Time Series Models Predicting Fine Particulate Matter in Bangkok by Influence of Meteorological Factors

Thidarat Ketjumpol[1*] and Arnond Sakworawich[2]

Received: 1 November 2019; Revised: 25 December 2019; Accepted: 25 December 2019

Abstract

This research applied time series models to investigate factors that related to fine particulate matter concentration in Bangkok. The factors are wind speed, wind direction, temperature, relative humidity and rain from 1 January 2017 to 31 December 2018 by Pollution Control Department of Thailand. The result revealed that there are relation between $PM_{2.5}$ concentration and hourly effect, monthly effect, the interaction effect between hourly effect and wind speed, the interaction effect between monthly effect and wind speed. Two-Pass Estimation is used to estimate the model because data characteristic is multi-seasonality times series. So dummy variables are needed. Monthly effect strongly influences on $PM_{2.5}$ concentration that relates to the season of Thailand and the highest months are January and February. Wind speed is a factor which has a high negative relation with $PM_{2.5}$ concentration because southwesterly wind blows and carries $PM_{2.5}$ particles. Furthermore, during winter from November to February, southwesterly wind is weak and Bangkok is covered with areas of high air pressure. This promotes $PM_{2.5}$ accumulation and shows hourly effect obviously in that month.

Keywords: PM2.5, Two-Pass Estimation, multi-seasonality, dummy variable

* Corresponding author: E-Mail: thidarat.ket@stu.nida.ac.th

[1] Graduate School of Applied Statistics National Institute of Development Administration

[2] Graduate School of Applied Statistics National Institute of Development Administration

1. Introduction

$PM_{2.5}$ dust particle is heard repeatedly in the past few years. It is a fine particulate matter smaller than 2.5 micrometers in diameter. It associates with adverse health effects. Fine particles cause the hazardous problems because they can get into your lungs and bloodstream. Several scientific studies found that pollution particles linked to human health, including nonfatal heart attack, irregular heartbeat, decreased lung function and increased respiratory symptoms (Brunekreef & Holgate, 2002; Cascio, 2018). It is known that $PM_{2.5}$ is one kind of carcinogen; cancer-causing substance that impacts on the human respiratory system, especially lung (Xing, Xu, Shi, & Lian, 2016).

The annual mean concentration of fine particulate matter in Bangkok exceeds World Health Organization (WHO) air quality guidelines levels (10 µg/m^3) since 2011 when it has been firstly monitored (Thailand Pollution Control Department, 2019). $PM_{2.5}$ in Bangkok is mainly from the vehicular exhaust and some are blown from agriculture crop residue burning from neighboring countries. The components of $PM_{2.5}$ in Bangkok are 51 elements of heavy metals. Most of them are much lower than the safe level but cadmium and lead are quite high. They originate from the agricultural ecosystem. (Pongpiachan et al., 2017). Agriculture crop residue burning induces air pollution with heavy metals. Inhaling cadmium affects the respiratory system severely such as shortness of breath, lung edema and destruction of mucous membranes. It promotes acute respiratory distress syndromes (Godt et al., 2006). It also has a correlation with death rates from hypertension and arteriosclerotic heart disease (Carroll, 1966). The studies in laboratory animal show lead might effect on reproduction (Hilderbrand, Der, Griffin, & Fahim, 1973) and liver toxicity (Mudipalli, 2007). Their values exceed the guidelines levels and might be higher if the government did not yet have a strong air pollution control policy.

Predicting models and studying related factors of $PM_{2.5}$ are advantageous for resolving root causes that promote $PM_{2.5}$ and likewise implementing short and long-term action plans to reduce and avoid $PM_{2.5}$ practically. For example, face mask preparation, safety procedure planning for moving patients to outdoor and traffic management policy launching.

2. Literature Review

Scientific studies about $PM_{2.5}$ concentration in several countries found that it related to meteorology and their meteorology depended on their seasonality and geography. So, this research focuses on using them to be the predictors.

2.1 Meteorology

The research in major cities in China found the relationship between $PM_{2.5}$ concentration and meteorological factors in terms of spatial and seasonal variations. In the same way, meteorological factors are associated with the location of the cities and season (Yang et al., 2017).

2.1.1 Wind

$PM_{2.5}$ particles can be carried by the wind from changes in wind direction. This also linked to daily pollution (Deryugina, Heutel, Miller, Molitor, & Reif, 2016). Same as in Great Cairo, wind direction has a correlation with pollutant concentrations and the high pollutant associated with low wind speed. Moreover, dust and sand from the surrounding desert were carried by the wind at higher wind speeds (Elminir, 2005). Another research in Beijing also found that wind speed decreased continually while the pollution level increased (Ma et al., 2017). In Sydney, temperature and wind speed reflected $PM_{2.5}$ concentration pattern during the time of day. It means that average temperature was cooler and wind speed was lower during mornings and evenings when air quality was poor (Virgilio, Hart, Jiang, & Physics, 2018).

2.1.2 Relative Humidity and Temperature

The relative humidity is a percentage of water vapor in the air. Relative humidity 100% means air completely saturated and cannot hold more water molecules at constant pressure. If air temperature increases, air can hold more water molecules. Therefore, temperature relates to the amount of moisture the atmosphere can hold. Furthermore, there is a positive correlation between daily average $PM_{2.5}$ concentration and relative humidity in traffic-based areas and a negative correlation in industry-based areas. The contrast of $PM_{2.5}$ concentration decreasing and relative humidity increasing might result from rain accumulation (Zalakeviciute, López-Villada, & Rybarczyk, 2018).

2.1.3 Rain

Periods of high and low heavy rain have a positive correlation between rainfall and relative humidity in Uyo (Umoh, Akpan, & Jacob, 2013). Rainfall and relative humidity control the daily variations of $PM_{2.5}$ (Hien, Bac, Tham, Nhan, & Vinh, 2002). Rain also effects on wiping out particles, especially fine particles which six times lower than the concentration before the rain (Hu et al., 2006).

2.2 Seasonality

$PM_{2.5}$ concentration was characterized by daily and seasonal concentration. The lowest value was measured in summer, followed by the spring and the highest one is in winter. It may correlate with the high level of precipitation high wind speed in winter, high temperature in summer and differences between atmospheric condition (Capatina et al., 2016). Another research found the same result that $PM_{2.5}$ concentration reached the highest level during winter and drop to the lowest level in summer (Wang, Zhou, Wang, Feng, & Hubacek, 2017). In Beijing, $PM_{2.5}$ concentration increased in winter haze periods (Ma et al., 2017) and also in California. Winter meteorology is probably key factors for the winter-nonwinter contrasts also (Chow et al., 2006).

2.3 Geography

$PM_{2.5}$ concentration depends on each geography. When considering the north of Thailand, it is a basin area covered by mountains. So $PM_{2.5}$ will be detained when high air pressure spreads from China. Even if Bangkok has much more vehicles and skyscrapers than Chiang Mai, but it is a plain area, nearby Gulf of Thailand. It receives more rain and wind from southwesterly monsoon in rainy season from May to October. They wipe out $PM_{2.5}$ particles. Moreover, the most rainfall volume is in September. In summer from March to April, there are no monsoon and heat low-pressure covers upper Thailand. It makes hot and stuffy weather but there are still southerly and southeasterly wind. Due to the above reasons, $PM_{2.5}$ will not be accumulated in Bangkok except in winter from November to February because northeasterly monsoon from China makes weather cooler, dry and high pressure (Thailand Climatological Center, 2017).

2.4 Times Series Model

Time Series with Multiple Seasonal can be predicted by various models. For example, in Kentucky Nonlinear Regression (NLR) and Back-Trajectory Model. The model is designed for use in summer when air quality is quite critical for human health (Cobourn, 2010). Multi-layer perception (MLP) artificial neural network (ANN) and Neuro-Fuzzy models are also used to predict $PM_{2.5}$ concentration (McKendry & Association, 2002; Mishra, Goyal, & Upadhyay, 2015). The classic and most used model is ARIMA model, also known as Box-Jenkins model. The model analyzes discrete time series, considers autoregressive integrated moving average and various extensions of the model (Box, Jenkins, Reinsel, & Ljung, 2015). ARIMA models are adjusted for a conditional variance to be ARCH and GARCH models. A two-pass estimation method is another method can be used to estimate GARCH model. It uses Regression Analysis with ordinary least square (OLS) and then saves its residuals to analyze as an observed time series. This estimation provides good approximations when the sample size is moderate or large (Tsay, 2014). Another complex seasonality model is TBATS, a combination of Fourier terms with an exponential smoothing state-space model and a Box-Cox transformation. It is allowed to change over time and suitable with long time series (De Livera, Hyndman, & Snyder, 2011).

