Two Stages Fitting Techniques using Generalized Lambda Distribution: Application on Malaysian Financial Return

MUHAMMAD FADHIL MARSANI* & ANI SHABRI

ABSTRACT

The underline distribution assumption used in the analysis of share market returns is crucial in risk management. An important aspect related to stock return modelling is to obtain accurate prediction. This paper presents an innovative fitting method called two stages (TS) method for modelling daily stock returns. The proposed approach by first establishing trend in the series, and then separately performing L-moment estimation on the generalized lambda distribution (GLD) parameter. The performance of the TS-GLD models had been evaluated using Monte Carlo simulation and Malaysian Kuala Lumpur Composite Index (KLCI) returns from year 2001 to 2015. Based on k-sample Anderson Darling goodness of fit test, the two stages GLD model in location parameter (GLD.1) performed well in all studied cases. The GLD.1 model benefits risk management by providing effective distribution fitting.

Keywords: Fat-tailed distributions; generalized lambda distribution; L-moment; risk management; stock returns

ABSTRAK

Andaian taburan yang digunakan dalam analisis pulangan pasaran saham adalah penting dalam pengurusan risiko. Isu utama dalam memodelkan pulangan saham adalah untuk mendapatkan anggaran yang tepat. Kajian ini membentangkan kaedah penyuaian inovatif iaitu kaedah dua peringkat (TS) dalam memodelkan pulangan saham harian. Pendekatannya ini dijalankan dengan cara mengenal pasti bentuk trend di dalam siri, kemudian melaksanakan anggaran L-momen pada parameter taburan generalisasi lambda (GLD). Prestasi model TS-GLD dinilai dengan menggunakan kaedah simulasi Monte Carlo dan data sebenar iaitu Indeks Komposit Kuala Lumpur Malaysia (KLCI) dari tahun 2001 hingga 2015. Berdasarkan ujian kebagusan k-sample Anderson darling, model dua peringkat (TS) GLD bagi parameter lokasi (GLD.1) menunjukkan prestasi yang lebih baik untuk semua kes yang dikaji. Model GLD.1 bermanfaat dalam pengurusan risiko dengan memberikan penyuaian taburan yang lebih baik.

Kata kunci: L-momen; pengurusan risiko; pulangan saham; taburan berekor tebal; taburan generalisasi lambda

INTRODUCTION

Stock market volatility is generally connected with risk measurement in finance. Economic crisis and natural disaster are phenomena that can drive extreme volatility on stock market series (Ben Slimane et al. 2013). Analysis of probability distribution is one approach to comprehend fundamental stochastic processes in these phenomena. Stock return modelling aims to yield the best distribution estimation that can explain the behaviour of stock returns, because accurate calculation is essential for risk management in financial investments.

The study on best fitting probability distribution performance in stock returns has been the subject of much systematic investigation (Gettinby et al. 2006; Hasan et al. 2012; Hussain & Li 2015; Longin 1996; Marsani & Shabri 2019; Marsani et al. 2017; Tolikas 2014, 2011, 2008; Tolikas & Gettinby 2009). The existing body of research frequently assumes that the stock return movement is stationary. However, this condition is erroneous in describing the real process due to the rising sign of the variability in the stochastic process of stock returns (Stărică & Granger 2005). Return movement follows non-stationary process (Marsani & Shabri 2019) as it possesses several common statistical characters such as volatility clustering (Dong & Wang 2013; Niu & Wang 2013a; Rizvi et al. 2014; Yu & Wang 2012), multifractality of volatility (Calvet & Fisher 2008; Fang & Wang 2012; Kantelhardt et al. 2002; Stošić et al. 2015; Suárez-García & Gómez-Ullate 2014), power law of logarithmic returns (Gabaix et al. 2003; Niu & Wang 2013b) and fat tails (Cont 2001; Ding et al. 1993; Mandelbrot 2013; Mantegna & Stanley 1995).

In stochastic processes, two underlying assumptions are usually used, namely stationary and non-stationary. The stationary process is an unconditional joint probability distribution of a series that does not change across time, which suggest that the parameters such as mean and variance are constant over time (Gagniuc 2017). Since ignoring the non-stationarity of the returns
series could provide inaccurate and bias risk estimates, therefore, the development of a model should provide benefits for determining risk (Acharya et al. 2012). The present study examines the behavior of the stock market in Malaysia by reflecting the dynamic progress of returns properties over time. The characteristics of non-stationary statistical features model are developed based on the weak assumption on time-invariant probability densities for location and scale parameters. The new technique is proposed based on GLD assumption, given that this distribution is competent to clarify the daily stock return behavior (Chalabi et al. 2009, 2012; Corlu et al. 2016; Corrado 2001; Marsani et al. 2017). The advantage of the two-stage fitting method over the traditional approach is twofold. Firstly, this new technique has successfully improved the accuracy of distribution fitting on extreme asset returns in the context of the non-stationarity setting. Secondly, this method is simple and straightforward as the calculations between the trend estimators and assumed probability distribution are independent. In this respect, the unique values of the probability distribution and trend estimators could be maintained without any interference in the statistical properties. The rest of this paper is arranged as follows: Next section describes the methods, consisting of GLD probability density function, non-stationary algorithm, and simulation design. Subsequent section deliberates the outcome for the best fitting model in simulation and real data application. Last section concludes the study.

