International Journal of Engineering & Technology, 7 (3.19) (2018) 119-124
International Journal of Engineering & Technology
Website: www.sciencepubco.com/index.php/IJET
Research paper

The Assessment of Time Series for an Entire Air Quality Control District in Southern Taiwan Using GARCH Model

Shu-Lung Kuo[1], Edward Ming-Yang Wu[2]*, Shu-Hsien Chang[3]

1.Kelee Environmental Consultant Corporation, Kaohsiung, Taiwan
Tel: +886-7-8150815; Fax: +886-7-8150816 e-mail:singsuey@ms28.hinet.net
2.* Department of Technology Management, Open University of Kaohsiung, No.436, Taiwan
*Corresponding author E-mail:Edward@isu.edu.tw
3.Pingtung Agriculture Biotechnology Park, Pingtung County, Taiwan

Abstract

The General Autoregressive Conditional Heteroskedastic (GARCH) model and 10 ordinary air quality monitoring stations in the entire air quality control district in Kaohsiung-Pingtung were used in this study. First, the factor analysis results within multivariate statistics were employed to select the main factor that affects air pollution, namely, the photochemical pollution factor. The characteristics of the GARCH model were discussed in terms of asymmetric volatility among the three air pollutants (PM_{10}, NO_2, and O_3) within the factor. In addition, this study also combined the multiple time series model VARMA to explore changes in the time series of the three air pollutants and to discuss their predictability.

The results showed that, although the coefficient of the GARCH model was negative when estimating the variance equation, the conditional variance would always be positive after taking the logarithm. The results also suggested that the GARCH model was quite capable of capturing the asymmetric volatility. In other words, if the condition that pollution factors might be subject to seasonal changes or outliers generated by the human contamination is not considered, the GARCH model had very good ability to verify the results and make predictions, regardless of whether it adopted any of the three risk concepts: normal distribution, t-distribution, and generalized error distribution. For example, under the trend of time series temporal and spatial distribution in various pollution concentrations of photochemical factors, the optimal model VARMA(2,0,0)-GARCH(1,1) selected in this study was used to conduct time series predictability after the verification procedure. After capturing the last 50 entries of data on O_3 concentrations in the sequence, the results showed that the predictability correlation (r) was 0.812, the predictability of NO_2 was 0.783 and the predictability of PM_{10} was 0.759. It can be learned from the results that under the sequence of the GARCH model with strong asymmetric volatility, the residual values of these three sequences as white noise were quite evident, and there was also a high degree of correlation in predictability.

Keywords: the entire air quality control district in Kaohsiung-Pingtung, GARCH model, asymmetric volatility, photochemical pollution factor

1. Introduction

In recent years, Taiwan has witnessed an ever-increasing number of factories and cars/scooters. Although the emission standards have been tightened, the air quality in districts where pollution sources are concentrated remains unlikely to enjoy significant improvement. Therefore, it is necessary to promote total quantity control strategies to further improve the air quality. The Environmental Protection Administration (EPA) in Taiwan has divided Taiwan into seven air quality districts: Northern Taiwan, Hsinchu-Miaoli, Central Taiwan, Yunlin-Chiayi-Tainan, Kaohsiung-Pingtung, Hualien-Taitung and Yilan. Also, the EPA announced the total control districts in stages based on demands. Priority is given to the entire air quality control district in Kaohsiung-Pingtung. For those districts that do not meet the air quality standards, their total amounts of emissions were reduced to comply with the control of allowable increased limits on pollutants conducted in districts that meet the air quality standards.