3. Research Methodology

3.1 Scope

This research is supported $PM_{2.5}$ concentration and meteorology data by Pollution Control Department of Thailand. Data period is from 1 January 2017 to 31 December 2018 by hourly. The meteorological factors include wind speed, wind direction, temperature, relative humidity and rain.

3.2 Data

There are 11 weather stations in Bangkok but only 4 stations have data from 1 January 2017 to 31 December 2018. They contain (05t) Thai Meteorological Department: Bangna, (52t) Metropolitan Electricity Authority Substation: Thonburi, (59t) The Government Public Relations Department: Phrayathai and (61t) Bodindecha (Sing Singhaseni) School: Ramkhamhaeng. Station 05t, 52t, 59t contain wind speed, wind direction, temperature, relative humidity and rain as meteorological factors but station 61t has only wind speed, wind direction and temperature. In addition, all stations have missing data.

Station	Variable	Valid	Missing	Mean	SD	Min	Max	Skewness	Kurtosis
05t	$PM_{2.5}$	15979	1541	23.30	17.40	1.00	173.00	1.56	3.39
	Wind speed	17433	87	1.41	0.86	0.10	5.00	0.36	-0.64
	Wind direction	17433	87	187.56	89.88	0.00	360.00	-0.11	-0.80
	Temperature	17432	88	28.45	2.85	16.10	37.70	-0.03	-0.11
	Relative humidity	17432	88	72.53	15.17	26.00	99.00	-0.16	-0.55
	Rain	17427	93	0.21	1.87	0.00	82.00	16.74	405.36
52t	$PM_{2.5}$	17024	496	29.95	19.55	1.00	148.00	1.58	3.21
	Wind speed	17387	133	0.78	0.72	0.00	4.10	0.75	-0.43
	Wind direction	17342	178	218.50	63.39	0.00	360.00	-1.53	1.16
	Temperature	17454	66	29.16	2.74	17.40	38.50	0.09	0.15
	Relative humidity	17454	66	66.57	14.87	20.00	99.00	-0.25	-0.43
	Rain	17447	73	0.22	2.02	0.00	78.00	16.71	379.69
59t	$PM_{2.5}$	14909	2611	22.91	16.12	1.00	168.00	1.55	4.19
	Wind speed	16862	658	0.60	0.31	0.00	2.20	0.78	0.51
	Wind direction	16858	662	184.87	110.68	0.00	360.00	-0.27	-1.47
	Temperature	16958	562	28.35	2.99	16.50	38.00	0.14	-0.18
	Relative humidity	16957	563	71.54	16.98	20.00	99.00	-0.12	-0.74
	Rain	16955	565	0.26	2.32	0.00	84.40	16.60	366.22
61t	$PM_{2.5}$	16046	1474	25.44	16.64	1.00	149.00	1.66	3.86
	Wind speed	17163	357	0.89	0.60	0.00	3.70	0.60	-0.05
	Wind direction	17288	232	173.74	51.72	58.00	319.00	0.83	-0.29
	Temperature	17287	233	29.02	3.01	12.70	38.00	-0.22	0.13

Table 1 Descriptive statistics of station 05t, 52t, 59t and 61t raw data

This research uses linear interpolation for one to three adjacent missing data to fill the gaps because these data points vary depends on the time that is next to its time. However, in case of more than three adjacent missing data, hot deck imputation by manually replacing missing data with previous data in the same period is used to support seasonal data patterns.

Station	Variable	Valid	Missing	Mean	SD	Min	Max	Skewness	Kurtosis
05t	$PM_{2.5}$	17520	0	23.24	17.33	1.00	173.00	1.56	3.39
	Wind speed	17520	0	1.41	0.86	0.10	5.00	0.36	-0.64
	Wind direction	17520	0	187.45	89.79	0.00	360.00	-0.11	-0.80
	Temperature	17520	0	28.46	2.85	16.10	37.70	-0.03	-0.11
	Relative humidity	17520	0	72.53	15.15	26.00	99.00	-0.16	-0.54
	Rain	17520	0	0.21	1.87	0.00	82.00	16.77	407.13
52t	$PM_{2.5}$	17520	0	30.02	19.57	1.00	148.00	1.56	3.09
	Wind speed	17520	0	0.78	0.72	0.00	4.10	0.76	-0.42
	Wind direction	17520	0	218.34	63.43	0.00	360.00	-1.52	1.14
	Temperature	17520	0	29.16	2.74	17.40	38.50	0.10	0.16

Station	Variable	Valid	Missing	Mean	SD	Min	Max	Skewness	Kurtosis
	Relative humidity	17520	0	66.58	14.86	20.00	99.00	-0.25	-0.43
	Rain	17520	0	0.22	2.01	0.00	78.00	16.75	381.27
59t	$PM_{2.5}$	17520	0	23.45	16.60	1.00	168.00	1.46	3.60
	Wind speed	17520	0	0.60	0.31	0.00	2.20	0.79	0.53
	Wind direction	17520	0	184.74	110.77	0.00	360.00	-0.27	-1.47
	Temperature	17520	0	28.37	2.98	16.50	38.00	0.14	-0.17
	Relative humidity	17520	0	71.62	16.92	20.00	99.00	-0.12	-0.74
	Rain	17520	0	0.25	2.28	0.00	84.40	16.83	376.99
61t	$PM_{2.5}$	17520	0	26.32	17.24	1.00	149.00	1.57	3.27
	Wind speed	17520	0	0.90	0.60	0.00	3.70	0.61	-0.03
	Wind direction	17520	0	173.48	51.70	58.00	319.00	0.83	-0.28
	Temperature	17520	0	29.02	3.01	12.70	38.00	-0.21	0.12

Table 2 Descriptive statistics of station 05t, 52t, 59t and 61t after data imputation

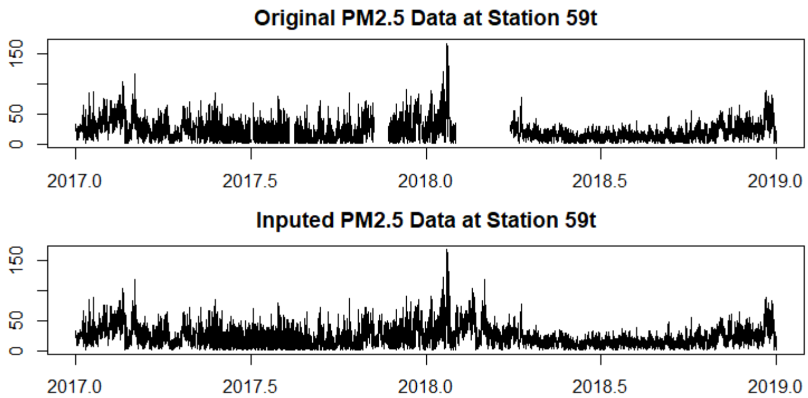


Figure 1 Comparison of original $PM_{2.5}$ data and imputed $PM_{2.5}$ data at station 59t

3.3 Methodology and Research Design

Two-Pass Estimation is used to estimate the model because the data characteristic is multi-seasonality times series so dummy variables are needed. Furthermore, there is no trend in $PM_{2.5}$ concentration (Figure 1) so it is not needed to do differencing.

First, this estimation uses Regression Analysis with ordinary least square (OLS) that is included main effect and interaction effect into Regression Model then saves its residuals as e_t.

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_t x_{pt} + \sum_{m=1}^{M} \gamma_m d_{m,t} + \sum_{m=1}^{M} \gamma_m d_{m,t} \cdot \overline{ws} + \sum_{h=1}^{H} \delta_h d_{h,t} \cdot \overline{ws} + e_t \tag{1}$$

Where y_t is the multiple time series of time $t = 1, 2, \ldots, T$ and predictor $p = 1, 2, \ldots, P$

γ_m is the constant of dummy variable of month $m = 1, 2, \ldots, M$

δ_h is the constant of dummy variable of hour $h = 1, 2, \ldots, H$

$d_{m,t} = 1$ if t is true for m and 0 otherwise

$d_{h,t} = 1$ if t is true for h and 0 otherwise

$\overline{ws}$ is the average of daily wind speed as interaction effect

e_t is the error term

Second, treats $\{e_t\}$ as an observed time series by using ARMA Model

$$e_t = \alpha_0 + \sum_{i=1}^{p} \emptyset_i e_{t-1} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{2}$$

The analysis accuracy of Regression Model uses T-test and R Square. In case of ARMA Model is compared by T-test, residual ACF, residual PACF and Mean Absolute Percentage Error (MAPE).