MATERIALS AND METHODS

GENERALIZED LAMBDA DISTRIBUTION (GLD)

A significant advantage of four parameter-GLD measured by Karian and Dudewicz (2000) is the wide flexibility in assessing symmetrical and asymmetrical distribution’s shape, which makes it feasible to be applied in many univariate applications. The GLD can only be expressed in terms of inverse distribution function (Ramberg & Schmeiser 1974).

\[ F^{-1}(q) = \mu + \frac{q^\alpha - (1-q)^\alpha}{\alpha}, \quad \text{(1)} \]

where \( \mu \) is location parameter; \( \alpha \) is scale parameter; and \( k, h \) represent the shape parameters. The scale parameter \( \alpha \) is denoted in numerator form. The quantile for time-independent random variable, \( X \) is expressed as \( F^{-1}(q) \), and \( F \) denotes the non-exceedance probability. GLD is valid if and only if

\[ \frac{\alpha}{kq^{\alpha} + h(1-q)^{\alpha}} \geq 0 \quad \text{for all } q \in [0,1]. \quad \text{(2)} \]

Accordingly, the GLD quantiles can be written as:

\[ F_{\text{GLD}}(q) = \mu + \alpha \left( q^\alpha - (1-q)^\alpha \right), \quad \text{(3)} \]

whose non-exceedance probability is \( 0 \leq q \leq 1 \). The four parameters of GLD using L-moments expressions have been described by Asquith (2007).

TWO-STAGES METHOD

The two-stage model proposed in this study is used to tackle the complex sampling moments in stock volatility. This complexity can be addressed by patterning the covariates location \( (\mu) \) and scale \( (\alpha) \) parameter proportion to the functions of time-dependent. After this, the two-stage model becomes a non-stationary model namely GLD.1, GLD.2, GLD.11, and GLD.21, co-existing with the stationary model GLD.0 as the original model. All four different non-stationary models can be expressed as:

\[ g_{\text{GLD.1}}(\mu_{(j)} = \mu_0 + \mu_1 t, \, \alpha, \, k, \, h), \quad \text{linear in location} \quad \text{(4)} \]
\[ g_{\text{GLD.2}}(\mu_{(j)} = \mu_0 + \mu_1 t + \mu_2 t^2, \, \alpha, \, k, \, h), \quad \text{quadratic in location} \quad \text{(5)} \]
\[ g_{\text{GLD.11}}(\mu_{(j)} = \mu_0 + \mu_1 t, \, \ln \alpha_{(j)} = \alpha_0 + \alpha_1 t, \, k, \, h), \quad \text{linear in location and scale} \quad \text{(6)} \]
\[ g_{\text{GLD.21}}(\mu_{(j)} = \mu_0 + \mu_1 t + \mu_2 t^2, \, \ln \alpha_{(j)} = \alpha_0 + \alpha_1 t, \, k, \, h), \quad \text{quadratic in location and linear in scale} \quad \text{(7)} \]

where \( t \) is time; and \( k \) and \( h \) are the shape parameters, respectively. The natural log in scale \( \ln \alpha_{(j)} \) is operated to restrain a positive value in the scale parameter. The time-dependent assumptions in location and scale parameters are described in the next section.

PROCEDURE FOR TWO-STAGES ANALYSIS

First, express the non-stationary sequence \( Q_n \) as trend component \( tr_{(j)} \) and a residual time-dependent \( \varepsilon_{(j)} \) that diverts from the trend in the location parameter.

\[ Q_n^* = tr_{(j)} + \varepsilon_{(j)} \quad \text{(5)} \]

Second, fit the non-stationary (linear or quadratic model) into the trend component \( tr_{(j)} \) by estimating the location parameter.

\[ \mu_{(j)} = \mu_0 + \mu_1 t + \mu_2 t^2 + K + \mu_3 t^3 \quad \text{(6)} \]

Third, estimate the de-trended residual component \( \varepsilon_{(j)} \) given by

\[ \varepsilon_{(j)} = Q_n^* - tr_{(j)} \quad \text{(7)} \]

Fourth, express the transformed residual \( \varepsilon'_{(j)} \) component from the residual component \( \varepsilon_{(j)} \) as

\[ \varepsilon'_{(j)} = \left| \varepsilon_{(j)} - \mu \right| \quad \text{(8)} \]
where $\mathcal{F}$ represents the mean of the residual component.

