Journal of Applied Mathematics & Bioinformatics, vol. x, no. x, xxxx, x-x
ISSN: 1792-6602(print), 1792-6939(online)

Modeling Underlying Assets Log-return in Merton Jump-Diffusion Framework

Megang Nkamga Junile Staures[1], Jane Akinyi Aduda[2] and Romuald Momeya[3]

Abstract

In this present paper we analyze two exponential Lévy models, the Black-Scholes model and the Merton Jump-Diffusion model from the perspective of the investigation of the skewness and excess kurtosis present in underlying assets log-returns distribution. Calibrating both models on real-world financial data and investigating their various moments and mean square error, we obtain results which show how the Merton jump-diffusion model performs better than the Black-Scholes model for modeling log-returns. This robust conclusion was also confirmed by using the Diebold-Mariano test to compare the forecast accuracy of the two models.

Mathematics Subject Classification : xxxx
Keywords: Black-Scholes (BS) model, Merton jump-diffusion (MJD) model, log-returns.

[1]Department of Mathematics, Pan African University, Institute of Basic Sciences, Technology, and Innovation, Kenya, e-mail: stauresmeg@gmail.com
[2]Department of Actuarial Science and Statistics, Jomo Kenyatta University of Agriculture, Technology and Innovation, e-mail: jaduda@jkuat.ac.ke
[3]Quantitative Analyst, CIBC Asset Management Inc. Montreal, QC, e-mail: momeya2008@gmail.com

Article Info: *Received* : xxxx. *Published online* : xxxx

1 Introduction

The key property in financial markets is extreme volatility. The prices of financial instruments such as stocks vary constantly and cause significant risk on businesses or organizations connected to such fluctuating prices. To mitigate this risk, modern finance establishes suitable models to assist investors or traders analyze real-world data to predict the future behavior of asset prices in the financial market. In this paper, we will focus on two exponential Lévy models (BS and MJD) to analyze the trend of the underlying asset prices log-return.

The BS model is a pioneer valuation model, proposed by [1] and a benchmark against which other models can be compared. BS model assumes that the log-return of the underlying assets is Gaussian, i.e., normally distributed. Unfortunately, assets in real markets rarely exhibit this behavior (see [2, 3]). Real asset prices and related market indexes show far more extreme fluctuations than predicted by Gaussian statistics. According to [2, 4], these extreme fluctuations leads to two empirical phenomena: *leptokurtic feature and volatility smile*. Therefore the need to propose another valuation model to handle this issue since the BS model experiences numerous inadequacies and is unable to explain empirical stylized facts in the financial markets.

In 1976 Robert C. Merton received the Nobel Prize award in Economics for developing a JDM called Merton jump-diffusion (MJD) model, a generalization of the BS model. Merton's main idea was to implant discontinuous jump processes in the classical BS model to help describe discontinuous price behavior in stock. The formula is the geometric Brownian motion (random walk) plus a compound Poisson process to account for "jumps" in the stock prices. The random walk together with the jump component are assumed to be sources of randomness in the stock prices. By adding three more parameters to the classical BS model, the MJD model captures the skewness and leptokurtic feature in the underlyings log-return distribution which differs from the BS normal log-return distribution. To compare the forecast accuracy of the two models, both models are calibrated and the results are used to compare the plot of their log-density functions, the values of their corresponding moments and mean square error. To boost our comparison, the Diebold-Mariano test [5] is applied and indicates the predictive performance of the MJD model over the BS model.

To investigate both models, some basic tools in probability and stochastic process theory will be introduced in section 2. Also in section 2, we will introduce the two exponential Lévy models and estimation of the model parameters by *Multinomial Maximum Likelihood function*. We conclude this section with a brief introduction to the Diebold-Mariano test. In section 3 and 4 we present the corresponding results and conclusion respectively.

2 Preliminary Notes

Definition 2.1 (log-return).
Asset is an investment instrument that can be bought and sold. Let X_t be the price of an asset at time t. For a time increment $\triangle t$, we define a log return $r_{\triangle t}$ as the natural logarithm of the simple gross return of an asset

$$r_{\triangle t} = \log\left(\frac{X_{t+\triangle t}}{X_t}\right)$$

Log return will be used in this study as their the statistical properties are more manageable.

Definition 2.2 (Stochastic Process).
A stochastic process X on a probability space $(\Omega, F, \mathbb{P})$ is a collection of random variables $\{X_t\}_{0\leq t\leq\infty}$.

Definition 2.3 (Lévy Process). *Let L be a stochastic process. Then L_t is a Lévy process if the following holds:*

- $L_0 = 0$
- *L has independent increments. That is $L_t - L_s$ is independent of F_s, $\quad 0 \leq s \leq t \leq \infty$.*
- *L has stationary increments. That is $\mathbb{P}(L_t - L_s \leq x) = \mathbb{P}(L_{t-s} \leq x), \quad 0 \leq s \leq t \leq \infty$*
- *L_t is continuous in probability, i.e $\lim_{t\longrightarrow s} L_t = L_s$*

Definition 2.4 (Poisson process).
A stochastic process $(N_t, t \geq 0)$ is called a Poisson process if

- *N_t only takes on values $0, 1, 2, \dots$ and $N_0 = 0$.*
- *If $s \leq t$, then $N_t - N_s$ is Poisson-distributed with mean $\lambda(t-s)$: this means that*

$$\mathbb{P}(N_t - N_s = n) = \frac{(\lambda(t-s))^n}{n!} e^{-\lambda(t-s)}, \qquad n = 0, 1, 2, \dots \tag{1}$$

  Let $t - s = dt$

- *If $u \leq s < t$, then N_u and $N_t - N_s$ are independent.*

Here $\lambda > 0$ is a constant which is called the intensity of the Poisson process.

