

## Article

# The Role of GARCH Effect on the Prediction of Air Pollution

Kai-Chao Yao <sup>1</sup>, Hsiu-Wen Hsueh <sup>1,\*</sup>, Ming-Hsiang Huang <sup>2</sup> and Tsung-Che Wu <sup>3,\*</sup>

<sup>1</sup> Department of Industrial Education and Technology, National Changhua University of Education, Changhua City 50007, Taiwan; kcyao@cc.ncue.edu.tw

<sup>2</sup> Department of Business Administration, National Changhua University of Education, Changhua City 50007, Taiwan; mhhuang@cc.ncue.edu.tw

<sup>3</sup> Department of Banking and Finance, National Chiayi University, Chiayi City 60054, Taiwan

\* Correspondence: pk121@cc.ncue.edu.tw (H.-W.H.); tcwujeff@mail.ncyu.edu.tw (T.-C.W.); Tel.: +886-4-723-2105 (ext. 7405) (H.-W.H.); +886-5-273-2866 (T.-C.W.)

**Abstract:** Air pollution prediction is an important issue for regulators and practitioners in a sustainable era. Air pollution, especially PM<sub>2.5</sub> resulting from industrialization, has fostered a wave of global weather migration and jeopardized human health in the past three decades. Taiwan has evolved as a highly developed economy and has a severe PM<sub>2.5</sub> pollution problem. Thus, the control of PM<sub>2.5</sub> is a critical issue for regulators, practitioners and academics. More recently, GA-SVM, an artificial-intelligence-based approach, has become a preferred prediction model, attributed to the advances in computer technology. However, hourly observation of PM<sub>2.5</sub> concentration tends to present the GARCH effect. The objective of this study is to explore whether the integration of GA-SVM with the GARCH model can build a more accurate air pollution prediction model. The study adopts central Taiwan, the region with the worst level of PM<sub>2.5</sub>, as the source of observations. The empirical implementation of this study took a two-step approach; first, we examined the potential existence of the GARCH effect on the observed PM<sub>2.5</sub> data. Second, we built a GA-SVM model integrated with the GARCH framework to predict the 8 h PM<sub>2.5</sub> concentration of the sample region. The empirical results indicate that the prediction performance of our proposed alternative model outperformed the traditional SVM and GA-SVM models in terms of both MAPE and RMSE. The findings in this study provide evidence to support our expectation that adopting the SVM-based approach model for PM<sub>2.5</sub> prediction is appropriate, and that prediction performance can be improved by integrating the GARCH model. Moreover, consistent with our prior expectation, the evidence further supports that taking the GARCH effect into account in the GA-SVM model significantly improves the accuracy of prediction. To the knowledge of the authors, this study is the first to attempt to integrate the GARCH effect into the GA-SVM model in the prediction of PM<sub>2.5</sub>. In summary, with regard to the development of sustainability for both regulators and practitioners, our results strongly encourage them to take the GARCH effect into consideration in air pollution prediction if a regression-based model is to be adopted. Furthermore, this study may shed light on the application of the GARCH model and SVM models in the air pollution prediction literature.

**Keywords:** GA-SVM; GARCH; PM<sub>2.5</sub> prediction; machine learning



**Citation:** Yao, K.-C.; Hsueh, H.-W.; Huang, M.-H.; Wu, T.-C. The Role of GARCH Effect on the Prediction of Air Pollution. *Sustainability* **2022**, *14*, 4459. <https://doi.org/10.3390/su14084459>

Academic Editors: Tsu-Ming Yeh, Ali Elkamel, Hsin-Hung Wu, Yuh-Wen Chen and Fan-Yun Pai

Received: 28 February 2022

Accepted: 25 March 2022

Published: 8 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Air pollution has long been an important concern for all countries. In particular, the level of air pollution is often closely related to the industrialization of each country. Although industrialization brings economic growth and prosperity to society, it is often accompanied by factors causing the deterioration of the ecological environment, such as waste and air pollution. For a long time, developed countries, in response to citizens' pursuit of a cleaner environment and better life quality, have had to deal with the air pollution problems associated with high levels of industrialization. In recent years, as a result of climate change and the rapid economic development of emerging economies, air

pollution has become an important issue of global concern. Accordingly, highly industrialized countries held the United Nations Framework Convention on Climate Change (UNFCCC) in 1997 and proposed the Kyoto Protocol, advocating the 55 participating member countries' commitment to save energy and reduce carbon emissions. The goal was to stabilize the concentration of greenhouse gases in the atmosphere, while balancing food production and economic development. In addition, the private sector also raised funds to set up the non-governmental organization the Green Peace Organization (GPO), which is dedicated to research on environmental protection and the promotion of air pollution prevention measures.

In Taiwan, economic development and air pollution show a considerable degree of correlation. In the 1950s, Taiwan implemented the first phase of economic construction plans to cultivate the development of labor-intensive light industrialization. The economic growth in this decade, with an 8.7% annual growth rate, did not cause serious air pollution problems. In the 1960s, Taiwan set up a manufacturing export zone and actively developed the export industry. The average annual economic growth rate during this period was 9.8%. In the 1970s, the world experienced an oil energy crisis. The Taiwan government adopted a domestic demand-oriented strategy and promoted the Ten Major Construction Projects, which increased public expenditure, improved infrastructure and actively developed the petrochemical industry. Although Taiwan is one of the four Asian Tiger countries, the development of the petrochemical industry has brought economic growth, with the consequence of Taiwan's current air pollution problem. In the 1980s, Taiwan actively developed technology-intensive industries and established its first science industrial park in Hsinchu, with an average annual economic growth rate of 8.5% throughout this decade. The development of the technology industry, especially the semiconductor supply chain, which contributed greatly to Taiwan's economic growth, involved immense power and energy consumption. In order to fully meet the power demand, the government increased the production capacity of coal-fired power generation for sufficient power supply. However, the seriousness of air pollution was also aggravated [1].

In response to people's demands for better life quality, the Taiwan government has worked on the prevention and control of air pollution, especially fine particles ( $PM_{2.5}$ ) [2]. However, in the past, when the government formulated air pollution control measures, it was often necessary to take into account the costs incurred by the industry and the need to promote domestic economic development. As a result, the effect of the early prevention and control of air pollution was not significant, especially for  $PM_{2.5}$ . The harm caused by  $PM_{2.5}$  to human health has attracted international attention, because the spread of  $PM_{2.5}$  results not only in a deterioration of air quality but also in fatal health problems [3]. In recent years, the GPO has invested a lot of resources toward empirical research on the harm caused by air pollution to human health. According to a National Policy Research Report from the National Policy Research Foundation [4], based on an air pollution research by Harvard University on Asian countries including Taiwan, South Korea, Japan, China and the Philippines, the GPO found that coal-fired power plants not only directly emit  $PM_{2.5}$ , but also emit sulfur dioxide ( $SO_2$ ), nitrogen oxides ( $NO_x$ ), soot and dust, etc., thus stimulating the formation and increase in  $PM_{2.5}$ . These fine particles enter the lungs of the human body through the respiratory system, spreading through blood circulation and causing diseases such as asthma, respiratory diseases, cancer or even death. As a result, the government's energy choices are not only closely related to carbon emissions, but also directly related to air pollution, which has a considerable negative impact on human health.

In fact, in recent years, there have been numerous cases of harm to human health caused by air pollution [3,5]. In 2010,  $PM_{2.5}$  from air pollution caused more than 3 million cases of premature death in the world. According to the 2013 annual report on air quality monitoring by the Environmental Protection Agency, the annual average concentration of  $PM_{2.5}$  in Taiwan was  $24 \mu\text{g}/\text{m}^3$ , which was 2.4 times higher than the World Health Organization (WHO) standard value [1]. With that  $PM_{2.5}$  concentration, the risk of lung cancer and childhood asthma increased by 15% and the risk of stroke, heart disease and

chronic respiratory disease increased by nearly 25%. The seriousness of Taiwan's air pollution problem did not come to the surface until the Changhua Guoguang Petrochemical Development Project Evaluation in 2010 [6]. With government's promotion of a non-nuclear policy, the supply of green energy is still insufficient, thus the demand for coal-fired power has continued to increase in recent years. The government is still planning to build new coal-fired thermal power plants even though thermal power generation has accounted for more than 80% of annual power generation structure in Taiwan. Therefore, in this study we show that, in order to maintain citizens' health, accurately predicting the level of  $PM_{2.5}$  and analyzing the sources of air pollution are extremely important issues for the government to formulate air pollution policies and to control target levels.

In terms of forecasting methodology, in order to deal with autoregressive characteristics of observed time series data, the General Autoregressive Conditional Heteroscedasticity model (GARCH) is the most widely used method. With recent progress of computer programming and development of big data analysis, besides traditional statistical-based forecasting methods, the machine-learning-based approach is preferred in industry and academia. Among them, neural network systems and support vector machine (SVM) models are the most commonly used. Recently, the application of methods has evolved towards integrating models. For example, to improve prediction accuracy, neural networks are combined with wavelet analysis [7], while in terms of SVM, the combination of genetic algorithm (GA-SVM) is preferred [8].

Regarding the improvement of prediction accuracy, in addition to considering the integration of methods, the attributes of observation data are also very important. The sample data examined in this study is the hourly air pollution index from the EPA monitoring stations, which may be affected by factors such as terrain, atmosphere and so on. The air pollution in the previous hourly period may not have dissipated yet, thus the deferred effects on air pollution index observed in the next period may lead to autoregression. Conventional prediction models have been built on the basis of regression analysis. In the classical regression model, residual autoregression and heteroscedasticity should not exist; therefore, their presence may produce bias and violate the underlying assumptions [9,10]. In addition, GA-SVM is a machine-learning-based AI approach. Accuracy of prediction depends not only on quality of data, but also on availability of input (correlated) variables. Thus, a sufficient number of input variables (for more training) will be helpful to generate better prediction results. If we can provide more related variables in the training process, we will be able to produce better prediction accuracy. Our intention is to examine variables, from the GARCH model, which can improve the accuracy of the prediction model for air pollution.

