Univariate Time Series Models For Fuel Price

M. Sulaiga Beevi , K. Senthamarai Kannan , S. Syed Ali Fathima

Abstract: any of the researcher, economist and a businessman currently exists interested in estimating the future of population, prices, national income etc. Accuracy of the future forecasts depends to a large extent on the success or failure. Hence the analysis of time series assumes just as great importance in the study of every single one economic problems.Here many methods are currently available in the literature to solve the problem of prediction. This study provides a detailed comparison of Fuzzy Time Series (FTS), Double Exponential Smoothing (DES) model and Auto Regressive Integrated Moving Average (ARIMA) model. Future values currently are forecasting using FTS, DES and ARIMA model. Forecasted values for Mean Square Error (MSE), Root Mean Square (RMSE), and Mean Absolute Percentage Error (MAPE) are calculated individually for all the three methods.

Index Terms : Fuzzy Time Series, Double Exponential Smoothing, ARIMA and Forecasting.

1 INTRODUCTION

Fuel price analysis to forecasting have an enormous worth for many professionals in the fields of finance. A significant part of the study currently exist the range of a suitable possibility form of the data. Analysis of Time series more essential in various fields of research, such as business, economics, medicine… etc. descriptive analysis, spectral analysis, forecasting, and explanative analysis carry out different goals. Nowadays several methods of forecasting first one is traditional (ARMA, ARIMA, SARIMA, Smoothing method, Holt-winters, etc), and other one is computational intelligence (neural networks, LSTM, etc). First one is finding the analysis used by statistical tool. Another one is computational tool.[1] R. G. Brown (1959)writing a book for Statistical Forecasting for Inventory Control, this book contains classical time series models. [2]L.Broze and G.Meland (1990) proposed to estimate some maximum likelihood of Exponential smoothing data. [3]C.Chat field and m.yar (1991) finding and prediction of intervals for multiplicative Holt's winter's methods. [4]Song and chissom (1993) proposed mamadani's method to compute fuzzy relation to evaluate the difference between the traditional time series along with fuzzy time series. [5]Song and chissom (1994) separate universe of discourse in to seven same lengths of intervals by max-min composition defuzzification of the output. Song and chissom take the applied fuzzy time series model to the forecast Enrollments University for Alabama, where that first-order invariant model is developed and that step-by Stepped procedure is provide further the researcher [6]chen (1996) used new model of max-min composition procedure with new less complexes followed by used max-min composition. Chen's algorithm has some drawbacks, specifically not considering the existence of duplication in addition to the absence of weighting which currently exists receiving lesser in the longer the observation. Some people followed by tried to improve chen's algorithm. According to[6] cheng, et al (1996), the differences in these methods currently are create the steps of fuzzy set pattern in addition to there currently are weights in all group of fuzzy relations.

- *M.SulaigaBeevi, ResearchScholar, Department of Statistics, ManonmaniamSundaranarUniversity,Abishekapatti,Tirunelveli-627012. ssulaiga@gmail.com*
- *Dr. K. Senthamarai Kannan, Senior Professor, Department of Statistics, Manonmaniam Sundaranar University, Abishekapatti Tirunelveli-627012.. senkannan2002@gmail.com*
- *Dr. S. Syed Ali Fathima, Assistant Professor, Department of Mathematics, Sadakathullah Appa College, Tirunelveli - 627011. syedalifathima2014@gmail.com*

[7]T. M. J. A. Cooray (2008) writing a book about Applied Time Series Analysis and Forecasting, in this book contain various forecasting methods.[8]Wang (2011) proposed a comparison method of forecasting by fuzzy time series in addition to arima model. For this object over here purpose, they have used the data related to the taiwan export data. [9]Yun – sheng hsu et al., (2011) discussed usual methods of forecasting in addition to heuristics model of forecasting in addition to then compare the two methods. [10]Niyimbanira(2015) examine between fuel price , Exchange rate ,New vehicle sales in south Africa.He compare two models one is Mathematical (cointegration) other one is econometric model (VAR) model. Econometric model produced a better solution. [11]Hansun (2016) proposed a new approach of brown's double exponential smoothing in time series analysis. The new approach merge the calculation of weighting factor in weighted moving average in addition to execute the results with brown's double exponential smoothing method. The proposed method tested on jakarta stock exchange composite index data. [12]Edward and Manoj (2016) proposed ARIMA two sector analysis model. They were used stock price of Automobile sector and partitioned in to two sector first one is 70% observations used to model development remaining 30% confirmation accuracy of the model developed. [13]Tularam and Saeed (2016) comparing different time series for oil prices one is ARIMA and other one is Holt-Winter Exponential smoothing. They were used Mat lab (R2014a) software used for forecasting. The ARIMA (2, 1, 2) provided more accurate than other models. [14] Tsai M-C, Cheng C-H et al., (2018) proposed a hybrid time-series model based on a feature selection method for forecasting the leading industry stock prices. They proposed stepwise regression is first adopted, and multivariate adaptive regression and kernel ridge regression used to select the key features. And then construct the forecasting model by a genetic algorithm to optimize the parameters of support vector regression and then evaluate the forecasting performance of the proposed models, this study collects five leading enterprise datasets in different industries from 2003 to 2012. K.Senthamarai Kannan et al., disgusted a method for Comparison of Fuzzy Time Series and ARIMA model for this purpose they have used the data set related to the daily petrol price.

