

Article



# Study on Air Pollution Behavior of VOCs with Photochemical Monitoring Stations Using EGARCH Model in Southern Taiwan

Edward Ming-Yang Wu<sup>1</sup> and Shu-Lung Kuo<sup>2,\*</sup>

- <sup>1</sup> College of Maritime, National Kaohsiung University of Science and Technology, Kaohsiung City 811, Taiwan; edmywu@isu.edu.tw
- <sup>2</sup> Department of Technology Management, The Open University of Kaohsiung, Kaohsiung City 812, Taiwan
- \* Correspondence: singsuey@ms28.hinet.net

Abstract: This study adopted the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model to examine the 10 ozone precursors of the highest concentrations among the 54 that were assessed over a number of years at the four photochemical assessment monitoring stations (PAMSs) in the Kaohsiung-Pingtung Area in Taiwan. First, the 10 ozone precursors, which were all volatile organic compounds (VOCs), were analyzed using the factor analyses in multiple statistical analyses that had the most significant impact on the area's ozone formation: mobile pollution factor, which included 1,2,4-Trimethylbenzene ( $C_9H_{12}$ ), toluene ( $C_7H_8$ ), and Isopropyl benzene ( $C_9H_{12}$ ). Then, taking into consideration that the number sequences might be affected by standardized residuals, this study applied the vector autoregressive moving average-EGARCH (VARMA-EGARCH) model to analyze the correlation between the three VOCs under different polluting activities. The VARMA-EGARCH model in this research included dummy variables representing changing points of variance structures in the variance formula to predict the conditional variance. This process proved able to effectively estimate the relevant coefficients of the three VOCs' dynamic conditions that changed with time. The model also helped to prevent errors from occurring when estimating the conditional variance. Based on the testing results, this study determined the VARMA(2,1)-EGARCH(1,0) as the most suitable model for exploring the correlation between the three VOCs and meteorological phenomena, as well as the interplay between them in regard to interaction and formation. With the most representative of the three, toluene (TU), as the dependent variable and 1,2,4-Trimethylbenzene (TB) and Isopropyl benzene (IB) as the independent variables, this study found it impossible to calculate the TU concentration with TB and IB concentrations in the same period; estimations of TB and IB concentrations with a period of lag time were required because TU was mainly contributed by automobiles and motorcycles in Kaohsiung. TB and IB resulted from other stationary pollution sources in the region besides cars and motorbikes. When TU was evenly distributed and stayed longer in the atmosphere, the TB and IB concentrations were lower, so distribution conditions and concentrations could not be used to effectively estimate the concentration of toluene. This study had to wait until the next period, or when stationary pollution sources started producing TB and IB of higher concentrations during the daytime, in order to estimate the TU concentrations in a better photochemical situation.

Keywords: EGARCH; photochemical monitoring station; toluene; mobile pollution factor; VOCs

# 1. Introduction

The term "volatile organic compound" (VOC) refers to chemical substances released from solid materials or liquids in the form of a gas. VOCs are volatile because they quickly evaporate under room temperature; they are organic because they contain carbon. VOCs play a significant role in the formation of ozone and particulates, which form smog. When exposed to sunlight and precursors, such as nitrogen oxide, VOCs form ozone and



**Citation:** Wu, E.M.-Y.; Kuo, S.-L. Study on Air Pollution Behavior of VOCs with Photochemical Monitoring Stations Using EGARCH Model in Southern Taiwan. *Atmosphere* **2021**, *12*, 1167. https:// doi.org/10.3390/atmos12091167

Academic Editors: Rosa Fitzgerald and William R. Stockwell

Received: 14 August 2021 Accepted: 8 September 2021 Published: 10 September 2021

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



**Copyright:** © 2021 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/). particulates that have a pronounced impact on outdoor air pollution in cities [1–3]. In addition, the VOC family consists of a wide range of chemical compounds with varying degrees of toxicity: benzene is an example of a highly toxic substance that is known to be a carcinogen; there are others whose impact on human health remains unknown. Similar to other pollutants, the influence of VOCs on people's health is contingent on the time of exposure [4–7]. O<sub>3</sub> is not released directly to the atmosphere; instead, it is formed by a series of photochemical reactions between NOx and VOCs [4], whose influences on the formation of O<sub>3</sub> vary. In light of this, we should gather data on the concentrations of ozone's precursors in areas with high levels of O<sub>3</sub>, and then look into the relations between the precursors' concentrations and the formation of O<sub>3</sub> in order to fully understand ozone pollution.

Motorbikes are one of the most significant mobile pollution sources. Although they are the major means of transportation in Taiwan, making life easier, they also cause air pollution [8,9]. Motorbikes are popular among Taiwanese people for the following reasons. Firstly, their cost is considerably cheaper than that of cars or public transport. Secondly, there is adequate parking space in Taiwan, and it is easy to find a spot to park a motorbike along the roadside. Thirdly, Taiwan is a small island with its population concentrated in the plain area in the west where it is convenient to move around by motorbike. Finally, the favorable tropic weather condition enables the year-round use of motorbikes since it does not snow in plain areas in winter. To make matters worse, long-term use of mobile vehicles, such as old bikes plagued with incomplete combustion, increases the amounts of CO and VOCs released into the atmosphere. To solve the problem of air pollution, the Taiwanese government established adjustable emission standards for bike production for different periods and requires that motorbikes undergo emissions testing regularly. In light of the policy, motorbike repair shops have to shoulder the responsibility of carrying out emissions testing, which increases their employees' chances of being exposed to VOCs [10]. Studies have pointed out that VOCs are mostly detrimental to human health. For instance, benzene impedes the growth of bone cells, and chronic benzene poisoning causes blood disorders, such as aplastic anemia and leukemia [9,11]. Brief exposure to a high concentration of benzene can even lead to death. Toluene, ethylbenzene, and xylene are toxic and irritating substances that cause an imbalanced immune system, damage the central nervous system, and cause dizziness, headaches, hypersomnia, fatigue, and tightness in the chest, and affect the digestive system, leading to loss of appetite or nausea [12]. These VOCs can cause severe cases of poisoning and even damage the liver or hematopoietic system. Benzene, toluene, ethylbenzene, and xylene, collectively known as BTEX, serve as indicators of various pollution sources (automobiles and motorbikes included) since the composition ratios of these four substances in air pollution remain stable [13,14].

