The Bandung Statistical Institute

Forecasting Methods: All What Policy Maker Needs

Maman Abdurachman Djauhari
The Bandung Statistical Institute, Bandung, Indonesia
&
Centre for Research, Consultation and Training in Statistical Analysis
Universitas Pasundan, Indonesia

Invited talk at MyStats 2016, organized by Bank Negara Malaysia, KL, November 15, 2016

CONTENT

- Policy maker's requirements in time series model building
- Characteristics of policy maker's data
- Fulfilling their requirements: **Richard Feynman philosophy**
- Results:
    1. **Theorem**
    2. **3 Corollaries** (GBM model)
    3. **A new procedure in time series modelling**
- Examples of real application (using R):
    1. Number of ***AirAsia*** passengers: NP
    2. Daily maximum electricity consumption
- Concluding Remarks
- References

Policy maker's requirements in time series model building

(in order of priority)

- Fast, fast, and fast model building
- Simple method; easy to digest & memorize
- Less computational effort
- Cheap modelling process
- Provide model with comparable accuracy

Characteristics of Policy maker’s data

Mostly positive

- This is a neglected aspect in time series modelling;
- No specific method is available in the literature

How to fulfil their requirements?

Approach:
Using Richard Feynman philosophy instead of Box-Jenkins' principle

Assumption:
Data represent log-normal process

Method:
Stochastic Differential Equation

Richard Feynman philosophy

Box-Jenkins' principle:
- Model identification
- Parameter estimation
- Model validation

Richard Feynman philosophy:
- Guess the law
- Derive the computational consequences
- Compare the guessed law with Nature

Result 1

Theorem. Consider the model $ln(X_{t+1}) = ln(X_t) + \varepsilon_1$ where ε_1 are i.i.n.d. If $\Delta t \to 0$, then $\{X_t\}$ satisfies

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

where dW_t is a Wiener process.

Result 2

Corollary 1. Since dW_t is a Wiener process, any stochastic process $\{X_t\}$ where X_t is positive and log-normally distributed is a GBM process.

Corollary 2. For any log-normal process $\{X_t\}$, the log-returns R_t are i.i.n.d. Thus, $R_{t+1} = R_t + \varepsilon_t$ where ε_t are i.i.n.d. with mean 0 and constant variance.

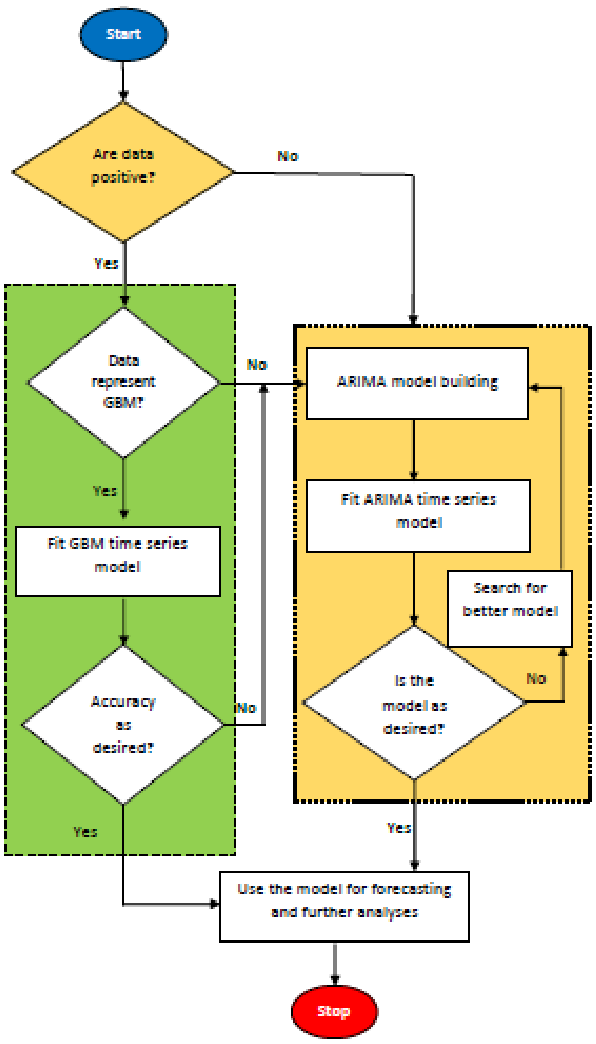
Result 2

Corollary 3. In general form, see Wilmott (2007) and Ross (2011), $\{X_t\}$ is an AR(1) process with constant term c, i.e., $R_t = c + \theta R_{t-1} + \varepsilon_t$. If $\hat{c}$ and $\hat{\theta}$ are the maximum likelihood estimates of c and θ, the predicted value of X_t is

$$\hat{X}_t = \exp(\hat{c}) \cdot X_{t-1} \left(\frac{X_{t-1}}{X_{t-2}} \right)^{\bar{\theta}} .$$

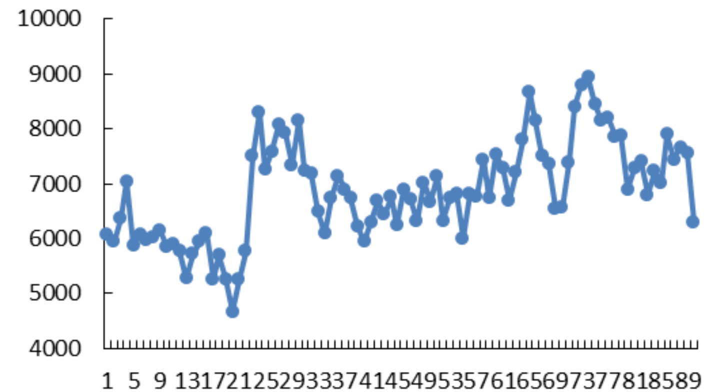
Result 3

A new procedure in time series modelling (Flow Chart):