2. MATERIALS AND METHODS

2.1 DATA SOURCE

The data on daily Prices for petrol have been collected from the web sites https://www.iocl.com/home.aspx for the period from 22st November, 2018 to 11th November, 2019 All the Procedures and Methods adopted in[1][4][5][6][7][15].

2.2 ARIMA MODEL

G.E.P.Box and G.M.Jenkins (1970) developed the procedure for autoregressive Integrated moving average (ARIMA).The term integrated, used because the differencing process can be reversed to obtain original series. They recommend differencing non-stationary series one or more times to obtain stationary. When the explanatory variables in a regression model are time-lagged values of the forecast variable, then the model is called an autoregressive (AR) model. The general form of an autoregressive model of order p denoted as AR (p), is

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_p Y_{t-p} + e_t \quad (1)$$

Where et is the error or residual term and p is an integer denoting the order of the in which the observations in the time series are correlated. When a time series is analyzed using its dependence relation with the past error terms, a moving average (MA) model is applied. The general form of the MA(q) model of order q is

$$Y_t = \varphi_0 + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \cdots + \varphi_q e_{t-q} + e_t \quad (2)$$

Autoregressive (Ar) model currently am able to exists effective coupled with moving average (ma) model to form a general in addition to useful class of time series models called autoregressive moving average ARMA (p, q) models. However, they currently are able to only exist used when the time series currently exists stationary. When a time series currently exists studied based on the dependence relationship among the time lagged values of the forecast variable in addition to the past error terms, an autoregressive integrated moving average (ARIMA) model currently exists more appropriate. It currently is able to exist used when the time series currently exists non-stationary. The general form of the ARIMA (p, d, q) model currently exists

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} \cdots + b_p Y_{t-p} + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \cdots + \varphi_q e_{t-q} + e_t \quad (3)$$

p, d , q represent respectively the order of an autoregressive part, the degree of differencing involved in the stationary time series which currently exists usually 0,1 or at most 2 in addition to the order of the moving average part. An ARIMA model currently is able to exist obtained by estimating its parameters. The values of p in addition to q currently are able to exist determined from the patterns in the plotting of the values of ACF in addition to PACF. The spikes falling above the time axis currently are used to estimate the value of p. the spikes falling below the time axis currently are used to estimate the value of q. For an AR (p) model, the spikes of ACF decay exponentially or there currently exists a sine wave pattern in addition to the spikes of PACF currently are close to zero beyond the time lag q whereas the spikes of PACF decay exponentially or there currently exists a sine wave pattern. [7]

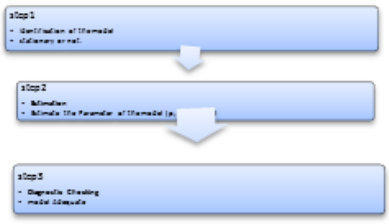


Figure (1) *Selection Procedure for ARIMA*

2.3 EXPONENTIAL SMOOTHING METHOD

Time series data occur frequently in many real world applications. One of the major vital parts in analyzing a time series data currently exist the selection of correct statistical model for the data. Due to the fact that it helps in prediction, hypothesis testing in addition to rule discovery. One of the most efficient forecasting methods now exist the exponential smoothing techniques. Also, it currently exists changed well to use effectively for time series with seasonal patterns. It currently exists also easy to adjust for past errors in addition to easy to prepare to follow on forecasts, ideal for situations at whatever place many forecasts must exists prepared, and numerous unusual ways of forms presently used depending on the presence of trend or cyclical variations. In other words, an exponential smoothing at this time exists an averaging technique that object over there uses unequal intervals; however, the intervals applied to past observations decline in an exponential manner. An exponential smoothing over an already smoothed time series currently exists called double exponential smoothing. In some other cases, it might exist necessary to extend it even to a triple exponential smoothing.