Generalized autoregressive conditional heteroskedasticity (GARCH) models are widely adopted to study economic development and pollution related to environmental engineering [15,16]. To illustrate, the autoregressive conditional heteroskedasticity (ARCH) model overcomes error terms in the second hypothesis of the time series data (constant variance) proposed in traditional econometrics or environmental engineering. A GARCH model, on the other hand, is an expanded version of the ARCH model, which was developed by Bollerslev [17]. In econometrics, univariate EGARCH models are applied to predict conditional variances by incorporating dummy variables that represent changes in variance structures into variance formulas. Applying the said models for forecasting means that researchers do not have to hypothesize the distribution of the return on assets and can assess whether the variance of the return brought by bad news on the market is higher than that brought by good news [18]. Fitting returns and a bivariate vector autoregression model can be applied to conduct a Granger causality test to see whether the spill-over effect occurs between the variances and returns between the two markets. The above review shows that the EGARCH model has been extensively applied in the study of the economic market. The high degree of variability among the various air quality variables is similar to that found in the economic market so that the EGARCH model would be applicable to studying the

quality problems. When the information of the second moment is completely revealed and controlled, and further coped with vector model test results, the most appropriate model can be developed. Hence, the mutual dependence among the various pollution parameters and predictions of the various air quality variables can be fully explained.

The EGARCH model is known for its prediction ability and regularity in time series data when analyzing air pollutants. Although the government is tightening up emission regulations as the number of factories and motor vehicles increases, we have not seen any noticeable improvement in the air quality of pollutant-concentrated areas. Therefore, we decided to conduct this study to help the government develop an emissions control policy. This study adopted the VARMA-EGARCH model to analyze three VOCs, all identified as mobile pollution factors, in exploring how seasonal factors and polluting activities affect the relation between the VOCs and air pollution development. In addition, this study also examined the interplay between the three VOCs and their concentrations in different seasons under the same time series. It is hoped that the results of this study can shed light on the correlation between the VOCs and ozone formation, providing reference to the authority when designing air quality simulation models, emissions restrictions, and policies for preventing air pollution. Furthermore, local governments can also refer to this study when assessing the improvement in local air quality.

#### 2. Methodology

# 2.1. Selection of Data and Location of PAMSs

The PAMSs chosen for this study are the Qiaotou Station, Xiaogang Station, Linyuan Station in Kaohsiung City, and Chaozhou Station in Pingtung County, all of which are set up by the Environmental Protection Administration (EPA) of Taiwan. This study assessed 10 VOCs of the highest concentrations in the atmosphere among the 54 (all of them are precursors of ozone) monitored at the stations. The 10 VOCs analyzed were acetylene (AT), propane (PP), n-butane (NB), 1,2,4-trimethylbenzene (TB), toluene (TU), isopentane (IP), 2-methylhexane (MH), ethane (EA), isopropyl benzene (IB), and *n*-undecane (NC). Figure 1 is a map showing the locations of the 4 PAMSs. Monitoring data from 1 January 2020 to 30 June 2021 were chosen; 24 data sets were collected per day, amounting to 12,000 data sets. This study collected 18 months' worth of data because concentrations of various air pollutants vary in different seasons and periods (for example, the concentrations during daytime and nighttime are different). In addition, the EGARCH model allowed the observation of the variables of each VOC and whether there are dependencies among data series. The verification tool adopted for this study was EVIEW 10.0. EVIEW is a widely adopted statistical analysis software that can be used for multiple purposes and is easy to use. It is most often adopted to analyze various applications of regression formulas under varying scenarios, such as descriptive statistical analysis, statistic regression analysis, and forecast analysis. Since it fulfills the requirements for the analysis in primary statistics and econometrics studies, the software is popular with researchers in business, management, social science, and engineering schools around the globe, who use EVIEW for statistical purposes and the analysis of time series.

# 2.2. Selection of Influence Factors by Multivariate Analysis

Before conducting a VARMA-EGARCH model, this study carried out a factor analysis on the 10 VOCs that contributed the most to photochemical pollution in the Kaohsiung– Pingtung area and divided them into 3 categories: mobile sources, stationary sources, and hydrocarbon sources, whose eigenvalues were 2.231, 1.482, and 1.160, respectively. The 3 types of polluting sources are the major contributing factors to the photochemical pollution in the region. Then, this study selected the most significant factor among the 3 (or the one with the highest eigenvalue), namely mobile pollution sources, to conduct a simulation with the VARMA-EGARCH model. Because the analysis process is highly complex and not the focus of the research, it will not be discussed in this paper.



**Figure 1.** Photochemical monitoring stations in Kaohsiung–Pingtung area, Taiwan. Source of image: google (http://maps. www.google.com accessed on 30 August 2021).

# 2.3. Development of the VARMA Model

In vector time series analyses, in addition to the regression among variables, the feedback and time difference among variables are considered and included in the analyses as weighted parameters. Hence, this feedback is effective for developing the dynamic relationship among variables and improving the accuracy of time series model prediction [19,20].

The VARMA model can be expressed as

in which

$$\phi(B) = \underbrace{I}_{-} - \phi B - \dots - \phi B^{P}_{-P}$$
(2)

$$\underline{\theta}_{-}(B) = \underbrace{I}_{--1} \underbrace{\theta}_{--1} B - \dots - \underbrace{\theta}_{-q} B^{q}$$
(3)

Equations (2) and (3) are polynomials of *B*;  $\Phi(B)$  and  $\theta(B)$  are K × K matrices; *C* is a K × 1 vectors with fixed values;  $\alpha$  is a series of independent random vibration vectors with normal distributions.