4. Result and Analysis

4.1 Data Exploration

Due to $PM_{2.5}$ characteristics, this research visualizes seasonal effect of all stations. However, due to many data dimensions, this research shows only significant data. From table 3, the highest averages of $PM_{2.5}$ concentration are observed in February but the lowest averages are found scattered randomly in rainy season. However, the lowest averages mostly found in June and the most rainfall volume is in September. So, this research focuses on data visualization in February, June and September at stations with low missing data.

2017	05t	52t	59t	61t	2018	05t	52t	59t	61t
Jan	40	42	30	37	Jan	32	50	38	40
Feb	**47**	**49**	**40**	**48**	Feb	**41**	**55**	**40**	**47**
Mar	29	31	29	29	Mar	31	32	29	32
Apr	28	27	24	24	Apr	21	23	22	28
May	15	22	23	19	May	**13**	17	14	19
Jun	14	**20**	18	15	Jun	13	**16**	**12**	16
Jul	**10**	21	18	15	Jul	16	18	14	17
Aug	11	21	18	**14**	Aug	14	16	13	16
Sep	14	23	**14**	16	Sep	16	19	13	**16**
Oct	20	34	22	25	Oct	20	24	18	18
Nov	23	39	26	31	Nov	30	35	27	30
Dec	28	46	32	40	Dec	35	42	32	41

Table 3 Average of $PM_{2.5}$ concentration by year and month

4.1.1 Hourly Effect

Hourly effect is obviously observed on several days that causes high $PM_{2.5}$ concentration in rush hour, moreover, it leverages at 8-9 AM and after office hours. This effect may correlate with work commuting of white-collar workers and a government policy that allows trucks to transport at night only. In addition, the hourly effect seems flat on days with low $PM_{2.5}$ concentration.

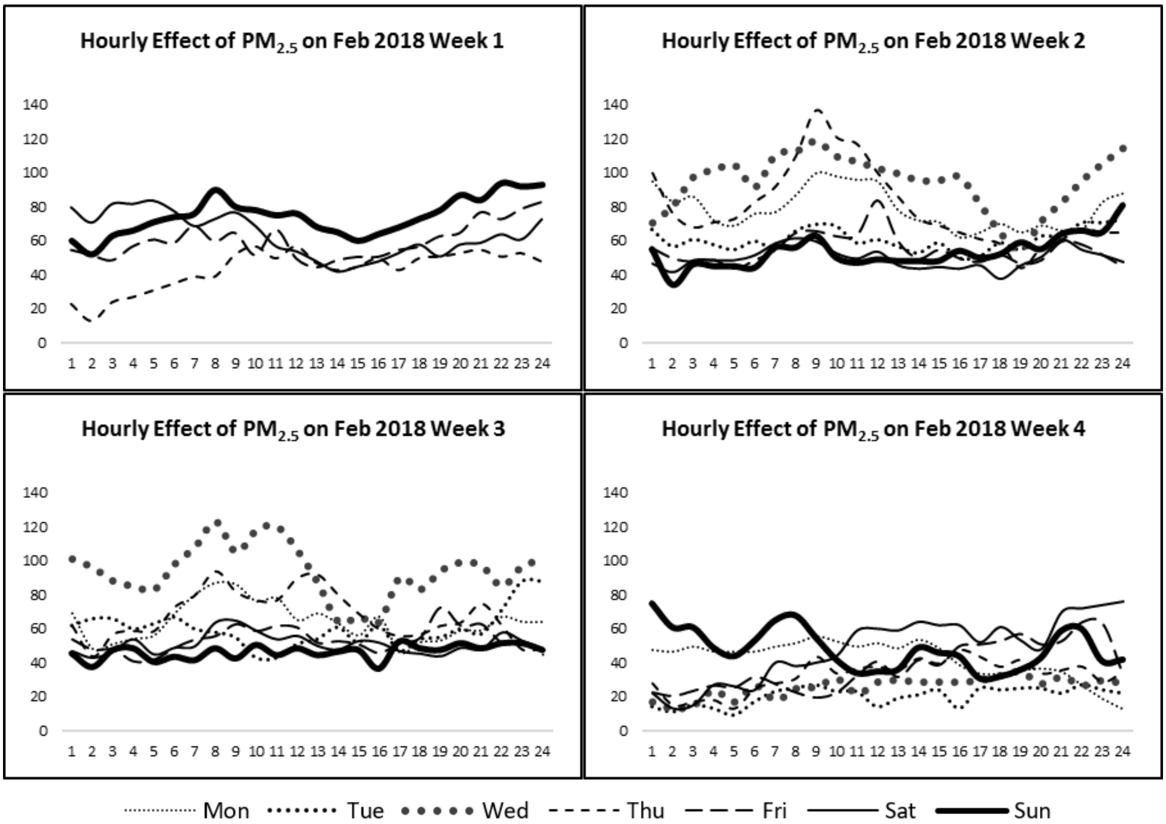


Figure 2 Hourly effect of $PM_{2.5}$ concentration in February 2018 at station 52t

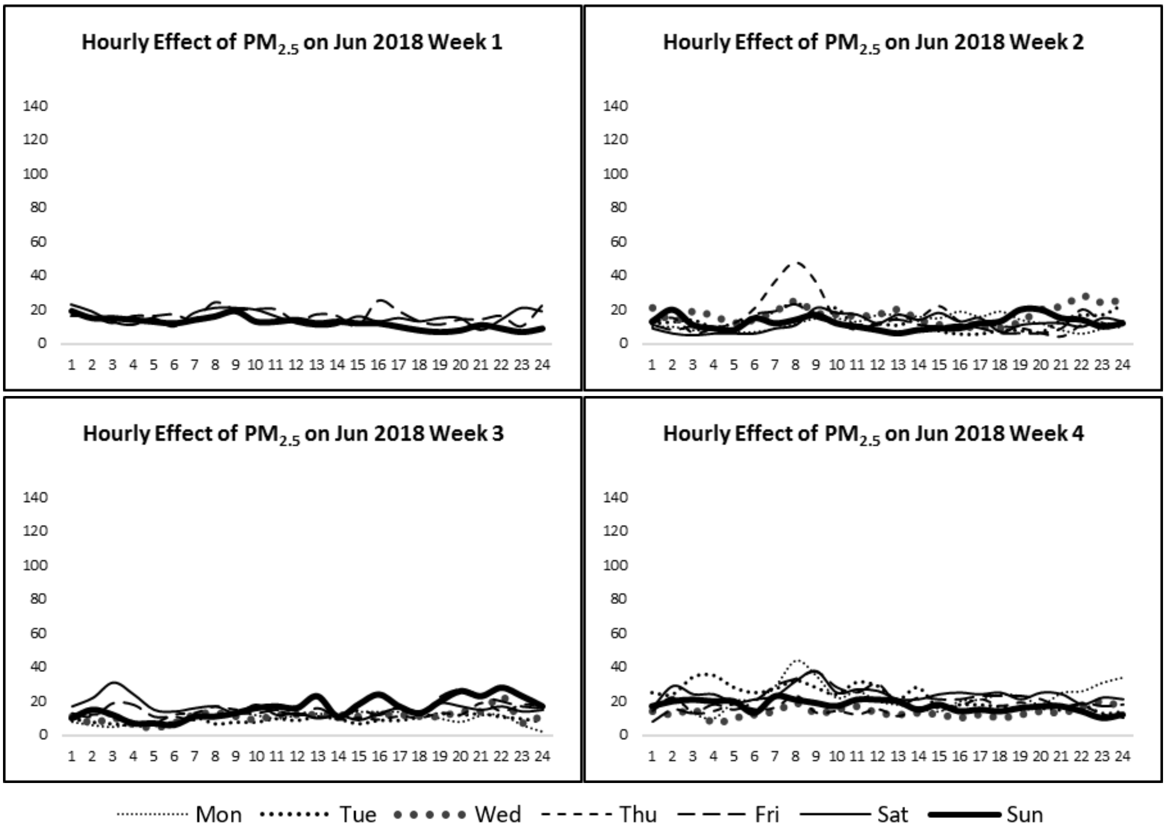


Figure 3 Hourly effect of $PM_{2.5}$ concentration in June 2018 at station 52t

4.1.2 Daily Effect

Daily effect is not clear in any weeks with both high and low $PM_{2.5}$ concentration.

