International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2024, 14(1), 418-426.

Forecasting PM10 Caused by Bangkok's Leading Greenhouse Gas Emission Using the SARIMA and SARIMA-GARCH Model

Tanattrin Bunnag*

Faculty of Science and Social Sciences, Burapha University, Thailand. *Email: ratanan@buu.ac.th

Received: 21 September 2023 **Accepted:** 22 December 2023 **DOI:** https://doi.org/10.32479/ijeep.15275

ABSTRACT

This paper analyzes the relationship between air pollutants and the amount of PM10 measured in Bangkok. It forecasts the amount of PM10 in Bangkok by using the SARIMA and SARIMA-GARCH models to formulate policies to reduce the occurrence of PM10 and guidelines for further prevention. PM's data is from January 2008 to July 2023. First, the process is to build the SARIMA Model and SARIMA-GARCH Model Estimation. We perform model comparisons that SARIMA (3,1,3)(1,1,2)12 and SARIMA(3,1,3)(1,1,2)12-GARCH(1,1), which model gives lower MAE and RMSE values, which indicates good prediction accuracy than another model. The results show that the MAE and RMSE predictions of the SARIMA (3,1,3) (1,1,2)12 model are 15.303 and 20.839 better than those of the SARIMA (3,1,3) (1,1,2)12-GARCH (1,1) model are 17.280 and 22.677. Therefore, the SARIMA (3,1,3) (1,1,2)12 forecast results are better precise. Thus, in summary, we will choose the first model to use in forecasting for policy making. Moreover, in the study results, we found the relationship between air pollutants and PM10 in Bangkok and found that the elements of NO_2 and O_3 will require quite a lot of attention because they affect the relationship with PM10 at a moderate level.

Keywords: PM10, The SARIMA Model, The SARIMA-GARCH, Air Pollutants
JEL Classifications: Q53, Q54

1. INTRODUCTION

Pollution is introducing substances harmful to humans and other living organisms into the environment. Pollutants are harmful solids, liquids, or gases produced in higher-than-usual concentrations that reduce the quality of our environment (Manisalidis et al., 2020).

Human activity harms the environment by polluting the air and soil on which plants grow. Although the Industrial Revolution achieved great success in technology, society, and the provision of a wide range of services, it also caused the production of enormous amounts of pollutants released into the air that are harmful to human health. Undoubtedly, global environmental pollution is a multifaceted international public health issue. Social, economic, and legal concerns and lifestyle habits are involved in this crucial issue. Urbanization and industrialization are reaching unprecedented proportions and causing global dissatisfaction in our time. Air pollution from human activities is considered one of the most important public health hazards worldwide. This is because approximately 9 million people die per year (Kumar et al., 2020).

Pollution has many health effects. The health of frail and sensitive individuals can be affected even on days with low air pollution. Short-term exposure to air pollution is closely related to chronic obstructive pulmonary disease (COPD), coughing, shortness of breath, wheezing, asthma, and respiratory diseases, and high rates of hospital admissions (it is a measure of illness).

The long-term effects of air pollution are chronic asthma, pulmonary insufficiency, cardiovascular diseases, and cardiovascular mortality. Moreover, air pollution has various malign health effects in early human life, such as respiratory, cardiovascular, mental, and perinatal disorders, leading to infant mortality or chronic disease in adulthood (Kelishadi et al., 2010).

National reports have noted an increased risk of illness and death. These studies were conducted in many locations worldwide and show the relationship between the daily concentration of particulate matter (PM) and the daily death rate. Climate change and global warming of planet earth may make the situation worse. Fine and ultra-fine particles are associated with more severe illness. This is because they can invade the deepest parts of the airways and gain more accessible access to the bloodstream (Thangavel et al., 2022).

Air pollution mainly affects people living in large urban areas. Emissions from roads are the most significant contributor to air quality deterioration. Air pollution and climate change are closely related. Climate is another side of the same coin that degrades our planet. Pollutants such as black carbon, methane, tropospheric ozone, and aerosols affect the amount of sunlight that enters. As a result, the world's temperature will rise. This causes ice, icebergs, and glaciers to melt.

The World Health Organization (WHO) in 2021 reports on six major air pollutants: particle pollution, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead. Air pollution can have a disastrous effect on all components of the environment, including groundwater, soil, and air. Additionally, it poses a serious threat to living organisms. In this vein, we are interested in forecasting these pollutants, as they are related to more extensive and severe problems in human health and environmental impact. Acid rain, global warming, the greenhouse effect, and climate change have important ecological implications for air pollution.

PM10 refers to naturally occurring or artificially occurring particles (PM) suspended in the atmosphere as solid particles or gases with a radius of $<10\ \mu g/m^3$, respectively. Naturally occurring PM accounts for only a fraction of that total PM concentration when PM floats in the atmosphere. It can sometimes increase the frequency of traffic accidents and deaths in the short term and harm people's health in the long run, causing heart and lung disease, lung cancer, respiratory diseases, and stroke. It is reported that for every $10\ \mu g/m^3$ increase in PM10 concentration, the death rate increases by 0.36% and 0.40%, respectively. High PM concentrations are especially harmful to children and the elderly over 75. Therefore, many countries are trying to control and reduce air pollution emissions at the national or regional level through real-time air quality monitoring and prediction (Kunt et al, 2023).

Particulate matter (PM10) is currently Thailand's most severe air pollution problem. Especially in Bangkok, greenhouse gas emissions from traffic are the leading cause of air pollution in Bangkok (Kanjanasiranont et al., 2022).

The combustion of fossil fuels is a significant cause of PM10. Also, the combustion of fossil fuels PM10 comes mainly from automobile emissions. In urban areas, PM10 is primarily derived from the fuel of transport vehicles, the primary source of air pollution. Many health effects result from exposure to PM10, such as heart disease, lung disease, and chronic bronchitis. Stroke and cancer PM10 are not only harmful to human health. But it also harms visibility in the atmosphere.

PM10 has the potential to carry many chemicals, including toxic chemicals. Chronic exposure to these components, linked to PM, has wide-ranging health effects.