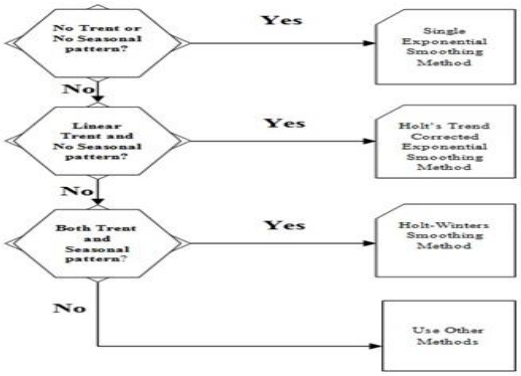


Figure: 2 *Selection Procedure for Exponential Smoothing method*

An exponential smoothing previously smoothed time series currently exists called double exponential smoothing. In some other cases, it currently exists necessary to extend it up to a triple-exponential smoothing. While SES requires stationary conditions, the DES currently evaluate in linear trends in addition to triple exponential smoothing presently handle both trend in addition to seasonal pattern in time series data. Figure (2) shows the selection procedure of different ES method now exists defined.

2.3.1 DOUBLE EXPONENTIAL SMOOTHING

If time series contains linear trend with linear regression equation $\hat{x}_t = a + bt$ where the estimates a and b represent the intercept

and slope of the model and t = 1, 2, 3. . . n, the correct procedure would be Double exponential smoothing. The argument and techniques for the double exponential smoothing method are similar in nature to that of single exponential smoothing. A form of smoothing equation:
Next period forecast = weight * (present period observation) + 1 – weight) * present period forecast
If we let $\hat{x}_t = a + b_t(\mathrm{T})$
Represent the updated forecast, then

at = $2S_t^{(1)} - S_t^{(2)}$ (updated intercept) (4)

$$b_t = \frac{\alpha}{1-\alpha}\left(S_t^{(1)} - S_t^{(2)}\right)$$ (Updated slope) (5)

T = number of the time period ahead. The $S_t^{(1)}$ and $S_t^{(2)}$ are the single and double smoothing statistics found by applying the smoothing equation

$$S_t^{(1)} = \alpha x_t + (1-\alpha)S_{t-1}^{(1)} \quad (6)$$

and

$$S_t^{(2)} = \alpha S_t + (1-\alpha)S_{t-1}^{(2)} \quad (7)$$

Once the single and double smoothed statistics are computed for a time period, the values may be substituted into the updating formula for the intercept and slope. To start the double smoothing process, initial values of the smoothed estimates must be obtained; this can be done by substituting the values for the estimated intercept and the slope from the linear regression analysis into the following equations

$$S_0 = a - \left[\frac{1-\alpha}{\alpha}\right](b) \quad (8)$$

And

$$S_0^{(2)} = a - 2\left[\frac{1-\alpha}{\alpha}\right](b) \quad (9)$$ [7]

2.4 Fuzzy Time Series Model

In [4][5][6] [15]The fuzzy logic, proposed by Lotfi zadeh (1965), state that object over there a certain element may belong in addition to simultaneously do not belong to the same set at certain levels, such that object over there the membership currently exists a value in the interval [0, 1]. The main objective in analyzing time series currently exists to know, understand in addition to estimate changes in the phenomena in the hope of more properly anticipating the course of future events. A good considerate of the mechanism generating the series may also help us to control the phenomena involved in the generating mechanism in addition to in this object over here manner control the future behavior of the process. Fuzzy Time series is assumed to be a fuzzy variable along with associated member ship function Song and Chissom (1993) have proposed a procedure for solving fuzzy time series model described as follows.
The time-variant and time-invariant fuzzy time series definitions are given below.

Definition 1:

Let Y(t), a subset of real numbers, be the universe of discourse on which fuzzy sets f(t) are defined. If F(t) is a collection of f1, f2,…. then F(t) is called a fuzzy time series defined on Y(t).

Definition 2:

Suppose F(t) is implied by F(t-1) only, that is F(t-1) F(t). Then this relation can be expressed as
F(t) = F(t-1) * R(t-1), where R(t-1) is the fuzzy relationship between F(t-1) and F(t) and is called the first order model of F(t). Here, "*" is maximum–minimum composition operator. The relation R is called first-order model of F(t).

