Hui, Wai Lun
MPhil Student
Department of Management Sciences
City University of Hong Kong

Time Series Analysis of Air Pollution Level in Hong Kong

Background

The levels of nitrogen dioxide (NO_2) and ozone (O_3) in Hong Kong have attracted much attention from the public because these 2 pollutants can affect the health of the local residents and they may affect the tourist industry and economy of Hong Kong.

Figure 1: The long-term trend of NO_2 in Hong Kong

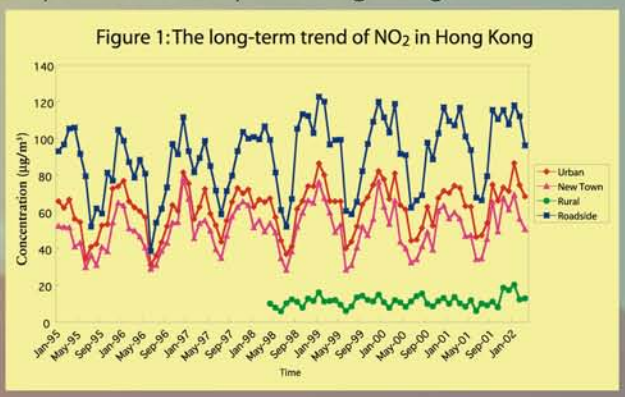
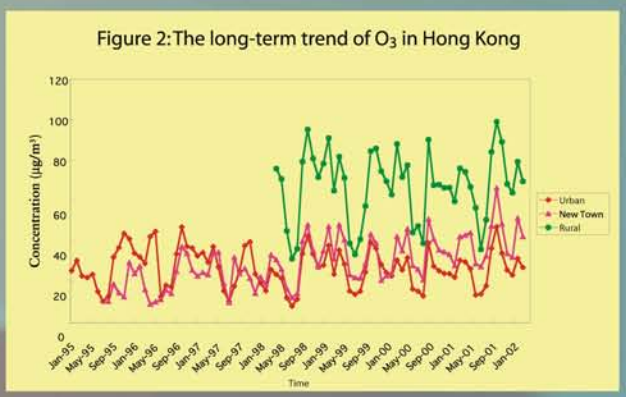


Figure 2: The long-term trend of O_3 in Hong Kong



The objectives of this study are to study the relationship and forecast the levels of the 2 major air pollutants in Hong Kong. The methods used include the following:

Unit Root Test

The augmented Dickey-Fuller test is carried out to test whether the time series of NO_2 and O_3 are stationary.

$$\Delta Y_t = \alpha + rY_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \varepsilon_t \qquad \Delta Y_t = \alpha + \beta_1 t + rY_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \varepsilon_t$$

If $r = 0$, the time series is nonstationary and may grow without bound.

Result:

The test shows that all the time series of the 2 pollutants are stationary and no differencing is required.

Structural Vector Autoregressive (VAR) Model

The chemical reactions of the pollutants in the atmosphere are highly dynamic: $NO_{(g)} + O_{3(g)} \leftrightarrow NO_{2(g)} + O_{2(g)}$

As the levels of pollutants are affected by each other, this kind of dynamic situation can be modeled adequately by the structural VAR model. The endogenous variables of the model are O_3, nitric oxide (NO) and NO_2.

Example:

The models estimated for the pollutants in Shatin are: $\Phi_0 Y_t = \delta + \sum_{i=1}^{8} \Phi_i Y_{t-i} + \varepsilon_t$

Model Performance (compare with univariate ARMA forecast):

NO_2:

Measure	Univariate	Structural VAR
MAE	0.24	0.16
RMSE	0.31	0.20
MAPE	6.59%	4.26%
MPE	0.73%	0.31%

O_3:

Measure	Univariate	Structural VAR
MAE	0.40	0.29
RMSE	0.52	0.39
MAPE	13.65%	10.00%
MPE	3.37%	2.06%

Improved forecast accuracies when compare with univariate forecast!!!

Figure 3: Actual and Predicted NO_2 Level

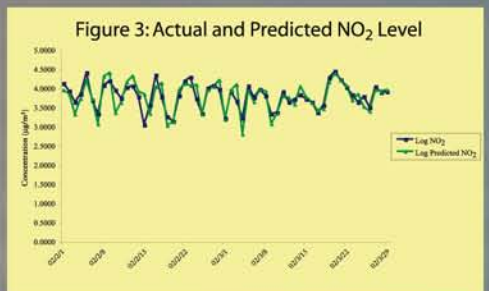
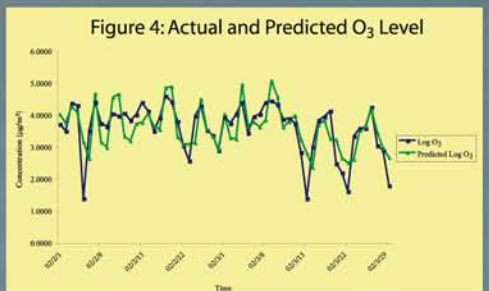


Figure 4: Actual and Predicted O_3 Level



Conclusion

The use of multivariate time series forecast can improve the forecast accuracy of air pollution data.