At present, three of the most popular models to capture the time-varying volatility in financial time series are the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model of Engle (1982) and Bollerslev (1986), the GJR model of Glosten et al., (1992), and the Exponential GARCH (EGARCH) model of Nelson (1991). Multivariate extensions of GARCH models are also available in the literature, such as the Constant Conditional Correlation (CCC) GARCH model Bollerslev (1990) and Ali (2013), Vector Autoregressive Moving Average GARCH (VARMA-GARCH) model of Ling and McAleer (2003), and VARMA Asymmetric GARCH (VARMA-AGARCH) model of Hoti et al., (2002). de Veiga and McAleer (2004) presented that the multivariate VARMA-GARCH model of Ling and McAleer (2003) and VARMA-AGARCH model of McAleer et al. (2008) provided better volatility than the nested univariate model, namely GARCH of Bollerslev (1986) and GJR of Glosten, et al., (1992), respectively. On the other hand, in order to capture the dynamics of time-varying conditional correlation, recently development model is generalized autoregressive conditional correlation (GARCC) of McAleer et al. (2008).

In addition, by using the GARCH model, Maher (1997) finds that the time-varying interest rate sensitivity renders tests over long periods inconclusive. To address the time-varying nature of the stock return generating process for banks, Song (1994) employs an ARCH-type methodology. Elyasiani and Mansur (2003) go further by employing an extended GARCH-M model, which includes an interest rate in the mean and interest rate volatility as an argument in the volatility of the bank stock return generating process. Carlson et al. (2004) and Zhang (2005) employ simulations to show that asset betas change over time in response to historical investments undertaken by companies as well as by changes in product market demand. They conclude that the book-to-market effect is related to operating leverage while the size effect relates to the relative riskiness of growth options to assets in place.

Based on this reason, this study selected 10 automatic air quality monitoring stations set up in the entire air quality control district in Kaohsiung-Pingtung, Taiwan. Firstly, the factor analysis within the multivariate statistical analysis was employed to select the most evident air pollutant factor that affected the air quality in the entire air quality control district in Kaohsiung-Pingtung, namely the photochemical pollution factor. Given that there is variability between air pollutants and that air pollutants have the ability to fully control second moment messages, this study adopted the VARMA-GARCH model to analyze the time-varying trend of individual air pollutants (parameters) among the relevant factors. Also, the predictability of air pollution parameters over time was verified. The results could immediately reflect the correlation between various air pollution parameters, and the government could formulate the application and examination of air quality models accordingly, along with the model simulation for allowable increased limits, in order to provide a reference and basis for the benefit of air quality improvement.

2. Experimental Method

2.1 The Selection of Air Quality Monitoring Stations

The entire air quality control district in Kaohsiung-Pingtung selected in this study mainly took 7 pollutants: sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), particulate matter with particle size below 10 microns (PM_{10}), ozone (O_3), total hydrocarbon compounds (THC) and methane (CH_4) in the 10 ordinary air quality monitoring stations (including Meinong Station, Nanzi Station, Cianjin Station, Renwu Station, Zuoying Station, Xiaogang station, Daliao station and Linyuan station in Kaohsiung City, and Pingtung Station and Chaozhou Station in Pingtung County) as the basis for an analysis. Also, a total of 303 entries of data on complete air pollutant monitoring between April 1, 2017 and April 30, 2018 were taken as the basis for a model analysis. The geographical location of each air quality monitoring station is shown in Figure 1.

2.2 General ARCH Model (GARCH)

According to the traditional ARMA model determination method, Bollerslev (1986) generalized the moving average (MA) by adding conditional variation of lag length to the ARCH model. As such, the conditional variation of current lag length was not only affected by the previous average squared residual items, but also the previous conditional variance, and thus became the GARCH model (Jiang, 2012).

2.3 ARMA-GARCH Model

The threshold GARCH model proposed by Glosten et al. (1992) included the threshold value in the GARCH model, making 0 the boundary point of the previous residual sequence a_{t-1}. In addition to retaining its advantages (such as explaining the clustering of fluctuations and describing the fat tail), GARCH further solved the asymmetry of time sequences that ARCH could not explain. The model is as follows:

$$\sigma_t^2 = \alpha_0 + (\alpha_1 + \gamma d_t)\varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (1)$$

2.4 Ljung-Box Sequence Test

It was necessary to test whether the residual items in the regression model have sequence correlation before estimating the ARCH and GARCH models. If the residual items have sequence correlation, the squared residual items will be examined to see if it has an ARCH effect. As such, it is very important to check if the residual items have sequence correlation before estimating the ARCH and GARCH models.