Fifth, estimate the trend $t r_{(i)}$ component from the transformed residual component $e_{(i)}$ using a linear or quadratic model in the scale parameter $a_{(i)}$

$$a_{(i)} = \exp (a_0 + a_1 t + a_2 t^2 + K + a_f r)$$  \hspace{1cm} (9)

Sixth, express the stationary sequence $\hat{Q}_{(i)}$ by eliminating the trend (scale $a_{(i)}$) from residual component $\hat{e}_{(i)}$ as given (Cudenlik & Burn 2003).

$$\hat{Q}_{(i)} = \begin{cases} e_{(i)} - a_{(i)} & \forall e_{(i)} \geq \mathcal{F} \\
  e_{(i)} + a_{(i)} & \forall e_{(i)} < \mathcal{F} \\
  e_{(i)} - a_{(i)} & \forall e_{(i)} < \mathcal{F}
\end{cases}$$  \hspace{1cm} (10)

Seventh, apply the stationary series $\hat{Q}_{(i)}$ to estimate parameters $\mu, \alpha, k$ and $h$ and get the quantile for GLD.

Last, re-trend the calculated stationary quantiles by reversing the step taken to obtain non-stationary fitted quantile.

The proposed two-stage models are described as follows.

**GLD.1 MODEL**

The location parameter is modelled using the linear function of time $\mu_{(i)} = \mu_0 + \mu t$ where $\mu_0$ represents the mean intercept at period $t = 0$, while $\mu t$ denotes the mean shift for every period. Sen’s non-parametrical robust slope estimator is employed to estimate $\mu_0$, as described by Sen (1968):

$$\mu_t = \text{Med} \left( S_y \right) \hspace{1cm} S_y = \frac{X_j - X_i}{j - i} \forall i > j$$  \hspace{1cm} (11)

where $X_i$ and $X_j$ represent random variables of $X$ at times $i$ and $j$ individually, the mean $\mu_0$ at $t = 0$ is computed as: $\mu_0 = \bar{X} - \mu \bar{T}$ where both $\bar{X}$ and $\bar{T}$ signify an average for random variable and period. The GLD moment ($t$) ascribed by Asquith (2007) is defined as

$$\mu_{(i)} = \xi_{(i)} + \alpha \left( \frac{1}{k+1} - \frac{1}{h+1} \right)$$  \hspace{1cm} (12)

by substituting the location parameter $\mu_{(i)}$ into (15), which is then rearranged as

$$\xi_{(i)} = \mu_0 + \mu t - \alpha \left( \frac{1}{k+1} - \frac{1}{h+1} \right)$$  \hspace{1cm} (13)

Meanwhile, the location parameter $\xi_{(i)}$ at period $t = 0$ is expressed as

$$\xi_{(0)} = \left\{ \frac{1}{k+1} - \frac{1}{h+1} \right\} \hspace{1cm} (14)$$

and the shift in the location parameter at period $t = 1$ is expressed as

$$\xi_{(1)} = \left\{ \frac{1}{k+1} - \frac{1}{h+1} \right\} + \mu t$$  \hspace{1cm} (15)

**GLD.2 MODEL**

The GLD.2 location parameter is modelled using a quadratic function of time, as

$$\mu_{(i)} = \mu_0 + \mu t + \mu t^2$$  \hspace{1cm} (16)

while the GLD moment ($t$) is modelled using quadratic function as

$$\xi_{(i)} = \mu_0 + \mu t + \mu t^2 - \alpha \left( \frac{1}{k+1} - \frac{1}{h+1} \right)$$  \hspace{1cm} (17)

Accordingly, the location parameter $\xi_{(0)}$ at period $t = 0$ is written as

$$\xi_{(0)} = \left\{ \frac{1}{k+1} - \frac{1}{h+1} \right\}$$  \hspace{1cm} (18)

where $\mu_0$ and $\mu t$ represent the shift in the location parameter at periods $t = 1$ and 2, respectively.

**GLD.11 MODEL AND GLD.21 MODEL**

The location parameter is estimated as linear function in GLD.11, and as quadratic functions in GLD.21 model, as

$$\mu_{(i)} = \mu_0 + \mu t \hspace{1cm} \text{linear}$$  \hspace{1cm} (19)

$$\mu_{(i)} = \mu_0 + \mu t + \mu t^2 \hspace{1cm} \text{quadratic}$$

An additional analysis needs to be conducted on scale parameter to models GLD.11 and GLD.21, where the log scale parameters for both GLD.11 and GLD.21 are estimated by using linear function, as

$$\ln \alpha_{(i)} = \alpha_0 + \alpha t \hspace{1cm} \text{linear}$$  \hspace{1cm} (20)

$$\alpha_{(i)} = e^{\alpha_0 + \alpha t}$$
The second moment (\( \mu^2 \)) of the GLD by following (Asquith 2007) written as,

\[
\mu^2 = \left( \frac{k}{(k+2)(k+1)} \right) \left( \frac{h}{(h+2)(h+1)} \right) \tag{21}
\]

where,

\[ c = \frac{1}{\left( \frac{k}{(k+2)(k+1)} \right) \left( \frac{h}{(h+2)(h+1)} \right)} \tag{22} \]