Therefore, a better understanding of the status and forecast of PM10 trends in Bangkok is critical to supporting both national and regional governments in policy formulation and implementation and improving tools for assessing and managing air quality.

2. LITERATURE REVIEW

The author has used SARIMA and SARIMA-GARCH to forecast the volatility of the growth of tourists arriving in Thailand (Bunnag, 2023) but has never used them to predict PM. The author is interested in environmental issues and the negative externality of consuming energy. From collecting research on PM10 forecasting, we first started by using the ARIMA or SARIMA model; after that, we will describe the part of the hybrid-GARCH or hybrid ARIMA and other methods.

Taneja et al. (2017) analyze the future trends of Particle Particular (PM). For this, the Box–Jenkins ARIMA (Autoregressive Integrated Moving Average) model has been used for simulating the monthly average Aerosol Optical Depth (AOD550 nm) retrieved from Terra MODIS (Moderate Resolution Imaging Spectroradiometer) over New Delhi, the urban capital of India. The satellite dataset has been collected for 10 years, from 2004 to 2014. The analysis of the autocorrelation function indicates the existence of seasonality in the AOD time series. After rigorous evaluation of the selected models, the ARIMA (1,0,0) (0,1,2)12 is identified as the best-fit model.

Uzair et al. (2021) studied the ambient air quality of Lahore city of Pakistan. A correlation study suggests a positive correlation between the particulate matter and other mass-concentration particles like Ozone (O_3), Nitrogen Oxide (NO), and Sulphur Dioxide (SO_2). Predicting future concentration of PM2.5 is predicted using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, which gives the increasing value of PM2.5 in the next year and provides the lowest and highest predictions (more than $100\mu g/m^3$). The study reveals that the particulate matter in the Lahore season PM2.5/PM10 exceeds Pakistan's National Environmental Quality Standards.

Borhani et al. (2022) present a time-series analysis of SO_2 air concentration and the effects of particulate (PM2.5 or PM10) concentrations and meteorological conditions on SO_2 trends in Tehran from 2011 to 2020. The source data was obtained from meteorological stations in Tehran. To predict the status of future concentration of SO_2, PM2.5, and PM10, a Box–Jenkins ARIMA approach was used to model the monthly time series. Considering the 10 years, a downward trend was noted for SO_2 air concentration, even though a slight rise was observed in 2020. Monthly sulfur dioxide concentrations were lowest in June and the highest in January. Seasonal concentrations were most lacking in spring and highest in winter. In the same year, Veleva et al. (2022) studies the concentrations of PM10 in Vidin, Bulgaria. The town of Vidin is in north-western Bulgaria, on the south bank of the river Danube in the north Bulgarian border with Romania. They use official PM10 concentration level measurements by the Bulgarian Ministry of Environment and Water for 2010-2021. Appropriate methods were used - classical time series decomposition and stochastic Box- Jenkins ARIMA. Models with good statistical indicators for training, fitting, and forecasting PM10 have been built. A declining trend has been established on an annual basis.

Next, the method of a new hybrid-GARCH (Generalized Autoregressive Conditional Heteroskedasticity) methodology is proposed by Wang et al. (2017) to integrate the individual forecasting models of the ARIMA (Autoregressive Integrated Moving Average) and SVM (Support Vector Machine). The hybrid-GARCH approach for time series prediction is tested by 10-day hourly PM2.5 concentration data in Shenzhen, China. Empirical results from six station data sets indicate that the PM2.5 concentrations of Shenzhen experience a regular fluctuation during the 24 h of the whole day with the peak value in working hours due to factory and vehicle emissions, and the proposed hybrid model generates a more reliable and accurate forecast capability.

Veleva and Zheleva (2018) presents an empirical study of air pollution in Bulgarian cities caused by PM10. Univariate ARIMA, hybrid ARIMA-GJR-GARCH, and hybrid ARIMA-EGARCH models are constructed and statistically evaluated. The comparison between the three models is made by widely used. The hybrid ARIMA-EGARCH model has smaller RMSE, MAE, and TIC values than the ARIMA-GJR-GARCH model. The main advantage of the hybrid models is that they directly interpret the original values of PM10 and simultaneously model the conditional variance of the process. The predictions for the conditional variance by the two hybrid models are agreed upon and capture well the increases in the volatility of the values of PM10.

Alexis et al. (2022) studied the SARIMA-GARCH combination as an excellent tool to forecast PM10 behavior in the Caribbean. The modeling results could be extended to the nearby Guadeloupe and Puerto Rico islands to better understand the seasonal impact of dust outbreaks on the environment and human health. In the same year, Zhao et al. (2022) make the forecasting of Beijing PM2.5 with a hybrid ARIMA model based on integrated AIC and improved GS fixed-order methods and seasonal decomposition, and finally, the reconstructed series is predicted. They used Beijing PM2.5 data for validation, and the results showed that the new hybrid ARIMA model improved values of RMSE 99.23%, MAE 99.20%, R2 118.61%, TIC 99.28%, NMAE 98.71%, NMSE 99.97%, OPC 43.13%, MOPC 98.43% and CEC 99.25% compared with the traditional ARIMA model. The results show that the method significantly improves prediction performance and provides a convincing policy formulation and governance tool.

Finally, is another method; in this study, Zhu et al. (2017) employed seven single models and ensemble learning algorithms and constructed a hybrid learning algorithm, the LSTM-SVR model, totaling eight machine learning algorithms, to predict the Air Quality Index in six major urban agglomerations in China. The results reveal that, in areas with higher levels of air pollution, the situation for model prediction is more complicated, leading to a decline in predictive accuracy. The constructed hybrid model LSTM-SVR demonstrated the best predictive performance, followed by the ensemble model RF, which effectively enhanced the predictive accuracy in heavily polluted areas.

Xiao et al. (2020) propose a weighted long short-term memory neural network extended model (WLSTME). Daily PM2.5 concentration and meteorological data on Beijing–Tianjin–Hebei from 2015 to 2017 were collected to train models and to evaluate their performance. Experimental results with three existing methods showed that the proposed WLSTME model has the lowest RMSE (40.67) and MAE (26.10) and the highest p (0.59). Further experiments showed that in all seasons and regions, WLSTME performed the best. This finding confirms that WLSTME can significantly improve PM2.5 prediction accuracy.