Accordingly, the purpose of this study is, using key observation data from local monitoring stations, to analyze whether the hourly  $PM_{2.5}$  presents an autoregressive conditional heteroscedasticity (ARCH) effect and whether incorporating the GARCH effect into the GA-SVM model can improve the performance of an air pollution prediction model and further identify factors affecting  $PM_{2.5}$ . In the first stage of analysis, the examination starts with an ADF test for stationary in our time series data, followed by an LM test and an ARCH test to investigate the existence of autoregression and ARCH effect in our dataset. We further estimate using a GARCH(1,1) model to confirm the existence of the GARCH effect and integrate the GARCH effect into the GA-SVM model with  $PM_{2.5}$  as the prediction variable in the second stage of analysis [11]. Empirical results indicate that the prediction performance of our proposed alternative model outperformed traditional SVM and GA-SVM models in terms of both MAPE and RMSE as accuracy measures.

Consistent with previous SVM literature, which suggests a trend to integrate various approach into SVM model [8], our empirical results provide evidence to support our expectation that adopting an SVM-based approach model for  $PM_{2.5}$  prediction is appropriate and that prediction performance can be improved by integrating models, such as incorporate the GARCH effect into a GA-SVM-based approach. Moreover, consistent with our prior expectation, evidence further supports that taking the GARCH effect into account, in a

GA-SVM model, clearly improves the accuracy of prediction. To the knowledge of the authors, this study is the first to integrate the GARCH effect into the GA-SVM model in the prediction of PM<sub>2.5</sub>. In summary, with regard to the development of sustainability for both regulators and practitioners, our results strongly encourage them to take the GARCH effect into consideration in air pollution prediction if the regression-based model is to be adopted. Furthermore, this study may shed light on the application of the GARCH model, as well as machine learning methods, in the air pollution prediction literature.

## 2. Materials and Methods

### 2.1. Literature Review

After the announcement of Kyoto Protocol, the prediction and management of air pollution has become an important common concern among industry, government and academics around the world. Scholars have put a lot of effort into empirical research and have obtained quite significant results. Nevertheless, the pursuit of the best predictive model has not reached a consistent conclusion. Here, we review key studies from the literature relevant to this study.

Zickus et al. employed daily average of PM<sub>10</sub> concentrations monitored in the Terro region of Helsinki, Finland from 1996 to 1999 as a sample and variables such as wind speed, wind direction, air pressure, humidity, precipitation, temperature, dew point temperature, and terrain to predict daily PM<sub>10</sub> concentration in 1999 with a three-year training period from 1996 to 1998 [12]. Empirical methods include logistic regression, decision tree, multiple adaptive regression splines (MARS) and artificial neural networks (ANN). Results show that logistic regression, multiple adaptive regression, and neural networks perform more consistently, while decision trees perform significantly worse. Dudot et al. employed a neural network combined with a neural classifier to predict hourly maximum ozone concentrations in central France [13]. The neural model is based on the MLP structure, and the sample data is collected from the French air quality agency LIG'AIR, which has 15 ground monitoring stations, and this study only focuses on three stations in Orleans. The daily maximum data for hourly ozone mean concentrations from 1999 to 2003 were adopted. Since the ozone observation peaks in summer, the authors only used data from April to September. Nonetheless, the model developed can be used to make valid forecasts throughout the year. Results showed that the use of neural networks for ozone peaks produced better predictions, with a 92% concordance index, MAE = RMSE = 15 µg/m<sup>3</sup>, MBE = 5 µg/m<sup>3</sup>, as compared to the European threshold for hourly ozone of 180 µg/m<sup>3</sup>. In order to improve the accuracy of prediction, the authors use a neural classifier with a sigmoid function in the output layer. The output range of the network was [0,1], which can be interpreted as the probability of exceeding the standard. Comparing this model with logistic regression shows that the prediction accuracy index using the neural classifier is 78%, compared to 65% to 72% for the classical MLP. Voukantsis et al. collected air quality monitoring data from 2001 to 2003 at Kallio and Vallia stations in Thessaloniki, Greece and at Sindos and Agias Sofias stations in Helsinki, Finland to predict daily average of PM<sub>10</sub> concentration in 2003 through a neural network model [14]. Research results show that the predicted R<sup>2</sup> values of PM<sub>10</sub> at the Kallio, Vallila, Sindos and Agias Sofias stations were 0.587, 0.639, 0.472 and 0.427, respectively. RMSE was 5.884, 7.128, 16.387 and 23.577 µg/m<sup>3</sup>, respectively. Higher values of R<sup>2</sup> did not correlate with lower RMSE. In addition, the concordance index between measured and modeled daily PM<sub>10</sub> concentrations was between 0.8 and 0.85, while the predicted PM<sub>10</sub> kappa index for both cities reached 60%. Empirical findings indicate that despite significant differences in the environment of the two sample cities, the performance of the PM<sub>10</sub> prediction model was not significantly different.

Russo et al. applied a neural network to predict air quality, with hourly NO<sub>2</sub> concentration data collected from Chelas and Avenida da Liberdade stations in Lisbon, Portugal from 1 January 2002 to 31 December 2006 [15]. The prediction is divided into two parts. In the first part, only the period from 2002 to 2005 was considered and the data of the first three years was used as training data to predict the year of 2005. Next, the data from 2002, 2003

and 2005 were adopted for training to predict 2004, and so on for each year. In the second part, the authors used the parameter values obtained in the first part to predict the NO<sub>2</sub> concentration value in 2006. Selected variables include NO<sub>2</sub>, NO, CO as well as temperature, wind speed, wind direction, humidity, air pressure, solar radiation, and atmospheric boundary height. Empirical results showed that the RMSE is about 20 or above. Tamas et al. used ANN to predict the next 24 h ozone concentration in Corsica, France [16]. Sample data were collected from four stations, Canetto, Giraud, Montesoro and Sposata, two in cities and two in suburbs, from 2008 to 2012 to establish an MLP-based ANN early warning model. The input variables include the concentrations of O<sub>3</sub>, NO<sub>2</sub>, MET and TI, and these variables were configured into five combined models for prediction, namely O<sub>3</sub>, O<sub>3</sub>+NO<sub>2</sub>, O<sub>3</sub>+NO<sub>2</sub>+MET, O<sub>3</sub>+NO<sub>2</sub>+TI and O<sub>3</sub>+NO<sub>2</sub>+MET+TI. Empirical results showed that for the Canetto station, the prediction of the MLP (O<sub>3</sub>+NO<sub>2</sub>+TI) model was more accurate, the RMSE was 17.31 µg/m<sup>3</sup>, and the MAE was 13.66 µg/m<sup>3</sup>; for the Sposata station, the MLP (O<sub>3</sub>+NO<sub>2</sub>+MET+TI) model was more accurate, the RMSE was 16.90 µg/m<sup>3</sup>, and the MAE was 13.31 µg/m<sup>3</sup>; for the Giraud station, the prediction of the MLP (O<sub>3</sub>+NO<sub>2</sub>+MET+TI) model was more accurate, the RMSE was 16.53 µg/m<sup>3</sup> and the MAE was 12.82 µg/m<sup>3</sup>; for the Montesoro station, the MLP (O<sub>3</sub>+NO<sub>2</sub>+MET+TI) model was more accurate in prediction, the RMSE was 15.90 µg/m<sup>3</sup> and the MAE was 12.30 µg/m<sup>3</sup>. Zhou et al. proposed an EEMD-GRNN (ensemble empirical mode decomposition—general regression neural network) model for predicting PM<sub>2.5</sub> concentrations [17]. EEMD decomposes the raw data of PM<sub>2.5</sub> into several intrinsic mode functions (IMFs), while GRNN is used to predict each IMF. This study adopted input variables obtained from a principal component regression (PCR) model to train a hybrid EEMD-GRNN model to remove redundancy. Training data was from 1 January to 1 November in 2013, and the model was tested with data from 2 November to 21 November 2013 in Xi'an Province, China. The resulting MAE values of the multiple linear regression (MLR) model, PCR model, traditional integrated moving average autoregression (ARIMA) model, GRNN model and EEMD-GRNN model were 29.30, 26.35, 23.55, 23.50 and 19.80, respectively. MAPE values were 37.00, 32.75, 29.67, 31.25 and 28.01, respectively, and MSE values were 37.42, 31.30, 28.98, 32.31 and 29.41, respectively. The EEMD-GRNN model outperformed the MLR model, PCR model, ARIMA model and GRNN model without EEMD and can be adopted to develop an early warning system for air quality. Elangasinghe et al. used neural networks combined with k-means clustering to grasp the complex time series of PM<sub>10</sub> and PM<sub>2.5</sub> concentrations in coastal New Zealand [18]. Wind speed, wind direction, solar radiation, temperature and relative humidity were used as variables and the adjacent pollution sources were used as references. Results support improved prediction accuracy based on the values of the correlation coefficient and RMSE. The correlation coefficient between observed and predicted PM<sub>2.5</sub> concentrations increased from 0.77 to 0.79, and PM<sub>10</sub> concentrations increased from 0.63 to 0.69. The RMSE index values of PM<sub>2.5</sub> and PM<sub>10</sub> were reduced from 5.00 to 4.74 and from 6.77 to 6.34, respectively. Kristiani et al. implemented short-term prediction of PM<sub>2.5</sub> in Taiwan using the long short-term memory (LSTM) deep learning method [19]. Results indicate that LSTM had the lowest RMSE value at 1.9, as compared to other models such as CNN at 3.5, Bi-LSTM at 2.5, Bi-GRU at 2.7 and RNN at 2.4.

Zheng et al. applied neural networks and linear regression as spatial and temporal prediction models and combined with regression trees [20]. The study included monitoring data from 43 cities in China from 1 May 2014 to 30 April 2015 and combined temporal predictors, spatial predictors, prediction aggregators and deformation predictors to forecast air quality at Beijing, Shanghai, Tianjin, and Guangzhou for the following 48 h. Results indicate that the model can achieve an accuracy of 0.75 in the first 6 h and 0.6 in the next 7 to 12 h. Even though forecast accuracy in Beijing was the worst among the four cities, it was still better than the weather forecasting model (WFM) adopted by the Beijing Environmental Protection Monitoring Center. Feng et al. proposed a hybrid model that combines air quality trajectory analysis and wavelet transformation with a neural network (ANN) to predict PM<sub>2.5</sub> concentrations beyond two days, and its accuracy was observed [7].