Definition 2.5 (Compound Poisson process)**.**
Let $\{U_i\}_{i\geq 1}$ be a sequence of iid random variables, and $\{N_t\}_{t\geq 0}$ be a Poisson process with intensity parameter λ. Then the compound Poisson process $\{C_t\}_{t\geq 0}$ is defined by

$$C_t = \sum_{i=1}^{N_t} (U_i - 1),$$

with jump intensity λ.

Definition 2.6 (Jump-diffusion processes)**.**
Suppose we have a filtered probability space $(\Omega, F_T, \mathbb{F}, \mathbb{P}^*)$. Let N_t be a Poisson Process with intensity λ and waiting times $\tau = T - t$, W_t a Brownian Motion and $\{C_t\}_{t\geq 0}$ be a compound Poisson process as define above, then

$$S_t = \gamma t + \sigma W_t + C_t$$

is a jump-diffusion process, where $\sigma \in \mathbb{R}^+$ and $\gamma \in \mathbb{R}$

2.1 The BS model specification

In this model the asset price is described as follows

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t, \tag{2}$$

where

- μ is the expected return of the asset
- σ is the volatility of the asset and

- W_t is a Wiener process.

The solution to equation 2 is given by

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{3}$$

where X_0 is the stock price at time t=0, μ, σ and W_t are define as in equation 2. The assets are modeled by 3. From 3, the log-return of the asset price is:

$$L_t = \ln(\frac{X_t}{X_0}) = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t$$

Under the BS model, definition 2.1 tells us that for a time increment $\triangle t$ the

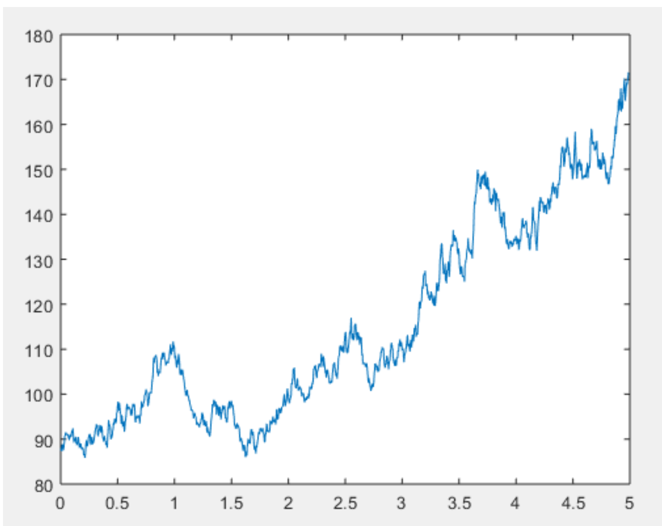

Figure 1: For the given parameters $\mu = 0.17, \sigma = 0.20$ and $X_0 = 365$ and simulation of the standard Brownian motion, we easily generate BS modeled asset prices from equation 3.

log-return of the asset price is given by:

$$\begin{aligned} L_{\triangle t} &= \ln(\frac{X_{t+\triangle t}}{X_t}) \\ &= (\mu - \frac{1}{2}\sigma^2)\triangle t + \sigma(W_{t+\triangle t} - W_t) \\ &\sim N\left((\mu - \frac{\sigma^2}{2})\triangle t, \sigma^2 \triangle t\right) \end{aligned} \tag{4}$$

which means

$$L_{\triangle t} - (\mu - \frac{\sigma^2}{2})\triangle t = \sigma \triangle W_t \sim N(0, \sigma^2 \triangle t)$$

The above equation shows that the log-returns follow the Gaussian distribution. It is no news that the BS model is complete, and in this framework there exists a unique risk-neutral measure $\mathbb{P}^*$ under which the discounted asset price is a martingale, see for example [6].

By the Girzanov theorem, the asset prices are described by the risk-neutral dynamic as follows

$$\frac{dX_t}{X_t} = rdt + \sigma dW_t^*,$$

where

$$W_t^* = \frac{\mu - r}{\sigma} t + W_t, \quad t \in [0, T]$$

is a $\mathbb{P}^*$-Brownian motion and $\frac{\mu - r}{\sigma}$ the market price of risk measured per unit of volatility. Under BS model, the characteristic function associated with the log-return asset price $L_{\triangle t}$ is given by:

$$\Phi(\theta) = \exp\left\{\triangle t \left(i(\mu - \frac{\sigma^2}{2})\theta - \frac{1}{2}\sigma^2\theta^2 \right)\right\} \tag{5}$$

Regarding the leptokurtic feature which is a measure of how heavy or fat the tail of the log-returns distribution is, we'll exploit the approach suggested in [7].The theorem below is a consequence of equation 4.

Theorem 2.7.
Let $\mathbb{E}^*_{BS}, M_i^{BS}, i = 2:4, S_{BS}$ and K_{BS} denote the mean, the central moments, the skewness and excess kurtosis for the BS-modeled log-returns respectively. We've

$$\begin{cases} \mathbb{E}^*_{BS} = (\mu - \frac{1}{2}\sigma^2)\triangle t \\ M_2^{BS} = \sigma^2 \triangle t \\ M_3^{BS} = 0 \\ M_4^{BS} = 3\sigma^4(\triangle t)^2 \\ S_{BS} = 0 \\ K_{BS} = 3 \end{cases}$$

Proof.
From equation 4, we clearly see that

$$\begin{aligned} \mathbb{E}^*_{BS} &= \mathbb{E}^*[L_{\triangle t}] \\ &= (\mu - \frac{1}{2}\sigma^2)\triangle t \quad \text{since } \triangle W_t = W_{t+\triangle t} - W_t \sim N(0, \triangle t). \end{aligned}$$

By the central moment definition

$$\begin{aligned} M_i^{BS} &= \mathbb{E}^*\left[(L_t - \mathbb{E}^*[L_t])^i\right] \\ &= \mathbb{E}^*\left[(\sigma\triangle W_t)^i\right] \\ &= \sigma^i\mathbb{E}^*\left[(\triangle W_t)^i\right] \end{aligned} \tag{6}$$