The vector ARMA model in (1) can be re-written as

$$\phi(B)\frac{\dot{Z}}{-t} = \frac{\theta}{-}(B)\frac{a}{-t} \tag{4}$$

in which

$$\dot{Z}_{-t} = Z_t - \mu \tag{5}$$

In addition, an inspection of autocorrelation function (ACF) and partial autocorrelation function (PACF) helps in identifying the orders AR (p) and MA(q). More objectively defined criterions such as Akaike information criterion (AIC), Hannan–Quinn information criterion (HIC), Bayesian information criterion (BIC) and final prediction error (FPE) can also be used to identify the correct orders p and q [19,21,22].

# 2.4. VARMA-EGARCH Modeling

Conditional mean formula

$$r_{t} = a_{0} + \sum_{i=1}^{m} a_{i} r_{t-i} + \sum_{j=1}^{n} b_{j} \mathcal{E}_{t-j}, \ \mathcal{E} \Big| \Omega_{t-1} \sim N(0, h_{t})$$
(6)

Conditional variance formula

$$\ln(\mathbf{h}_{t}) = \alpha_{0} + \sum_{i=1}^{q} \left( \gamma_{i} \frac{\boldsymbol{\varepsilon}_{t-i}}{\sqrt{h_{t-i}}} + \alpha_{i} \left( \frac{\left|\boldsymbol{\varepsilon}_{t-i}\right|}{\sqrt{h_{t-i}}} - E\left(\frac{\left|\boldsymbol{\varepsilon}_{t-i}\right|}{\sqrt{h_{t-i}}}\right) \right) + \sum_{j=1}^{p} \beta_{j} \ln(h_{t-j}) \quad (7)$$

 $\frac{\mathcal{E}_{t-i}}{\sqrt{h_{t-i}}} > 0$  represents good news on the air pollution condition while  $\frac{\mathcal{E}_{t-i}}{\sqrt{h_{t-i}}} < 0$  reflects bad news on the air pollution condition, where  $\gamma_i = 0$  indicates that the degree of air pollution has a symmetrical effect in reaction to news impact.

If  $\gamma_i < 0$ , this suggests that the degree and fluctuation of air pollution incurred by air pollution shock are more significant than the situation when air pollution shock does not result in air pollution; in other words, a leverage effect can be observed when  $\gamma_i < 0$ . To determine whether the 3 GARCH models are applicable to the time series analysis, we first needed to confirm whether the time series had an ARCH effect; this was tested in this study using the LM test proposed by Engle [23,24]. Since it is necessary for the residual of the conditional mean formula in the GARCH model to follow the white noise process, all possible orders (p,q) of the GARCH model must be compared with each other via trial and error. To determine whether or not residuals reached white noise, we later adopted the modified Q-statistic proposed by Ljung and Box [25] to analyze the residual of each series. However, the *p* and q orders of the model selected via trial and error might be subject to over-fitting. To solve this issue, we used the Akaike information criteria (AIC) [26] and Schwartz Bayesian criterion (SBC) [21,27] to analyze the model under the principle of parsimonious parameterization.

# 2.5. Essence of EGARCH Model Development

The model used to simulate mobile pollution factor was calibrated using statistical principle and method in order to expound the significance of fat tail test, Ljung–Box sequence examination, and EGARCH examination.

# 2.5.1. Fat Tail Test

A fat tail test is a probability distribution that exhibits a large skewness, kurtosis, and Jarque–Bera, relative to that of either a normal distribution or an exponential distribution. In common usage, the terms fat-tailed and heavy-tailed are sometimes synonymous; fat-tailed is sometimes also defined as a subset of heavy-tailed. Different research communities favor one or the other, largely for historical reasons, and may have differences in the precise definition of each of the terms. The investigation of time series empirical distribution often leads to characteristics of fat tail test. Hence, the assumption of normal distribution is not the optimal choice. The results of examining the skewness, kurtosis, and Jarque–Bera normal distribution can be used for determining whether the distribution of modeling errors has fat tails.

#### 2.5.2. Ljung–Box Series Examination

The residual needs to be examined to determine whether it shows the sequential correlation before evaluating the EGARCH models. The square of a residual that shows sequential correlations will have the EGARCH effect [28]. Ljung–Box q test is applied to examine the randomness and independence of any time series. If data in a time series are not independently distributed, they exhibit serial correlation after a certain period (k), a phenomenon known as autocorrelation. Autocorrelation might compromise the accuracy of a forecast model in a time series, such as a time series graph, and lead to erroneous interpretations.

# 2.5.3. Examination of the EGARCH Effectiveness

Before conducting simulations using the time series combination of ARCH and GARCH models, the process of model calibration must be carried out in order to confirm that the residual series is not related to the first order series, or white noise, to assure an appropriate model. Next, the residual square examination is used to determine whether the model has EGARCH effect. In this paper, the Q-statistic proposed by Ljung–Box was used to examine whether the residual had high order autocorrelation. Only when the model has EGARCH effectiveness, can it be used to carry out iterative non-linear calculations for estimating model parameters [29].

# 2.6. Outline of This Research

This paragraph provides an overview of this paper. This study gathered data on 54 ozone precursors, all VOCs, from 4 photochemical assessment monitoring stations in the Kaohsiung–Pingtung area. Ten VOCs of the highest concentrations over a number of years were then chosen and studied with multivariate analysis. Moreover, this study analyzed the mobile pollution sources that had the most significant impact on ozone formation in the Kaohsiung–Pingtung area, i.e., TU, TB, and IB. The study adopted a VARMA-EGARCH model to examine whether the 3 VOCs' residual terms are correlated and whether they have the EGARCH effect. Furthermore, this study also applied the optimal model, VARMA(2,1)-EGARCH(1,0), to explore interplays between the 3 VOCs and how their concentrations in the atmosphere change under different polluting activities.

#### 3. Results and Discussion

#### 3.1. Selection and Standardization of VOCs' Data

Before using the VARMA-EGARCH model for analysis, 54 precursors of ozone were chosen from four PAMSs located in the Kaohsiung–Pingtung area. Among the 54 precursors, 10 VOCs of the highest concentrations in the atmosphere were selected and a multivariate analysis was conducted to find the most significant factor, i.e., the mobile pollution factor. The mobile pollution factor consisted of three VOCs: TU, TB, and IB. The sequence needed to be standardized due to the seasonality and periodicity of the chosen three VOCs; the standardized sequence was as follows

$$Z_{\nu,t} = \frac{Y_{\nu,t} - \mu_t}{\sigma_t}, t = 1, ..., w$$
 (8)

In the above equation,  $Y_{v,t}$  represented a primeval sequence, and  $\mu_t$  and  $\sigma_t$  represented the periodic average and standard deviation of 1 to w, respectively.