The sample data was collected from 13 air quality monitoring stations in Tianjin and Hebei, China from 1 September 2013 to 31 October 2014. Wind speed and wind direction were set as parameters affecting air quality. The prediction results show that this hybrid model can effectively improve prediction accuracy of  $PM_{2.5}$ , and its RMSE value can be reduced by as much as 40% on average. In particular, the days with high  $PM_{2.5}$  concentration can almost be predicted by wavelet decomposition, and the detection rate (DR) can reach 90% on average for the alert threshold set by the hybrid model.

Wang et al. established an urban air quality prediction system based on the weather research forecast and chemistry (WRF-Chem) model and a regional haze weather forecast system based on the Regional Atmospheric Environment Modeling System (RegAEMS) and applied to Shanghai and Nanjing in the Yangtze River Delta region of China [21]. The study conducted a one-year forecast in Shanghai from May 2009 to April 2010 and a one-month test in Nanjing in October 2007. Results show that WRF-Chem performs well in the prediction of  $SO_2$ ,  $NO_2$ , and  $PM_{10}$ , with the prediction accuracy of API index in Shanghai and Nanjing of 50–83% and 80%, respectively. RegAEMS performed well in haze weather forecasting in terms of RH,  $PM_{2.5}$  and visibility. The accuracy rates of Shanghai and Nanjing were 77% and 58%, respectively. The authors developed new classification criteria by taking relative humidity,  $PM_{2.5}$  and visibility as key parameters. Saide et al., applied WRF-Chem model combined with a two-kilometer grid to build a forecasting system to predict  $PM_{2.5}$  concentration for the next one to three days [22]. The test period was from April to August in 2014 and the sample included hourly  $PM_{2.5}$  observations at nine cities in Chile and the United States: Santiago, Rancagua, Curico, Talca, Chillan, Los Angeles, Temuco, Valdivia and Osorno. Empirical results show that the prediction accuracy ranged from 50 to 70%, while the optimal initialization was 61 to 76%.

Delavar et al. established an air pollution prediction model to predict  $PM_{10}$  and  $PM_{2.5}$  concentrations in Tehran [8]. The day of the week, month, topography, meteorology and pollution rates of two neighboring areas were adopted as input parameters for the machine learning methods adopted including SVR (support vector regression), NARX (nonlinear autoregressive exogenous), ANN and GWR (geographically weighted regression). Cross validation was applied on results to evaluate the best method for modeling air pollution predictions. Empirical results show that, SVR, NARX, ANN and GWR can reduce the RMSE of  $PM_{10}$  by 53%, 47%, 47% and 94%; and predict the RMSE of  $PM_{2.5}$  by 58%, 57%, 61% and 94%, respectively. The best prediction method was NARX with external input. Using the proposed prediction model, its RMSE value reached 1.79. In addition, using a genetic algorithm (GA), the authors found that variables such as day of the week, month, topography, wind direction, maximum temperature and pollution rates in two neighboring areas were the most effective parameters for predicting air pollution. Hu et al. used the hourly CO concentration values of four stations, namely, Liverpool, Chullora, Rozelle and Prospect, in Sydney, Australia from May 2009 to May 2016 as samples [23]. Using SVR as the method, CO concentration values were predicted and compared with the prediction results of ANN. Empirical results show that when MAE is used as evaluation index, prediction accuracy of SVR and ANN are 0.314 and 0.435, respectively; when RMSE is used, prediction accuracy of SVR and ANN are 0.414 and 0.677, respectively. In summary, the prediction results of SVR are more accurate than those of ANN.

Davis and Speckman adopted the generalized additive model (GAM) method to establish an air quality prediction system to estimate the next-day maximum and the average ozone concentration over an 8 h period (10 a.m. to 5 p.m.) in Houston [24]. The study collected ozone data from 10 stations in the Houston area from 1983 to 1991, as well as meteorological data at international airports. Data from April to October from 1983 to 1987 and from 1989 to 1990 were used as the training period, and 1988 and 1991 as the forecasting period. Empirical results indicate that wind direction, opaque cloud cover factor, the previous day's maximum ozone concentration, current day's maximum temperature and morning mixing depth were all very important variables in the model. In addition, the 8 h prediction results of the average ozone concentration at each station

showed that the RMSE ranged from 13.2 to 16.3 ppb ( $R^2$  is ranged from 0.66 to 0.73); the prediction results of the maximum average ozone concentration indicated an error range from 18.5 to 22.0 ppb ( $R^2$  is ranged from 0.61 to 0.68). Siwek et al. took data collected in southern Warsaw from 2005 to 2007 as a sample to predict the  $PM_{10}$  concentration [25]. Three machine learning networks were adopted: Multilayer Perceptron (MLP), Radial Basis Function (RBF), and SVM. Other models proposed include wavelet-transformed MLP, RBF, and SVM models and a model integrating Blind Source Separation (BSS) and another neural network structure. Empirical results showed that the MAE values of MLP, RBF and SVM models were 6.47, 6.99 and 7.07  $\mu\text{g}/\text{m}^3$ , respectively, and the MAPE values were 26.43, 28.49 and 27.05%, respectively. After wavelet transformation, the MAE values of the MLP, RBF and SVM models were 4.37, 5.76 and 4.93  $\mu\text{g}/\text{m}^3$ , respectively, and the MAPE values were 18.04, 23.43 and 20.93%, respectively. The MAE values of the models integrated by BSS and SVM were 3.89 and 4.03  $\mu\text{g}/\text{m}^3$ , respectively, and the MAPE values were 15.78 and 15.96%, respectively. Results indicate that accuracy of the prediction was improved, and the prediction performance of the model integrated by SVM was the best, which was over 12% higher than the SVM model transformed by wavelet, and higher than the pure RBF model, which was the worst at over 44%.

Sotomayor-Olmedo et al. took monthly air quality monitoring data, including  $O_3$ ,  $NO_2$  and  $PM_{10}$ , from Mexico City as a sample and applied SVM to predict the air pollution quality of each month in 2009 [26]. Parameters were adjusted through three kernel functions and the performance of prediction results were compared. Empirical results showed that in the prediction of  $O_3$  and  $NO_2$ , the SVM model applying the Gaussian kernel function had higher accuracy. The empirical results also indicated that the prediction accuracy of the three kernel functions was lower in the last couple months of the year, especially in December. As for the prediction of  $PM_{10}$ , the Gaussian kernel function mode performed better with a large number of SVMs and the polynomial and spline kernel function modes were relatively accurate with a small number of SVMs.

Song et al. explored a more accurate model in the prediction of power usage load spikes. The study proposed an FKM-ASVM-GARCH ECM model, which integrates GARCH and SVM models, as an alternative model to be compared with traditional FKM-ASVM model which does not include GARCH-modified errors [27]. The study adopted China's electricity supply as an observation sample, using the daily load capacity from June to July 2014 as the training period of SVM and the daily load capacity in August as the test period. Results indicate that MAPE, the evaluation criterion for prediction accuracy, was reduced from 1.72 in traditional method without GARCH to 0.74 in the alternative GARCH model. Therefore, it was suggested that the FKM-ASVM-GARCH ECM is superior to the FKM-ASVM. Integration of the GARCH model can indeed improve the accuracy of the SVM prediction model. Ishak et al. applied SVR and random forest (RF) models to establish a prediction model for the daily maximum ozone concentration at three monitoring stations in Tunisia, namely, Gabes, Ghazela and Manouba [28]. Using the station data of the National Environmental Protection Agency (ANPE) from 20 June 2014 to 30 September 2014 as the observation sample, 36 explanatory variables, including daily maximum ozone concentration ( $\text{max}O_3$ ) and other pollutant concentrations ( $SO_2$ ,  $NO_2$ ,  $NO$  and  $PM_{10}$ ), were adopted to explain daily maximum ozone concentration. The experimental results showed that prediction performance of the RF model was better than that of the support vector regression model. The RMSE values of the RF model at the Gabes, Ghazela and Manouba stations were 2.26, 4.16, and 6.71, respectively; the MAE values were 1.85, 3.18 and 5.29 and the MAPE values were 4.08, 3.51 and 8.63, respectively. Lin et al. adopted three machine learning methods, including decision tree regression (DTR), gradient boosted tree regression (GBTR) and SVR to predict  $PM_{2.5}$  concentration in the next hour at 67 locations in Taiwan through a big data platform, with RMSE and MAE as the accuracy evaluation criteria [29]. Results showed that the RMSE of DTR, GBTR and SVR methods were 8.52, 5.17 and 4.68, respectively, and the MAE indicators were 6.25, 3.63

and 3.46, respectively. A preliminary conclusion suggests that SVR is considered to be the better prediction model.

Altogether, with recent evolution of quantitative methods, methodology of the air pollution quality prediction in the literature has evolved from traditional Logistic regression analysis to the application of machine-learning-based approaches. The most widely used methods include neural network analysis (ANN) and SVM. The methodology of ANN has evolved from a traditional single-layer input and output to a multi-layer recurrent neural network analysis method (RNN). However, the need for huge volume of (big) data to improve accuracy in RNN method has become a challenge for empirical study with constrained data collection.

Another development trend is the use of hybrid prediction models combined with other methods, such as wavelet analysis, in pursuit of higher prediction accuracy. In terms of SVM methodology, it is moving towards combining other algorithms, such as the combination of the genetic algorithm (GA) model and machine learning methods, such as GA-SVM model. There is still a lack of consensus about which methodology can provide the best prediction accuracy. Finally, the choice of air pollution predictors is also inconclusive among the literature. The intention of this study is to examine predictors and alternative prediction models that integrate the GARCH effect into the GA-SVM model, to improve prediction accuracy for air pollution.

## 2.2. Dataset

Data for our empirical study was retrieved from the Environmental Protection Agency (EPA) database in Taiwan. The study adopts the central Taiwan region, which has the worst PM<sub>2.5</sub> density, as the observed sample. The choice of predictive variables refers to previous literature related in Section 2.1. However, due to the availability of data, nine variables were adopted for our examination: fine particles (PM<sub>2.5</sub>), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO<sub>2</sub>), nitrogen oxide (NO<sub>x</sub>), ozone (O<sub>3</sub>), suspended particulate matter (PM<sub>10</sub>), sulfur dioxide (SO<sub>2</sub>), wind direction (WindDirection), and wind speed (WindSpeed).