Let $\Phi_{\triangle W_t}(\theta)$ denote the characteristic function of $\triangle W_t$. Upon calculation we obtain

$$\begin{aligned} \Phi_{\triangle W_t}(\theta) &= \mathbb{E}^*(e^{i\theta\triangle W_t}) \\ &= \int_{-\infty}^{\infty} e^{i\theta x}.\frac{1}{\sqrt{2\pi\triangle t}}e^{-\frac{x^2}{2\triangle t}}dx \\ &= \frac{1}{\sqrt{2\pi\triangle t}}\int_{-\infty}^{\infty} e^{-\frac{(x-i\theta\triangle t)^2+(\theta\triangle t)^2}{2\triangle t}}dx \\ &= e^{-\frac{\theta^2\triangle t}{2}} \quad \text{since } \frac{1}{\sqrt{2\pi\triangle t}}\int_{-\infty}^{\infty} e^{-\frac{(x-i\theta\triangle t)^2}{2\triangle t}}dx{=}1 \\ &= \exp\{\Psi(\theta)\} \end{aligned}$$

where $\Psi(\theta) = e^{-\frac{\theta^2\triangle t}{2}}$.

Assuming that the characteristic function is sufficiently differentiable,

$$\begin{aligned} &\Longrightarrow \mathbb{E}^*[(\triangle W_t)^n] = \frac{1}{i^n}\left[\frac{d\Psi(0)}{d\theta}\right] \\ &\Longrightarrow \mathbb{E}^*[(\triangle W_t)^2] = \triangle t \end{aligned}$$

From equation 6 we get

$$M_2^{BS} = \sigma^2\triangle t$$

Also we have that

$$\mathbb{E}^*[(\triangle W_t)^3] = 0$$

leading to

$$\begin{aligned} M_3^{BS} &= 0 \quad \text{and} \\ M_4^{BS} &= 3\sigma^4\triangle t \end{aligned}$$

which yields

$$\begin{aligned} S^{BS} &= \frac{M_3^{BS}}{(M_2^{BS})^1.5} = 0 \\ K_{BS} &= \frac{M_4^{BS}}{(M_2^{BS})^2} = 3 \end{aligned}$$

□

2.1.1 Estimation of the parameters μ and σ for the BS model

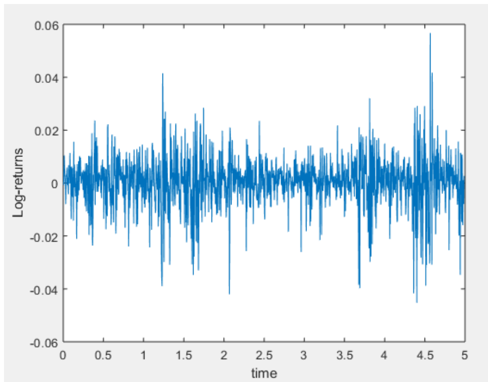

Figure 2: Empirical log-returns r_{dt} of NASDAQ index.

To find estimates of μ and σ based on empirical data, let us consider the empirical log-returns r_{dt} of NASDAQ index data from 2014-02-06 to 2019-05-23, with 1332 trading days in figure 2 above.

Estimation of model parameters (μ, σ) is required for us to fit empirical data to the BS model. From theorem 2.7:

$$\begin{cases} \mathbb{E}^*_{BS} = (\mu - \frac{1}{2}\sigma^2)\triangle t \\ M_2^{BS} = \sigma^2 \triangle t \end{cases}$$

so that

$$\begin{cases} \hat{\mu} = \frac{2\widehat{\mathbb{E}^*_{BS}} + \widehat{\mathrm{Var}}[L_{\triangle t}]\triangle t}{2\triangle t} \\ \hat{\sigma} = \sqrt{\frac{\widehat{\mathrm{Var}}[r_{dt}]}{\triangle t}} \end{cases}$$

where $\widehat{\mathbb{E}^*_{BS}}$ and $\widehat{\mathrm{Var}}[L_{\triangle t}]$ are the sample mean and sample variance of the empirical log-returns respectively.

2.2 The MJD model specification

The model propose that the underlying asset price evolves according

$$\frac{dX_t}{X_{t-}} = \underbrace{(\mu - \lambda k)dt + \sigma dW_t}_{\text{Continuous part}} + \underbrace{d\left(\sum_{i=1}^{N_t}(J_i - 1)\right)}_{\text{Discontinuous or jump part}} \tag{7}$$

where: μ is the instantaneous expected return on the stock; $k = \mathbb{E}[J_t - 1]$ is the expected percentage change in the stock price if a Poisson event occur; σ is the instantaneous variance of the return, condition on no arrival of important new information (i.e. a Poisson process does not occur) that one assumes constant. X_{t-} the asset price before a jump occurs at time t; dW_t is a standard Brownian motion; N_t is a Poisson process represents the arrival of new information (events) which has a significant effect on the stock price with parameter λ which stand for the average number of jump arrivals per unit of time; J_t the jump sizes are *i.i.d.* Because of the jump component provided by the compound Poisson process, the model considered in 7 is a process not purely Gaussian representing a particular case of Lévy process. The solution to 7 is given by:

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2 - \lambda k)t + \sigma W_t + \sum_{i=1}^{N_t} U_i} \tag{8}$$

where $\sum_{i=1}^{N_t} U_i$ is a normally distributed compound Poisson process with mean $N_t \mu_j$, variance $N_t \sigma_j^2$ and intensity λ. In other-words, for $n = N_t, \sum_{i=1}^{n} U_i \sim \mathcal{N}(n\mu_j, n\sigma_j^2)$. The index j represents the jump part of MJD model.