2.5 ARCH Effect Test

Before combining the ARCH and GARCH models for time sequence, it is necessary to go through the steps of model testing to confirm that the sequence residual items do not have a first order correlation, i.e. white noise; at this stage, the model is an appropriate model. Secondly, the test of squared residual items will serve as a means of determining if the model has a(n) (G)ARCH effect. This study used the Q statistics proposed by Ljung-Box to test whether the residuals have high order autocorrelation. After the model is shown to have an ARCH effect, it can then perform parameter estimations of the repeated nonlinear operation.

3. Empirical Analysis and Discussion

3.1 An Analysis of the Basic Characteristics of Photochemical Pollution Factors

The photochemical pollution factors in this study included three air pollutants, NO_2, O_3 and PM_{10}. Table 1 shows the basic characteristics of these three air pollutants, including mean, standard deviation, skewness, kurtosis, and the Jarque-Bera test (Chen and Li, 1999). In terms of skewness, said air pollutants skewed on the right (with positive skewness). NO_2 had the highest skewness of 3.68, indicating a sudden increase in numerous entries in this sequence. PM_{10} only had a skewness of 0.86, and O_3 1.03. The main sources of air pollution in the entire air quality control district in Kaohsiung-Pingtung, as suggested in this study, were PM_{10} and O_3. Therefore, the higher the concentrations of PM_{10} and O_3, the higher the degree of air pollution! The concentrations of PM_{10} and O_3 in this district were often high, which often led to poor air quality, especially in the winter time. Consequently, they did not exhibit evident skewness. As for O_2, it was known from the original data that it only had a high concentration during some periods in winter and early spring, but it did not contribute much to the deterioration of air quality in the atmosphere. Therefore, it had higher skewness. In terms of kurtosis, it was larger than the normal distribution coefficient for said air pollutants (normal distribution was 3), showing that each sequence had the characteristics of seasonal time sequence. In addition, through the Jarque-Bera test, said air pollutants were greater than the critical value (the degree of freedom was 2, and $\chi^2_{0.05,2} = 5.99$) at the 5% significant level, which showed the hypothesis of refusal of normal distribution. Said air pollutants had the property of thick-tailed distribution. In other words, the concentration of each sequence was indeed affected by seasonality, and therefore had a different concentration value.

3.2 Ljung-Box Sequence Test for Photochemical Pollution Factors

This study used the Ljung-Box test to conduct the sequence test of photochemical pollutant factors, and the result is shown in Table 2. Table 2 shows that the statistical values of L-B-Q(K) were all less than the critical value, and they could not reject the null hypothesis. This characteristic indicated that the residual items of each sequence did not have sequence correlation, which complied with the white noise phenomenon. The model configuration was quite appropriate.

3.3 ARCH Effect Test for Photochemical Pollution Factors

If a sequence is to test whether the model has an ARCH effect, the LM (Lagrangian Multiplier) statistics can be used for verification. The LM statistics is $T*R^2$, where T is the number of samples and R^2 is the coefficient of determination using OLS regression. $T*R^2$ is the chi-square distribution with a degree of freedom P. When the LM statistics of the model is significant, it means that the sequence has an ARCH effect. Table 3 shows the statistical results concerning the three air pollutants within the photochemical pollution factors. The results showed that the conditional variations of said air pollutants all had a strong ARCH effect (that is, all $T\times R^2$ values were significant at a 5% significant level). Therefore, it is appropriate to use the ARCH effect to explain photochemical pollution factors.