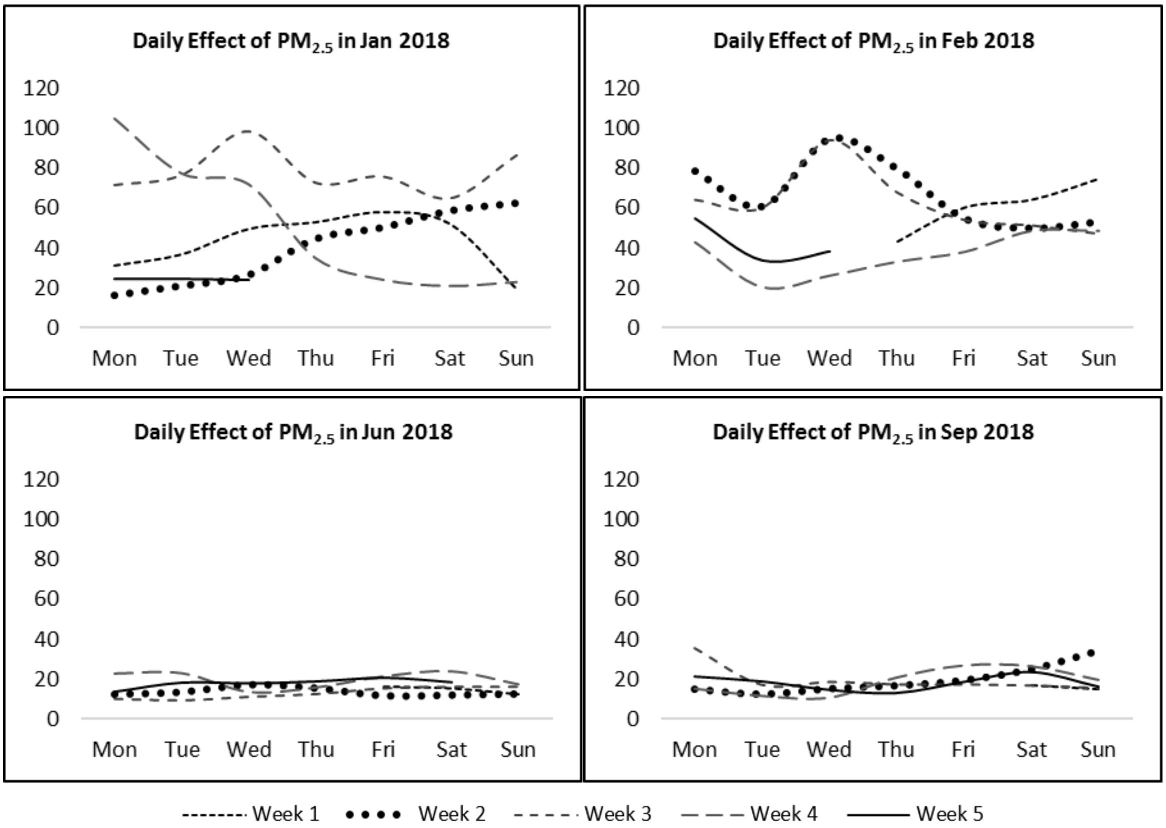


Figure 4 Daily effect of $PM_{2.5}$ concentration at station 52t

4.1.3 Weekly Effect

Weekly effect is also not clear in any months with both high and low $PM_{2.5}$ concentration. Then, separates weekly effect line chart into three charts by season to visualize easily.

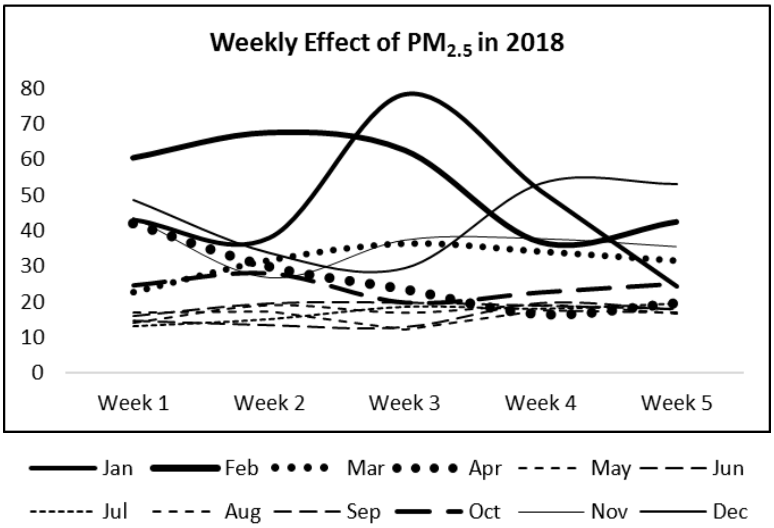


Figure 5 Weekly effect of $PM_{2.5}$ concentration at station 52t

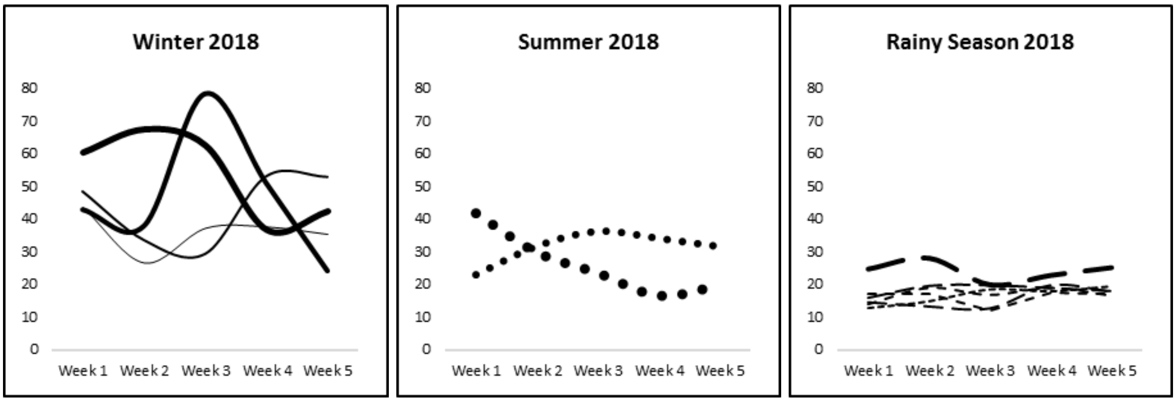


Figure 6 Weekly effect of $PM_{2.5}$ concentration at station 52t by season

4.1.4 Monthly Effect

Monthly effect causes a similar pattern for each station. It dramatically rises from November to February, definitely drops in March to April and slightly goes up and down around May to October. This effect may correlate with the season of Thailand; winter, summer and rainy season respectively.