Sampling for this study includes hourly observation data from five air pollution monitoring stations, including FongYen, SaLu, DaLi, ChungMin and SeaTun stations, in central Taiwan. The observation period for our study was from 20 October 2020 to 16 December 2020. The dataset was split into two samples, training data and testing (holdout) data, to evaluate the prediction performance. We held the testing data as the out-of-sample test and used the holdout test to predict hourly PM<sub>2.5</sub> concentration in next 8 h in January 2021. In Table 1, we present basic features of the data, including the locations of five monitoring stations, duration and frequency of data and observation numbers in each station. There are three major reasons for choosing the research period from 20 October 2020 to 16 December 2020. First of all, seasonal characteristics are distinct among the four seasons in Taiwan. Especially, the effects of subtropical island weather, temperature and continental air mass on air quality are unique in winter season. The most serious air pollution problems occur around wintertime. Thus, we adopted observation data from this period for examination considering the seasonal effect. Second, in this dataset with over 1200 hourly observations in each station, from a statistical point of view, the mean behavior of the sample is closer to observation data, which makes it more representative for splitting hourly observation data into training data and testing (holdout) data. Third, the data was retrieved from the EPA database in Taiwan. There is redundant overlap in later period observations and clean data is not easy to retrieve, which is why we adopted sample data from the 20 October to 16 December period.

**Table 1.** Basic features of observation data.

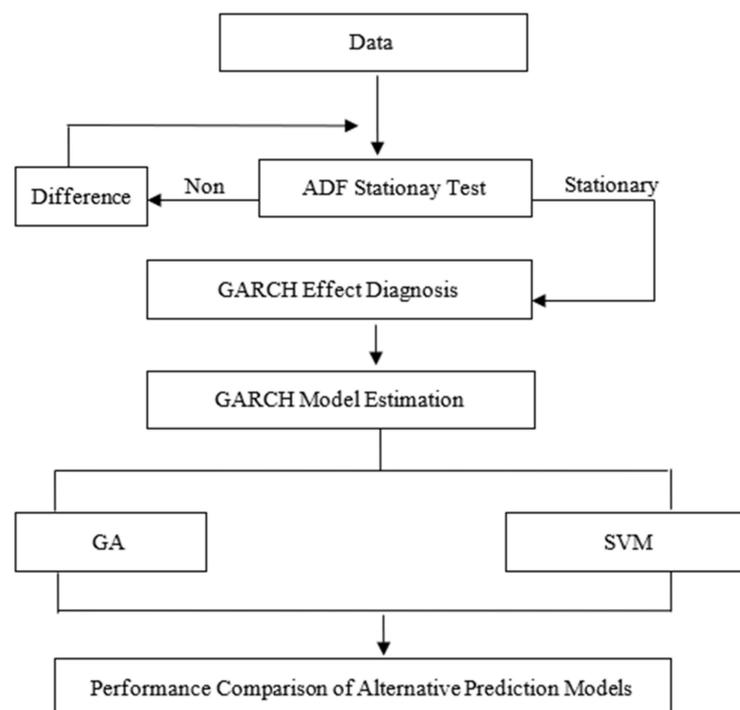
	Station 1	Station 2	Station 3	Station 4	Station 5
Location	FongYen	SaLu	DaLi	ChungMin	SeaTun
Duration	58 days				
Frequency	Hourly	Hourly	Hourly	Hourly	Hourly
Observations	1211	1211	1215	1211	1211

Note: Stations 1, 2, 3, 4 and 5 represent monitoring stations in the counties of FongYen, SaLu, DaLi, ChungMin and SeaTun, respectively. Data was retrieved from the EPA database in Taiwan. The observation period was from 20 October 2020 to 16 December 2020.

### 2.3. Methodology

In this study, we attempting to apply machine learning methods to estimate the degree of  $PM_{2.5}$  based on hourly observation data. Due to the influence of terrain, temperature, humidity, wind direction and so on, the  $PM_{2.5}$  in the previous period may not have completely dissipated and will affect the  $PM_{2.5}$  concentration in the next period, thus the phenomena of autoregression and heteroscedasticity, which are common in time series data, might exist in the dataset. Therefore, it is necessary to check the time series features before the prediction. Where the phenomena of autoregression and heteroscedasticity exist, we applied the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model to capture the time series characteristics, and then examined various cross-model integration methods, including integrating the GARCH effect in the GA-SVM model, to establish a method that provides the highest prediction accuracy.

A two-stage approach will be performed in this study. In the first stage, as shown in the upper part of Figure 1, we start with an Augmented Dickey–Fuller (ADF) unit root test for stationary, followed by the GARCH effect diagnosis, including LM test and ARCH test, and end with GARCH model estimation.

**Figure 1.** The Analysis Process Flow Chart.

#### 2.3.1. Unit Root Test

Granger and Newbold suggested that, if an analysis is carried out with time series data in a non-stationary state, the results will be biased and lead to spurious regression [30]. In such cases, difference should be carried out till the time series data is stationary for

further analysis. For the stationary test method, we follow Engle and Granger and adopt the Augmented Dickey–Fuller (ADF) test to check whether the PM<sub>2.5</sub> concentration in the dataset is a stationary time series [31]. Depending on the inclusion of intercept or trend, or not, the three variations of ADF test model are:

1. Model with neither intercept nor trend

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=2}^p \beta_i \Delta y_{t-i+1} + \varepsilon_t \quad (1)$$

2. Model with intercept but without trend

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=2}^p \beta_i \Delta y_{t-i+1} + \varepsilon_t \quad (2)$$

3. Model with both intercept and trend

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \alpha_2 t + \sum_{i=2}^p \beta_i \Delta y_{t-i+1} + \varepsilon_t \quad (3)$$

where  $\alpha_0$  is the intercept;  $t$  is the trend for time;  $\gamma$ ,  $\beta_i$  and  $\alpha_2$  are parameters to be estimated;  $p$  is the optimal lag order; the residual is  $\varepsilon_t \sim iid(0, \sigma^2)$  and fit white noise; the null hypothesis is  $H_0 : \gamma = 0$ . If the test statistics are significant and the null hypothesis is rejected, this time series data does not have a unit root phenomenon and belongs to a stationary time series.

### 2.3.2. ARCH Test

When the conditional variance of the regression residuals is not uniform, the estimated coefficient is not valid. Therefore, in traditional quantitative empirical analysis, testing whether the model has heterogeneous variance (Heteroscedasticity) has become the main step in diagnosing the model. In empirical analysis, before fitting the GARCH correlation model, it is necessary to check whether the sample time series data has the feature of heterogeneous variation, that is, whether there is an ARCH effect, as the basis for whether the ARCH model can be configured. For the test method, we applied the Lagrange multiplier (LM) test proposed by Engle to test whether the ARCH effect was present [9]. The testing steps were as follows.

(1) We first run the OLS regression to estimate the appropriate mean equation:  $y_t = x_t \hat{\alpha}$ , where  $\hat{\alpha}$  is the regression coefficient estimated by OLS, and the residual  $\hat{\varepsilon}_t = y_t - x_t \hat{\alpha}$  is calculated accordingly, and then save the residual square  $\hat{\varepsilon}_t^2$  as another time series.

(2) Regress the residual square estimate  $\hat{\varepsilon}_t^2$  on intercept and  $q$  lagging terms to calculate the coefficient of determination,  $R^2$ , of this regression analysis. The estimation function is as follows:

$$\hat{\varepsilon}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\varepsilon}_{t-1}^2 + \hat{\alpha}_2 \hat{\varepsilon}_{t-2}^2 + \dots + \hat{\alpha}_q \hat{\varepsilon}_{t-q}^2 \quad (4)$$

(3) Multiply the determined coefficient,  $R^2$ , by the number of samples,  $T$ , to calculate the LM test statistic,  $LM = T \times R^2 \sim \chi^2(q)$ , where the LM statistic approaches the chi-square distribution with the degree of freedom  $q$ . If the resulting LM test statistic significantly rejects the null hypothesis:  $H_0 : \hat{\alpha}_1 = \hat{\alpha}_2 = \dots = \hat{\alpha}_q = 0$ , it means that the time series data inspected has an ARCH effect, and an ARCH or GARCH model should be further fitted.

### 2.3.3. GARCH Model

If the ARCH effect exists in the hourly PM<sub>2.5</sub> concentration data, the conditional heterogeneity variance model will fit. Econometricians have proposed correction meth-

ods to improve the heteroscedasticity of time series data. Among them, Engle [9] and Bollerslev [10] are the most popular.

Engle considered the conditional variance to change over time and included it in the autoregressive conditional heterogeneous variance (ARCH) model, allowing the conditional variance to be a function of the squared term of residual in previous period [9]. Thus, previous volatility will affect the subsequent volatility, which is in line with the phenomenon of volatility clustering. The model specification of ARCH( $q$ ) can be described as follows:

$$\begin{aligned} y_t &= ax_t + \varepsilon_t \\ h_t = \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \\ \varepsilon_t | \Omega_{t-1} &\sim N(0, h_t) \end{aligned} \quad (5)$$

where  $\alpha_0 > 0$ ,  $\alpha_i > 0$ ,  $i = 1, 2, \dots, q$ ;  $y_t$  is the time series data;  $ax_t$  is the conditional mean of  $y_t$ ;  $\Omega_{t-1}$  is the information collected up to period  $t - 1$  and  $h_t$  is the conditional heterogeneity variance of  $y_t$ .

Bollerslev further added the lag conditional variance to the ARCH model, so that ARCH conforms to the traditional ARMA process, which is called the generalized autoregressive conditional heterogeneous variation (GARCH) model [10]. The conditional variance is not only affected by the squared term of residual in previous period, but also by the conditional variance in previous period. The GARCH ( $p, q$ ) model is stated as follows:

$$\begin{aligned} y_t &= ax_t + \varepsilon_t \\ h_t = \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \\ \varepsilon_t | \Omega_{t-1} &\sim N(0, h_t) \end{aligned} \quad (6)$$

where  $\alpha_0 > 0$ ,  $\alpha_i > 0$ ,  $i = 1, 2, \dots, q$ ;  $\beta_j > 0$ ,  $j = 1, 2, \dots, p$ ;  $y_t$  is the time series data;  $ax_t$  is the conditional mean of  $y_t$ ;  $\Omega_{t-1}$  is the information collection up to period  $t - 1$ ;  $h_t$  is the conditional heterogeneity variance of  $y_t$ .