In addition to that, Chen et al. (2023) study integrated the advantages of convolutional neural network (CNN) feature extraction and random forest (RF) regression ability to propose a novel CNN-RF ensemble framework for PM2.5 concentration modeling. The observational data from 13 monitoring stations in Kaohsiung in 2021 were selected for model training and testing. First, CNN was implemented to extract critical meteorological and pollution data. Subsequently, the RF algorithm was employed to train the model with input factors. The findings demonstrated that the proposed CNN–RF model had better modeling capability than the independent CNN and RF models: the average improvements in root mean square error (RMSE) and mean absolute error (MAE) ranged from 8.10% to 11.11%, respectively. In addition, the proposed CNN–RF hybrid model has fewer excess residuals at thresholds of 10 $\mu g/m^3$, 20 $\mu g/m^3$, and 30 $\mu g/m^3$. The results revealed that the proposed CNN–RF ensemble framework is a stable, reliable, and accurate method that can generate superior results compared with the single CNN and RF methods.

However, a review of research on using tools to predict PM values in each study found that forecasting accuracy and the popularity of the tools used still prefer SARIMA and hybrid-GARCH, so the author chooses to use these methods. Then, each procedure is compared to find the best way to forecast.

2.1. The Objectives of this Research Are

1. Finding the relationship between air pollutants and the amount of PM10 measured in Bangkok.
2. To forecast the amount of PM10 in Bangkok by using the SARIMA and SARIMA-GARCH models to formulate policies to reduce the occurrence of PM10 and guidelines for further prevention.

2.2. PM10 Data Collection in Bangkok

PM10's data comes from http://air4thai.pcd.go.th of Thailand's Pollution Control Department, Ministry of Natural Resources and Environment. It is a monthly average of PM10 data starting from January 2008 to July 2023. The data set was selected from

Bansomdejchaopraya Rajabhat University Station (O2T) in Bangkok, as shown in Figure 1. The reason for choosing the data from that station. Because it is a station that has quite a lot of traffic congestion problems. This causes some issues with air quality, and the completeness of the data is very high compared to data from other stations, which is a good representative of the data for predicting the amount of PM10 that occurs in Bangkok.

3. SARIMA MODEL

Box and Jenkins proposed a complete set of methods for time series analysis, prediction, and control, known as the Box-Jenkins modeling method (Naylor et al., 1972). The ARIMA model is divided into a simple seasonal model (P=D=0) and a seasonal model according to the difficulty of extracting seasonal effects. When there are both short-term correlations and seasonal effects are in the sequence, a more complex between the two can be used to fit the sequence model. In this study, the product seasonal model [denoted as SARIMA(p,d,q)(P, D, Q)s] describes the autocorrelation between a group of time-dependent random variables. The general expression of the ARIMA seasonal model is ARIMA (p, d, q), where and represent continuity and seasonal auto-regression differences, respectively. The order of the moving average means the length of the seasonal cycle.

3.1. Model Implementation

For stationary time series data, an autoregressive moving average ARMA (p, q) model can be established in the form of

$$X_t = \varphi_0 + \varphi_1 X_{t-1} + \ldots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

Among them, X_t is the sequence value of the first period, ε_t refers to the residual of the t period, and ϕ_1,θ are the parameters to be estimated by the model which can also be written as

$$X_t = \frac{\theta(B)}{\varphi(B)} \varepsilon_t$$

where *B* is a backward shift operator, which satisfies;

$$X_{t-1} = BX_t$$

For non-stationary time series with short-term trends, if a difference of order *d* is used to achieve stationary, then a differential autoregressive moving average model is established, which is denoted as ARIMA (p, d, q) model.

$$\Delta^d X_t = \varphi_0 + \varphi_1 \Delta^d X_{t-1} + \ldots + \varphi_p \Delta^d X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where $\Delta^d X_t$ represents the *t*-t sequence value after the d-th order difference.

The form expressed by the back shift operator is:

$$\Delta^d X_t = \frac{\theta(B)}{\varphi(B)} \varepsilon_t$$

For the ARIMA model with seasonal effects, the seasonal difference can be converted into a stationary sequence model. The seasonal effect and other effects in the sequence are additive relationships. A simple seasonal model can be established as

$$\Delta_D \Delta^d X_t = \frac{\theta(B)}{\varphi(B)} \varepsilon_t,$$

where $\Delta_D \Delta^d X_t$ represents the t-th sequence value after d-step D-step difference.

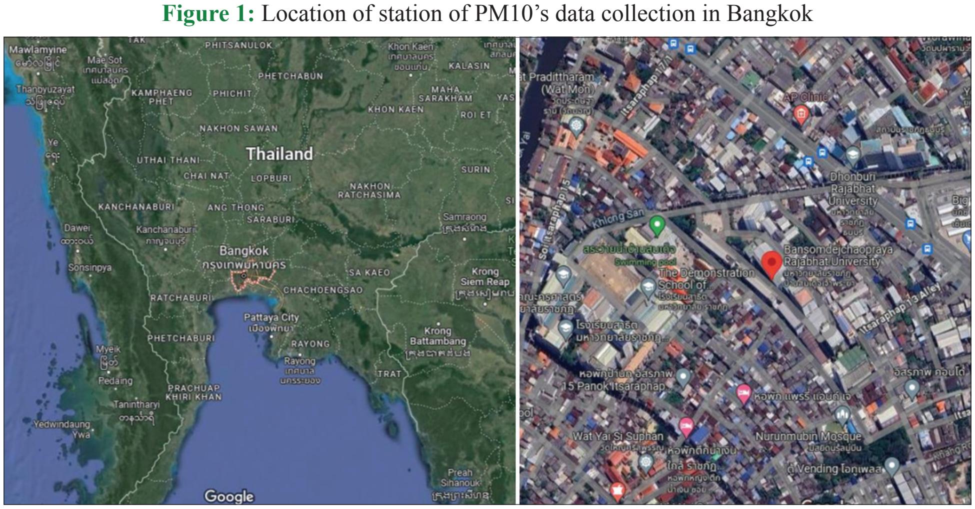
If the seasonal effects, long-term trend effects, and random fluctuations of the sequence have complex correlations, and the simple seasonal model cannot fully extract the correlations among them, the seasonal product model should be used, and the ARMA (p, q) model short-term correlation, using ARMA (p, q) model with the period step S as the unit to extract seasonal correlation, the model form is:

$$\Delta_D \Delta^d X_t = \frac{\theta(B)\theta_S(B)}{\varphi(B)\varphi_S(B)} \varepsilon_t$$

The above theory shows that, according to the characteristics of data stability, seasonality, trend, etc., an appropriate method should be selected for modeling.

