Time Series Analysis of Air Pollution Level in Hong Kong

Background
The levels of nitrogen dioxide (NO₂) and ozone (O₃) 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.

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 NO₂ and O₃ are stationary.

\[ \Delta Y_t = \alpha + \beta Y_{t-1} + \sum_{j=1}^{r} \beta_j \Delta Y_{t-j} + \epsilon_t \]

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: \( NO_2 + O_3 \rightarrow NO_2 + O_2 \). 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 O₃, nitric oxide (NO) and NO₂.

Example:
The models estimated for the pollutants in Shatin are:

\[ \Phi \cdot Y_t = \delta + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \epsilon_t \]

Model Performance (compare with univariate ARMA forecast):

<table>
<thead>
<tr>
<th>Measure</th>
<th>NO₂ Univariate</th>
<th>Structural VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.24</td>
<td>0.16</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.31</td>
<td>0.20</td>
</tr>
<tr>
<td>MAPE</td>
<td>6.59%</td>
<td>4.26%</td>
</tr>
<tr>
<td>MPE</td>
<td>0.73%</td>
<td>0.31%</td>
</tr>
</tbody>
</table>

Improved forecast accuracies when compare with univariate forecast!!!

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