# 3.2. Discussion on the Effects of the VARMA-EGARCH Model on Mobile Pollution Factor in the Simulation

3.2.1. Fundamental Property Analysis of the Three VOCs

The mobile pollution factor consisted of three VOCs: TU, TB, and IB. Table 1 shows the fundamental property analysis of the three VOCs' sequences from 1 January 2010 to 30 June 2021, including kurtosis, skewness, mean, standard deviation, and Jarque–Bera normal distribution. Regarding the mean of the VOCs, TU had the highest value, which means

that its concentration was higher in the atmosphere. The high concentration of TU resulted from more and longer exposure to various complicated pollutants, such as the organic pollutants from automobile and motorbike emissions and heavy industry emissions, given that the Kaohsiung–Pingtung area is a heavy industry district [8]. In Table 1, all three VOCs were skewed to the right, which means that the skewness was positive. Among the three VOCs, IB had the lowest average concentration in the atmosphere because it skewed the most to the right. A sudden spike, or an outlier, may appear because when skewness increased, information asymmetry became more influential. Hence, the volatility of IB concentration was less predictable because it was less likely to be influenced by seasonal changes than TU and TB, a result which was consistent with that of Wu and Kuo's study [30]. TU concentration in the atmosphere was higher and stable with even distribution as its skewness was remarkably less than the other two VOCs. This result explained why TU had the largest factor loading in mobile pollution factor: when TU concentration remained high due to the long-term high automobile and motorbike emissions, TU in the atmosphere was easily accumulated and hard to disperse [9]. The kurtosis statistics of the three VOCs were higher than when they are normally distributed (which would be two), which clearly shows that the concentration distribution of the three VOCs changes according to seasonal and periodic factors.

VOC			
	TU	ТВ	IB
Parameters			
Mean	2.48	1.63	1.06
Median	2.31	1.44	0.98
Maximum	8.16	4.03	3.92
Minimum	0.05	0.05	0.01
Std. Dev.	1.40	0.92	1.01
Skewness	1.23	2.46	3.55
Kurtosis	5.99	8.16	7.07
Jarque–Bera	1517.59	1810.24	988.54
Probability	0.000	0.000	0.000
Observations	12,000	12,000	12,000

Table 1. Fundamental properties of the 3 VOCs in mobile pollution factor.

From the analysis of the Jarque–Bera normal distribution, this study was able to determine whether seasonal changes impacted the concentration distribution of the three VOCs' sequences. In Table 1, all three VOCs were higher than the critical value (degree of freedom is  $2, X_{0.05,2}^2 = 5.99$ ) at a 5% significance level. The results also proved that seasonal changes impact the concentrations of the three VOCs because they all rejected the hypothesis of normal distribution; the distribution had fat tails.

# 3.2.2. Results of Ljung-Box Serial Examination

Before adopting the EGARCH model, the residuals of the three VOCs in the regression model must first be examined for serial correlation because serial correlation can cause a false EGARCH effect in the residual sum of squares. Hence, it is necessary to run a serial correlation examination before using the EGARCH model. In this study, the Ljung–Box serial examination was used, and the results are shown in Table 2. Table 2 shows that the examination results of L–BQ(K) were all smaller than the critical value, while the null hypothesis could not be rejected, indicating that the residuals of the three VOCs were not serially correlated. This finding conformed to the white noise phenomenon; hence, the model configuration chosen for this study was quite suitable.

L–BQ(K)	TU	ТВ	IB	<b>Critical Value</b> $x^2_{(0.05,k)}$
1	0.43	0.86	0.38	3.84
2	1.09	1.69	1.09	5.99
3	2.38	2.29	1.87	7.82
4	4.87	4.67	3.35	9.49
5	5.52	5.16	4.43	11.07
6	7.69	7.03	5.60	12.59
7	9.46	8.45	7.59	14.07
8	12.01	11.64	9.56	15.51
9	13.97	12.50	10.74	16.92
10	15.88	14.62	12.76	19.68
11	17.94	16.58	14.92	21.96
12	20.60	19.41	17.40	24.05
16	23.95	24.47	22.79	30.30
20	29.66	30.79	28.74	36.41
24	35.51	36.17	34.33	43.42

Note:  $R_t = c + \theta R_{t-1} + \varepsilon_t$ ;  $\alpha = 0.05$ .

# 3.2.3. Examination of EGARCH Effectiveness

The Lagrange multiplier (LM) statistics were used to examine whether the model had the EGARCH effect with a sequence [29]. The LM statistics were shown as  $T \times R^2$ , which obeyed a chi-square distribution with degree of freedom being P. T represented the sample size of the sequence (which is 12,000 in this study), and  $R^2$  represented the coefficients of determination obtained using OLS regression. When the LM statistic of a model was significant, the sequence had the EGARCH effect. When a series has a strong EGARCH effect, the explanatory power of the regression is very large, as is the coefficient of determination  $R^2$ . Table 3 showed the three VOCs' sequences in the mobile pollution factor had strong EGARCH effects. In other words, all  $T \times R^2$  values were significant at a 5% level of significance. Therefore, it was suitable to use EGARCH effectiveness to explain the characteristic of the mobile pollution factor.

Q (Lagged Variables)	TU	ТВ	IB	Critical Value $x^2_{(0.05,k)}$
1	9.15	20.59	17.20	3.84
2	16.47	30.65	22.54	5.99
3	25.66	38.91	32.68	7.82
4	29.71	53.06	35.89	9.49
5	33.73	68.56	44.57	11.07
6	38.49	80.34	56.40	12.59
7	46.72	83.67	70.63	14.07
8	51.64	90.48	77.85	15.51
9	55.67	103.26	86.71	16.92
10	61.08	109.60	98.45	19.68
11	68.58	116.82	104.68	21.96
12	79.40	123.60	111.12	24.05
16	91.55	137.02	128.19	30.30
20	98.78	150.16	143.06	36.41
24	109.56	172.44	163.47	43.42

**Table 3.** Results of EGARCH(q) examination of 3 VOCs.

Note: All T  $\times$  R<sup>2</sup> values are less than 5% indicating "obviousness".