The modeling steps of SARIMA(p,d,q)(P, D, Q)s model by Xiang (2022) are as

Figure 1: Location of station of PM10's data collection in Bangkok



Source: www.google.map.com, 2023

1. The stationarity test is carried out on the original time series. Suppose the series does not meet the stationarity condition. In that case, the difference transformation is needed to make the series meet the stationarity condition to obtain the value of d in the model.
2. The values of p, q and P,Q in the model are determined using ACF and PACF.
3. The SARIMA model parameter estimation methods include maximum likelihood to have been used to estimate model parameters and test their significance. The ARCH test is essential in the research and analysis of time series. Further investigation can be undertaken only if the residual sequence passes the ARCH test. If it does not pass the ARCH test, then the analysis process must be repeated from the model recognition stage. The Akaike Information Criterion (AIC), Bayesian Information Criteria (BIC), and Hannan-Quinn Criterion (HQ) are used to evaluate the model's goodness of fit. The model with the relatively smallest statistic value has the best-fitting effect, which is then used as the optimal model.
4. Predict the future value of time series.

3.2. ARCH Model

$$\{x_t = f\left(t, x_{t-1}, x_{t-2}, \ldots\right) + \varepsilon_t, \varepsilon_t = \sqrt{h_t}\, e_t, h_t = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, e_t \sim IID(0,1),$$

where α_i is nonnegative and $f(t, x_{t-1}, x_{t-2}, \ldots)$ is the deterministic information fitting model of $\{x_t\}$.

3.3. GARCH Model

$$\{x_t = f\left(t, x_{t-1}, x_{t-2}, \ldots\right) + \varepsilon_t, \varepsilon_t = \sqrt{h_t}\, e_t,$$

$$h_t = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \gamma_j h_{t-j}, e_t \sim IID(0,1),$$

where α_i and γ_j are nonnegative and $f(t, x_{t-1}, x_{t-2}, \ldots)$ is the deterministic information fitting model of $\{x_t\}$. It is an extension of the ARCH model and claims that h_t has AR $\sum_{j=1}^{p} \gamma_j h_{t-j}$ and ARCH term is $\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$. In general, the GARCH model is easier to identify and estimate, and the GARCH model can capture the flat period and fluctuation period series.

3.4. Evaluation Metrics

Mean Absolute Error (MAE) and RMSE (Root Mean Square Error) have widely been used in evaluating the accuracy of a recommender system (Wang and Lu, 2018), given by:

$$MAE = \frac{\sum_{n=1}^{N} \left| \hat{r}_n - r_n \right|}{N} \quad and \quad RMSE = \sqrt{\frac{\sum_{n=1}^{N} \left(\hat{r}_n - r_n \right)^2}{N}}$$

where r_n means the prediction rating; n means the true rating in testing data set; N is the number of rating prediction pairs between the testing data and prediction result.

4. THE RELATIONSHIP BETWEEN PM10 AND AIR POLLUTANTS

Table 1 shows the relationship between air pollutants. Correlation of monthly averages of nitrogen oxides (NO_2) and particulate matter (PM10) from January 2008 to July 2023.

First, since diesel combustion (from heavy vehicles) is the primary source of nitrogen oxides and particulate matter (PM10), Particulate matter (PM10) was correlated with nitrogen oxides (NO_2) (r^2 = 0.575 with moderate positive correlation). It can be inferred from Figure 2 that it contributes to the primary and secondary particulate matter (PM10) in the atmosphere (because of the correlation of nitrogen oxides (NO_2) mainly coming from automobiles). Therefore, it is interesting to check the relationship of these substances in the surrounding air, especially in the urban environment, where photochemical conversion (including removal mechanisms) can be ignored, and then check these relationships against the emission inventory.

Second, the relationship between particulate matter (PM10) and ground ozone (O_3) was positively correlation (r^2 = 0.529 with moderate positive correlation). Unlike the other pollutants mentioned above, surface ozone (O_3) is not released directly into the atmosphere but is a secondary pollutant created by the reaction of nitrogen dioxide (NO_2), hydrocarbons, and sunlight (see Zhu et al., 2019).

Third, the relationship between particulate matter (PM_{10}) and sulfur dioxide (SO_2) is positively correlated (r^2= 0.425 with low positive correlation). Fossil fuels contain traces of sulfur compounds; sulfur dioxide is produced when burned. Most SO_2 emitted into the air comes from power generation, with little contribution from transportation sources. The sulfuric acid produced by the reaction of SO_2 in the atmosphere is the main component of acid rain and the ammonium sulfate. Particles are the secondary particles with the highest content in the air.

Fourth, the relationship between particulate matter (PM10) and Carbon monoxide (CO) is positively correlated (r^2= 0.292 with low positive correlation). Carbon monoxide (CO) is one of the most widely distributed and commonly occurring air pollutants. It results from the incomplete fuels. Thus, transport is the primary sector responsible for the emission of these species. The primary concern regarding CO pollution is its adverse health effects. When inhaled, CO is absorbed in the lungs and combines irreversibly with hemoglobin (Hb) in the blood to form carboxyhemoglobin.

Table 1: The correlation between PM10 and air pollutants in Bangkok from January 2008–July 2023

Air pollutants/PM_{10}	CO	NO_2	O_3	SO_2	PM_{10}
CO	1	0.549	0.434	0.454	0.292
NO_2	0.549	1	0.476	0.647	0.575
O_3	0.434	0.476	1	0.347	0.529
SO_2	0.454	0.647	0.347	1	0.425
PM_{10}	0.292	0.575	0.529	0.425	1

Hence, the principal toxic properties of CO arise from the resulting lack of oxygen in tissues (hypoxia). Therefore, carbon monoxide pollution is of particular concern in urban locations with heavy traffic and moderate or weak atmospheric dispersion.