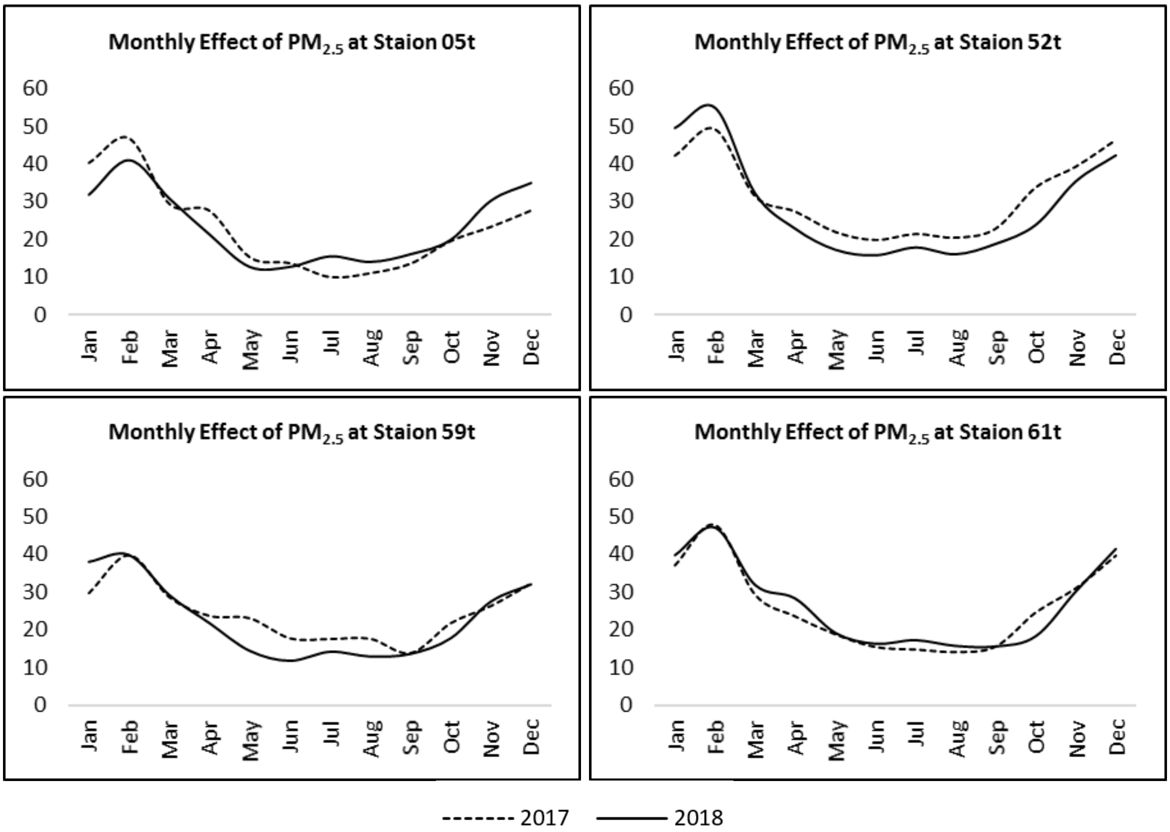


Figure 7 Monthly effect of $PM_{2.5}$ concentration

4.1.5 Dummy Variable

Create seven dummy variables to encode eight categories of hourly effect following $PM_{2.5}$ concentration pattern.

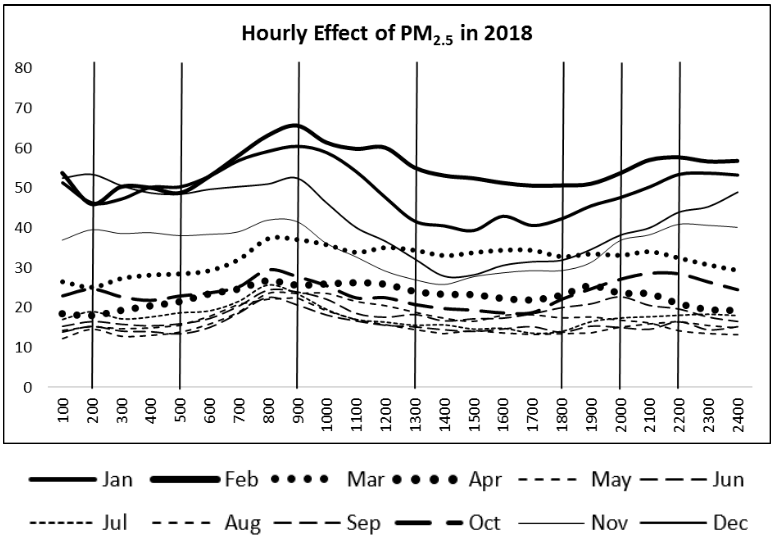


Figure 8 Dummy variables of hourly effect

Then, create six dummy variables to encode seventh categories of monthly effect by using season of Thailand. Summer starts around mid of Februry to mid of May, rainy season starts from mid of May to mid of October and winter starts from mid of October to mid of February. (Thailand Climatological Center, 2017).

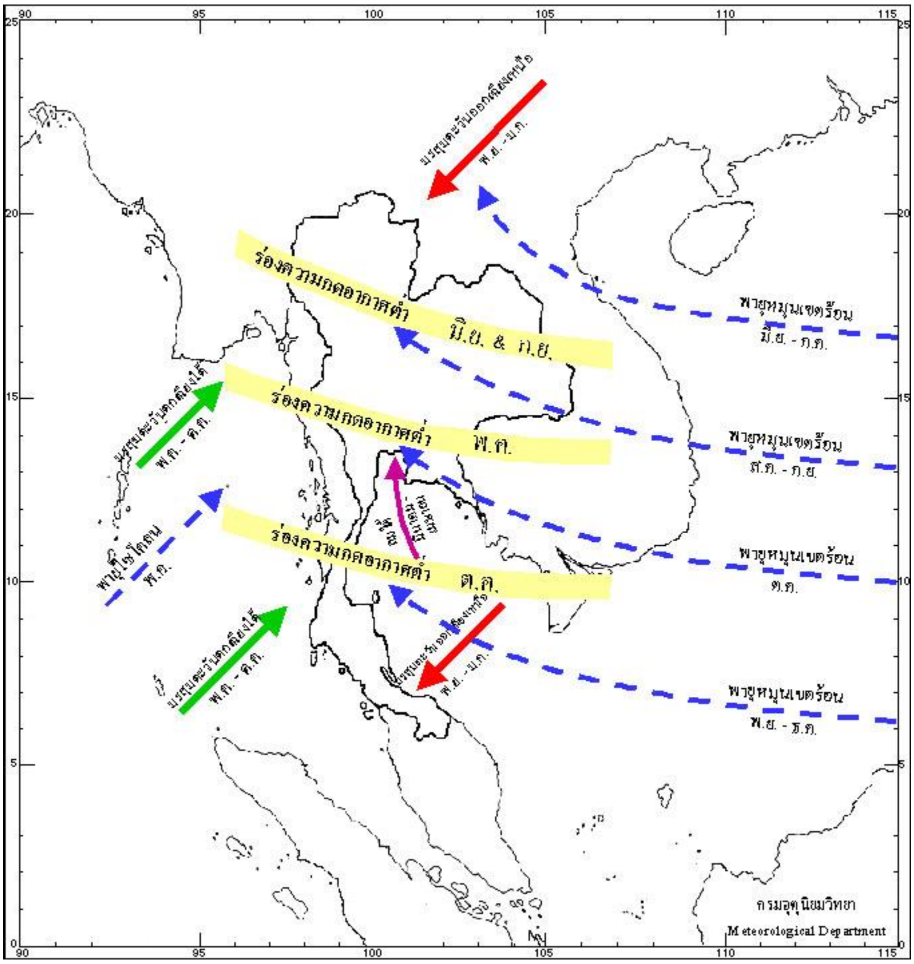


Figure 9 Intertropical Convergence Zone direction, Monsoon direction and Cyclone direction

Month	Season	Monsoon direction	Cyclone	Intertropical Convergence Zone	Dummy Variable
Jan	Winter	NE	-	-	M1
Feb	Winter	-	-	-	M2
Mar	Summer	-	-	-	M3
Apr	Summer	-	-	-	M3
May	Rainy	SW	Yes	Yes	M4
Jun	Rainy	SW	Yes	Yes	M4
Jul	Rainy	SW	Yes	-	M5
Aug	Rainy	SW	Yes	-	M5
Sep	Rainy	SW	Yes	Yes	M6
Oct	Rainy	SW	Yes	Yes	M6
Nov	Winter	NE	Yes	-	Otherwise
Dec	Winter	NE	Yes	-	Otherwise

Table 4 Season of Thailand separated by month and its characteristics

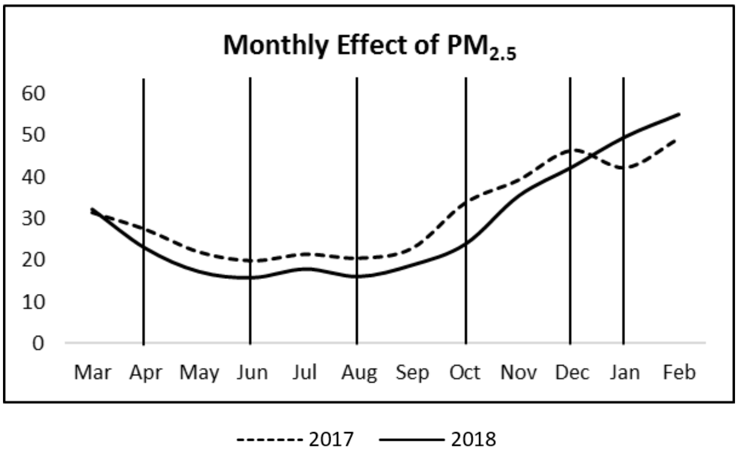


Figure 10 Dummy variables of monthly effect

4.1.6 Interaction Effect

$PM_{2.5}$ concentration of each station highly correlates with different meteorological factors by considering correlation matrix of all stations.