We follow the GARCH (1,1) specification since it represents most popular specification based on prior literature [10,11,32,33].

After the GARCH model was estimated, to make sure the GARCH (1,1) model specification is at its optimal level, we applied the ARCH-LM test to check the model fit as well.

The estimation function of ARCH-LM test is as follows:

$$\hat{\varepsilon}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\varepsilon}_{t-1}^2 + \hat{\alpha}_2 \hat{\varepsilon}_{t-2}^2 + \dots + \hat{\alpha}_q \hat{\varepsilon}_{t-q}^2 + v_t \quad (7)$$

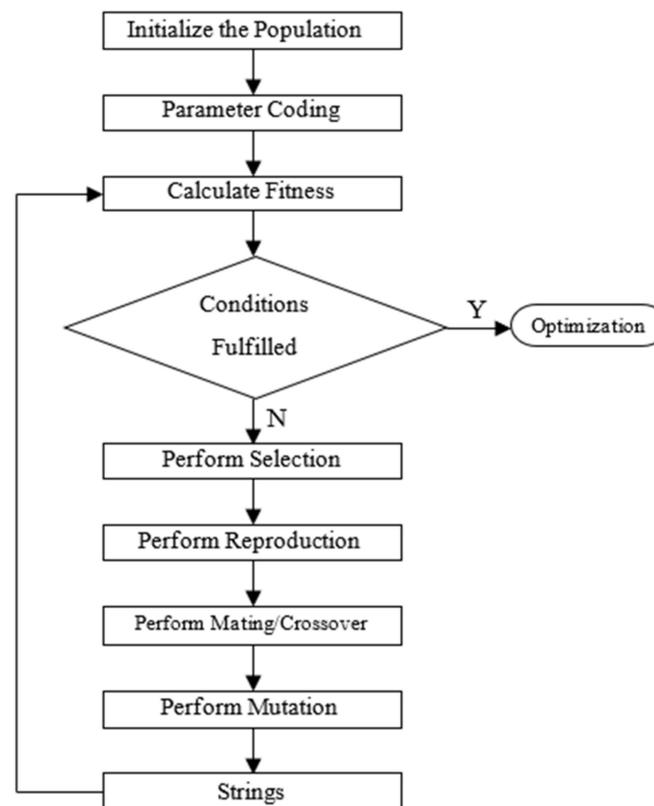
Since the SVM model is a regression-based model, when autoregression exists in the dataset, the prediction result will be biased. If the GARCH effect was confirmed in the first stage, in the second stage of analysis, we integrate the GARCH effect into the GA-SVM model with PM<sub>2.5</sub> as the prediction variable. We add the PM<sub>2.5</sub> in the previous period (PM<sub>2.5</sub>( $t-1$ )), the conditional heterogeneity variance in the previous period ( $h_{t-1}$ ) and the squared term of the residual in the previous period,  $\varepsilon^2(t-1)$ , from the GARCH model estimation into the GA-SVM model to establish an alternative PM<sub>2.5</sub> prediction model and to compare the prediction accuracy with traditional GA-SVM model which does not take the GARCH effect into consideration.

#### 2.3.4. GA-SVM Model

The Genetic Algorithm (GA) was proposed by John Holland and is based mainly on Darwin's theory of evolution to simulate the "natural selection" in the evolution of the biological world [34]. The natural elimination mechanism of "survival of the fittest" is widely applied in solving optimization problems, data search, artificial intelligence and machine learning [35]. In the financial field, many scholars have also applied the genetic

algorithm to examine various topics, such as: trading systems, stock or portfolio selection, bankruptcy prediction, credit evaluation, budget allocation, etc. [36].

The genetic algorithm mainly operates through three processes: reproduction, crossover and mutation. During the reproduction process, an initial population is randomly generated, and each individual is coded in binary and substituted into a fitness function. Then, based on the obtained fitness value, individuals with high fitness are selected and reproduced to the mating pool. Two individuals are selected randomly in the mating pool for mating each time and the algorithm decides whether the resulting offspring should undergo further mutation. The process of reproduction, crossover and mutation is repeated until the most resilient population is produced. A processing flow chart of the genetic algorithm is shown in Figure 2 [36].



**Figure 2.** Processing Flow Chart of the Genetic Algorithm (GA).

SVM is a learning method widely applied in classification-related topics. SVM was proposed in 1995 by Vladimir Naumovich Vapnik and the AT&T laboratory team [37]. SVM is a machine learning system developed based on the Structural Risk Minimization (SRM) method in statistical learning theory. The main concept of SVM is to use a separating hyperplane to divide data into two or more classes and to deal with the problem of classification in data mining.

### 2.3.5. Evaluation Indicators for Prediction Models

Regarding the performance evaluation of prediction models, four indicators are generally adopted, including mean-absolute percentage error (MAPE), root mean squared error (RMSE), mean absolute error (MAE) and correlation coefficient (CC) [38]. Specifications of the four indicators are presented in Table 2 based on Witten, Frank and Hall [38]. MAPE is the most commonly used criteria for prediction performance evaluation [36]. We further added RMSE to reinforce our evaluation and will present results in Section 3.3.

**Table 2.** Performance Measures for Numeric Prediction.

Mean absolute percentage error (MAPE)	$\frac{100\%}{n} \times \left[ \frac{\rho_1 - a_1}{a_1} + \dots + \frac{\rho_n - a_n}{a_n} \right]$
Root mean squared error (RMSE)	$\sqrt{\frac{(\rho_1 - a_1)^2 + \dots + (\rho_n - a_n)^2}{n}}$
Mean absolute error (MAE)	$\frac{ \rho_1 - a_1  + \dots +  \rho_n - a_n }{n}$
Correlation coefficient (C. C.)	$\frac{S_{PA}}{\sqrt{S_P S_A}} \quad S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n-1}$ $S_P = \frac{\sum_i (p_i - \bar{p})^2}{n-1} \quad S_A = \frac{\sum_i (a_i - \bar{a})^2}{n-1}$ Here, $\bar{a}$ is the mean value over the test data.

Lewis stated that MAPE is often applied to evaluate the predictive ability of a model [39]. The smaller the MAPE, the better the prediction performance. The denominator of percentage termed MAPE is the actual value, thus there is no problem of unstable comparison basis due to the size of the value. When MAPE is the measure, the value of (1-MAPE) represents the accuracy of prediction; thus, the lower the MAPE value, the better the predictive ability. RMSE mainly measures the degree of deviation between the predicted value and the actual value. The degree of deviation is standardized with the actual value of the variable, so the predictive ability of each variable can be compared. When RMSE is the measure, the closer its value is to 0, the better the predictive ability.

In Table 3, we present the interpretation of MAPE values based on Lewis [39]. For example, if the MAPE value is less than 10%, the prediction performance will be classified as “highly accurate forecasting”. When the MAPE value is above 50%, the prediction performance will be classified as “inaccurate forecasting”.

**Table 3.** Interpretation of MAPE values.

MAPE	Interpretation
MAPE < 10%	Highly accurate forecasting
10% < MAPE < 20%	Good forecasting
20% < MAPE < 50%	Reasonable forecasting
50% < MAPE	Inaccurate forecasting

### 3. Results and Discussion

#### 3.1. GARCH Effect Diagnosis

In Table 4, we present the results of the GARCH effect diagnosis in each of five air pollution monitoring stations including FongYen, SaLu, DaLi, ChungMin and SeaTun stations in Taichung, Taiwan, which are labeled as Stations 1, 2, 3, 4 and 5, respectively, in Table 3. Before the diagnosis, we start with an ADF unit root stationary test, as suggested by Engle and Granger [31], and find that the test statistics in all five stations are significant with probability of chi-square equal to 0.0000. The results significantly reject the null hypothesis of non-stationary at alpha = 0.01 level of confidence, thus the PM<sub>2.5</sub> time series data in each station are all stationary and can be used for further estimation.

As shown in Table 4, in all five monitoring stations, F-statistics of the OLS model are all statistically significant at 0.01 level and the adjusted R<sup>2</sup> values are at least 0.7773. The coefficient estimates and the t-statistics of each variable at each station are presented in the table. The coefficient estimates of PM<sub>2.5</sub> concentration in previous period (PM<sub>2.5(t-1)</sub>) at Station 1 to 5 are 0.5408, 0.5828, 0.4945, 0.5563 and 0.5543, respectively; all are statistically significant at 0.01 level. Similar results can be found on the coefficient estimates of PM<sub>10</sub> and sulfur dioxide (SO<sub>2</sub>); both are statistically significant at 0.01 level for all five stations. The coefficient estimates of ozone (O<sub>3</sub>) were not significant for all five stations.