Real Application 1: using R
(Number of AirAsia passengers: NP)

1. Q1-2009



GBM model:

$$\hat{X}_t = \exp(0.001) \cdot X_{t-1} \cdot \left(\frac{X_{t-1}}{X_{t-2}} \right)^{-0.1519}$$

MAPE = 6.82%, running time 0.14 sec

ARIMA model: **ARIMA(1,1,1)**

$$\hat{X}_t = X_{t-1} + 0.7288\left(X_{t-1} - X_{t-2}\right) - 0.9449 e_{t-1}$$

MAPE = 6.29%, running time 4.95 sec

2. Q1-2012

GBM model:

$$\hat{X}_t = \exp(0.0012) \cdot X_{t-1} \cdot \left(\frac{X_{t-1}}{X_{t-2}} \right)^{-0.1942}$$

MAPE = 4.44%, running time 0.16 sec

ARIMA model: **ARIMA(1,0,0)**

$$\hat{X}_t = 4602.567 + 0.539 X_{t-1}$$

MAPE = 4.10%, running time 5.62 sec

During those periods NP can be considered as a realization of GBM process. Is this also true for all periods? If "yes", then GBM process is a characteristic of NP which must be considered by AirAsia management in marketing development program.

Real Application 2: using R

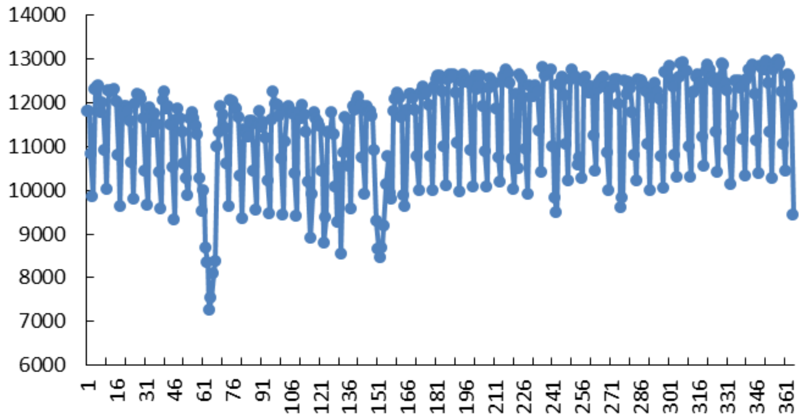
(Maximum daily electricity consumption)

1. WITHOUT CONSIDERING SEASONAL EFFECT

GBM model:

$$\hat{X}_t = \exp(-0.0006)X_{t-1}\left(\frac{X_{t-1}}{X_{t-2}}\right)^{-0.0522}$$

MAPE = 6.89%, running time 0.12 sec



ARIMA model: **ARIMA(4,1,5)**

$$\hat{X}_t = 1.7959X_{t-1} - 2.2243X_{t-2} + 2.2171X_{t-3} - 1.7743X_{t-4} + 0.9858X_{t-5}$$
$$-1.2101e_{t-1} + 1.6516e_{t-2} - 1.3333e_{t-3} + 1.1442e_{t-4} - 0.4502e_{t-5}$$

MAPE = 4.17%, running time 5.24 sec

2. SEASONAL EFFECT IS INCORPORATED

SARIMA model: **SARIMA(1,0,0)(0,1,1)-7**

$$\hat{X}_t = 0.7991X_{t-1} + X_{t-7} - 0.7991X_{t-8} - 0.9663e_{t-7}$$

MAPE = 2.66%

GBM model: $\hat{X}_t = 0.9993A_t X_{t-1}\left(\frac{X_{t-1}}{X_{t-2}}\right)^{-0.1255}$ with MAPE = 2.90%

$$A_t = \left(Adj(F_t)/Adj(F_{t-1})\right)\left(Adj(F_{t-2})/Adj(F_{t-1})\right)^{-0.1255}$$

and $Adj(F_t)$ is the Adjusted Seasonal Factor at time *t*.

ARIMA model: $\hat{X}_t = Adj(F_t)\left(2252.9566 + 0.8030\frac{1}{Adj(F_{t-1})}X_{t-1}\right)$

with MAPE = 2.81%

Concluding Remarks

- From data go directly to construct the model in Corollary 3;
- No need to check whether data are governed by GBM process.
- In general only **logarithmic transformation** & parameter estimation in **Simple Linear Regression** are required.
- Only if GBM model accuracy is not desirable, go to other method to find the desired one.
- We need 2 – 3 minutes to get GBM model using MS Excel.

References

- **Principal reference**:
  Djauhari, M.A., and Lee, S.L. (2016). *Forecasting methods: What all policy makers need*. UPM Press (In process of publication)*.

- **Additional references**:
  See the full paper

** The paper is part of this reference book*

Now, we can yell ABCDEF !
(Any Body Can Do Enjoyable Forecast !)

THANK YOU
FOR YOUR INTEREST

FOR FURTHER DISCUSSION:

Please visit The Bandung Statistical Institute

or

Contact

maman_djauhari@yahoo.com