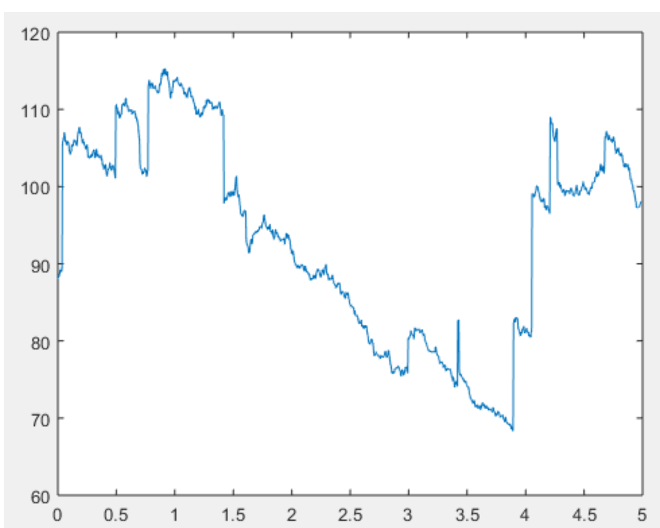

Figure 3: For the assumed parameters $\mu = 0.21, \sigma = 0.056, \lambda = 4, \mu_j = 0.051$ and $\sigma_j = 0.097$, we present simulated sample path of MJD-modeled asset prices.

Equation 8 describes stock price behavior under MJD model. From equation 8, the log stock prices are given by

$$\ln X_t = \ln X_0 + (\mu - \frac{1}{2}\sigma^2 - \lambda k)t + \sigma W_t + \sum_{i=1}^{N_t} U_i$$

Then by definition 2.1 for a time increment $\triangle t$, we get the log-returns $L_{\triangle t}$ of the asset prices modeled by MJD model as

$$\begin{aligned} L_{\triangle t} &= \ln(\frac{X_{t+dt}}{X_t}) \\ &= \gamma\triangle t + \sigma\triangle W_t + \sum_{i=1}^{\triangle N_t} U_i \end{aligned} \tag{9}$$

where $\gamma = \mu - \frac{1}{2}\sigma^2 - \lambda k$, $U_i^{'s}$ are $i.i.d$ and $U_i \sim \mathcal{N}(\mu_j, \sigma_j^2)$; $\triangle W_t = W_{t+\triangle t} - W_t$ is the standard Brownian motion increment; and lastly $\triangle N_t = N_{t+\triangle t} - N_t$ a Poisson process with mean $\lambda\triangle t$. [7] proof that the assumption about the log-return jump size distribution being normal facilitates the derivation of the probability density function of the log-returns $L_{\triangle t}$ which is a converging series of the following form:

$$\begin{aligned} g_t = \mathbb{P}(L_{\triangle t} \in B) &= \sum_{h=0}^{\infty} \mathbb{P}(\triangle N_t = a)\mathbb{P}(L_{\triangle t}|\triangle N_t = a) \quad \text{for } B \subset \mathbb{R} \\ &= \sum_{a=0}^{\infty} \frac{e^{-\lambda\triangle t}(\lambda\triangle t)^a}{a!} \mathcal{N}\left(L_{\triangle t}; \gamma\triangle t + \mu_j a, \sigma^2\triangle t + \sigma_j^2 a\right) \end{aligned}$$

where

$$\mathcal{N}\left(L_{\triangle t}; \gamma\triangle t + \mu_j a, \sigma^2\triangle t + \sigma_j^2 a\right) = \frac{1}{\sqrt{2\pi(\sigma^2\triangle t + \sigma_j^2 a)}} e^{-\frac{\left(L_{\triangle t} - [\gamma\triangle t + \mu_j a]\right)^2}{2(\sigma^2\triangle t + \sigma_j^2 a)}}$$

is the normal density function of $L_{\triangle t}$ assuming that the asset price jumps a times in the time interval $\triangle t$ and $\mathbb{P}(\triangle N_t = a)$ is the probability that the asset price jumps a times in the time interval $\triangle t$. g_t is expressed as the weighted sum of normal densities. In such a framework, the market model is incomplete, hence the existence of more than one risk-neutral measure $\mathbb{P}^*$. To have a risk-neutral measure, we need to replace the compound Poisson process in equation 7 by a compensated compound Poisson process.The latter can be done in several ways. Let us consider the discounted asset price $e^{-rt}X_t$ such that

$$\begin{aligned}
d(e^{-rt}X_t) &= -re^{-rt}X_t dt + e^{-rt}dX_t \\
&= -re^{-rt}X_t dt + e^{-rt}X_t(\mu - \lambda k)dt + e^{-rt}\sigma dW_t X_t + e^{-rt}X_t d\left(\sum_{i=1}^{N_t}(J_i - 1)\right) \\
&= e^{-rt}X_t(\mu - r - \lambda k + \lambda^* k^*)dt + e^{-rt}\sigma d(W_t^* - ut)X_t + e^{-rt}X_t d\left(\sum_{i=1}^{N_t}(J_i - 1) - \lambda^* k^*\right) \\
&= e^{-rt}X_t(\mu - r - \lambda k + \lambda^* k^* - \sigma u)dt + e^{-rt}\sigma dW_t^* X_t + e^{-rt}X_t d\left(\sum_{i=1}^{N_t}(J_i - 1) - \lambda^* k^*\right)
\end{aligned}$$

For $e^{-rt}X_t$ to be a martingale we suppose

$$\mu - r - \lambda k + \lambda^* k^* - \sigma u = 0$$

where u is such that

$$W_t^* = ut + W_t \quad t \in [0, T]$$

is a $\mathbb{P}^*$-Brownian motion thanks to the Girzanov theorem. $\lambda^* > 0$ is the new intensity and $k^* = \mathbb{E}^*[J - 1]$. Merton proposes the following for the change of measure

$$\lambda^* = \lambda$$

So that

$$f_U^*(j) = f_U(j) \Longrightarrow k = k^*$$

$$\Longrightarrow u = \frac{\mu - r}{\sigma}$$

For simplicity, if we let $u = 0$, then $\mu = r$. Since the jump risk is diversify and no risk premium is attached to it, we can leave the jump part unchanged. So the new asset price dynamic under $\mathbb{P}^*$ is

$$\frac{dX_t}{X_{t-}} = (r - \lambda k)dt + \sigma dW_t^* + d\left(\sum_{i=1}^{N_t}(J_i - 1)\right) \tag{10}$$