	$PM_{2.5}$	-.36	-.16	-.14	-.26	-.04	
	-.20	**Wind speed**	.32	.38	-.14	.02	
Station 05t	-.07	-.09	**Wind direction**	.04	.18	-.02	Station 52t
	-.18	.53	.03	**Temp**	-.67	-.13	
	-.22	-.49	.10	-.60	**Relative humidity**	.16	
	-.05	.04	.04	-.13	.16	**Rain**	

	$PM_{2.5}$	-.03	.05	-.19			
	-.09	**Wind speed**	-.08	.35			
Station 59t	-.21	.20	**Wind direction**	-.07			Station 61t
	-.16	.32	.22	**Temp**			
	-.15	-.36	.10	-.64	**Relative humidity**		
	-.03	-.04	.02	-.11	.14	**Rain**	

Table 5 Correlation matrix by each station

This research also considers the correlation between $PM_{2.5}$ concentration and meteorological factor and found that there is a correlation between $PM_{2.5}$ concentration and wind speed (Table 6 and Table 7). Therefore, the interaction effect between dummy variables and wind speed is considered in this research.

Dummy	H1	H2	H3	H4	H5	H6	H7	Otherwise
(Hour)	01:00-02:00	03:00-05:00	06:00-09:00	10:00-13:00	14:00-18:00	19:00-20:00	21:00-22:00	23:00-24:00
Station 05t	-.20	-.15	-.08	-.16	-.21	-.33	-.24	-.23
Station 52t	-.38	-.35	-.30	-.42	-.34	-.43	-.44	-.41
Station 59t	-.01	.01	-.04	-.11	-.05	-.09	-.09	-.07
Station 61t	-.07	.01	.08	.09	.08	-.11	-.09	-.07

Table 6 Correlation matrix between $PM_{2.5}$ concentration and wind speed by hourly

Dummy	M1	M2	M3	M4	M5	M6	Otherwise
(Month)	Jan	Feb	Mar-Apr	May-Jun	Jul-Aug	Sep-Oct	Nov-Dec
Station 05t	-.43	-.29	-.24	-.16	-.22	-.12	-.35
Station 52t	-.41	-.39	-.31	-.14	-.14	-.25	-.16
Station 59t	-.30	-.18	-.12	-.20	-.13	-.10	-.02
Station 61t	-.44	-.25	-.21	-.02	-.10	-.06	-.27

Table 7 Correlation matrix between $PM_{2.5}$ concentration and wind speed by monthly

4.2 Model

$y_t = PM_{2.5} concentration$

Block 1: Main effect of Time Series

$$wind\ speed, wind\ direction, temperature, relative\ humidity\ and\ rain$$

Block 2: Main effect of month

$$six\ dummy\ variables\ of\ monthly\ effect$$

Block 3: Interaction effect of Monthly Effect and wind speed

$$six\ dummy\ variables\ of\ monthly\ effect \times average\ of\ daily\ wind\ speed$$

Block 4: Interaction effect of hourly effect and wind speed

$$seven\ dummy\ variables\ of\ hourly\ effect \times average\ of\ daily\ wind\ speed$$

Block 5: ARMA(p,q)

Regression Model of each station has different coefficients but similar ARMA Model which is ARMA(1,1) and ARMA(24,1). It may correlate with hourly data; 24 hours per day.

4.3 Model Summary

Model	Variable	Station 05t					Station 52t					Station 59t					Station 61t				
		B	S.E.	t	*p*		B	S.E.	t	*p*		B	S.E.	t	*p*		B	S.E.	t	*p*	
	(Constant)	96.09	2.11	45.61	.000		106.15	2.55	41.65	.000		67.25	2.37	28.44	.000		43.39	1.35	32.23	.000	
Block 1	Wind speed	-4.06	.18	-22.65	.000	R^2	-3.64	.25	-14.52	.000	R^2	-.30	.51	-.59	.558	R^2	-3.48	.24	-14.25	.000	R^2
	Wind direction	.01	.00	6.58	.000	.25	.03	.00	17.12	.000	.31	.00	.00	.95	.340	.15	.03	.00	13.97	.000	.04
	Temperature	-1.22	.06	-21.39	.000	ΔR^2	-1.66	.07	-24.50	.000	ΔR^2	-.87	.07	-13.32	.000	ΔR^2	-.31	.04	-6.91	.000	ΔR^2
	Rel. humidity	-.40	.01	-38.12	.000	.25	-.39	.01	-29.97	.000	.31	-.19	.01	-16.89	.000	.15					.04
	Rain	.04	.05	.77	.444	***P*** .000	-.03	.06	-.58	.564	***P*** .000	-.05	.05	-1.09	.275	***P*** .000					***P*** .000
Block 2	M1 (Jan)	37.94	1.20	31.58	.000	R^2	14.41	.64	22.37	.000	R^2	38.54	1.64	23.43	.000	R^2	27.77	1.30	21.30	.000	R^2
	M2 (Feb)	31.97	1.37	23.35	.000	.41	24.51	.84	29.25	.000	.41	12.91	1.51	8.54	.000	.26	13.55	1.40	9.68	.000	.39
	M3 (Mar-Apr)	13.94	.95	14.71	.000	ΔR^2	4.67	.73	6.44	.000	ΔR^2	1.53	1.00	1.54	.125	ΔR^2	.27	.87	.31	.760	ΔR^2
	M4 (May-Jun)	-1.27	.91	-11.27	.000	.16	-12.76	.77	-16.57	.000	.10	-3.27	.79	-4.13	.000	.11	-2.20	.80	-25.23	.000	.35
	M5 (Jul-Aug)	-15.97	1.05	-15.23	.000	***P***	-13.53	1.18	-11.51	.000	***P***	-6.98	1.04	-6.72	.000	***P***	-2.62	.85	-24.16	.000	***P***
	M6 (Sep-Oct)	-5.56	.89	-6.26	.000	.000	-2.55	.62	-4.10	.000	.000	-2.84	1.13	-2.50	.012	.000	-22.38	.85	-26.34	.000	.000
Block 3	M1xAvgWS	-25.90	.92	-28.03	.000	R^2	-24.36	1.23	-19.88	.000	R^2	-56.59	2.62	-21.59	.000	R^2	-25.52	1.27	-2.03	.000	R^2
	M2xAvgWS	-12.06	.94	-12.85	.000	.45	-19.79	1.09	-18.10	.000	.44	-5.22	2.29	-2.28	.022	.28	-.18	1.24	-.14	.888	.41
	M3xAvgWS	-4.61	.53	-8.74	.000	ΔR^2	-1.74	.71	-15.22	.000	ΔR^2	-3.49	1.37	-2.55	.011	ΔR^2	-4.00	.70	-5.75	.000	ΔR^2
	M4xAvgWS	.82	.59	1.38	.167	.04	-3.52	.78	-4.50	.000	.03	-12.52	1.30	-9.60	.000	.02	2.87	.93	3.09	.002	.02
	M5xAvgWS	3.91	.68	5.77	.000	***P***	-2.76	.98	-2.82	.005	***P***	-5.43	1.40	-3.88	.000	***P***	-.23	1.03	-.22	.824	***P***
	M6xAvgWS	-.95	.68	-1.39	.164	.000	-14.45	.90	-16.10	.000	.000	-14.76	2.23	-6.61	.000	.000	6.47	1.00	6.45	.000	.000
Block 4	H1xAvgWS	-2.26	.31	-7.32	.000		-1.66	.57	-2.92	.003		.90	.81	1.12	.264		-2.17	.49	-4.41	.000	
	H2xAvgWS	-3.04	.28	-1.78	.000	R^2	-1.67	.52	-3.19	.001	R^2	-.08	.74	-.11	.916	R^2	-3.32	.45	-7.36	.000	R^2
	H3xAvgWS	-.88	.27	-3.33	.001	.46	4.36	.49	8.87	.000	.45	.52	.70	.74	.457	.29	-.30	.42	-.70	.485	.42
	H4xAvgWS	-1.78	.30	-6.02	.000	ΔR^2 .01	3.42	.52	6.64	.000	ΔR^2 .01	-7.82	.75	-1.42	.000	ΔR^2 .01	-.16	.46	-.34	.734	ΔR^2 .01
	H5xAvgWS	-2.89	.29	-9.82	.000	***P***	1.16	.52	2.26	.024	***P***	-5.79	.74	-7.79	.000	***P***	-1.91	.45	-4.26	.000	***P***
	H6xAvgWS	-2.51	.31	-8.04	.000	.000	1.44	.57	2.50	.012	.000	-2.58	.81	-3.18	.001	.000	-1.34	.50	-2.70	.007	.000
	H7xAvgWS	-1.43	.31	-4.64	.000		1.63	.57	2.86	.004		-.57	.80	-.71	.475		.49	.49	.99	.320	
Block 5	AR Lag 1	.84	.01	167.92	.000	R^2 .84	.85	.01	182.84	.000	R^2 .86	.76	.01	137.81	.000	R^2 .84	.86	.00	195.70	.000	R^2 .87
	AR Lag 24	.07	.00	16.12	.000	ΔR^2 .38	.07	.00	17.12	.000	ΔR^2 .41	.10	.01	19.66	.000	ΔR^2 .55	.07	.00	18.28	.000	ΔR^2 .45
	MA Lag 1	.07	.01	7.22	.000	***P*** .000	.06	.01	6.68	.000	***P*** .000	-.29	.01	-34.62	.000	***P*** .000	.07	.01	7.94	.000	***P*** .000

Table 8 Model summary of all stations

All blocks of each station are significant at the .01 level but some variables in each block are not.