**Table 4.** Results of GARCH Effect Diagnosis.

Variable	Station 1	Station 2	Station 3	Station 4	Station 5
PM <sub>2.5(t-1)</sub>	0.5408 25.80 ***	0.5828 31.70 ***	0.4945 24.83 ***	0.5563 30.10 ***	0.5543 28.47 ***
CO	-0.9842 -1.94	0.5406 1.14	-0.1548 -0.47	-0.2233 -0.51	0.9440 1.70
NO	1.2822 1.55	-0.0410 -0.56	-0.0046 -0.01	0.2324 0.47	0.2220 0.34
NO <sub>2</sub>	1.2568 1.54	0.1181 2.19 **	-0.1277 -0.23	0.2902 0.59	0.1788 0.27
NO <sub>x</sub>	-1.0688 -1.31	-0.0139 -0.31	0.1713 0.30	-0.2036 -0.42	-0.1715 -0.26
O <sub>3</sub>	-0.0102 -0.82	0.0090 0.75	0.0068 0.51	0.0030 0.25	0.0016 0.13
PM <sub>10</sub>	0.1547 13.28 ***	0.1734 16.86 ***	0.1970 18.01 ***	0.1927 17.60 ***	0.1801 16.73 ***
SO <sub>2</sub>	0.4612 9.84 ***	0.5779 9.84 ***	0.6934 11.05 ***	0.5627 8.72 ***	0.3856 6.07 ***
WindDirection	0.0004 0.30	0.0012 1.30	0.0019 1.82	-0.0003 -0.42	-0.0026 -2.52 **
WindSpeed	0.3763 -2.15 **	-0.1501 -1.86*	-0.0007 -0.09	-0.4826 -3.14 ***	-0.5618 -5.11 ***
C	0.8794 1.71 *	-1.3351 -2.75 ***	-1.2469 -3.56 ***	-0.1998 -0.39	2.0251 3.43 ***
Observation	1211	1211	1215	1211	1211
F-statistic	423.2258	739.5768	710.4244	740.6729	602.5055
Prob. (F-stat.)	0.0000	0.0000	0.0000	0.0000	0.0000
Adj. R <sup>2</sup>	0.7773	0.8592	0.8532	0.8594	0.8325
LM Test					
F-statistic	16.2423	20.4623	21.2019	21.6132	20.8101
Prob. (F-stat.)	0.0000	0.0000	0.0000	0.0000	0.0000
Obs*R <sup>2</sup>	31.9702	40.0022	41.4075	42.17369	40.6592
Prob. (Chi <sup>2</sup> )	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH test					
F-statistic	231.4378	24.6600	31.1825	96.8117	148.2826
Prob. (F-stat.)	0.0000	0.0000	0.0000	0.0000	0.0000
Obs*R <sup>2</sup>	194.5480	24.2067	30.4505	89.7771	132.2895
Prob. (Chi <sup>2</sup> )	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Test statistics of ADF test in five stations are significant with the chi-square probability equal to 0.0000; thus, the data in each station does not have a unit root phenomenon and belongs to a stationary time series. Stations 1, 2, 3, 4 and 5 represent monitoring station in the counties of FongYen, SaLu, DaLi, ChungMin and SeaTun, respectively. Variables examined include fine-particle observation in previous period (PM<sub>2.5(t-1)</sub>), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO<sub>2</sub>), nitrogen oxide (NO<sub>x</sub>), ozone (O<sub>3</sub>), suspended particulate matter (PM<sub>10</sub>), sulfur dioxide (SO<sub>2</sub>), wind direction (WindDirection), and wind speed (WindSpeed). t-statistic values are presented below each coefficient estimates. \*\*\*, \*\* and \* indicate statistical significance at 1%, 5% and 10%, respectively.

As we expected, autocorrelation in the PM<sub>2.5</sub> data does exist, and the result of the LM test reconfirms such a phenomenon. As shown in the lower part of Table 4, the F-statistics are all significant at the 0.01 level. Test statistics, Obs\*R<sup>2</sup>, in five stations are 31.9702, 40.0022, 41.4075, 42.1737 and 40.6592, respectively, are all significant at the 0.01 level and all significantly reject null hypothesis of no autocorrelation phenomenon. We further performed an ARCH test to investigate whether a GARCH effect existed in our sample data. As shown at the bottom of Table 4, the F-statistics are all significant at 0.01 level. Test statistics, Obs\*R<sup>2</sup>, in five stations are 194.5480, 24.2067, 30.4505, 89.7771 and 132.2895, respectively, and are all significant at 0.01 level, significantly rejecting the null hypothesis of no GARCH effect, implying that the GARCH effect does exist in our sample data.

The results of the LM test and ARCH test indicate that the PM<sub>2.5</sub> time series data is autoregressive and does have an ARCH effect. A bias would be expected if we further estimated with a regression-based model, such as the SVM model. Thus, an ARCH or GARCH model should be further fitted.

### 3.2. GARCH Estimation

In Table 5, we present the GARCH estimation and model specification for each of the five monitoring stations. We follow the GARCH (1,1) specification since it represents the most popular specification, based on prior literature [29–32]. As shown in the upper section of Table 4, for all five stations, the coefficient estimates of both PM<sub>2.5</sub>(*t*−1), the PM<sub>2.5</sub> observation in previous period and PM<sub>10</sub> were statistically significant at an alpha level of 0.01. Coefficient estimates of PM<sub>2.5</sub>(*t*−1) at stations one to five were 0.5210, 0.5528, 0.4402, 0.5577 and 0.4786, respectively. Coefficient estimates of PM<sub>10</sub> at stations one to five were 0.1634, 0.1635, 0.2327, 0.1755 and 0.2001, respectively. Coefficient estimates of SO<sub>2</sub> at stations one to five were 0.2115, 0.5537, 0.6148, 0.2355 and 0.2485, respectively, and all were statistically significant at  $\alpha = 0.01$  level. The coefficient estimates of Ozone (O<sub>3</sub>) were not significant for all five stations.

**Table 5.** Result of GARCH (1,1) Estimation.

Variable	Station 1	Station 2	Station 3	Station 4	Station 5
GARCH (1,1)					
PM <sub>2.5</sub> ( <i>t</i> −1)	0.5210 27.31 ***	0.5528 33.80 ***	0.4402 31.17 ***	0.5577 41.47 ***	0.4786 28.22 ***
CO	2.9016 6.28 ***	2.4174 8.16 ***	0.5319 1.47	2.5439 9.13 ***	3.3907 5.75 ***
NO	0.9557 1.38	−0.0513 −0.34	0.6952 2.68 ***	0.1474 0.36	0.1618 0.21
NO <sub>2</sub>	1.0488 1.53	0.0779 0.53	0.5710 2.24 **	0.1766 0.44	0.1172 0.16
NO <sub>x</sub>	−0.9170 −1.33	−0.0024 −0.02	−0.5645 −2.21 **	−0.1312 −0.32	−0.1266 −0.17
O <sub>3</sub>	−0.0025 −0.28	−0.0064 −0.64	−0.0153 −1.26	0.0064 0.75	0.0060 0.58
PM <sub>10</sub>	0.1634 18.24 ***	0.1635 26.84 ***	0.2327 31.04 ***	0.1755 29.07 ***	0.2001 26.42 ***
SO <sub>2</sub>	0.2115 5.64 ***	0.5537 16.23 ***	0.6148 9.04 ***	0.2355 5.42 ***	0.2485 3.71 **
WindDirection	−0.0012 −1.28	−0.0006 −0.95	−0.0001 −0.10	−0.0002 −0.28	−0.0024 −3.40 ***
WindSpeed	−0.3804 −2.73 ***	−0.2110 −3.40 ***	0.0008 0.07	−0.2507 −2.02 **	−0.4646 −5.44 ***
C	1.1559 2.76 ***	−0.6811 −1.66 *	−0.4225 −1.69 *	−0.0385 −0.10	1.8648 4.10 ***
Variance Equation					
C	2.8001 7.03 ***	0.5845 6.24 ***	7.1652 9.53 ***	1.3838 5.47 ***	0.9128 4.01 ***
$\varepsilon^2$ ( <i>t</i> −1)	0.3133 8.81 ***	0.1689 8.53 ***	0.4319 10.44 ***	0.2426 8.46 ***	0.1741 10.03 ***
<i>h</i> ( <i>t</i> −1)	0.5639 14.08 ***	0.8122 44.93 ***	0.2937 5.72 ***	0.6820 20.11 ***	0.7891 35.78 ***

Note: Stations 1, 2, 3, 4 and 5 represent monitoring stations in the counties of FongYen, SaLu, DaLi, ChungMin and SeaTun, respectively. Variables examined include fine particle observation in the previous period (PM<sub>2.5</sub>(*t*−1)), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO<sub>2</sub>), nitrogen oxide (NO<sub>x</sub>), ozone (O<sub>3</sub>), suspended particulate matter (PM<sub>10</sub>), sulfur dioxide (SO<sub>2</sub>), wind direction (WindDirection) and wind speed (WindSpeed). t-statistic values are presented below each coefficient estimates. \*\*\*, \*\* and \* indicate statistical significance at 1%, 5% and 10%, respectively.

We also found that the coefficient estimates of nitric oxide (NO), nitrogen dioxide (NO<sub>2</sub>), and nitrogen oxide (NO<sub>x</sub>) were statistically significant at the 0.01, 0.05 and 0.05 levels, respectively, only at station 3 (DaLi) but not significant at the other four stations. Conversely, coefficient estimates of wind speed (WindSpeed) were significant at station 1, 2, 4 and 5 but not significant at station 3 (DaLi). A similar result is found in the coefficient estimates of carbon monoxide (CO), which were significant at Station 1, 2, 4 and 5, all at the 0.01 level, but not significant at station 3 (DaLi).

In the lower section of Table 5, we present the variance estimates of the GARCH model including the squared term of the residual in the previous period ( $\varepsilon^2_{(t-1)}$ ) and the conditional heteroscedasticity in previous period ( $h_{t-1}$ ) at five stations. Coefficient estimates of  $\varepsilon^2_{(t-1)}$  at station one to five are 0.3133, 0.1689, 0.4319, 0.2426 and 0.1741, respectively; all are statistically significant at 0.01 level. Coefficient estimates of  $h_{t-1}$  at stations one to five are 0.5639, 0.8122, 0.2937, 0.6820 and 0.7891, respectively; all are statistically significant at  $\alpha = 0.01$  level.

Empirical results do support our expectation that the PM<sub>2.5</sub> observations in the previous period (PM<sub>2.5(t-1)</sub>), conditional heteroscedasticity ( $h_{t-1}$ ) and residual square of GARCH model ( $\varepsilon^2_{(t-1)}$ ) are appropriate to be incorporated as prediction variables into a GA-SVM model in our second stage procedure to establish an alternative PM<sub>2.5</sub> prediction model. After the GARCH model was estimated, to make sure the GARCH (1,1) model specification was at an optimal level, we perform an ARCH-LM test to check the model fit as well. Test statistics in all five monitoring stations except Station 3 (DaLi) were significant, rejecting the null hypothesis at 0.01 level, indicating that the GARCH model fit is optimized.