5. EMPIRICAL ANALYSIS AND RESULTS

In this part of this study, we analyze and estimate variables, but generally, in time series analysis, the primary stage is to investigate the integrated order of the study variables. ADF (see Dickey and Fuller, 1981) and PP (see Phillips and Perron, 1988) approaches have been used in this study to check the order of integration. We want to test the hypothesis of the existence of a unit root. The null and alternative hypotheses can be formulated as follows:

H_0: $\alpha = 1$ (unit root)

H_1: $\alpha < 1$ (Integrated of order zero)

These two tests are based on the null of non-stationarity, which indicates the presence of a unit root, and the alternative hypothesis of the non-existence of a unit root, which means that the variable examined is stationary.

5.1. Results of the unit root test

All data for variables of PM10 are shown in Table 2. The different results of the stationarity test are indicated in Table 2. The results show that all variables are stationary in the first difference I(1), which is integrated into order one I(1). In this case, we can reject the null hypothesis of the presence of unit root, and we can accept the alternative hypothesis. After determining the order of integration, we will verify the existence of integration between variables using the SARIMA (p,1,q)(P,1,Q)s next for estimation and forecasting the PM10 in Bangkok.

5.2. Build the SARIMA Model

The time series of PM10 has transferred as a stationary series after being differencing at a time, so we need to ensure the value of p and q, P and Q. Thus, we observe first the difference figure of ACF and PACF, as shown in Figure 3. We can judge that the value of p is 3, the value of q is 3, the value of P is 1, and the value of Q is 2. So, we can build the SARIMA(3,1,3)(1,1,2)12 model and be the model with the lowest AIC value at 8.480 (goodness-of-fit) along with the value of the coefficients as below in Table 3.

From Table 4, after we have the model, the next issue is that the completeness of the model will require testing of problems of residual autocorrelation and the problem of heteroscedasticity. Testing the problem, the residual autocorrelation found that the value of Q-statistic from lag 1-6 was found to reject H_1 at the 1% significance level, which means there is no residual autocorrelation, including in testing problems heteroscedasticity using the ARCH test (ARCH effect) at lag 4 and lag 8, it encountered to reject H_1 at the 1% significance level, respectively. It means no heteroscedasticity.

Furthermore, we can test the normality considering the skewness, the kurtosis, and the Jarque-Bera statistics; we found that this model is statistically significant at a 1% level (reject H_0), thereby implying that the distribution is not normal. In summary, this model can be used to estimate and predict PM10 in Bangkok, which is suitable to a certain extent.

5.3. SARIMA-GARCH Model Estimation

The build of the SARIMA-GARCH model first needs to create the GARCH model of PM10. The conditional mean that we choose is SARIMA (3,1,3)(1,1,2)12, and the conditional volatility selected is GARCH(1,1); estimated parameters are listed in Table 5, and the models are as below:

Figure 2: The relationship between PM10 and air pollutants

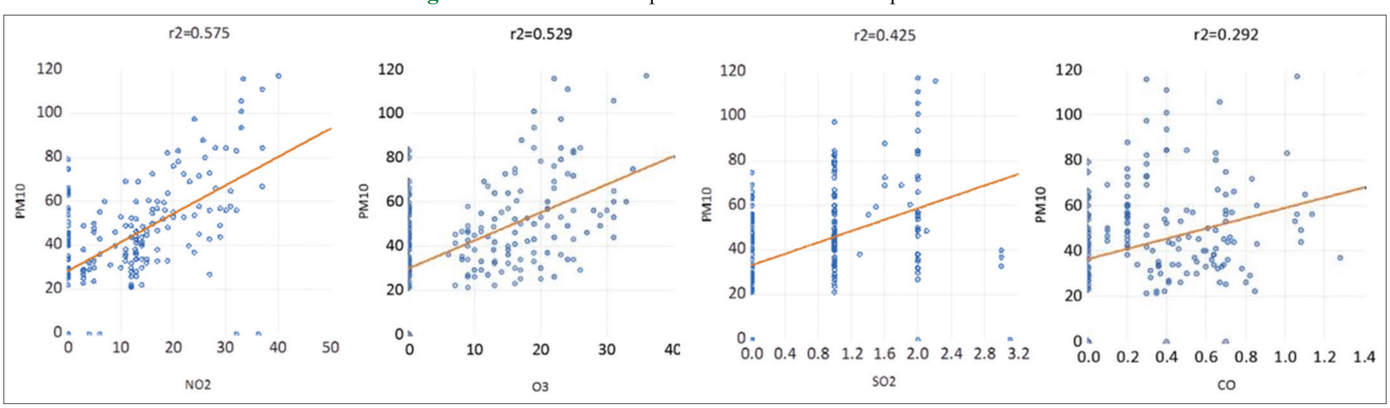


Table 2: Unit root test for PM10 in Bangkok

Variables	ADF test statistic				PP test statistic			
	intercept		Intercept and trend		Intercept		Intercept and trend	
	I (0)	I (1)	I (0)	I (1)	I (0)	I (1)	I (0)	I (1)
PM_{10}	−3.102**	−7.522***	−2.898	−7.518***	−5.829***	−16.446***	−5.933***	−16.322***

**5% significant level, ***1% significant level

Figure 3: The ACF (Autocorrelation) and PACF (Patial autocorrelation) of the SARIMA(p,1,q)(P,1, Q)s model