By substituting (20) into (21),

\[ e^{\frac{k}{2} + \frac{1}{2}h} = c \sigma^2 \tag{23} \]

Therefore, the scale parameter is given as follows,

\[ \sigma = e^{\frac{k}{2} + \frac{1}{2}h} \tag{24} \]

SIMULATION LAYOUT

To accurately portray the real data comportment, the sampling properties of the non-stationarity had been investigated by applying GLD as known parent distribution function. According to Fournier et al. (2006) GLD parameters (0, 0.19, 0.14, 0.14) is appropriate to study a symmetric distribution which close to the standard Gaussian. Figure 1 illustrates the known parent GLD for different skewness level. The values of location and scale parameters applied in this study were \( \mu = 0 \) and \( \alpha = 0.08 \), respectively. Six different GLD shape parameters, namely kh1(k=0.05, h=0.23), kh2(k=0.08, h=0.20), kh3(k=0.11, h=0.17), kh4(k=0.17, h=0.11), kh5(k=0.20, h=0.08) and kh6(k=0.23, h=0.05) were used to represent different levels of non-stationary processes portrayed, using tail-fatness of the distribution. The combination of kh1, kh2, and kh3 was skewed to the left, while the combination of kh4, kh5, and kh6 was skewed to the right. L-moment estimation method was employed to estimate all the GLD parameters. In this study, the best fit GLD model was chosen from the model that could minimize the K-Sample Anderson Darling (k-ad) statistics (Scholz & Stephens 1987), expressed as

\[
AD_k = \sum_{i=m}^{n} \left( \frac{1}{n} \int_{-\infty}^{\infty} \left( \frac{\hat{F}_k(x) - H'(x)}{H'(x)} \right)^2 dH'(x) \right) \tag{25}
\]

where \( n \) is the sample size of \( x \), and \( H'(x) \) denotes the empirical distribution function of the pooled sample of all \( \hat{F}_k(x) \), where \( 0 \leq i \leq k-1 \). k-ad test statistic signifies the difference between experimental and pooled samples value. The studied GLD model could properly fit the data as the model could minimize the k-ad test statistics. The performance of the k-ad test statistics was assessed using average k-ad value, given by,

\[
\text{Average } AD_k = \frac{\sum_{i=1}^{N_{sim}} AD_{k,i}}{N_{sim}} \tag{26}
\]

where \( AD_k \) represents k-ad statistics and \( N_{sim} \) is the number of generated samples. This simulation was repeated for 5000 simulation runs with samples sizes, \( n = 100, 300 \) and 1000 to represent small, medium, and large samples.

RESULTS AND DISCUSSION

SIMULATION RESULTS

Table 1 presents the k-ad simulation results for traditional stationary and proposed model at different combination of the shape parameter (k and h) and sample size (n), respectively. The best fitting model should yield a value which minimizes k-ad statistics. Overall, the k-ad statistics for the stationary and non-stationary were close to each other. The results for different combinations of the shape parameters k and h of distribution tail fatness were fairly similar. However, even though the proportion of the tail fatness and sample size in the data had been increased, GLD.1 seemed to outperform the other models. In order to get a clear picture of performances comparison, the results as presented in Table 1 had been simplified in Table 2. As shown in Table 2, GLD.0, GLD.2, GLD.11, GLD.21 model produced higher values of k-ad compared to GLD.1, indicating that for all choices of the estimation of GLD shape parameters (k and h), GLD.1 model surpassed all the other models for best fitting performance.
Consequently, the GLD studied model was applied to Malaysian daily KLCI stock price. The data of 14 year-daily stock returns from 2001 until 2015 were obtained from Yahoo Finance and calculated using formula

\[ R_t = \ln \left( \frac{P_t}{P_{t-1}} \right) \]

where \( R_t \) is return index at \( t \) period, \( P_t \) is stock price index in term of \( t \), while \( P_{t-1} \) is stock price index at time of \( t-1 \). Note that, daily interval sample in this study had been divided evenly into five periods, assigned for every three years starting from 2001 until 2015 to avoid any external bias.