### 3.3. Evaluations of the Prediction Models

In the second stage of our analysis, we integrate the GARCH effect into the GA-SVM model by adding PM<sub>2.5</sub> observations in the previous period (PM<sub>2.5(t-1)</sub>), GARCH effect ( $h_{t-1}$ ) and residual of the GARCH model ( $\varepsilon^2_{(t-1)}$ ) to establish an alternative PM<sub>2.5</sub> prediction model (GA-SVM-GARCH) and compare the prediction performance of two traditional approaches, the SVM and the GA-SVM model, to our proposed alternative model. The observation period for our study was from 20 October 2020 to 16 December 2020. The dataset was split into two sample sets, training data and testing (holdout) data, to evaluate the prediction performance. We held the testing data as the out-of-sample test and used the holdout test to predict hourly PM<sub>2.5</sub> concentration in next 8 h in January 2021.

In Table 6, we present the MAPE and RMSE values of the SVM, GA-SVM and GA-SVM-GARCH models. MAPE is the most commonly used criteria for prediction performance evaluation [35]. The prediction accuracy was calculated by subtracting MAPE value from one. For example, the MAPE values of SVM, GA-SVM and GA-SVM-GARCH model in Station 2 (SaLu) are 35.94%, 33.14% and 0.68%, respectively, indicating prediction accuracies of 64.06%, 66.86% and 99.32%, respectively. The results indicate a significant increase in prediction accuracy with the GA-SVM-GARCH model by over 30% as compared to the traditional SVM and GA-SVM models. We further added RMSE as the second evaluation indicator to reinforce our evaluation. When the RMSE value is closer to 0, the predictive ability of the model is better. For example, in Station 2 (SaLu), RMSE values of SVM, GA-SVM and GA-SVM-GARCH models were 4.7724, 4.4938 and 0.0950, respectively, further proving the better prediction performance with the GA-SVM-GARCH model.

In terms of performance comparison, Song et al. adopted China's electricity supply as an observation sample and found that the MAPE was reduced from 1.72% in the FKM-ASVM model without GARCH, to 0.74% in the FKM-ASVM-GARCH ECM model with GARCH [27]. In our proposed alternative model, the MAPE value of Station 4 reduced from 26.89% in GA model and 26.64% in GA-SVM model to 0.14% in GA-SVM-GARCH model. The results indicate that integration of the GARCH model can indeed improve the accuracy of the SVM prediction models and our proposed GA-SVM-GARCH model provides the highest accuracy.

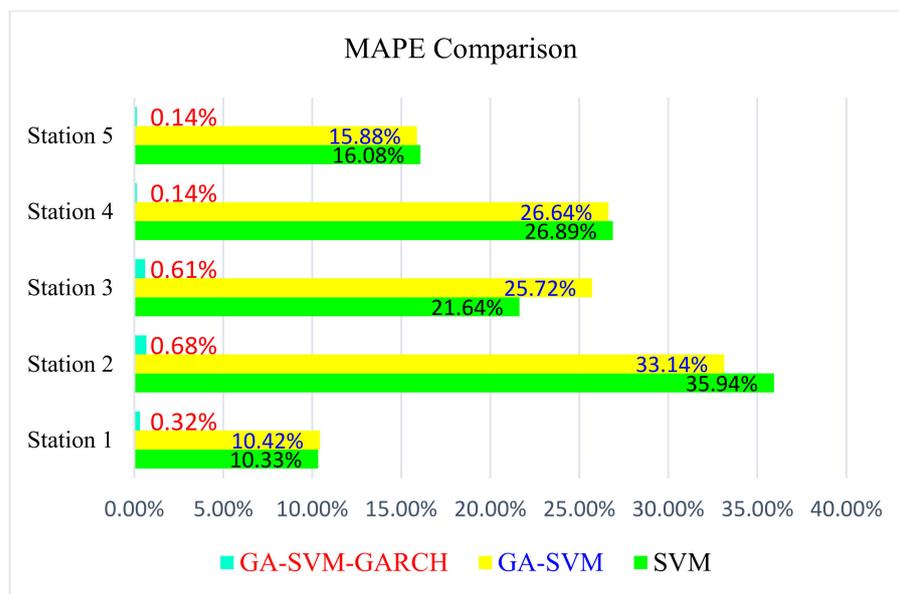
**Table 6.** Performance of Alternative Prediction Models.

Model	Station 1	Station 2	Station 3	Station 4	Station 5
SVM					
MAPE	10.33%	35.94%	21.64%	26.89%	16.08%
RMSE	3.9405	4.7224	2.9855	7.2913	4.2309
GA-SVM					
MAPE	10.42%	33.14%	25.72%	26.64%	15.88%
RMSE	4.0778	4.4938	3.1561	7.3327	4.2144
GA-SVM-GARCH					
MAPE	0.32%	0.68%	0.61%	0.14%	0.14%
RMSE	0.0596	0.0950	0.0632	0.0313	0.0255

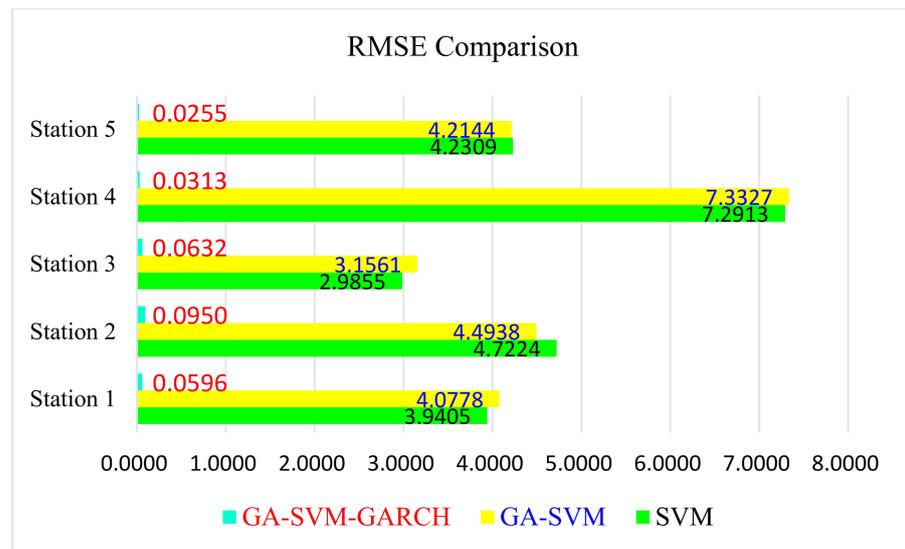
Note: Stations 1, 2, 3, 4 and 5 represent monitoring stations in the counties of FongYen, SaLu, DaLi, ChungMin and SeaTun, respectively.

In Figures 3 and 4, we present graphical abstracts of MAPE and RMSE comparison of three SVM models above in five monitoring stations. As shown in Figure 3, the MAPE values are within a range from 0.14% to 0.68% in the GA-SVM-GARCH model, as compared to 10.42% to 33.14% in the GA-SVM model and 10.33% to 35.94% in the GA model. In Figure 4, the RMSE values are within a range from 0.0225 to 0.0950 in the GA-SVM-GARCH model, as compared to 3.1561 to 7.3327 in the GA-SVM model and 2.9855 to 7.2913 in the GA model.

As shown in Table 6, Figures 3 and 4, the MAPE and RMSE value of the GA-SVM-GARCH model are the lowest compared to those values in the SVM and GA-SVM models. Overall, the performance of our proposed alternative model outperformed traditional SVM and GA-SVM models in terms of both MAPE and RMSE. When we integrated the GARCH effect into the GA-SVM model, the improvement in predicting accuracy exceeded our expectations.



**Figure 3.** MAPE comparison of three SVM models in five monitoring stations.



**Figure 4.** RMSE comparison of three SVM models in five monitoring stations.

#### 4. Conclusions

Air pollution, especially that of  $PM_{2.5}$ , resulting from industrialization has fostered a wave of global weather migration and jeopardized human health in the past three decades. The prediction and control of air pollution, especially  $PM_{2.5}$ , has been a critical issue for regulators, practitioners and academics in Taiwan. Much research has been carried out searching for better prediction models, yet there is no consensus on which is the most accurate approach. In this study we conducted a two-stage analysis and explored whether, by integrating the GA-SVM model with the GARCH effect, we can construct a more accurate air pollution prediction model. The study adopted the region with the worst  $PM_{2.5}$  density, central Taiwan, as the sample source.

In the first stage of analysis, the examination started with an ADF test for stationary in our time series data, followed by an LM test and ARCH test to investigate if autoregression and the ARCH effect existed in our dataset. We further estimated with a GARCH (1,1) model to confirm the existence of a GARCH effect and integrated the GARCH effect into the GA-SVM model with  $PM_{2.5}$  as the predictive variable in the second stage of analysis. Empirical results indicate that the prediction performance of our proposed alternative model outperformed traditional SVM and GA-SVM models in terms of both MAPE and RMSE as the accuracy indicators.

Consistent with previous SVM literature [7,23,26,27], which shows a trend of integrating various approaches into the SVM model, our empirical results provide evidence to support our expectation that adopting an SVM-based approach model for  $PM_{2.5}$  prediction is appropriate and that prediction performance can be improved by integrating models, such as incorporate the GARCH effect into the GA-SVM-based approach. Moreover, consistent with our prior expectation, evidence further support that taking the GARCH effect into account, in the GA-SVM model, clearly improves the accuracy of prediction. To the knowledge of the authors, this study is the first to attempt to integrate the GARCH effect into the GA-SVM model in the prediction of  $PM_{2.5}$ . It is possible to compare the empirical results of prediction in this study to findings in the recent  $PM_{2.5}$  literature, although different methods were adopted. Studies on the Taichung region by Chen et al. [2] and Kristiani et al. [19], cross country (Kusuma et al. [3]), and sub-district in China (Long et al. [5]) are comparable.

In summary, this study has implications for sustainability management by both government and industry. As long as a regression-based approach is adopted, ignoring possible autoregressive characteristic in time series dataset tends to result in lower prediction efficiency and accuracy. Furthermore, if variance clustering exists in a time series dataset, the

choice of prediction model should account for this phenomenon. It is highly recommended to take the GARCH effect into consideration in air pollution prediction to capture variance clustering and to improve prediction efficiency and accuracy. Moreover, this study may shed light on the application of the GARCH model, as well as machine learning methods, in the air pollution prediction literature.