At station 05t: Thai Meteorological Department: Bangna

Wind speed, wind direction, temperature and relative humidity are significant at the .01 level. Monthly effect strongly influences $PM_{2.5}$ concentration, The highest values are M1 and M2 which represent January and February. The regression model can describe $PM_{2.5}$ concentration around 46% and reach 84% when using along with ARMA model.

At station 52t: Metropolitan Electricity Authority Substation: Thonburi

The result of station 52t is likely similar to station 05t. Wind speed, wind direction, temperature and relative humidity are significant at the .01 level. Monthly effect strongly influences $PM_{2.5}$ concentration in February. The regression model can describe $PM_{2.5}$ concentration around 45% and reach 86% when using along with ARMA model.

At station 59t: The Government Public Relations Department: Phrayathai

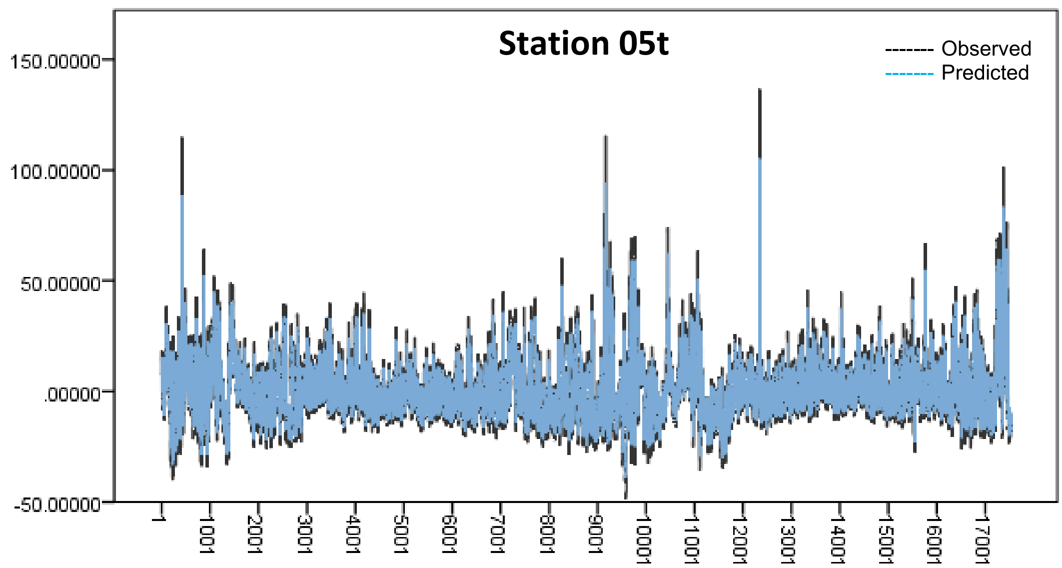
The result of station 59t is quite different from the other, maybe a result of missing data 15% high. Temperature and relative humidity are significant at the .01 level. Wind speed seems does not effect $PM_{2.5}$ concentration but the temperature still has an effect to. Monthly effect strongly influences $PM_{2.5}$ concentration in January. The regression model can describe $PM_{2.5}$ concentration just 29% but ARMA model reaches the remains 84%.

At station 61t: Bodindecha (Sing Singhaseni) School: Ramkhamhaeng.

This station does not have relative humidity and rain data. Wind speed, wind direction and temperature are significant at the .01 level. Monthly effect strongly influences $PM_{2.5}$ concentration in January. The regression model can describe $PM_{2.5}$ concentration around 42% and reach 87% when using along with ARMA model.

4.4 Fit Data

These charts are comparisons between observed data and predicted data at each station. By looking at these charts, it seems predicted data fit properly on observed data but range of predicted data is shorter than observed data.



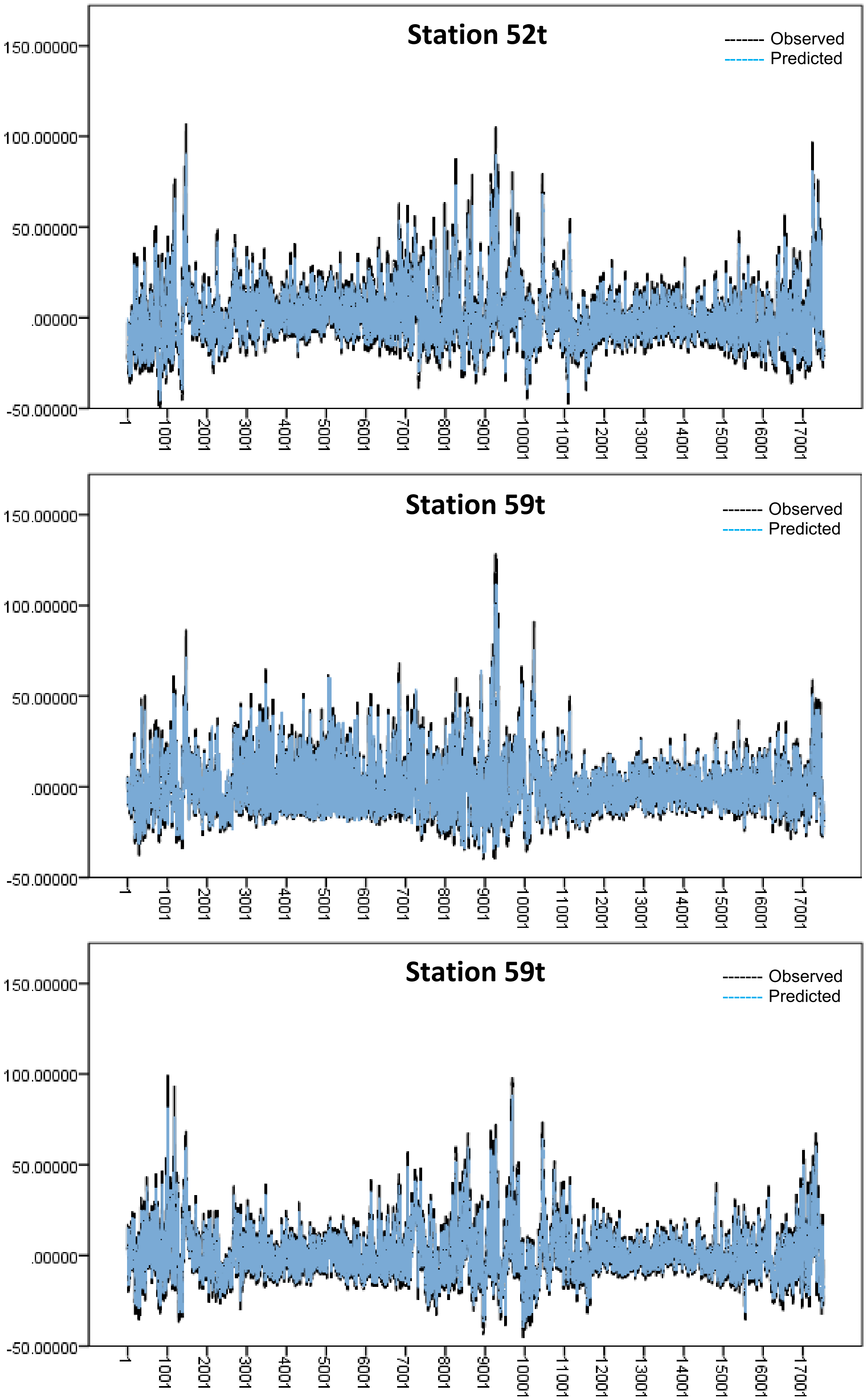


Figure 11 Comparison between observed data and fit data of each station

5. Conclusion and Recommendation

Wind speed, wind direction, temperature and relative humidity are significantly related to $PM_{2.5}$ concentration. They predict $PM_{2.5}$ concentration around 25% (R^2 = 0.25). Wind speed, temperature and relative humidity have a negative relation with $PM_{2.5}$ concentration. Wind speed is a meteorological factor that has high negative relation than other meteorological factors because southwesterly wind blows which carries $PM_{2.5}$ particles out of Bangkok.