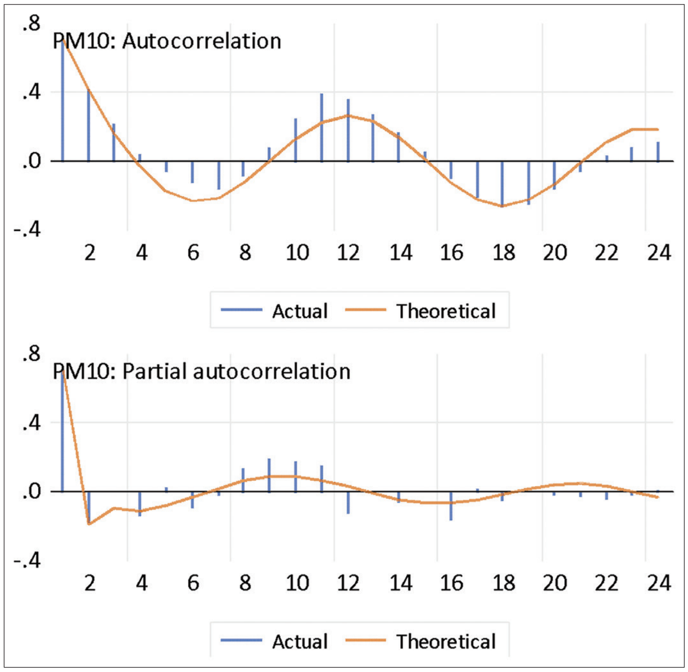


Table 3: The coefficient of the SARIMA (3,1,3)(1,1,2) 12 model

Variable	Coefficient	SE	P
Constant	42.894	2.201	0.000
AR (1)	2.278	0.003	0.000
AR (2)	−1.950	0.001	0.000
AR (3)	0.550	0.001	0.000
SAR (12)	0.773	0.193	0.000
MA (1)	−1.554	0.142	0.000
MA (2)	0.715	0.192	0.000
MA (3)	0.162	0.089	0.072
SMA (12)	−0.855	0.446	0.056
SMA (24)	−0.117	0.096	0.227
Diagnostic check	**Value**		
Log-likelihood	−781.901		
AIC	8.480		
BIC	8.670		
HQ	8.557		

AIC, BIC, and HQ denote the AIC: Akaike information criterion, BIC: Bayesian information criteria, HQ: Quinn criterion

However, to consider the validity of conditional volatility, we must check the coefficients of the equation as follows which α is the parameter of ε_{t-1}^2 to be small (equal to 0.164), which is close to 0, and α+ γ is equal to 0.982, <1 (γ is the parameter of h_t).

From Table 6, testing the problem revealed that residual autocorrelation found in the values of Q-statistic from lags 1-6 were found to reject H_1 at the 1% significance level, meaning there was no residual autocorrelation. In addition, in testing the heteroscedasticity problem using the ARCH test (ARCH effect) at lag 4 and 8, H_1 was rejected at the 1% significance level, respectively. It means that there is no heteroscedasticity. All are the same as the first model.

Table 4: residual diagnostic test for the SARIMA (3,1,3) (1,1,2) 12 model

The SARIMA (3,1,3)(1,1,2) 12 model			
Test	**lags**	**Value**	**P**
Residual tests	1	0.006	0.938
for Autocorrelations	2	3.300	0.192
H_0=no residual	3	3.434	0.329
autocorrelation	4	4.299	0.367
(Q-stat)	5	5.763	0.330
The ARCH test	6	5.780	0.448
H_0=no			
Heteroskedasticity			
ARCH	4	0.811	0.519
ARCH	8	0.969	0.461
Residual normality test			
H_0=normal distribution			
Skewness	-	0.457	0.000
Kurtosis	-	5.899	0.000
Jarque-Bera	-	72.043	0.000

Table 5: The coefficient of the SARIMA (3,1,3)(1,1,2) 12-GARCH (1,1) model

Variable	Coefficient	SE	P
Conditional mean			
Constant	38.705	2.201	0.000
AR (1)	−0.459	0.003	0.000
AR (2)	−0.238	0.001	0.000
AR (3)	0.541	0.001	0.000
SAR (12)	−0.742	0.193	0.000
MA (1)	1.200	0.142	0.000
MA (2)	1.116	0.192	0.000
MA (3)	0.114	0.089	0.072
SMA (12)	1.760	0.446	0.056
SMA (24)	0.794	0.096	0.227
Conditional volatility			
Constant (ω)	5.013	6.015	0.404
α	0.164	0.056	0.003
γ	0.818	0.060	0.000
Diagnostic check	**Value**		
α+ γ	0.982		
Log-likelihood	−675.851		
AIC	8.009		
BIC	8.247		
HQ	8.106		

AIC, BIC, and HQ denote the AIC: Akaike information criterion, BIC: Bayesian information criteria, HQ: Quinn criterion

Additionally, we can test the normality based on skewness, kurtosis, and the Jarque-Bera statistic. We found that the model is statistically significant at the 1% level (rejecting H_0), which means that the distribution is not normal. In summary, this model can estimate and forecast PM10 in Bangkok, which is appropriate at a certain level. But before considering whether the first or second model is more accurate in forecasting. We must consider the discrepancies in the predictions of the two models, which will be explained in the following order.

Comparison of predictive accuracy between the SARIMA(3,1,3)(1,1,2)12 and SARIMA (3,1,3)(1,1,2)12-GARCH(1,1) are shown in Figures 4 and 5. It is impossible to clearly distinguish from the prediction numbers whether SARIMA(3,1,3)(1,1,2)12 and SARIMA(3,1,3)(1,1,2)12-GARCH(1,1) is better for prediction.