3.4 Photochemical Pollution Factor Model Simulation Results

To select the best matching model, this study used vector model EACF and GARCH to choose different combination tests. A VARMA(p,d,q)-GARCH(p,q) model was combined, and used to test the most suitable one, in order to conduct a simulation analysis. Through the results obtained from tens of tests, Table 4 shows the optimal parameter estimation for VARMA(2,0,0)-GARCH(1,1).

From the simulation results in Table 4, it can be observed that when the current PM_{10} concentration was generated, the concentration of O_3 produced in the current period could not be directly estimated from the concentration of PM_{10} (t-statistic of b0 is 1.39, <1.96, not significant). However, the concentration of PM_{10} in one time lag and two times lag could affect the generation of O_3 concentration in the current period (the t-statistic of b_1 and b_2 were 3.08 and 2.54, respectively, and >1.96, which is significant). In terms of NO_2, the current NO_2 concentration could not estimate the current O_3 concentration from its current concentration (the t-statistic of c_0 was 0.41, <1.96, not significant). However, the NO_2 concentration of one time lag could affect the current generation of O_3 concentration (the t-statistic of c_1 was 3.50, >1.96, significant), while the two times lag became insignificant again (t-statistic of c_2 was 1.13, <1.96, not significant). It could be learned from the above analysis that the current O_3 concentration would be affected by the PM_{10} one time lag and two times lag, as well as the NO_2 one time lag. It could be explained that when the concentrations of PM_{10} and NO_2 in the atmosphere were generated, these pollutants did not immediately produce a photochemical reaction when spreading into the air (Chen and Li, 1999). Instead, they would generate a photochemical reaction with sunlight during the one time lag or even the two times lag, resulting in a photochemical product O_3. Moreover, since PM_{10} could remain longer in the atmosphere (Chou, 2010), O_3 could still be generated due to photochemical reactions after the two times lag. Consequently, it could be learned that there was at least one time lag or even two times lag during the photochemical reactions and the generation of pollutants. In other words, the PM_{10} concentration in the one time lag and two times lag could affect the generation of current O_3, and the NO_2 concentration in the one time lag could affect the generation of current O_3. In terms of O_3 concentration, it was also significantly affected by its own one time lag (the t-statistic of a_1 was 3.75, >1.96, significant), but not significantly affected by the two times lag (the t-statistic of a_2 was -0.37, <1.96, not significant). This result could be explained in that the current O_3 concentration would be affected by the one time lag concentration itself, but less affected by the two times lag itself.

3.5 VARMA(2,0,0)-GARCH(1,1) Model Time Sequence Predictability Results

Based on the three air pollution parameters of the photochemical pollution factors, this study used the optimal model, VARMA(2,0,0)-GARCH(1,1), to predict the last 50 entries of data on three air pollutants, and the results are shown in Figures 2-4. The results showed that the predictability correlation (r) of O_3 is 0.812, that of NO_2 is 0.783, and that of PM_{10} is 0.759. The correlation coefficient of these three entries of predictability is at least greater than 0.75. As such, since the GARCH model is quite evident in terms of the asymmetric volatility phenomenon and the interpretation of the fluctuation clustering, it is quite capable of making predictions.

4. Conclusion

the conditional heteroskedasticity model is an extremely effective tool for a time sequence analysis. When discussing air pollution, no matter if it is photochemical pollution factor or fuel factor, an ARCH and GARCH effect exists because each air pollution variable tends to change with the seasons. When the former period had great (small) changes, the current one would change accordingly.