For the MJD model, the central moment is given by:

$$\begin{aligned}
M_i^{MJD} &= \mathbb{E}^*\left[(L_t - \mathbb{E}^*[L_t])^i\right] \\
&= \mathbb{E}^*\left[(\sigma \triangle W_t + U \triangle N_t - \lambda \mu_j \triangle t)^i\right]
\end{aligned}$$

with $W_t, U, \triangle N_t$ being *i.i.d.* The characteristic function of $L_{\triangle t}$ is obtained by applying Fourier transform to g_t as

$$\Phi(w) = e^{\triangle t \Psi(w)}$$

where

$$\Psi(w) = \lambda \exp\left\{(iw\mu_j - \frac{1}{2}\sigma_j^2 w^2) - 1\right\} + iw\left(\mu - \frac{1}{2}\sigma^2 - \lambda k\right) - \frac{1}{2}\sigma^2 w^2$$

is the characteristic exponent (cumulant generating function) and

$$\begin{aligned} k &= \mathbb{E}[e^U - 1] \\ &= e^{\mu_j + \frac{1}{2}\sigma_j^2} - 1. \end{aligned}$$

Theorem 2.8.

Considering the central moment and the characteristic exponent defined above, we've

$$\begin{aligned}
\mathbb{E}^*_{MJD} &= (\gamma + \lambda\mu_j)\triangle t \\
M^2_{MJD} &= \left(\sigma^2 + \lambda(\sigma_j^2 + \mu_j^2)\right)\triangle t \\
M^3_{MJD} &= \lambda(3\sigma_j^2 + \mu_j^3)\triangle t + 6\mu_j\sigma_j^2(\lambda\triangle t)^2 \\
M^4_{MJD} &= \lambda(3\sigma_j^4 + \mu_j^4 + 6\sigma_j^2\mu_j^2)\triangle t + 3(\sigma^2\triangle t)^2 + (3\mu_j^4 + 21\sigma_j^4 + 30\mu_j^2\sigma_j^2)(\lambda\triangle t)^2 \\
&\quad + 6\sigma^2\triangle t(\sigma_j^2 + \mu_j^2)\lambda\triangle t + (6\mu_j^2\sigma_j^2 + 18\sigma_j^4)(\lambda\triangle t)^3 + (6\sigma^2\sigma_j^2\triangle t(\lambda\triangle t)^2 + 3\sigma_j^4(\lambda\triangle t)^4 \\
S_{MJD} &= \frac{M^3_{MJD}}{[M^2_{MJD}]^{1.5}} \\
K_{MJD} &= \frac{M^4_{MJD}}{[M^2_{MJD}]^2}
\end{aligned}$$

For the proof of the above theorem, the reader is referred to [7, 8].

2.2.1 Some Important Properties of the PDF g_t of Merton log-returns

1. The sign of skewness is determined by the size of $\mathbb{E}[J] = \mu_j$. "Figure" 4 shows that g_t is asymmetric if $\mu_j \neq 0$ and symmetric if $\mu_j = 0$.

2. The value of λ makes the density fatter-tailed as illustrated in "Figure" 4. Note that the excess kurtosis in the case $\lambda = 50$ is much smaller than in the case $\lambda = 2$ or $\lambda = 4$. This is because excess kurtosis is a standardized measure (by standard deviation)

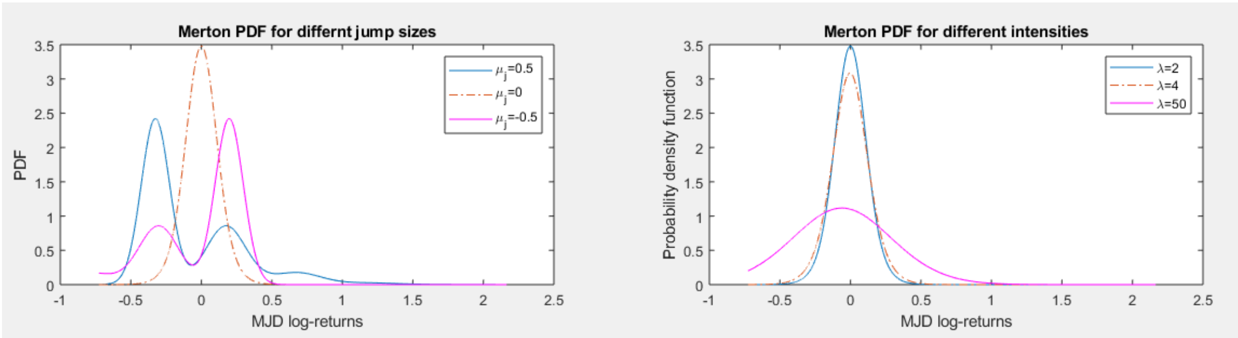

Figure 4: Pdf of Merton log-returns with different values of μ_j and λ

2.2.2 Estimation of model parameter $\Theta = (\mu, \sigma, \lambda, \mu_j, \sigma_j)$

Estimation of model parameter Θ is required for us to fit empirical data to the MJD model. From theorem 2.8, we see that the five parameters $\mu, \sigma, \lambda, \mu_j$, and σ_j do not have explicit expressions. To address this issue, we will consider the situation when a single jump occurs *i.e.* $N_t = 1$. Let us consider the empirical log-returns $L_{\triangle t}$ of NASDAQ index data in "Figure" 2.

We propose a decision rule about the occurrence of jumps whereby a threshold $\epsilon > 0$ is chosen by observing the plot of the empirical log-returns, such that a jump occurs only if the absolute value of the log-returns is greater than ϵ.