Figure 4: Forecasting average monthly PM10 in Bangkok using the SARIMA (3,1,3)(1,1,2)12 model

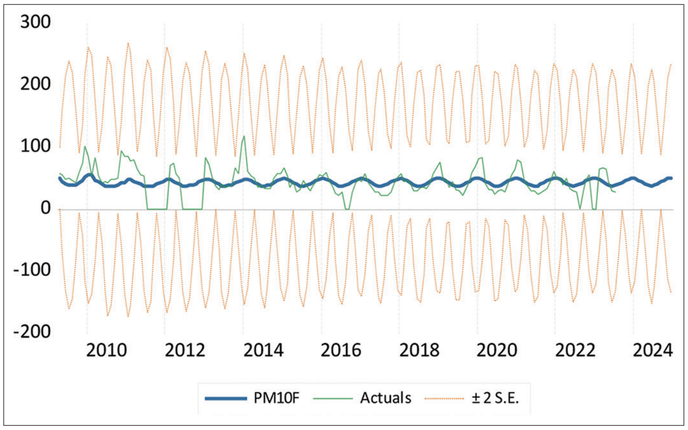


Figure 5: Forecasting average monthly PM10 in Bangkok using the SARIMA (3,1,3)(1,1,2)12 GARCH(1,1) model

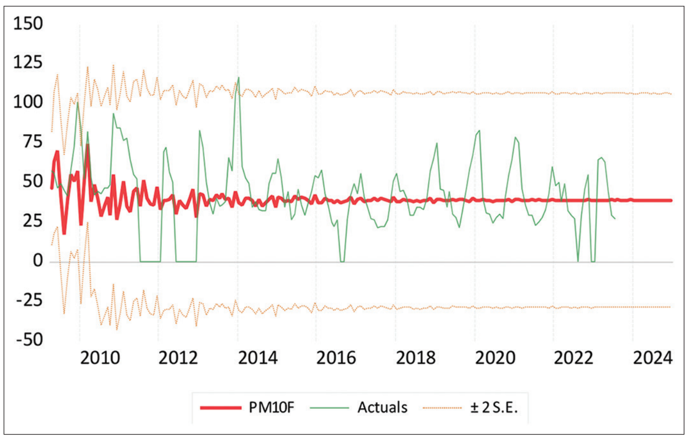


Table 6: The diagnostic check for the SARIMA (3,1,3)(1,1,2) 12-GARCH (1,1) model

The SARIMA (3,1,3)(1,1,2) 12-GARCH (1,1) model			
Test	**Lags**	**Value**	**P**
Residual tests	1	0.007	0.930
For Autocorrelations	2	1.054	0.590
H_0=No residual autocorrelation	3	2.626	0.453
(Q-stat)	4	3.280	0.512
The ARCH test	5	4.320	0.504
H_0=no heteroskedasticity	6	4.430	0.619
ARCH	4	0.755	0.556
ARCH	8	0.514	0.844
Residual normality test			
H_0=Normal distribution			
Skewness	-	−0.188	0.000
Kurtosis	-	3.811	0.000
Jarque-Bera	-	5.737	0.000

However, the forecast portion is scheduled from August 2023 to December 2024. It is used to forecast future PM10 data. We perform model comparisons that SARIMA (3,1,3)(1,1,2)12 and SARIMA(3,1,3)(1,1,2)12-GARCH(1,1), which model gives lower MAE and RMSE values, which indicates good prediction accuracy than another model. The forecast results can be shown in Figures 4 and 5. The results show that the MAE and RMSE predictions of the SARIMA (3,1,3) (1,1,2)12 model are 15.303 and 20.839 better than those of the SARIMA (3,1,3) (1,1,2)12-GARCH (1,1) model are 17.280 and 22.677. Therefore, the forecast results of the SARIMA (3,1,3) (1,1,2)12 are better precise, as shown in Table 7. Thus, in summary, we will choose the first model to use in forecasting for policy making.

Table 7: The prediction value of the SARIMA (3,1,3)(1,1,2) 12 and SARIMA (3,1,3)(1,1,2) 12-GARCH (1,1) model

Periods	Prediction value of the SARIMA (3,1,3)(1,1,2) 12 (unit of PM=μg/m³)	Prediction value of the SARIMA (3,1,3)(1,1,2) 12-GARCH (1,1) (unit of PM=μg/m³)
August 2023	39.426	38.989
September 2023	42.433	38.274
October 2023	45.654	38.864
November 2023	48.060	38.857
December 2023	48.994	38.594
January 2024	48.364	38.675
February 2024	46.184	38.684
March 2024	42.970	38.944
April 2024	39.831	38.716
May 2024	37.464	38.907
June 2024	36.535	38.798
July 2024	37.370	38.704
August 2024	39.700	38.493
September 2024	42.780	39.026
October2024	45.960	38.584
November 2024	48.256	38.591
December 2024	49.041	38.788
Evaluation Metrics	**Value**	**Value**
MAE	15.303	17.280
RMSE	20.839	22.677

MAE: Mean Absolute Error, RMSE: Root mean square error

Table 8: Recommended AQG levels and interim targets

Pollutant	Averaging time	Interim target				AQG level
		1	2	3	4	
PM10 (μg/m³)	Annual	70	50	30	20	15
	24 h	150	100	75	50	45

Source: The WHO, 2021. WHO: World Health Organization

6. CONCLUSION

From the study results, we can find the relationship between air pollutants and PM10 in Bangkok and found that the elements of NO_2 and O_3 will require quite a lot of attention. Because it affects the relationship PM10 at a moderate level, we forecast the monthly average of the PM10 amount. It was found that the average yearly PM10 amount (we calculate the annual average PM10 amount from the average monthly PM10 amount from Table 8) exceeded 30 μg/m³ for the interim target 3, which is the average that was determined by the World Health Organization (WHO) according to Table 8. However, the AQG level (air quality guideline level) is 15 μg/m³ (The World Health Organization (WHO), 2021); the forecasting PM10 amount exceeds both levels. Therefore, it is necessary to create understanding with the people to take care of their health by wearing protective masks, including long-term planning that will happen next.

As you know, the PM problem occurred in Bangkok. It is a problem we call the environmental impacts of transport (negative externality).

- Greenhouse gas emissions: The transportation sector significantly contributes to Bangkok's greenhouse gas emissions, making it a critical area for reducing emissions.
- Pollution: The pollution created by automobiles and other forms of transportation can have serious adverse health consequences for individuals and nearby communities.
- Resource depletion: The production and transportation of fossil fuels can be resource-intensive, leading to the need for more natural resources in affected areas.