The GARCH model adopted for this study is quite capable of capturing the asymmetric volatility. In other words, if the condition that pollution factors might be subject to seasonal changes or outliers generated by the human contamination is not considered, the GARCH model had very good ability to verify the results and make predictions, regardless of whether it adopted any of the three risk concepts: normal distribution, t-distribution, and generalized error distribution; such can be fully seen from the analysis results in Chapter 3. With the use of VARMA(2,0,0)-GARCH(1,1) to simulate the three air pollutant parameters among the photochemical pollution factors in the entire air quality control district in Kaohsiung-Pingtung, the results showed that the current O_3 concentration would be affected by the one time lag and two times lag of PM_{10}, as well as the one time lag of NO_2, which meant that the concentrations of PM_{10} and NO_2 from the previous period would affect the concentration of current O_3. This result could be explained in that when pollutants produced PM_{10} and NO_2, they would not immediately produce photochemical oxidation when spreading into the atmosphere, but would produce O_3 in the next period due to their interaction with sunlight. The PM_{10} could stay even longer in the atmosphere and produce ozone due to photochemical effects after two times lag. Therefore, the photochemical reaction and the time of pollutant generation have at least one time lag.

References

[1] Ali, G. (2013). EDARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH Models for Pathogens at Marine Recreational Sites. Journal of Statistical and Econometric Methods, 2, 57-73

[2] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307-327.

[3] Bollerslev, T. (1990). Modelling the Coherence in Shortrun Nominal Exchange Rate: A Multivariate Generalized ARCH Approach, Review of Economics and Statistics, 72, 498-505.

[4] Carlson, M., A. Fisher, and R. Giammarino. 2004. Corporate Investment and Asset Price Dynamics: Implications for the Cross-Section of Returns. *Journal of Finance* 59: 2577-2603.

[5] Chen, W.K. and Li, T.D. (1999). The Establishment of Aerosol Control Mode Based on the Photochemical Reaction Mechanism. National Science Council, Executive Yuan, Taiwan. NSC-1999-EPA-Z-231-001. (in Chinese)

[6] Chou, T.G. (2010). The Research of Air Quality Management in Fine Particulate Matters. Central Office of Administration, Academia Sinica Research weekly 1276, Central Office of Administration, Academia Sinica, Taiwan. (in Chinese)

[7] de Veiga, B. and McAleer, M. (2004), Testing multivariate volatility and spillover effects in financial markets, in C. Pahl-Wostl (ed.), Proceedings of the International

[8] Environmental Modelling and Software Society Conference, Osnabruck, Germany, June 2004.

[9] Elyasiani, E., and Mansur, I., 2003, International Spillover of Risk and Return among Major Banking Institutions: A Bivariate GARCH Model, *Journal of Accounting, Auditing, and Finance*, 18(2): 303-330.

[10] Engle, R. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K.Inflation. *Econometrica* 50, 987-1007.

[11] Glosten, L., R. Jagannathan and D. Runkle (1992), On the Relation Between the Expected Value and Volatility of Nominal Excess Returns on Stocks, Journal of Finance, 46, 1779-1801.

[12] Hoti, S., F. Chan and M. McAleer (2002), Structure and asymptotic theory for multivariate asymmetric volatility: Empirical evidence for country risk ratings, paper presented to the Australasian Meeting of the Econometric Society, Brisbane, Australia, July 2002.

[13] Jiang, W. (2012), Using the GARCH Model to Analyze and Predict the Different Stock Market. Master Thesis in Statistics Department of Statistics Uppsala University, Sweden.

[14] Ling, S. and M. McAleer (2003), Asymptotic Theory for a Vector ARMA-GARCH Model, *Econometric Theory*, 19, 278-308.

[15] McAleer, M., Chan, F., Hoti, S. and Lieberman, O. (2008), Generalized autoregressive conditional correlation. *Econometric Theory*, 24, 1554-1583.

[16] Maher, M., 1997, Bank Holding Company Risk From 1976-1989 With a Two-Factor Model, *Financial Review*, 32: 357-371.

[17] Nelson, D.B. (1991), Conditional heteroscedasticity in asset returns: A new approach, *Econometrica*, 59, 347-370.

[18] Song, E, 1994, A Two Factor ARCH Model for Deposit-Institution Stock Returns, *Journal of Money, Credit, and Banking*, 26: 323-340.