### 3.3. Selection of the Best Combination for EGARCH

Different VARMA and EGARCH models were tested to produce the best combination and configure an appropriate time series statistical model, namely the VARMA(p,q)-EGARCH(p,q) model. Finally, the most suitable model was selected to simulate and describe the correlation among the three VOCs in the mobile pollution factor. Table 4 shows the analyses of different VARMA(p,q)-EGARCH(p,q) combinations. By testing dozens of different vector model configurations, a set of minimum values of AIC and SC were chosen to simulate the correlation among the three VOCs in the atmospheric environment and changes in the concentrations of the three VOCs in the atmosphere under various kinds of polluting activities in the researched area.

EGA	ARCH EGAR	EGARCH(0,1)		CH(1,1)	EGAR	CH(1,0)	EGARCH(2,0)		
VARMA Types	AIC	SC	AIC	SC	AIC	SC	AIC	SC	
VARMA(0,1)	) 8.748	8.806	8.687	8.816	8.425	8.449	8.524	8.604	
VARMA(0,2)	8.523	8.572	8.569	8.650	8.268	8.302	8.332	8.387	
VARMA(1,0)	8.466	8.505	8.382	8.505	8.343	8.387	8.359	8.442	
VARMA(1,1)	) 8.514	8.599	8.561	8.858	8.301	8.414	8.380	8.416	
VARMA(2,0)	8.400	8.487	8.610	8.724	8.216	8.405	8.345	8.413	
VARMA(2,1)	8.621	8.669	8.659	8.772	8.179	8.306	8.317	8.445	
VARMA(3,0)	8.457	8.512	8.498	8.650	8.283	8.341	8.402	8.426	
VARMA(3,1)	8.502	8.537	8.496	8.573	8.316	8.378	8.468	8.580	

Table 4. Combinations of VARMA(p,q)-EGARCH(p,q) for mobile pollution factor.

# 3.4. Simulation Results of Mobile Pollution Factor Using the VARMA(2,1)-EGARCH(1,0) Model

As mentioned in Chapter 2 of this study, a multivariate analysis was performed before the VARMA-EGARCH model simulation was carried out, which showed that mobile pollution factor was the major determinant of photochemical pollutants affecting the atmosphere in the Kaohsiung-Pingtung area. In terms of factor loading, TU had the highest value, TB came second, and IB was at the bottom. Therefore, this study set TU, the factor with the highest concentration in the atmosphere, as the dependent variable, and TB and IB as independent variables. Along with VARMA(p,q)-EGARCH(p,q), the variation trend and dependency of these VOCs in the time series analysis were examined. Table 5 demonstrates the simulation results of the optimum model: VARMA(2,1)-EGARCH(1,0), which was used to examine how TU, TB, and IB interact in a dynamic and complex atmospheric environment. According to Table 5, the current TU concentration could hardly be estimated based on the TB concentration because the t-statistic of  $b_0$  was -1.03 (not statistically significant since it was below 1.96). However, the TB concentrations after one and two time lags (the t-statistics of  $b_1$  and  $b_2$  were 2.31 and 2.80, respectively) were above 1.96; the statistical significance allowed the estimation of the TU concentration in the atmosphere. Similarly, the current TU concentration could hardly be estimated based on the current IB concentration because the t-statistic of  $b_0$  was 0.64 (not statistically significant since it was below 1.96). However, the IB concentrations after one and two time lags (the t-statistics of  $c_1$  and  $c_2$  were 2.32 and 1.99, respectively) were above 1.96; the statistical significance allowed the estimation of the TU concentration in the atmosphere. According to the above simulation results, the current TU concentration in the atmosphere in the Kaohsiung-Pingtung area should be estimated based on the TB and IB concentrations after one and two time lags instead of the current TB and IB concentrations. The following adequately explains the simulation results.

T, K. T, K. T															
Parameters	30	<b>a</b> 1	an	h.	h.	b.	60	61	62	d.	Ω.	<b>Q</b> 1	(Y2)	ßı	<b>V</b> 1
EGARCH Types	0	-1	•2	0	<i>v</i> 1	02	-0	<b>v</b> 1	-2	u	•••0	•••1	<u>-</u>	PΙ	71
VARMA(0,1)-	-1.54	2.34	2.56	3.47	2.03	-0.41	2.06	1.67	0.06	0.96	1.59	2.36	-0.95	3.04	-0.367
EGARCH(0,1)	0.92	2.42	1.04	-0.22	1.56	2.39	1.98	-2.48	0.42	2.51	0.67	1.19	2.01	-1.14	2.52
VARMA(0,1)-	0.99	3.30	1.91	3.06	-0.87	0.08	2.97	3.41	1.85	3.34	-2.46	1.03	2.61	0.93	1.87
EGARCH(1,1)	-2.16	1.81	2.04	-1.04	2.65	0.56	0.99	-2.23	1.32	-0.78	1.65	1.63	0.54	-2.16	2.63
VARMA(1,0)-	3.55	-1.23	2.20	1.98	3.06	0.95	3.34	1.15	-0.90	1.55	2.06	-1.09	0.87	-0.53	-1.53
EGARCH(1,0)	1.83	0.57	-0.54	2.29	-0.64	1.92	2.63	0.99	2.03	1.98	3.44	1.08	2.36	0.76	3.01
VARMA(1,1)-	1.99	-2.54	3.09	1.01	2.77	-0.69	0.88	-1.60	2.06	1.08	-0.96	2.40	3.11	1.57	-1.36
EGARCH(2,0)	2.77	3.06	1.78	1.46	-3.26	2.33	-1.48	1.98	-2.08	1.35	0.59	2.69	0.97	-1.43	3.04
VARMA(2,0)-	-0.96	1.77	-0.86	3.15	1.84	0.98	2.33	0.56	3.06	-1.86	-1.11	3.03	1.66	3.18	-0.96
EGARCH(0,1)	1.50	2.15	2.03	0.88	-2.06	3.01	2.04	3.03	1.57	2.50	3.13	2.06	1.54	2.15	2.08
VARMA(2,1)-	1.81	2.06	0.87	-0.49	3.03	2.88	-2.01	2.34	4.03	1.69	0.89	-0.87	0.96	1.19	-1.01
EGARCH(1,1)	2.04	1.59	3.06	1.68	1.86	2.03	1.52	1.63	2.19	2.03	1.57	-1.16	2.05	2.39	2.00
VARMA(2,1)-	2.01	3.55	1.59	-0.98	1.72	2.39	0.21	3.01	2.27	0.86	4.30	1.29	0.70	-1.07	-0.589
EGARCH(10)	1 56	4 14	2 67	-1.03	2 31	2 80	0.64	2.32	1 99	1 55	2.66	-0.98	1 58	5 55	4 43