Table 3 presents the descriptive statistics for five different intervals of daily KLCI stock price return. Daily return series recorded the lowest at -4.812% and the highest 5.210%. The mean average for all intervals was positive, except for the first period which was -0.00629%. The standard deviation recorded the highest value of 1.1094 in the third period. Skewness to measure distribution symmetries was negative for all periods, except for the first period which expressed the tail inclined to the left. Jarque-Bera test (JB) was performed to see the normality of the data dispersions. Immense JB value and significant p-value indicated that the data series for all periods did not follow a normal distribution. The test for stationarity KPSS showed significant p-value at all periods, indicating the series was non-stationary. Also, the existence of the trend had been inspected using Mann-Kendal test, which reported that all the series had a positive trend.

| Model      | k-ad | pval  | k-ad | pval  | k-ad | pval  |
|------------|------|-------|------|-------|------|-------|
| GLD.0.kh1  | 0.24974 | 0.976 | 0.14335 | 1 | 0.50487 | 0.747   |
| GLD.1.kh1  | 0.24974 | 0.9785 | 0.19604 | 0.9895 | 0.50394 | 0.75 |
| GLD.2.kh1  | 0.25085 | 0.9685 | 0.19331 | 0.9925 | 0.13179 | 0.999 |
| GLD.11.kh1 | 0.33301 | 0.92 | 0.2015 | 0.9895 | 0.499 | 0.7515 |
| GLD.21.kh1 | 0.25085 | 0.9715 | 0.19374 | 0.9895 | 0.13223 | 0.999 |
| GLD.0.kh2  | 0.1457 | 0.999 | 0.25606 | 0.9715 | 0.58405 | 0.661 |
| GLD.1.kh2  | 0.19286 | 0.9935 | 0.26588 | 0.9695 | 0.57795 | 0.6705 |
| GLD.2.kh2  | 0.152 | 0.9995 | 0.26247 | 0.9605 | 0.54369 | 0.6915 |
| GLD.11.kh2 | 0.16002 | 0.998 | 0.26128 | 0.9625 | 0.58133 | 0.6605 |
| GLD.21.kh2 | 0.14629 | 0.9985 | 0.25994 | 0.9675 | 0.54818 | 0.7055 |
| GLD.0.kh3  | 0.16679 | 0.9965 | 0.40688 | 0.8605 | 0.69157 | 0.5465 |
| GLD.1.kh3  | 0.086812 | 1 | 0.39061 | 0.856 | 0.69528 | 0.5515 |
| GLD.2.kh3  | 0.092371 | 1 | 0.39907 | 0.8395 | 0.69425 | 0.5735 |
| GLD.11.kh3 | 0.15209 | 0.998 | 0.39978 | 0.859 | 0.68019 | 0.5775 |
| GLD.21.kh3 | 0.092371 | 1 | 0.39943 | 0.855 | 0.69772 | 0.5575 |
| GLD.0.kh4  | 0.16464 | 0.997 | 0.22135 | 0.988 | 0.47866 | 0.767 |
| GLD.1.kh4  | 0.16916 | 0.998 | 0.12537 | 0.999 | 0.46447 | 0.7845 |
| GLD.2.kh4  | 0.19009 | 0.9945 | 0.22013 | 0.9795 | 0.46792 | 0.7835 |
| GLD.11.kh4 | 0.16929 | 0.9965 | 0.12744 | 0.9995 | 0.47373 | 0.778 |
| GLD.21.kh4 | 0.17205 | 0.9975 | 0.10406 | 1 | 0.46606 | 0.772 |
| GLD.0.kh5  | 0.19596 | 0.992 | 0.16424 | 0.9965 | 0.49408 | 0.7465 |
| GLD.1.kh5  | 0.20025 | 0.9925 | 0.16056 | 0.9985 | 0.50005 | 0.745 |
| GLD.2.kh5  | 0.27406 | 0.964 | 0.15682 | 0.999 | 0.47357 | 0.77 |
| GLD.11.kh5 | 0.20083 | 0.991 | 0.15941 | 0.997 | 0.50038 | 0.7465 |
| GLD.21.kh5 | 0.2706 | 0.967 | 0.1575 | 0.9985 | 0.47293 | 0.7915 |
| GLD.0.kh6  | 0.10989 | 1 | 0.22808 | 0.9865 | 0.24005 | 0.979 |
| GLD.1.kh6  | 0.10331 | 1 | 0.17179 | 0.996 | 0.23697 | 0.9775 |
| GLD.2.kh6  | 0.28254 | 0.9545 | 0.18102 | 0.998 | 0.17027 | 0.996 |
| GLD.11.kh6 | 0.10331 | 1 | 0.16148 | 0.999 | 0.23739 | 0.976 |
| GLD.21.kh6 | 0.10386 | 1 | 0.18142 | 0.9965 | 0.99826 | 0.3615 |
### TABLE 2. Average k-ad of simulation data

| Model  | 100    | 300    | 1000   |
|--------|--------|--------|--------|
| GLD.0  | 0.17212| 0.23666| 0.49888|
| GLD.1  | **0.167022** | **0.218375** | **0.496443** |
| GLD.2  | 0.206985| 0.23547| 0.413582|
| GLD.11 | 0.186425| **0.218482** | **0.495337** |
| GLD.21 | 0.17267 | 0.216015| 0.552563|

