Job Candidate: Chi-Sheng Lo

1. Objective

The objective of this analysis is to conduct forecasting for the handset and component markets with quarterly data.

2. Computational and statistical toolset

Language: Python 3.9

Packages: matplotlib, numpy, pandas, statsmodels, unitroot

IDE: Spyder 5.2.2

3. Data and methodology

3.1 Data and Data Generating Process

There are mainly four sets of time series data under two separate sections. The first section is the shipments summary of Nokia which has three sub-sections: total, smartphone, and feature phone. The second section is the AP shipments summary of MediaTek. The dataset provided is represented with both quarterly and annual data for only two years.

Since I only need to forecast the series for the next one quarter, I omit the annual data and only estimate the quarterly data in my forecasting model. Furthermore, I make sure the analyzed data is detrended if it contains unit-root and incorporate the seasonal variation into my time series forecasting model.

3.2 Econometric model

Because the demand for handset can be affected by the seasonal pattern, I use the Seasonal Autoregressive Integrated Moving Average (SARIMA) to estimate the future expected handset shipments (Total, Smartphone, Feature phone) of Nokia in 1Q13 as well as the expected component shipments of MediaTek in 1Q16.

This type of ARIMA model can be shown as:

$$\mathrm{A(B)A_s(B^s)(1-B)^d(1-B^s)^D Z_t = b(B)b_s(B^s)a_t}$$

Where $\mathrm{a_t}$ is a sequence of independently distributed random variables with mean zero and constant variance or the so-called white noise, B is the backward shift operator, $\mathrm{B^s}$ is the seasonable backward shift operator with s equal to 4, $\mathrm{(1-B)^d}$ is the dth difference, and finally, $\mathrm{(1-B^s)^D}$ is the Dth difference for the values s period apart.

Moreover, $\mathrm{A_s(B^s)}$ is the seasonal autoregressive polynomial in $\mathrm{B^s}$ of order P, $\mathrm{b_s(B^s)}$ is a seasonal moving average polynomial in B^s of order Q, A(B) is a non-seasonal autoregressive polynomial of order p, b(B) is a non-seasonal moving average polynomial of order q.

The model is a multiplicative seasonal multiplicative seasonal mode of order (p, d, q) x (P, D, Q, M).

Where p is the number of autoregressive term, d is the necessary number of differences for stationary process, q is the number of lagged forecast errors in the prediction equation, D is the necessary number of differences for stationary time series with seasonal component, P is the autoregressive order for the seasonal component, Q is the moving average order for the seasonal component, M is the number of periods in a season.

3.3 Stationary check

To avoid spurious regression which may cause misleading relationship, we must make sure that the variables are represented in stationary data. I use the Dickey-Fuller test (ADF) to conduct testing for the unit root. If the data is nonstationary, I will first try taking the first difference; usually, the first differenced data can be stationary with quarterly data. I write a separate Python function which is named as unitroot.py.

3.4 Optimal lag selection

The optimal lag selection is based on looking at the autocorrelation graph. The lag is chosen with the maximum lag that can eliminate the serial autocorrelation to avoid inaccurate standard error of forecast and find the best fitted model. Table 1 illustrates the optimal lag selection for each model below:

Table 1.

	p	**d**	**q**	**P**	**D**	**Q**	**M**
Nokia smartphone	1	0	0	0	0	0	4
Nokia feature phone	1	0	0	0	0	0	4
MediaTek AP	0	0	0	0	2	1	4

4. Concluding remark

4.1 Result

The result is generated from the econometric model coded in Python. You can find the innerworkings in GFK.py. The snapshot of the forecasted result is shown in table 2 below:

Table 2.

Nokia	Total	85.2m
Nokia	Smartphone	6.4m
Nokia	Feature phone	78.8m
MediaTek	AP	108m

4.2 Limitation and potential improvement

The data provided is definitely not enough for doing ARIMA with seasonal component. There are only eight quarters (two years) of data; therefore, the model may not fully capture the seasonal effect from the shipment of handset and its component due to lack of sample size.