[19] Zhang, L. 2005. The Value Premium. *Journal of Finance* 60: 67-103.

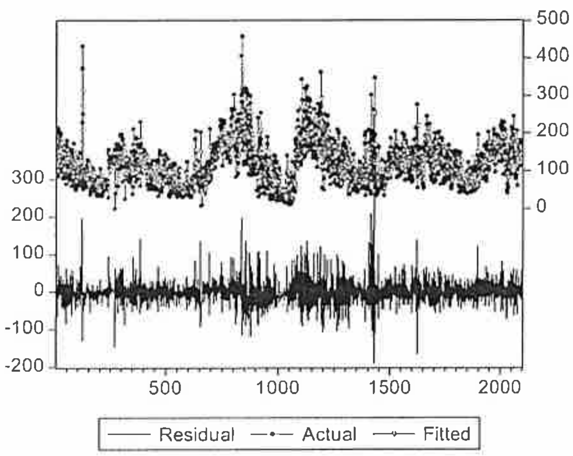

Figure 2 PM_{10} VARMA(2,0,0)-GARCH(1,1) Simulation Results

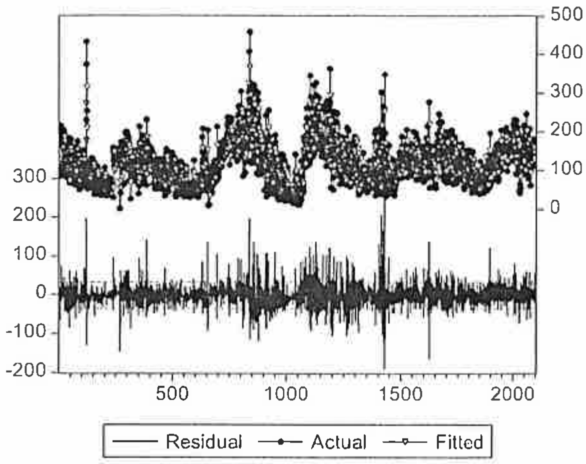

Figure 3 O_3 VARMA(2,0,0)-GARCH(1,1) Simulation Results

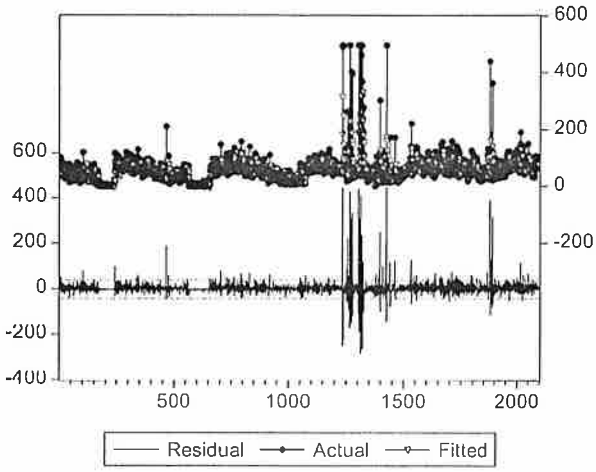

Figure 4 NO_2 VARMA(2,0,0)-GARCH(1,1) Simulation Results

Table 1 Basic Characteristics of Photochemical Pollution Factors

	PM_{10}	O_3	NO_2
Mean	110.928735	43.984532	76.038430
Median	107.400000	39.600000	73.040000
Maximum	509.100000	227.500000	511.037495
Minimum	22.800000	9.600000	20.180000
Std. Dev.	0.799028	0.903849	0.304678
Skewness	0.867267	1.032892	3.684204
Kurtosis	4.281771	10.530963	15.920375
Jarque-Bera	727.287394	3037.204	1015.686
Probability	0.000000	0.000000	0.000000
Sum	1283.400	267.0000	117.0990
Sum Sq. Dev.	165.289345	175.936212	72.735449
Observations	303	303	303