*non-stationary model that is superior then stationary model marked in bold*

### TABLE 3. Descriptive statistics of daily KLCI stock price return

|                  | period.1 year 2015-2013 | period.2 year 2012-2010 | period.3 year 2009-2007 | period.4 year 2006-2004 | period.5 year 2003-2001 |
|------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| n                | 794                      | 771                      | 778                      | 776                      | 619                      |
| min(%)           | -2.8185                  | -2.54474                 | -4.59222                 | -2.49268                 | -4.81267                 |
| average(%)       | -0.00629                 | 0.036022                 | 0.038787                 | 0.039847                 | 0.050836                 |
| max(%)           | 5.210395                 | 2.631346                 | 4.704941                 | 2.19047                  | 4.370158                 |
| std.deviation(%) | 0.599784                 | 0.575057                 | 1.1094                   | 0.600644                 | 0.893152                 |
| variance(%)      | 0.003597                 | 0.003307                 | 0.012308                 | 0.003608                 | 0.007977                 |
| skewness         | 0.464491                 | -0.4268                  | -0.14628                 | -0.04544                 | -0.02312                 |
| kurtosis         | 9.089633                 | 2.446088                 | 2.128506                 | 1.531641                 | 3.835696                 |
| jarque.bera      | 2723.07                  | 211.8389                 | 146.6964                 | 74.38701                 | 371.5278                 |
| p.value          | 0                        | 0                        | 0                        | 1.11E-16                 | 0                        |
| KPSS             | 9.212016                 | 9.216348                 | 9.410213                 | 9.637996                 | 8.626569                 |
| p.value          | 0.01                     | 0.01                     | 0.01                     | 0.01                     | 0.01                     |
| Mann-Kendal      | 0.997371                 | 0.998331                 | 0.998434                 | 0.998634                 | 0.998443                 |
| p.value          | 0                        | 0                        | 0                        | 0                        | 0                        |
Figure 2 shows the CDF plot curve of each GLD model, which clarifies the upper and lower tail event at four-period returns. The CDF curves for GLD.0, GLD.1, GLD.2, GLD.11, and GLD.21 models overlapped,
indicating a very similar pattern by each of the models. All the GLD models seemed to adequately fit the extreme interval, specifically the upper and lower tail of the daily KLCI return. The CDF curve for GLD.2 in period three slightly deviated from the data. Thus, it was difficult to determine the best-fitted model based on the graphs.

Table 4 presents the k-ad goodness of fit test for each of the model. The k-ad goodness of fit test for all five periods gave an acceptable fit ($p$ - value $\geq 5\%$), indicating the empirical and the fitted data were homogenous. In general, it can be observed that the GLD.1 operated better than the other models, as GLD.1 provided the finest fitting at all periods, with evidence of the low k-ad values. The average value for overall cases period confirmed that non-stationary model GLD.1 is an excellent model to explain the behavior of daily return.

| Model | period.1 | pval | period.2 | pval | period.3 | pval | period.4 | pval | period.5 | pval | Average k-ad |
|-------|----------|------|----------|------|----------|------|----------|------|----------|------|--------------|
| GLD.0 | 0.4599   | 0.7876 | 0.4384   | 0.8096 | 0.2471   | 0.9725 | 0.3339   | 0.9110 | 0.3007   | 0.9383 | 0.3560       |
| GLD.1 | 0.4442   | 0.8037 | 0.4293   | 0.8189 | 0.2383   | 0.9769 | 0.3170   | 0.9253 | 0.2939   | 0.9434 | 0.3445       |
| GLD.2 | 0.5159   | 0.7302 | 0.3548   | 0.8922 | 1.6690   | 0.1406 | 0.3485   | 0.8980 | 0.2875   | 0.9480 | 0.6352       |
| GLD.11| 0.4669   | 0.7804 | 0.4379   | 0.8102 | 0.2392   | 0.9764 | 0.3214   | 0.9217 | 0.2935   | 0.9437 | 0.3517       |
| GLD.21| 0.5154   | 0.7306 | 0.3545   | 0.8923 | 0.2533   | 0.9691 | 0.3485   | 0.8979 | 0.2873   | 0.9480 | 0.3519       |

The non-stationary model that is superior then stationary model marked in bold

### ANALYSIS OF TAIL DISTRIBUTION

Next, Value at risk (VaR) analysis was conducted to determine the best GLD model that could explain the stock return behavior at the tail distribution. VaR can be a useful instrument to inquire about potential losses of information in term of probability, as investors are often concerned with the downside risk (Ab Razak & Ismail 2019). In this section, we consider analysis at the tail distribution on GLD.0, GLD.1, GLD.2, GLD.11, and GLD.21 to investigate which of the model give excellent estimation at the tail.

Table 5 shows the probability of getting a daily KLCI stock return within the intervals and the coefficient of $R^2$ for every GLD model. The studied intervals were $[\mu - (i+1)sd, \mu - (i)sd]$ signify lower and upper tails, respectively, where $\mu$ represents the mean, and $sd$ denotes the standard deviation calculated from the daily return. In this study, the actual probability returns ($obs$) had been compared with the fitted probability return for each model. The best GLD model was determined based on the ability of the GLD model in capturing risk at specified interval and the values of coefficient of $R^2$. The $R^2$ can be explained as $\text{cor}(\hat{x}_n, \bar{x}_n)^2$ where $\hat{x}_n$ and $\bar{x}_n$ are the actual and fitted ($n$th) sample returns. The model was adequate in explaining the entire daily return when $R^2$ was close to one.