Splitting the empirical log-return data into two categories $\mathbb{D}$ and $\mathbb{J}$ for a given ϵ, the category D includes log-returns with no jump. The category $\mathbb{J}$ includes log-returns with jumps. We will refer to the model parameters estimated when one jump occur as *primary estimates*.

Case 1: Occurrence of one jump

Here the first and second moments of the log-return are

$$\begin{aligned} \mathbb{E}^*[L^{\mathbb{J}}_{\triangle t}] &= \mathbb{E}^*[L_{\triangle t} | \triangle N_t = 1] \\ &= \left(\mu - \frac{\sigma^2}{2} - k\right) \triangle t + \mu_j \end{aligned} \tag{11}$$

and

$$\begin{aligned} Var[L^{\mathbb{J}}_{\triangle t}] &= Var[L_{\triangle t} | \triangle N_t = 1] \\ &= \sigma_j^2 + \sigma^2 \triangle t \end{aligned} \tag{12}$$

respectively. From equation 11 and 12, we estimate μ_j and σ_j as follows

$$\begin{cases} \hat{\mu_j} = \widehat{\mathbb{E}}^*[L^{\mathbb{J}}_{\triangle t}] - (\hat{\mu} - \frac{\hat{\sigma}^2}{2} - k)\triangle t \\ \hat{\sigma_j} = \sqrt{\widehat{\text{Var}}(L^{\mathbb{J}}_{\triangle t}) - \hat{\sigma}^2\triangle t} \end{cases}$$

where $\widehat{\mathbb{E}}^*[L^{\mathbb{J}}_{\triangle t}]$ and $\widehat{\text{Var}}[L^{\mathbb{J}}_{\triangle t}]$ are the sample mean and the sample variance of the empirical log-returns in category $\mathbb{J}$.

Case 2: No occurrence of jumps

Here the first and second moments of the log-return are

$$\begin{aligned} \mathbb{E}^*[L^{\mathbb{D}}_{\triangle t}] &= \mathbb{E}^*[L_{\triangle t}|\triangle N_t = 0] \\ &= (\mu - \frac{\sigma^2}{2})\triangle t \end{aligned} \tag{13}$$

and

$$\begin{aligned} Var[L^{\mathbb{D}}_{\triangle t}] &= Var[L_{\triangle t}|\triangle N_t = 0] \\ &= \sigma^2\triangle t \end{aligned} \tag{14}$$

respectively. From equation 13 and 14, we estimate μ and σ as follows:

$$\begin{cases} \hat{\mu} = \frac{2\widehat{\mathbb{E}}^*[L^{\mathbb{D}}_{\triangle t}] + \widehat{\text{Var}}(L^{\mathbb{D}}_{\triangle t})\triangle t}{2\triangle t} \\ \hat{\sigma} = \sqrt{\frac{\widehat{\text{Var}}(L^{\mathbb{D}}_{\triangle t})}{\triangle t}} \end{cases}$$

where $\widehat{\mathbb{E}}^*[L^{\mathbb{D}}_{\triangle t}]$ and $\widehat{\text{Var}}[L^{\mathbb{D}}_{\triangle t}]$ are the sample mean and the sample variance of the empirical log-returns in category $\mathbb{D}$.

Supposing time is measured in years, the parameter λ is estimated as follows

$$\begin{aligned} \hat{\lambda} &= \text{number of jumps per year} \\ &= \frac{\text{Total number of jumps}}{\text{Total length in years}} \end{aligned}$$

From "Figure" 2, if we take for example $\epsilon = 0.01$, we obtain the values for our primary estimators as follows: $\hat{\lambda} = 60.74, \hat{\mu} = 0.1899, \hat{\sigma} = 0.0741, \hat{\mu_j} = -0.0012$ and $\hat{\sigma_j} = 0.0182$.

To get optimal estimators let us consider the case below:

Case 3: occurrence of $N_t = a$ jumps

For the estimates of $\Theta = (\mu, \sigma, \lambda, \mu_j, \sigma_j)$ when $N_t = a$, we will use the *Multinomial Maximum Likelihood Approach* proposed by [7]. To be implement this approach, we will do the folowing:

1. Divide the empirical data set into categories of length $n <$ length (empirical data), which has already been achieved above. *i.e.* $\mathbb{D}$ and $\mathbb{J}$.

2. Minimize the objective function by finding an optimal $\widehat{\Theta}$ that minimizes the likelihood function

$$L(\Theta; x) = -\prod_{i=0}^{n} \log(\mathbb{P}(L^i_{\triangle t}) \in B) \quad B \subset \mathbb{R}$$

where the $L^i_{\triangle t}$ represent the empirical log-returns.

The primary estimators obtained in the previous section is used to numerically minimize the MJD model objective function. The presence of the sum of accumulated jumps in the log-return evolution makes it non-normal. In section 2.2, we saw that the probability density function of $L_{\triangle t}$ is given by

$$\mathbb{P}(L_{\triangle t} \in B) = \sum_{a=0}^{\infty} \frac{e^{-\lambda \triangle t}(\lambda \triangle t)^a}{a!} \frac{1}{\sqrt{2\pi(\sigma^2 \triangle t + \sigma_j^2 a)}} e^{-\frac{\left(L_{\triangle t} - [\gamma \triangle t + \mu_j a]\right)^2}{2(\sigma^2 \triangle t + \sigma_j^2 a)}}$$

Some authors like [9] and [10, 11] highlight that minimization problem is easily obtained through the *regime switching technique*. In this paper, the MATLAB code *fminsearch* will ease the process of finding an optimal $\widehat{\Theta}$ that minimizes $L(\Theta; x)$.

2.3 Diebold-Mariano Test

This test is often used to compare time series models. Suppose we denote the empirical data set by $\{y_t, t = 1, \ldots, N\}$. Let the samples from BS model and MJD model be denoted by $\{\widehat{y}_{1t}, t = 1, \ldots, N\}$ and $\{\widehat{y}_{2t}, t = 1, \ldots, N\}$ respectively. We refer to the BS model as model one and the MJD model as model two. The question we ask ourselves is which forecasting model is actually good to better represent the empirical data.