Table 2 Ljung-Box Sequence Test for Photochemical Pollution Factors

L-BQ(K)	PM_{10}	O_3	NO_2	Critical value $\chi^2_{(0.05,k)}$
1	1.27	2.04	2.64	3.84
2	3.09	4.03	4.49	5.99
3	6.01	5.83	6.06	7.82
4	7.00	7.08	8.28	9.49
5	10.28	9.18	10.23	11.07
6	11.44	10.06	11.04	12.59
7	13.25	12.72	12.58	14.07
8	14.23	13.55	13.83	15.51
9	15.62	14.71	14.42	16.92
10	16.77	15.58	16.14	18.31
16	21.98	22.74	23.65	26.30
20	26.65	27.90	28.57	31.41
24	32.04	33.76	34.53	36.42

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

Table 3 Photochemical Pollution Factor ARCH(q) Effect Verification

Q (lagged variables)	PM_{10} (TR^2)	O_3 (TR^2)	NO_2 (TR^2)	Critical value $\chi^2_{(0.05,k)}$
1	401.28	5.75	87.74	3.84
2	438.03	7.11	104.89	5.99
3	472.09	9.12	107.93	7.82
4	524.30	16.66	132.17	9.49
5	550.21	20.08	143.32	11.07
6	572.24	21.77	150.07	12.59
7	583.46	23.85	155.98	14.07
8	592.80	27.06	167.31	15.51
9	604.07	31.45	173.46	16.92
10	615.23	36.77	186.38	19.68

Note: All TR^2 values are less than 5% indicating "significance"

Table 4 Parameter Estimation of Photochemical Pollution Factors Vector Model with GARCH(1,1) Process

	a_0	a_1	a_2	b_0	b_1	b_2	c_0	c_1	c_2	d_1	α_0	α_1	α_2	β_1
VARMA(1,0,0)	1.61	0.89		0.07	0.13		0.20	0.57			3.09	0.37	0.35	0.57
t-statistic	4.73	1.94		3.94	0.20		-1.28	0.89			5.06	3.06	-0.87	1.48
VARMA(1,1,1)	1.91	0.45	-0.09	0.25	1.56	3.06	0.4	1.09	-0.62		4.12	0.04	0.33	0.56
t-statistic	0.57	-2.08	0.17	5.09	-0.20	5.09	1.99	2.94	2.05		6.94	2.38	1.09	-3.11
VARMA(0,0,1)	1.55			0.30			0.34			0.56	4.07	2.50	-0.54	0.58
t-statistic	0.39			0.49			-0.03			3.06	5.94	6.06	-3.94	2.96
VARMA(2,0,1)	3.55	1.22		0.28	0.33		0.04	2.03		-0.08	4.03	1.05	-1.94	1.09
t-statistic	10.46	-3.46		5.03	1.06		1.56	1.55		3.57	1.04	1.07	0.90	-0.88
VARMA(2,0,0)	**2.33**	**2.85**	**-1.12**	**0.19**	**1.64**	0.30	-1.32	0.47	**0.48**		**5.04**	**1.83**	**0.50**	**1.36**
t-statistic	**5.56**	**3.75**	**-0.37**	**1.39**	**3.08**	2.54	0.41	3.50	**1.13**		**9.98**	**4.05**	**-3.94**	**3.06**

$O_3 = a_0 + a_1 O_{3(t-1)} + a_2 O_{3(t-2)} + b_0 PM_{10(t)} + b_1 PM_{10(t-1)} + b_2 PM_{10(t-2)} + c_0 NO_{2(t)} + c_1 NO_{2(t-1)} + c_2 NO_{2(t-2)} + d_1 \varepsilon_{t-1}$ $\quad h_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} + \alpha_2 \varepsilon^2_{t-2} + \beta_1 h_{t-1}$