**Table 5.** Parameters of VARMA(p,q)-EGARCH (p,q) for mobile pollution factor.

Note: (1)  $TU = a_0 + a_1TU_{(t-1)} + a_2TU_{(t-2)} + b_0TB_{10(t)} + b_1TB_{10(t-1)} + b_2TB_{10(t-2)} + c_0IB_{(t)} + c_1IB_{(t-1)} + c_2IB_{(t-2)} + d_1\varepsilon_{t-1}$   $h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \beta_1h_{t-1}$  (2)  $\gamma_{i,t} = c_{i0} + \sum_{j=1}^{p_i} b_{ij}\gamma_{i,t-j} + u_{i,t}$  (3) If  $\gamma_i = 0$ , this suggests that air quality has a symmetrical effect on information impact. If  $\gamma_i < 0$ , this suggests that air quality has a leverage effect on information impact. (4) If the t-statistic is  $\geq 1.96$ , this means that the parameter reaches statistical significance.

First, Kaohsiung and Pingtung, both located in southern Taiwan, lie within Pingtung Plain. With alternating warm and cold spells, as well as weak east and southeast winds, sea winds took the pollutants inland in the daytime, while the on-shore winds took them back to Pingtung Plain in the evening. Meanwhile, the pollutants continued to accumulate and were barely dispersed, thus polluting the atmosphere for a long time. The Kaohsiung-Pingtung area is located in the wake zone of the northeast monsoon, and the northeast monsoon took the pollutants south, leading to severe pollution cases. In the meantime, the east wind started to blow in the area, which was on the leeward side, and there was lower average wind speed and rainfall. The above conditions made it difficult for pollutants to disperse and, thus, perpetuated the pollution of the atmosphere. In other words, poor air quality in the Kaohsiung–Pingtung area occurs more often in early spring (from January to April) and winter (from October to December) because pollutants accumulate quickly in the region in these two time periods. During winter, northeast seasonal winds bring pollutants from central and northern Taiwan across the Central Mountain Range to the Kaohsiung–Pingtung area. After crossing the mountain range, seasonal winds sink to the bottom of the mountains, bringing down with them pollutants that stay in the area instead of naturally dispersing due to temperature inversion. Thus, the pollutants from the north, coupled with emissions produced by factories and cars in Kaohsiung, exacerbate air pollution in winter. On the other hand, air quality during the summer is usually better, thanks to southwest seasonal winds that facilitate diffusion and disperse pollutants more quickly. Refineries, power plants, and petrochemical plants in the Kaohsiung-Pingtung area regularly released VOCs into the atmosphere in the daytime; these VOCs included TU, TB, and IB. The main sources of TU concentration in the Kaohsiung-Pingtung area were local mobile pollution sources, such as large diesel cars, which frequently accessed the Port of Kaohsiung, the 15th largest port worldwide. Compared to stationary pollution sources, mobile pollution sources were more influential in terms of time and space. Moreover, mobile pollution sources in the Kaohsiung–Pingtung area caused greater impacts than stationary pollution sources regarding air pollution and photochemical reaction [8]. When concentrations of VOCs coming from mobile pollution sources were high, those of TB and IB in the atmosphere were low, which was unfavorable for VOCs to disperse. If automobiles and motorbikes happened to emit a high concentration of TU under such circumstances, TU concentration in the atmosphere could hardly be estimated promptly. Moreover, mobile and stationary pollution sources continued to release high concentrations of TU into the atmosphere after VOCs, such as TB and IB, were dispersed in the atmosphere for a certain period of time. As a result, the concentrations of TB and IB in the atmosphere reached a point where the estimation of TU concentration in all seasons was possible regardless of the conditions for dispersion.

Moreover, this study adopted VARMA(2,1)-EGARCH(1,0), the optimum model, to explore the atmospheric concentration characteristics of the Kaohsiung–Pingtung area using  $\gamma_{i,t}$  of the EGARCH model. According to Table 5, the t-statistic of  $\gamma_1$  was 4.43, reaching statistical significance with a coefficient of -0.589 (<0). The statistics demonstrated that pollutants, including inorganic pollutants and TU, TB, and IB, accumulated easily in Kaohsiung in winter, the season characterized by unfavorable conditions for dispersion, which contributed to local air pollution incidents and violations of the air quality standard. For example, VOCs were active in photochemical reactions and oxidized easily under sunlight, which contains ultraviolet rays [5,6]. Hydroxyl radicals produced in the oxidation process reacted with other chemical compositions in the atmosphere, such as NO, and formed high concentrations of ozone and peroxides, such as PAN (peroxyacetyl nitrates). The researched area was influenced by various external factors, including a large population, a large number of automobiles and motorbikes, and a high density of stationary pollution sources. More importantly, there was less rain in spring, autumn, and winter, while the mobile pollution sources had a greater impact for a long time. Therefore, there tended to be a high concentration of TU in the atmosphere. In the summer, the southwest monsoon and subtropical high affected the Kaohsiung–Pingtung area, whereby atmospheric pollutants

were often sent northwards. Furthermore, atmospheric pollutants spread and fell easily due to rainfall in the plum rain season and typhoons [2,12]. Compared to other seasons, there were fewer days that fell short of the air quality standard. All of the above phenomena corresponded to the leverage effect in the EGARCH model (when  $\gamma_1$  was below 0) [31].