Table 5 shows that the models performed well in capturing risk at all intervals, as the fitted and actual probability return displayed almost similar results. However, GLD.1 was better in performance compared to traditional GLD.0 model when the probability of the estimated price returns was nearer to the actual data. For example, in period 1, the probability of the actual data at interval Inr.3 (mu - 3sd, mu - 2sd) was 0.0252%, almost the same with GLD.1, which was 0.0202%. Also, GLD.1 give better prediction at interval Inr.5 and Inr.6 for each studied period by effectively capturing the extreme returns. The R-squared value supports this claim as the R-squared for GLD.1 model was higher compared to GLD.0 models.
| period | model | Inr.1 | Inr.2 | Inr.3 | Inr.4 | Inr.5 | Inr.6 | Inr.7 | Inr.8 | Inr.9 | Inr.10 | R-sq. |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|-------|
| 1 obs  |       | 0.0025 | 0.0076 | 0.0252 | 0.0806 | 0.1322 | 0.1184 | 0.0982 | 0.0202 | 0.0038 | 0.0000 |       |
|        | GLD.0 | 0.0025 | 0.0063 | 0.0189 | 0.0856 | 0.1310 | 0.1398 | 0.0869 | 0.0189 | 0.0050 | 0.0025 | 0.9912 |
|        | GLD.1 | 0.0025 | 0.0063 | 0.0202 | 0.0844 | 0.1322 | 0.1360 | 0.0869 | 0.0189 | 0.0050 | 0.0025 | 0.9919 |
|        | GLD.2 | 0.0025 | 0.0063 | 0.0189 | 0.0793 | 0.1322 | 0.1360 | 0.0932 | 0.0202 | 0.0050 | 0.0013 | 0.9909 |
|        | GLD.11| 0.0025 | 0.0063 | 0.0189 | 0.0856 | 0.1310 | 0.1411 | 0.0869 | 0.0189 | 0.0050 | 0.0025 | 0.9908 |
|        | GLD.21| 0.0025 | 0.0063 | 0.0189 | 0.0793 | 0.1322 | 0.1360 | 0.0932 | 0.0202 | 0.0050 | 0.0013 | 0.9909 |
| 2 obs  |       | 0.0026 | 0.0052 | 0.0311 | 0.0843 | 0.1154 | 0.1466 | 0.1077 | 0.0169 | 0.0039 | 0.0000 |       |
|        | GLD.0 | 0.0026 | 0.0078 | 0.0220 | 0.0895 | 0.1245 | 0.1582 | 0.0960 | 0.0169 | 0.0039 | 0.0013 | 0.9909 |
|        | GLD.1 | 0.0026 | 0.0078 | 0.0220 | 0.0895 | 0.1245 | 0.1530 | 0.0960 | 0.0169 | 0.0039 | 0.0013 | 0.9918 |
|        | GLD.2 | 0.0026 | 0.0065 | 0.0220 | 0.0856 | 0.1258 | 0.1530 | 0.1012 | 0.0182 | 0.0026 | 0.0013 | 0.9900 |
|        | GLD.11| 0.0026 | 0.0078 | 0.0220 | 0.0895 | 0.1245 | 0.1582 | 0.0960 | 0.0169 | 0.0039 | 0.0013 | 0.9909 |
|        | GLD.21| 0.0026 | 0.0065 | 0.0220 | 0.0856 | 0.1258 | 0.1530 | 0.1012 | 0.0182 | 0.0026 | 0.0013 | 0.9900 |
| 3 obs  |       | 0.0000 | 0.0116 | 0.0167 | 0.0964 | 0.1298 | 0.1465 | 0.0925 | 0.0257 | 0.0026 | 0.0013 |       |
|        | GLD.0 | 0.0013 | 0.0051 | 0.0219 | 0.0951 | 0.1375 | 0.1440 | 0.0964 | 0.0193 | 0.0051 | 0.0026 | 0.9940 |
|        | GLD.1 | 0.0013 | 0.0051 | 0.0219 | 0.0951 | 0.1375 | 0.1440 | 0.0964 | 0.0193 | 0.0051 | 0.0026 | 0.9946 |
|        | GLD.2 | 0.0000 | 0.0000 | 0.0411 | 0.0990 | 0.0925 | 0.0964 | 0.1144 | 0.0347 | 0.0000 | 0.0000 | 0.9694 |
|        | GLD.11| 0.0013 | 0.0051 | 0.0219 | 0.0951 | 0.1375 | 0.1440 | 0.0964 | 0.0193 | 0.0051 | 0.0026 | 0.9945 |
|        | GLD.21| 0.0013 | 0.0051 | 0.0219 | 0.0964 | 0.1362 | 0.1440 | 0.0964 | 0.0193 | 0.0051 | 0.0026 | 0.9945 |
| 4 obs  |       | 0.0000 | 0.0077 | 0.0219 | 0.1044 | 0.1456 | 0.1353 | 0.1057 | 0.0206 | 0.0077 | 0.0000 |       |
|        | GLD.0 | 0.0013 | 0.0052 | 0.0206 | 0.1044 | 0.1456 | 0.1418 | 0.1031 | 0.0219 | 0.0052 | 0.0013 | 0.9956 |
|        | GLD.1 | 0.0013 | 0.0052 | 0.0219 | 0.1031 | 0.1443 | 0.1405 | 0.1018 | 0.0232 | 0.0052 | 0.0013 | 0.9966 |
|        | GLD.2 | 0.0013 | 0.0064 | 0.0206 | 0.0979 | 0.1469 | 0.1366 | 0.1082 | 0.0232 | 0.0039 | 0.0013 | 0.9969 |
|        | GLD.11| 0.0013 | 0.0052 | 0.0219 | 0.1031 | 0.1443 | 0.1405 | 0.1018 | 0.0232 | 0.0052 | 0.0013 | 0.9963 |
|        | GLD.21| 0.0013 | 0.0064 | 0.0206 | 0.0979 | 0.1469 | 0.1366 | 0.1082 | 0.0232 | 0.0039 | 0.0013 | 0.9969 |
CONCLUSION
A non-stationary method of GLD models is proposed in this paper to interpret the appearances of significantly changing financial markets. This method transforms a non-stationary time series into stationary series by decomposing the trends in both mean and standard deviation of the original series. Manipulating the advantage of GLD, a new method is added into GLD to improve estimation accuracy by considering the non-stationarity in data series. The developed methods had been successfully implemented by carrying out simulation and real data analysis. The performance of this method had been investigated through Monte Carlo simulation. The simulation study was conducted on GLD1, GLD2, GLD11, and GLD21 models using six different shape parameters to portray different levels of extreme values. Malaysia daily KLCI returns had been used to represent the actual figures.