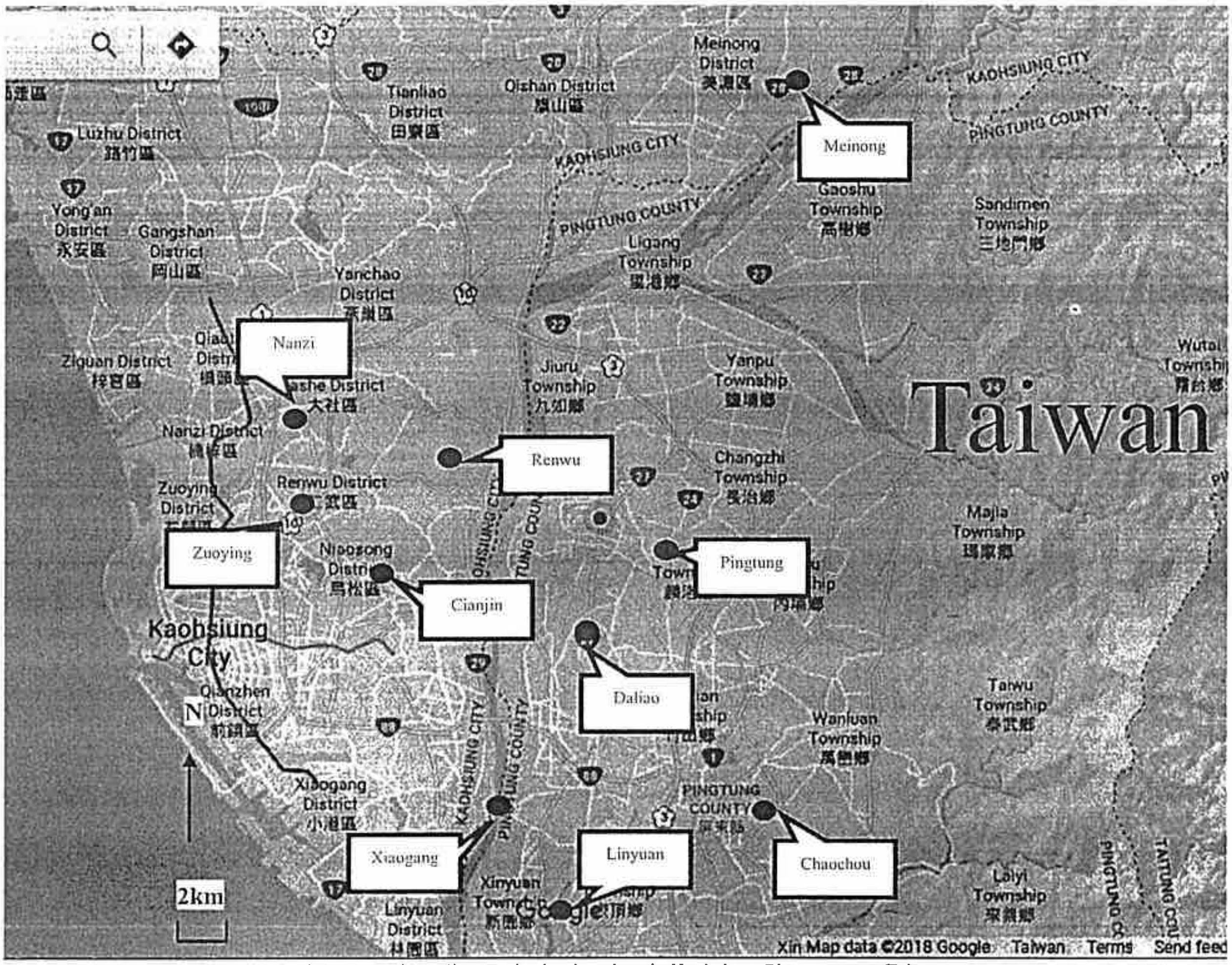

Figure 1 Air quality monitoring locations in Kaohsiung-Pingtung area, Taiwan