Define the forecast errors as

$$e_{it} = \widehat{y}_{it} - y_t, \quad i = 1 : 2$$

The loss associated with forecast i is assumed to be a function of the forecast error, e_{it}, and is denoted by $g(e_{it})$. The function $g(.)$ is a loss function, that is a function such that:

. takes the value zero when no error is made;

. is never negative;

. is increasing in size as the errors become larger in magnitude.

Typically, $g(e_{it})$ is the square (squared-error loss) or the absolute value (absolute-error loss) of e_{it}. We define the loss differential between the two forecasts by

$$d_t = g(e_{1t}) - g(e_{2t})$$

and say that the two forecasting models have equal accuracy if and only if the loss differential has zero expectation $\forall t$. So, we would like to test the null hypothesis

$$H_0 : \mathbb{E}[d_t] = 0, \quad \forall t$$

versus the alternative hypothesis

$$H_1 : \mathbb{E}[d_t] \neq 0, \quad \forall t$$

The null hypothesis is that the two models have the same accuracy. The alternative hypothesis is that the two models have different levels of accuracy.

In the case where both models have different accuracy, we perform another test to detect which model is more accurate than the other using the hypothesis below:

$$H_0 : \text{Model one and two have the same accuracy}$$

versus the alternative hypothesis

$$H_1 : \text{Model two is more accurate than model one}$$

For the level of significance $\alpha = 0.05$, we fail to reject the null hypothesis if the P-value $> \alpha$

3 Main Results

Models are rough and wrong approximation of real world phenomenon. The two option pricing models (BS and MJD) introduced in section 2 holds the potential to reproduce their real world counter parts. Here we discuss the modeling results obtained as describe in section 2.

The model parameters are estimated from NASDAQ index. Table 1 presents results under the BS model. Under the MJD model, we use the primary parameters estimated in section 2.2.2 and apply the MATLAB *fminsearch* code to obtain optimal parameters for different thresholds ϵ. See results in table 2.

Table 1: Parameter estimates for BS model.

Parameters	Estimates
μ	0.0005
σ	0.0101

Table 2: Relationship between MLE parameters and ϵ

Parameter	$\epsilon = 0.01$		$\epsilon = 0.02$		$\epsilon = 0.03$		$\epsilon = 0.06$	
	P.E	E_{mle}	P.E	E_{mle}	P.E	E_{mle}	P.E	E_{mle}
μ	0.1899	0.3615	0.2424	0.3615	0.2170	0.3615	0.1291	0.12914
σ	0.0741	0.0655	0.1176	0.0655	0.1411	0.0655	0.1602	0.1601
λ	60.74	210.8795	15.6372	210.8795	4.2100	210.8795	0	-0.8072
μ_j	-0.0012	-0.0028	-0.0076	-0.0028	-0.0216	-0.0028	NaN	NaN
σ_j	0.0182	0.0097	0.0267	0.0097	0.0310	0.0097	NaN	NaN
L	-3.3085×10^3		-3.3085×10^3		-3.3085×10^3		-3.3085×10^3	

From table 2, P.E and E_{mle} represent the model's primary and MLE estimates respectively. We observe that the MLE estimates of the parameters $\hat{\mu}, \hat{\sigma}, \hat{\lambda}, \hat{\mu_j}$ and $\hat{\sigma_j}$ are the same for $\epsilon \in \{0.01, 0.02, 0.03\}$. For $\epsilon = 0.06$ we obtain different values of θ because the maximum number of function evaluations has been exceeded. The values of ϵ are chosen by taking a close look at figure 2. It is logical to think that if there are no jumps like the case of $\epsilon = 0.06$, the jump parameters would not have estimated values. From this same table 2, we conclude that the MLE method is independent of the threshold ϵ.

From section 2.1.1 and 2.2.2 and considering the MLE estimates of the model parameters, the values for the moments and mean square error under the BS and MJD model are given in table 3 and 4 respectively.

From table 3, estimated means and variances of log-returns for the BS model coincides with the real log-returns. The estimated variances of log-

Table 3: Comparing moments of the BS model, MJD model with that of real data.

	Mean	Variance	Skewness	excess Kurtosis
Empirical	0.0005	0.0001	-0.4658	9.1592
BS model	0.0005	0.0001	0	3
MJD model	0.0014	0.0001	-0.6494	5.4869

returns for both the BS model and the real log-returns coincides with MJD-modeled log-returns.The MJD model captures the negative skewness of the empirical log-returns since its skewness coefficient is closer to the real data. Also, the kurtosis coefficient for the MJD model seems to be closer to the empirical kurtosis than for the BS model indicating more pronounce fatter tails for the log-returns distribution. The mathematical features of the MJD model makes it possible to handle the large spikes in empirical log-returns, see figure 2.

Table 4: Comparing mean square error of BS model and MJD model.

	Mean Square Error
BS model	1.8734×10^5
MJD model	1.6821×10^3

Investigating the performance of both models in terms of mean square error, table 4 clearly shows that the MJD model is better than the BS model since its mean square error is smaller. Hence, we are tempted to conclude at least for our empirical NASDAQ data that the MJD model is significantly more suitable than the BS model for modeling log-returns.

For NASDAQ data considered in this paper, ”Figure” 5 and 6 clearly shows that the distribution of the empirical log-returns is not Gaussian, and the MJD model seems to reflect reality on the market since its density happens to be above the kernel density of the log returns. On the other hand, the BS model happens not to be very suitable in modeling log-returns as it assumes that the log-returns of the underlying asset is Gaussian and our empirical log-returns show far more extreme fluctuations than predicted by Gaussian statistics.

