429–441, 2002.
- [21] T. Forsyth, "Public concerns about transboundary haze: A comparison of Indonesia, Singapore, and Malaysia," *Global Environmental Change*, vol. 25, pp. 76–86, Mar. 2014.
- [22] Y. F. Liu, "Haze in Southeast Asia," in *Proc. IRC Conference on Science, Engineering, and Technology*, NUS, Singapore, 2015.
- [23] A. C. Spessa *et al.*, "Seasonal forecasting of fire over Kalimantan, Indonesia," *Natural Hazards and Earth System Sciences Discussions*, vol. 2, no. 8, pp. 5079–5111, Aug. 2014.
- [24] Division of Air Quality Data, Pollution Control Department, Air Quality and Noise Management Bureau, Thailand. (August 2016). Air quality data. [Online]. Available: <http://www.aqmthai.com/aqi.php>
- [25] Airport weather for Thailand. (January 2018). [Online]. Available: <https://www.wunderground.com/global/TH.html>
- [26] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.
- [27] T. Kohonen, "The self-organizing map," in *Proc. the IEEE*, 1990, vol. 78, no. 9, pp. 1464–1480.
- [28] Department for Environment Food & Rural Affairs, U.K., "What is the daily air quality index? (January 2018). [Online]. Available: <https://uk-air.defra.gov.uk/air-pollution/daqi?view=more-info&pollutant=pm10#pollutant>

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Kacha Chansilp is an assistant professor at the School of Computer Engineering, Suranaree University of Technology (SUT), Thailand. He received his bachelor degree in computer science from Queens College, U.S.A., in 1993, master degree in graphic design from New York Institute of Technology, U.S.A., in 1995 and doctoral degree in interactive multimedia technologies from Edith Cowan University, Australia, in 2003. His current research includes web databases, educational systems, environmental modeling, and computational applications in engineering.



Paradee Chuaybamroong is currently an associate professor in environmental science with the Department of Environmental Science, Thammasat University, Thailand. She received her bachelor degree in Public Health from Mahidol University, Thailand, in 1990, master degree in environmental science from Colorado School of Mines, U.S.A. in 1997 and doctoral degree in environmental science from University of Florida, U.S.A., in 2002. Her current research includes

environmental science and engineering, geophysics, bioaerosols, and environmental photocatalysis.



Kittisak Kerdprasop is an associate professor and chair of the School of Computer Engineering, SUT. He received his bachelor degree in mathematics from Srinakarinwirot University, Thailand, in 1986, master degree in computer science from the Prince of Songkla University, Thailand, in 1991 and doctoral degree in computer science from Nova Southeastern University, Florida, U.S.A., in 1999. His current research includes machine learning and artificial intelligence.



Nittaya Kerdprasop is an associate professor and the head of Data and Knowledge Engineering Research Unit, School of Computer Engineering, SUT. She received her bachelor degree in radiation techniques from Mahidol University, Thailand, in 1985, master degree in computer science from the Prince of Songkla University, Thailand, in 1991 and doctoral degree in computer science from Nova Southeastern University, U.S.A., in 1999. Her research of interest includes data mining, artificial intelligence, logic programming, and machine intelligence.