Regarding how the three VOCs were produced and how their amounts fluctuated in the atmosphere, as mentioned above, TU in the atmosphere of the Kaohsiung–Pingtung area mainly came from local mobile pollution sources and TB and IB from local stationary pollution sources, with only small amounts from mobile pollution sources. In terms of their concentration in the atmosphere, TU was the highest, TB came second, and IB was at the bottom. The conditions for dispersion in the Kaohsiung–Pingtung area were unfavorable in spring, autumn, and winter. In addition, the concentrations of TB and IB were relatively low when they were first emitted; they were mostly emitted in the daytime when stationary pollution sources came into play. According to Table 5, the TU concentration in the atmosphere could barely be estimated promptly under such circumstances (proven by the t-statistics of  $b_0$  and  $c_0$ , both falling below 1.96). As stationary and mobile pollution sources continued to produce TB and IB, the TU concentration in the atmosphere could be estimated after one and two time lags (proven by t-statistics of  $b_1$ ,  $b_2$ ,  $c_1$ ,  $c_2$ , all above 1.96). Not meeting the total suspended particulates (TSP) standard was the top reason why the Kaohsiung–Pingtung area fell short of the air quality standard, while not meeting the  $O_3$  concentration standard was the second reason. Photochemical reactions of  $O_3$  in the atmosphere can be very complicated [1,3,7]. Many people travel within the Kaohsiung metropolitan area by automobiles and motorbikes on weekends. The amounts of VOCs released by stationary pollution sources, such as TB and IB, did not increase as the aforementioned pollution sources released a high concentration of TU into the atmosphere; hence, the concentrations of TB and IB in the atmosphere did not immediately reflect the concentration of TU. As mobile pollution sources continued to emit TB and IB after one time lag, while stationary pollution sources released TB and IB into the atmosphere, the concentration of TU in the atmosphere could be accurately estimated despite unfavorable conditions for dispersion (the t-statistics of  $b_1$ ,  $b_2$ ,  $c_1$ ,  $c_2$  were above 1.96).

#### 4. Conclusions

The three VOCs discussed in this study, i.e., TU, TB, and IB, had EGARCH effects, suggesting that they change with the seasons. In addition, the distributions of the three VOCs were positively skewed, with that of IB being skewed remarkably to the right, indicating that its average concentration in the atmosphere was smaller than the other two VOCs. Nevertheless, sudden spikes in the concentration, or outliers of IB, occur more often than TU and TB. Furthermore, TU had the least skewed distribution among the three VOCs because it had the highest factor loading. Moreover, automobiles and motorbikes in the Kaohsiung–Pingtung area continued to produce emissions, leading to more TU in the atmosphere; its concentration was especially high in spring, autumn, and winter, during which it accumulated quickly but was dispersed slowly due to weather changes. A simulation was carried out using the optimum VARMA(2,1)-EGARCH(1,0) model to analyze the interplay among, and changes in, TU, TB, and IB in a complicated and volatile environment. The results indicated that the current TU concentration could not be estimated based on the current TB and IB concentrations; it could be estimated based on the TB and IB concentrations after one and two time lags. In other words, the concentrations of TB and IB were relatively low when they were first emitted; they were mostly emitted in the daytime when stationary pollution sources came into play. Therefore, only after one and two time lags, when the concentrations of TB and IB increased as stationary and mobile pollution sources kept producing them, could the TB and IB concentration data be used to predict the TU concentration. In addition, this study used  $\gamma_{i,t}$  of the EGARCH model to analyze the atmospheric concentration characteristics of the Kaohsiung–Pingtung area. The result revealed that the t-statistic of  $\gamma_1$  was 4.43, reaching statistical significance with a coefficient of -0.589 (<0). The statistics demonstrated that pollutants, including inorganic

pollutants and TU, TB, and IB, accumulated easily in Kaohsiung in spring, autumn, and winter, during which conditions for dispersion were unfavorable, which contributed to local air pollution incidents and violations of the air quality standard. The findings of this study can serve as reference for handling air pollution caused by various sources and should be especially helpful for the development of ozone reduction mechanisms and the assessment of air quality improvement schemes. In addition, the results of this study present the correlation among different air pollutants data that governments can refer to when creating tools for the air pollution toolbox, such as an air quality simulation model, emissions regulations, and air pollution prevention schemes.

**Author Contributions:** E.M.-Y.W. is the first author of this paper. He was responsible for defining the topic and organizing the research. In addition, he also collected and compiled data from the four photochemical assessment monitoring stations in the Kaohsiung–Pingtung area. Meanwhile, the corresponding author (S.-L.K.) was in charge of the execution of EVIEWS 10.0 software and studied the interdependence between the three VOCs in the time series and how their concentrations change in different seasons. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