The findings of this paper are highlighted as follows: In the case of the non-stationarity in the data series, KPSS and Mann-Kendal tests had confirmed the existence of trends and non-stationarity in all periods from year 2001 to 2015. In the case of simulation data set, the performance of developed non-stationary GLD.1 model was superior to the stationary GLD.0, GLD.2, GLD.11 and GLD.21 models. GLD.1 produced lower k-ad on average. For the application of real data sets, the CDF curve had been used as a graphical tool to clarify the upper and lower tails risk event. Data analysis of tail distribution has explained the benefits of our proposed model in terms of tail behavior. The performance of the VaR using lower and upper tail interval analyses for each of the GLD model computed in this study is reasonably close to each other. Generally, the proposed model performance GLD.1 has been found better compared to the traditional model at the beginning part of lower and upper extreme distribution period precisely (mu-sd, mu-0.5*sd) and (mu+0.5*sd, mu+sd) interval, as the modeling technique emphasizes the center part of the distribution. Also, the $R^2$ of GLD.1 model was the highest in all cases, indicating that GLD.1 was the best in estimating the entire sample for all studied periods.

In general, on the basis of these results, it can be concluded that the proposed method by GLD1 model is the most accurate in explaining daily stock return in the environment of non-stationary. A simulation exercise has added further strength in this study. These findings provide new knowledge in the literature by improving the accuracy of the stock market projection as the ability of such risk measures is vital for investment and financial risk protection.

ACKNOWLEDGEMENTS
The authors would like to thank Universiti Sains Malaysia and the Ministry of Higher Education, Malaysia for the scholarship awarded.

REFERENCES
Ab Razak, R. & Ismail, N. 2019. Dependence modeling and portfolio risk estimation using GARCH-Copula approach. Sains Malaysia 48(7): 1547-1555.
Acharya, V., Engle, R. & Richardson, M. 2012. Capital shortfall: A new approach to ranking and regulating systemic risks. American Economic Review 102(3): 59-64.
Asquith, W.H. 2007. L-moments and TL-moments of the generalized lambda distribution. Computational Statistics & Data Analysis 51(9): 4484-4496.
Yu, Y. & Wang, J. 2012. Lattice-oriented percolation system applied to volatility behavior of stock market. *Journal of Applied Statistics* 39(4): 785-797.

Muhammad Fadhil Marsani* & Ani Shabri
Department of Mathematics
Universiti Teknologi Malaysia
81310, Johor Darul Takzim
Malaysia

Muhammad Fadhil Marsani*
School of Mathematical Sciences
Universiti Sains Malaysia
11800 Minden, Pulau Pinang
Malaysia

*Corresponding author; email: fadhilmarsani@gmail.com

Received: 24 June 2019
Accepted: 24 January 2020