**Author Contributions:** Conceptualization, K.-C.Y., H.-W.H., M.-H.H. and T.-C.W.; methodology, formal analysis, K.-C.Y., H.-W.H., M.-H.H. and T.-C.W.; resources, writing—original draft preparation, K.-C.Y., H.-W.H., M.-H.H. and T.-C.W.; writing—review and editing, K.-C.Y., H.-W.H., M.-H.H. and T.-C.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors would like to acknowledge that the APC was funded by Ministry of Science and Technology, Taiwan (MOST 109-2511-H-018-018-MY3).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data are available on request.

**Acknowledgments:** The authors would like to express our deep appreciation for the insightful comments offered by two anonymous reviewers to enhance the quality and readability of our paper. The authors would also like to acknowledge the financial support provided by Ministry of Science and Technology, Taiwan (MOST 109-2511-H-018-018-MY3).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wikipedia. Air Pollution in Taiwan. Available online: [https://en.wikipedia.org/wiki/Air\\_pollution\\_in\\_Taiwan](https://en.wikipedia.org/wiki/Air_pollution_in_Taiwan) (accessed on 27 February 2022).
2. Chen, H.L.; Li, C.P.; Tang, C.S.; Lung, S.C.C.; Chuang, H.C.; Chou, D.W.; Chang, L.T. Risk assessment for people exposed to PM<sub>2.5</sub> and constituents at different vertical heights in an urban area of Taiwan. *Atmosphere* **2020**, *11*, 1145. [CrossRef]
3. Kusuma, W.L.; Wu, C.D.; Zeng, Y.T.; Hapsari, H.H.; Muhamad, J.L. PM<sub>2.5</sub> pollutant in Asia—A comparison of metropolis cities in Indonesia and Taiwan. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4924. [CrossRef] [PubMed]
4. Xie, M.; Zhou, X. Taiwan Air Pollution Problem and Prevention Policy, National Policy Research Foundation. Available online: <https://www.npf.org.tw/2/18414> (accessed on 27 February 2022).
5. Long, Y.; Wang, J.; Wu, K.; Zhang, J. Population exposure to ambient PM<sub>2.5</sub> at the subdistrict level in China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2683. [CrossRef] [PubMed]
6. Chou, K.T. From anti-pollution to climate change risk movement: Reshaping civic epistemology. *Sustainability* **2015**, *7*, 14574–14596. [CrossRef]
7. Feng, X.; Li, Q.; Zhu, Y.; Hou, J.; Jin, L.; Wang, J. Artificial neural networks forecasting of PM<sub>2.5</sub> pollution using air mass trajectory based geographic model and wavelet transformation. *Atmos. Environ.* **2015**, *107*, 118–128. [CrossRef]
8. Delavar, M.R.; Gholami, A.G.; Shiran, R.; Rashidi, Y.; Nakhaeizadeh, G.; Fedra, R.K.; Afshar, S.H. A novel method for improving air pollution prediction based on machine learning approaches: A case study applied to the capital city of Tehran. *Int. J. Geo-Inf.* **2019**, *8*, 99. [CrossRef]
9. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]
10. Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]
11. Hansen, P.R.; Lunde, A. A forecast comparison of volatility models: Does anything beat a garch(1,1)? *J. Appl. Econom.* **2005**, *20*, 873–889. [CrossRef]
12. Zickus, M.; Greig, A.J.; Niranjana, M. Comparison of four machine learning methods for predicting PM<sub>10</sub> concentrations in Helsinki, Finland. *Water Air Soil Pollut. Focus* **2002**, *2*, 717–729. [CrossRef]
13. Dudot, A.L.; Rynkiewicz, J.; Steiner, F.E.; Rude, J. A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions. *Environ. Model. Softw.* **2007**, *22*, 1261–1269. [CrossRef]
14. Voukantsis, D.; Karatzas, K.; Kukkonen, J.; Räsänen, T.; Karppinen, A.; Kolehmainen, M. Intercomparison of air quality data using principal component analysis, and forecasting of PM<sub>10</sub> and PM<sub>2.5</sub> concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Sci. Total Environ.* **2011**, *409*, 1266–1276. [CrossRef] [PubMed]
15. Russo, A.; Raischel, F.; Lind, P.G. Air quality prediction using optimal neural networks with stochastic variables. *Atmos. Environ.* **2013**, *79*, 822–830. [CrossRef]
16. Tamas, W.; Notton, G.; Paoli, C.; Voyant, C.; Nivet, M.L.; Balu, A. Urban ozone concentration forecasting with artificial Neural Network in Corsica. *Math. Model. Civil. Eng.* **2014**, *10*, 29–37. [CrossRef]

17. Zhou, Q.; Jiang, H.; Wang, J.; Zhou, J. A hybrid model for PM<sub>2.5</sub> forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Sci. Total Environ.* **2014**, *496*, 264–274. [[CrossRef](#)] [[PubMed](#)]
18. Elangasinghe, M.A.; Singhal, N.; Dirks, K.N.; Salmond, J.A.; Samarasinghe, S. Complex time series analysis of PM<sub>10</sub> and PM<sub>2.5</sub> for a coastal site using artificial neural network modelling and kmeans clustering. *Atmos. Environ.* **2014**, *94*, 106–116. [[CrossRef](#)]
19. Kristiani, E.; Lin, H.; Lin, J.R.; Chuang, Y.H.; Huang, C.Y.; Yang, C.T. Short-term prediction of PM<sub>2.5</sub> using LSTM deep learning methods. *Sustainability* **2022**, *14*, 2068. [[CrossRef](#)]
20. Zheng, Y.; Yi, X.; Li, M.; Li, R.; Shan, Z.; Chang, E.; Li, T. Forecasting fine-grained air quality based on big data. In Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 2267–2276.
21. Wang, T.; Jiang, F.; Deng, J.; Shen, Y.; Fu, Q.; Wang, Q.; Fu, Y.; Xu, J.; Zhang, D. Urban air quality and regional haze weather forecast for Yangtze River Delta region. *Atmos. Environ.* **2012**, *58*, 70–83. [[CrossRef](#)]
22. Saide, P.E.; Mena-Carrasco, M.; Tolvett, S.; Hernandez, P.; Carmichael, G.R. Air quality forecasting for winter-time PM<sub>2.5</sub> episodes occurring in multiple cities in central and southern Chile. *J. Geophys. Res. Atmos.* **2016**, *121*, 558–575. [[CrossRef](#)]
23. Hu, K.; Sivaraman, V.; Bhugubanda, H.; Kang, S.; Rahman, A. SVR based dense air pollution estimation model using static and wireless sensor network. In Proceedings of the 2016 IEEE Sensors, Orlando, FL, USA, 30 October–3 November 2016; pp. 1–3. [[CrossRef](#)]
24. Davis, J.M.; Speckman, P. A model for predicting maximum and 8 hr average ozone in Houston. *Atmos. Environ.* **1999**, *33*, 2487–2500. [[CrossRef](#)]
25. Siwek, K.; Osowski, S.; Garanty, K.; Sowinski, M. Ensemble of predictors for forecasting the PM<sub>10</sub> pollution. In Proceedings of the VXX International Symposium on Theoretical Engineering, Lübeck, Germany, 22–24 June 2009; pp. 1–5.
26. Sotomayor-Olmedo, A.; Aceves-Fernandez, M.A.; Gorrostieta-Hurtado, E.; Pedraza-Ortega, C.; Ramos-Arreguin, J.M.; Vargas-Soto, J.E. Forecast urban air pollution in Mexico City by using support vector machines: A kernel performance approach. *Int. J. Intell. Sci.* **2013**, *3*, 126–135. [[CrossRef](#)]
27. Song, Z.; Niu, D.; Xiao, X.; Wu, H. Daily peak load forecasting based on fast k-medoids clustering, GARCH error correction and SVM model. *J. Appl. Sci. Eng.* **2016**, *19*, 249–258.
28. Ishak, A.B.; Daoud, M.B.; Trabelsi, A. Ozone concentration forecasting using statistical learning approaches. *J. Mater. Environ. Sci.* **2017**, *8*, 4532–4543.
29. Lin, K.M.; Chang, Y.S.; Zeng, Y.R.; Huang, C.X. Air pollution forecasting using machine learning methods on big data platform. In Proceedings of the TANET—Taiwan Academic NETwork Conference, Taoyuan, Taiwan, 20–24 October 2018; pp. 740–745.
30. Granger, C.W.J.; Newbold, P. Spurious regressions in econometrics. *J. Econom.* **1974**, *2*, 111–120. [[CrossRef](#)]
31. Engle, R.F.; Granger, C.W.J. Co-integration and error correction: Representation, estimation and testing. *Econometrica* **1987**, *55*, 251–276. [[CrossRef](#)]
32. Christoffersen, P.F.; Diebold, F.X. Further results on forecasting and model selection under asymmetric loss. *J. Appl. Econom.* **1996**, *11*, 561–571. [[CrossRef](#)]
33. Gerlach, R.; Chen, C.W.S.; Lin, D.S.Y.; Huang, M.H. Asymmetric responses of international stock markets to trading. *Phys. A Stat. Mech. Its Appl.* **2006**, *360*, 422–444. [[CrossRef](#)]
34. Holland, J.H. *Adaptation in Natural and Artificial Systems*; The University of Michigan Press: Ann Arbor, MI, USA, 1975.
35. Goldberg, D.E.; Korb, B.; Deb, K. Messy genetic algorithms: Motivation, analysis, and first results. *Complex Syst.* **1989**, *3*, 493–530.
36. Shin, K.S.; Lee, Y.J. A genetic algorithm application in bankruptcy prediction modeling. *Expert Syst. Appl.* **2002**, *23*, 321–328. [[CrossRef](#)]
37. Vapnik, V. *Estimation of Dependences Based on Empirical Data*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006; pp. 430–438.
38. Witten, I.H.; Mark, E.F.; Hall, A. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed.; Morgan Kaufmann: Burlington, MA, USA, 2011; p. 180.
39. Lewis, C.D. *Industrial and Business Forecasting Methods*; Butterworths: London, UK, 1982; p. 40.