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

# References

- 1. Pandis, S.N.; Harley, R.A.; Cass, G.R.; Seinfeld, J.H. Secondary organic aerosol formation and transport. *Atmos. Environ. Part A Gen. Top.* **1992**, *26*, 2269–2282. [CrossRef]
- 2. Karle, M.M.; Mahmud, S.; Sakai, R.; Fitzgerald, R.M. Investigation of the successive ozone episodes in the El Paso–Juarez region in the summer of 2017. *Atmosphere* **2010**, *11*, 532. [CrossRef]
- 3. Steward, D.; Saunders, E.; Perea, R.; Fitsgerald, R.M.; Campbell, D.E.; Stockwell, W.R. Linking air quality and human health effects models: An application to the Los Angeles air basin. *Environ. Health Insights* **2017**, *11*, 1–13. [CrossRef]
- 4. Algrim, L.B.; Ziemann, P.J. Effect of the hydroxyl group on yields and composition of organic aerosol formed from OH radicalinitiated reactions of alcohols in the presence of NOx. *ACS Earth Space Chem.* **2019**, *3*, 413–423. [CrossRef]
- 5. Kaipara, R.; Rajakumar, B. Oxidative degradation kinetics and mechanism of two biodiesel constituents (ethyl-2-methyl propionate (E2MP) and ethyl-2,2-dimethyl propionate (E22DMP)) initiated by Cl atoms. *ACS Earth Space Chem.* **2020**, *4*, 142–156. [CrossRef]
- Shi, X.; Qiu, X.; Cheng, Z.; Chen, Q.; Rudich, Y.; Zhu, T. Isomeric identification of particle-phase organic nitrates through gas chromatography and time-of-flight mass spectrometry coupled with an electron capture negative ionization source. *Environ. Sci. Technol.* 2020, 54, 707–713. [CrossRef] [PubMed]
- Fujita, E.M.; Campbell, D.E.; Stockwell, W.R.; Lawson, D.R. Past and future ozone trends in California's south coast air basin: Reconciliation of ambient measurements with past and projected emission inventories. *J. Air Waste Manag. Assoc.* 2013, 63, 54–69. [CrossRef] [PubMed]
- 8. Chou, T.G. The research of air quality management in fine particulate matters. In *Central Office of Administration, Academia Sinica Research Weekly* 1276; Central Office of Administration, Academia Sinica: Taipei City, Taiwan, 2010. (In Chinese)
- 9. Gouriou, F.; Morin, J.P.; Weill, M.E. On-road measurements of particle number concentrations and size distributions in urban and tunnel environments. *Atmos. Environ.* 2004, *38*, 2831–2840. [CrossRef]
- 10. Chen, W.K.; Li, T.D. *The Establishment of Aerosol Control Mode Based on the Photochemical Reaction Mechanism*; NSC-1999-EPA-Z-231-001; National Science Council, Executive Yuan: Taipei City, Taiwan, 1999. (In Chinese)
- 11. Fu, L.; Hao, J.; He, D.; He, K. Assessment of vehicular pollution in China. J. Air Waste Manag. Assoc. 2001, 51, 658-668. [CrossRef]
- 12. Kuo, H.W.; Wei, H.C.; Liu, C.S. Exposure to 16 volatile organic compounds while commuting in Taichung Taiwan. *Atmos. Environ.* **2000**, *34*, 3331–3336. [CrossRef]
- 13. Geiss, O.; Giannopoulos, G.; Tirendi, S.; Barrero-Moreno, J.; Larsen, B.R.; Kotzias, D. The airmex study-VOC measurements in public buildings and schools/kindergartens in eleven European cities: Statistical analysis of the data. *Atmos. Environ.* **2011**, *45*, 3676–3684. [CrossRef]
- 14. Cheng, J.H.; Hsieh, M.J.; Chen, K.S. Characteristics and source apportionment of ambient volatile organic compounds in a science park in central Taiwan. *Aerosol Air Qual. Resour.* **2016**, *16*, 221–229. [CrossRef]
- 15. Wu, E.M.Y.; Kuo, S.L. Air quality time series based GARCH model analyses of air quality information for a total quantity control district. *Aerosol Air Qual. Res.* **2012**, *12*, 327–339. [CrossRef]

- 16. Wu, E.M.Y.; Kuo, S.L.; Liu, W.C. Generalized autoregressive conditional heteroskedastic model for water quality analyses and time series investigation in reservoir watersheds. *Environ. Eng. Sci.* **2012**, *29*, 227–237. [CrossRef]
- 17. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. J. Econom. 1986, 31, 307–327. [CrossRef]
- 18. Nelson, D.B. Conditional heteroscedasticity in asset returns: A new approach. Econometrica 1991, 59, 347–370. [CrossRef]
- 19. Brockwell, J.B.; Davis, R.A. Introduction to Time Series and Forecasting; Springer-Verlag Inc: New York, NY, USA, 2002.
- 20. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications- with R Examples*; Springer Science+Business Media, LLC: Berlin/Heidelberg, Germany, 2006.
- 21. Ling, S.; McAleer, M. Asymptotic theory for a vector ARMA-GARCH model. Econom. Theory 2003, 19, 278–308. [CrossRef]
- 22. Kumar, U.; Prakash, A.; Jain, V.K. A multivariate time series approach to study the interdependence among O<sub>3</sub>, NOx and VOCs in ambient urban atmosphere. *Environ. Modeling Assess.* **2009**, *14*, 631–643. [CrossRef]
- 23. Elyasiani, E.; Mansur, I. International spillover of risk and return among major banking institutions: A bivariate GARCH model. *J. Account. Audit. Financ.* 2003, *18*, 303–330. [CrossRef]
- 24. Ali, G. Edarch, Gjr-Garch, Tgarch, Avgarch, Ngarch, Igarch and Aparch Models for Pathogens at Marine Recreational Sites. *J. Stat. Econom. Methods* **2013**, *2*, 57–73.
- 25. Glosten, L.R.; Jagannathan, R.; Runkle, D. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Financ.* **1993**, *48*, 1779–1801. [CrossRef]
- 26. Akaike, H. A new look at the statistical model identification. IEEE Trans. Autom. Control 1974, 19, 716–723. [CrossRef]
- 27. Schwartz, G.E. Estimating the dimension of a model. Ann. Stat. 1978, 6, 461–464. [CrossRef]
- 28. Jiang, W. Using the GARCH Model to Analyze and Predict the Different Stock Market. Master's Thesis, Statistics Department of Statistics Uppsala University, Uppsala, Sweden, 2012.
- Lee, C.L.; Lee, R.K.; Kao, Y.M. Lagrange multiplier optimization synthesis of long-period fiber gratings. In *National Tsing Hua* University Institutional Repository, Institute of Electrical and Electronics Engineers, EE Conference Papers; Institute of Electrical and Electronics Engineers: Hsinchu City, Taiwan, 2007; p. 826.
- 30. Wu, E.M.Y.; Kuo, S.L. VARMA-EGARCH model for air-quality analyses and application in aouthern Taiwan. *Atmosphere* **2020**, *11*, 1096. [CrossRef]
- 31. Forbes, K.; Rigobon, R. No contagion, only interdependence: Measuring stock market co-movements. J. Financ. 2002, 5, 2223–2261. [CrossRef]