  Therefore, understanding how these problems relate to economic principles is critical in finding practical solutions.
- Promoting efficient modes of transport: Government support for more efficient methods of transport, such as rail or water, that can reduce transportation costs and environmental harm, promotion of the demand for ride-sharing services, and the emergence of electric cars in creating new opportunities for alternative power sources. Especially for Thailand's energy policy, Insan et al. (2022) explained that Thailand has encouraged investment in producing and importing electric vehicles to replace fossil fuel combustion. They also promoted the development of electric cars to be more efficient and run longer distances. The cumulative number of electric vehicles from 2017 to now is increasing, making business opportunities for EV charging stations available in Thailand, moreover supporting the rise of autonomous cars, which are poised to significantly impact the transportation industry by improving safety and reducing traffic congestion.
- Investing in infrastructure: Government investment in transportation infrastructure can reduce congestion by expanding roads, implementing tolls, or building public transit systems.

REFERENCES

Alexis, E., Plocoste, T., Nuiro, S.P. (2022), Analysis of particulate matter (PM10) behavior in the Caribbean area using a coupled SARIMA-GARCH model. Atmosphere, 13(6), 862.

Borhani, F., Shafiepour, M.M., Rashidi, Y., Ehsani, A.H. (2022), Estimation of short-lived climate forced sulfur dioxide in Tehran, Iran, using machine learning analysis. Stochastic Environmental Research and Risk Assessment, 36(9), 2847-2860.

Bunnag, T. (2023), Guidelines for Econometrics and Application. Emphasis in Tourism and Financial Economics. In Book Series Socio-Economics, Research, Innovation and Technologies (SERITHA). Craiova: Ritha Publishing.

Chen, M.H., Chen, Y.C., Chou, T.Y., Ning, F.S. (2023), PM2.5 Concentration prediction model: A CNN-RF ensemble framework. International Journal of Environmental Research and Public Health, 20(5), 4077.

Dickey, D.A., Fuller, W.A. (1981), Likelihood ratio statistics for autoregressive time series with unit root. Econometrica, 49, 1057-1072.

Insan, D., Rakwichian, W., Rachapradit, P., Thanarak, P. (2022), The business analysis of electric vehicle charging stations to power environmentally friendly tourism: A case study of the Khao Kho Route in Thailand. International Journal of Energy Economics and Policy, 12(6), 102-111.

Kanjanasiranont, N., Butburee, T., Peerakiatkhajohn, P. (2022), Characteristics of PM10 levels monitored in Bangkok and its vicinity areas, Thailand. Atmosphere, 13(2), 239.

Kelishadi, R., Poursafa, P. (2010), Air pollution and non-respiratory health hazards for children. Archives of Medical Science, 6, 483-495.

Kim, B.Y., Lim, Y.K., Cha, J.W. (2022), Short-term prediction of particulate matter (PM10 and PM2.5) in Seoul, South Korea using tree-based machine learning algorithms. Atmospheric Pollution Research, 13(10), 101547.

Kumar, P., Vishwakarma, A.K., Kumar, D., Yadav, S., Pandey, D., Ram, S., Arora, S. (2020), Impact of Air Pollution on Environment and Human Health. In book: Emerging Trends in Environmental Science. New Delhi: Asiatech Publishers Inc.

Kunt, F., Ayturan, Z.C., Yümün, F., Karagönen, İ., Semerci, M., Akgün, M. (2023), Modeling and assessment of PM10 and atmospheric metal pollution in Kayseri province, Turkey. Atmosphere, 14, 356.

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., Bezirtzoglou, E. (2020), Environmental and health impacts of air pollution: A review. Frontiers in Public Health, 8, 505570.

Naylor, T.H., Seaks, T.G., Wichern, D.W. (1972), Box-Jenkins methods: An alternative to econometric models. International Statistical Review/Revue Internationale de Statistique, 40(2), 123-137.

Phillips, P.C., Perron, P. (1988), Testing for a unit root in time series regression. Biometrika, 75(2), 335-346.

Taneja, K., Ahmad, S., Ahmad, K., Attri, S.D. (2016), Time series analysis of aerosol optical depth over New Delhi using Box-Jenkins ARIMA modeling approach. Atmospheric Pollution Research, 7(4), 585-596.

Thangavel, P., Park, D., Lee, Y.C. (2022), Recent insights into particulate matter (PM2.5)-mediated toxicity in humans: An overview. International Journal of Environmental Research and Public Health, 19(12), 7511.

Veleva, E., Zheleva, I. (2018), GARCH models for particulate matter PM10 air pollutant in the city of Ruse, Bulgaria. Application of Mathematics in Technical and Natural Sciences. 10th International Conference for Promoting the Application of Mathematics in Technical and Natural Sciences. Vol.2025. AIP Publishing.

Veleva, E., Filipova, M., Zheleva, I. (2022), Statistical Study of Particulate Matter (PM10) Air Contamination in the City of Vidin, Bulgaria. In AIP Conference Proceedings. Vol.2522. AIP Publishing.

Wang, P., Zhang, H., Qin, Z., Zhang, G. (2017), A novel hybrid-GARCH model based on ARIMA and SVM for PM2.5 concentrations forecasting. Atmospheric Pollution Research, 8 (5), 850-860.

Wang, W., Lu, Y. (2018), Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. IOP Conference Series: Materials Science and Engineering. Vol.324. IOP Publishing.

World Health Organization. (2021), WHO Global Air Quality Guidelines. Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. WHO European Centre for Environment and Health, Platz der Vereinten Nationen 1 D53113 Bonn. (Electronic Version).

Xiang, Y. (2022), Using ARIMA-GARCH model to analyze fluctuation law of international oil price. Mathematical Problems in Engineering, 2022, 1-7.

Xiao, F., Yang, M., Fan, H., Al-Qaness, M. (2020), An improved deep learning model for predicting daily PM2.5 concentration. Scientific Report, 10, 20988.

Zhao, L., Li, Z., Qu, L. (2022), Forecasting of Beijing PM2.5 with a hybrid ARIMA model based on integrated AIC and improved GS fixed-order methods and seasonal decomposition. Heliyon, 8(12), e12239.

Zhu, J., Chen, L., Liao, H., Dang, R. (2019), Correlations between PM2.5 and Ozone over China and associated underlying reasons. Atmosphere, 10, 352.

Zhu, S., Lian, X., Liu, H., Hu, Y., Wang, J., Che, J. (2017), Daily air quality index forecasting with hybrid models: A case in China. Environmental Pollution, 231, 1232-1244.