When considering monthly effect, it increases the prediction 16% (ΔR^2 = 0.16). It strongly influences on $PM_{2.5}$ concentration that relates to the season of Thailand and the highest months are January and/or February. Furthermore, during winter from November to February southwesterly wind is weak and Bangkok is covered with high air pressure areas. This promotes $PM_{2.5}$ accumulation and shows hourly effect obviously in that month especially at rush hours and nighttime because of work commuting of white-collar workers and a government policy that allows trucks to transport at night only.

Relative humidity and temperature are related together. They also have a negative relation with $PM_{2.5}$ concentration because a collision between relative humidity and $PM_{2.5}$ particles typically results in coagulation. However, water-spraying into the atmosphere cannot reduce $PM_{2.5}$ concentration because a nozzle forms water aerosol with a more than 10 micrometers particle size (Azarov, Zhukova, & Antonov, 2017) so its size is too large to coagulate with $PM_{2.5}$ particle. This reason is supported by rain and is not significantly related to $PM_{2.5}$ concentration.

All stakeholders i.e. public sector, private sectors and individuals should participate to solve $PM_{2.5}$ issues with both quick-win and long-term action plans. The quick-win solutions should be preparing face masks and air purifiers before November. From December to February, there should be carpool and truck transportation reducing policy and announcement to refrain from outdoor activities. It may start solving the problems to targeted sensitive groups (including children, pregnant women, patients and weakness people especially cancer patients and chronic-respiratory-disease patients) with an initial plan such as having a safety procedure for moving patients to outdoor. There should be long-term solutions also such as having a tax credit for electric vehicles, managing traffic flow to avoid congestions by using traffic data, applying Certificate of Entitlement and encouraging people to use public transportation simultaneously. For the farmer, the public sector should provide financial assistance and infrastructure facility to encourage them not to burn crop residue to prevent air pollution.

In addition, monitoring current $PM_{2.5}$ concentration to compare with the model for signaling an unusual situation if the concentration is higher especially from December to February, $PM_{2.5}$ must be monitored intentionally. In case of emergency, all stakeholders must make a decision immediately, schools should be closed to prevent children's health risks and crop residue burning should be prohibited.

This research considers only variables and temporal autocorrelation. For the future work, Spatial-Temporal Regression best suits for $PM_{2.5}$ concentration predicting because it considers all variables, temporal autocorrelation and spatial autocorrelation. However, the model required more sample size for spatial autocorrelation.

6. References

English references

Azarov, A., Zhukova, N., & Antonov, F. J. P. E. (2017). **Water-spray systems reducing negative effects of fine-dispersion dust at operator’s workplaces of machine-building industries**. 206, 1407-1414.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). **Time series analysis: forecasting and control**. John Wiley & Sons.

Brunekreef, B., & Holgate, S. T. J. T. l. (2002). **Air pollution and health**. 360(9341), 1233-1242.

Capatina, C., Simonescu, C. M., Dadalau, N., & Cirtina, D. J. R. d. C. (2016). **Comparative Study of Air Pollution with PM2.5 and PM10 in Targu-Jiu**. 67(7), 1247-1254.

Carroll, R. E. J. J. (1966). **The relationship of cadmium in the air to cardiovascular disease death rates**. 198(3), 267-269.

Cascio, W. E. J. S. o. t. t. e. (2018). **Wildland fire smoke and human health**. 624, 586-595.

Chow, J. C., Chen, L. W. A., Watson, J. G., Lowenthal, D. H., Magliano, K. A., Turkiewicz, K., & Lehrman, D. E. J. J. o. G. R. A. (2006). **PM2.5 chemical composition and spatiotemporal variability during the California Regional PM10/PM2.5 Air Quality Study (CRPAQS)**. 111(D10).

Cobourn, W. G. J. A. E. (2010). **An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations**. 44(25), 3015-3023.

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. J. J. o. t. A. S. A. (2011). **Forecasting time series with complex seasonal patterns using exponential smoothing**. 106(496), 1513-1527.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2016). **The mortality and medical costs of air pollution: Evidence from changes in wind direction**. Retrieved from

Elminir, H. K. J. S. o. t. T. E. (2005). **Dependence of urban air pollutants on meteorology**. 350(1-3), 225-237.

Godt, J., Scheidig, F., Grosse-Siestrup, C., Esche, V., Brandenburg, P., Reich, A., . . . toxicology. (2006). **The toxicity of cadmium and resulting hazards for human health**. 1(1), 22.

Hien, P. D., Bac, V. T., Tham, H. C., Nhan, D. D., & Vinh, L. D. (2002). **Influence of meteorological conditions on PM2. 5 and PM2. 5‒ 10 concentrations during the monsoon season in Hanoi, Vietnam. Atmospheric Environment**, 36(21), 3473-3484.

Hilderbrand, D. C., Der, R., Griffin, W. T., & Fahim, M. (1973). **Effect of lead acetate on reproduction**. 115(8), 1058-1065.

Hu, M., Liu, S., Wu, Z., Zhang, J., Zhao, Y., Wehner, B., & Wiedensolher, A. J. H. j. k. x. H. k. (2006). **Effects of high temperature, high relative humidity and rain process on particle size distributions in the summer of Beijing**. 27(11), 2293-2298.

Ma, Q., Wu, Y., Zhang, D., Wang, X., Xia, Y., Liu, X., . . . Wang, Y. J. S. o. t. t. e. (2017). **Roles of regional transport and heterogeneous reactions in the PM2.5 increase during winter haze episodes in Beijing**. 599, 246-253.

McKendry, I. G. J. J. o. t. A., & Association, W. M. (2002). **Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting**. 52(9), 1096-1101.

Mishra, D., Goyal, P., & Upadhyay, A. J. A. E. (2015). **Artificial intelligence based approach to forecast PM2.5 during haze episodes: A case study of Delhi, India**. 102, 239-248.

Mudipalli, A. J. I. J. o. M. R. (2007). **Lead hepatotoxicity & potential health effects**. 126(6), 518.

Pongpiachan, S., Ho, K., & Cao, J. (2014). Effects of biomass and agricultural waste burnings on diurnal variation and vertical distribution of OC/EC in Hat-Yai City, Thailand. **Asian Journal of Applied Sciences**, 7(5), 360-374.

Pongpiachan, S., Liu, S., Huang, R., Zhao, Z., Palakun, J., Kositanont, C., . . . toxicology. (2017). **Variation in day-of-week and seasonal concentrations of atmospheric PM 2.5-bound metals and associated health risks in Bangkok**, Thailand. 72(3), 364-379.

Tsay, R. S. (2014). **An introduction to analysis of financial data with R**. John Wiley & Sons.

Umoh, A. A., Akpan, A. O., & Jacob, B. B. J. I. J. o. E. (2013). **Rainfall and relative humidity occurrence patterns in uyo metropolis, Akwa Ibom state, South-South Nigeria**. 3(8), 27-31.

Virgilio, G. D., Hart, M. A., Jiang, N. J. A. C., & Physics. (2018). **Meteorological controls on atmospheric particulate pollution during hazard reduction burns**. 18(9), 6585-6599.

Wang, S., Zhou, C., Wang, Z., Feng, K., & Hubacek, K. J. J. o. c. p. (2017). **The characteristics and drivers of fine particulate matter (PM2.5) distribution in China**. 142, 1800-1809.

Xing, Y.-F., Xu, Y.-H., Shi, M.-H., & Lian, Y.-X. J. J. o. t. d. (2016). **The impact of PM2.5 on the human respiratory system**. 8(1), E69.

Yang, Q., Yuan, Q., Li, T., Shen, H., Zhang, L. J. I. j. o. e. r., & health, p. (2017). **The relationships between PM2.5 and meteorological factors in China: Seasonal and regional variations**. 14(12), 151.

Zalakeviciute, R., López-Villada, J., & Rybarczyk, Y. J. S. (2018). **Contrasted Effects of Relative Humidity and Precipitation on Urban PM2.5 Pollution in High Elevation Urban Areas**. 10(6), 2064.

Thai references

Thailand Climatological Center. (2017). **Bangkok climate**. Retrieved from http://climate.tmd.go.th/

Thailand Climatological Center. (2017). **Seasons of Thailand**. Retrieved from https://www.tmd.go.th/info/info.php?FileID=23

Thailand Pollution Control Department. (2019). **Air quality data 2557-2562 BE**. Retrieved from http://air4thai.pcd.go.th/webV2/download.php