The Diebold-Mariano test carried out numerically we obtained a P-value

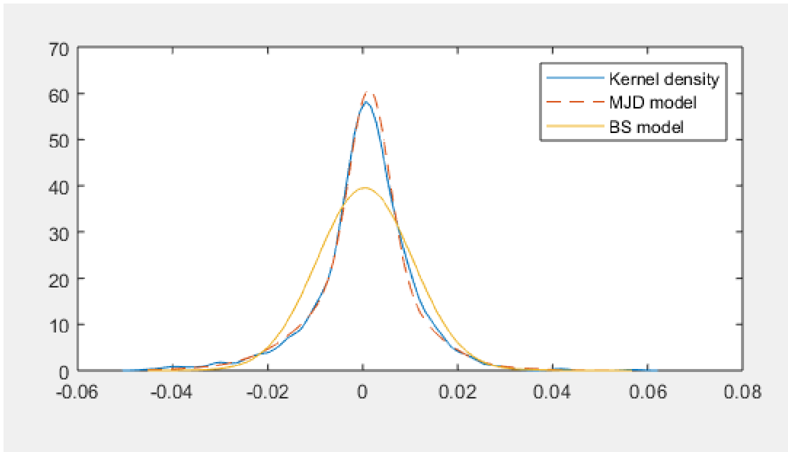

Figure 5: Density for IXIC1 time series. Solid blue line: Kernel density estimator applied directly on data. Dashed line: MJD model simulation with estimated parameters. Solid yellow line: BS model simulation with estimated parameters

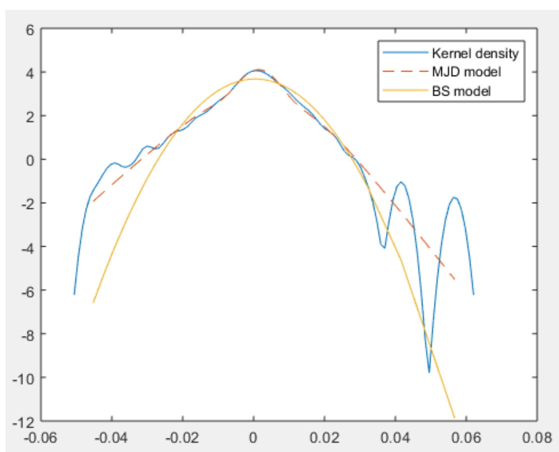

Figure 6: Logarithm density for IXIC1 time series.Kernel density estimator applied directly on data. Dashed line: MJD model simulation with estimated parameters. Solid yellow line: BS model simulation with estimated parameters

of 5.711×10^{-6} for the first test which is less than 0.05 showing that both models have different accuracy. For the second test, we obtain a P-value of 7.664×10^{-4}. This simply indicates that the MJD model is more accurate than the BS model.

4 Conclusion

In this paper we have compared the fitness to the data performances characterizing the BS model and the MJD model, which is obtained from the BS by adding a compensated compound Poisson process to the main stochastic differential equation. The financial assumptions behind these market models makes them similar in many ways. Nevertheless, the BS market model is complete, while the Merton model is incomplete. The latter fact is due to the impossibility to completely mitigate the risk carried by the introduction of sudden and unpredictable moves in the stock price. Hence, even if one can consider the latter as an advantage carried by the BS-approach, at least in terms of mathematical simplicity and numerical tractability, the Merton model turns out to outperform the BS model, when one takes into account the performances of the two with respect to real financial data. In particular, moving from a theoretical comparison to an empirical one, the addition of the jump parameters results in a great improvement in simulation of log-returns distribution. Also, the log-returns leptokurtic feature is much more evident using the Merton model approach instead of the BS model when we compare their density functions with the kernel density estimation for the NASDAQ empirical data. To boost our conclusion, we perform two tasks: first we compare the moments and density functions of both models with empirical moments and kernel density respectively and secondly we use the mean square error and the Diebold-Mariano test to compare both models. Despite the advantages of MJD model, it does not incorporate the volatility clustering effect. In future works, we plan to apply the Vasicek model approach similar to the one proposed by Merton, namely taking random jumps into consideration, to what concerns the interest rates financial frameworks.

ACKNOWLEDGEMENTS. First of all I want to thank God almighty for his support in my life. I also want to express my huge and sincere gratitude to my supervisors Dr.Jane Akinyi and Dr.Romuald Momeya for all their wise suggestions. I must thank them for their time, it was great pleasure to write this paper under their direction. I also appreciate very much all my teachers for the precious knowledge which I got from them and my friends for their support. Special thanks to the Pan African University Institute for Basic Sciences, Technology and Innovation for the good conditions and support.

References

[1] F. Black, M. Scholes, The pricing of options and corporate liabilities, Journal of political economy 81 (1973) 637–654.

[2] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues (2001).

[3] N. Gugole, Merton jump-diffusion model versus the black and scholes approach for the log-returns and volatility smile fitting, International Journal of Pure and Applied Mathematics 109 (2016) 719–736.

[4] V. G. Kulkarni, Modeling and analysis of stochastic systems, Chapman and Hall/CRC, 2016.

[5] F. X. Diebold, Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold–mariano tests, Journal of Business & Economic Statistics 33 (2015) 1–1.

[6] M. Kijima, Stochastic processes with applications to finance, Chapman and Hall/CRC, 2016.

[7] F. B. Hanson, Z. Zhu, Comparison of market parameters for jump-diffusion distributions using multinomial maximum likelihood estimation, in: 2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601), volume 4, IEEE, pp. 3919–3924.

[8] K. Matsuda, Introduction to merton jump diffusion model, Department of Economics. The Graduate Center, The City University of New York (2004).

[9] C. Kim, C. Nelson, State-space models with regime switching: Classical and gibbs-sampling approaches with applications (vol. 1)(0262112388), 1999.

[10] L. Di Persio, M. Frigo, Maximum likelihood approach to markov switching models, WSEAS Transactions on Business and Economics 12 (2015) 239–242.

[11] L. Di Persio, M. Frigo, Gibbs sampling approach to regime switching analysis of financial time series, Journal of Computational and Applied Mathematics 300 (2016) 43–55.