Definition3:

Suppose F(t) is calculated by F(t-1) only, and F(t) = F(t-1) R (t-1, t). For any t, if R (t-1, t) is independent of t, then F(t) is considered a time-invariant fuzzy time series. Otherwise, F(t) is time-variant.

The analysis of time series of observation consists of the followings:

Membership function μ, we can define a fuzzy set F on a universe of discourse U as μF(x): U→ [0,1]which is nothing but a mapping from the universe of discourse U into the unit interval [0,1] and μF(x) represent the scope to which x containing F. The notion of membership functions allows any element within the universe of discourse to have partial membership to a specific fuzzy set and also to have partial membership to other fuzzy sets Song and Chissom applied time invariant and time variant models to forecast the enrollment at the University of Alabama.
The time variant and invariant model includes the following steps:

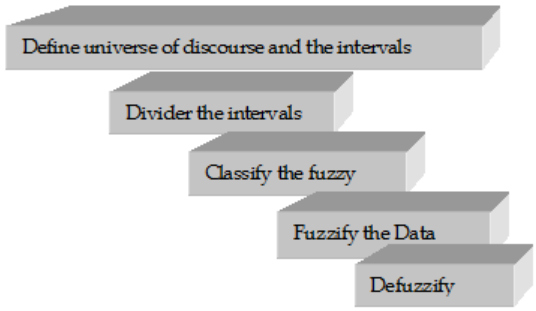


***Figure: 3** Selection PROCEDURES for Fuzzy time series Define universe of Discourse and the intervals*

Universe of Discourse S based on ranges available data in the historical time series data, by the following rule:

S=[Dmax+D1,Dmin-D2] (10)

Where, D1 and D2 are two suitable positive numbers.

Divider the Intervals

Divider the universe of Discourse into equal length of intervals s1, s2……sn. The number of intervals will be in the unity with the number of linguistic variable (Fuzzy sets) A1, A2 …A n to be consider. If we take into relation the fact that forecasting with fuzzy time series exhibits the least average error, it's required to find the middle point of the intervals

Classify the fuzzy set

Classify the fuzzy set into linguistic variablesA1=very poor, A2=poor, A3=very good, A4= good and so on. To every linguistic value here corresponds a fuzzy variable which, according to a definite rule is assigned beside a corresponding fuzzy set formative the significance of this variable. If the value of variable U in method is accepted as the midpoint of the corresponding interval of fuzzy set defined as,
Mi= (μMi (si)/si) si ∈S μMi (si)∈ [0, 1] is a fuzzy set. (11)

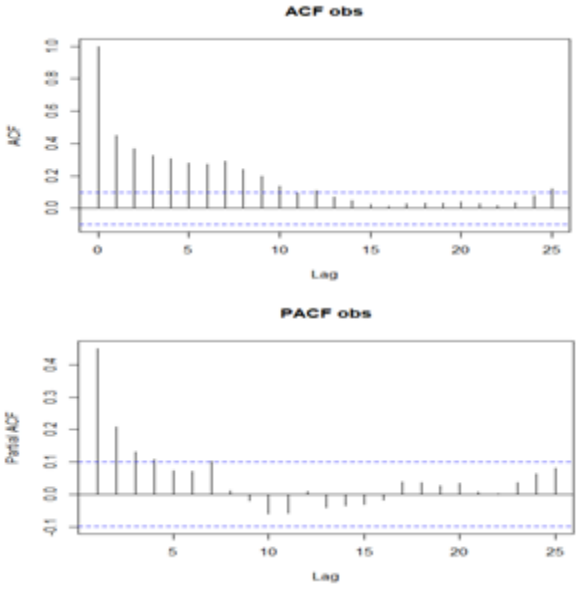
Fuzzify the data
Fuzzify the past data and create the fuzzy logical relationships by the following rule:
If Ai is the fuzzy price of the day n and Aj the fuzzy price of the day n + 1, then the fuzzy logical relation is denoted as Ai→ Aj, here Ai is called current state and Aj is called next state.

Defuzzify
Fuzzy membership functions to crisp formats this process called defuzzification. Many methods have been defuzzifying fuzzy output function Max membership principle, centroid method, weighted average method, mean max membership, center of sums and so on. (Some times this step is not necessary)

First finding mean of the data
Find the sum of the data values and divide the sum by the number of data values Find the absolute value of the difference each data value and the mean Find the sum of absolute value of the difference

3. RESULTS AND DISCUSSIONS

By[15] IDENTIFICATION OF THE MODEL
ACF and PACF of correlogram Figure (4) determine the variable is stationary or not.



FIGURE(4) *ACF AND PACF FOR PETROL PRICE*

Estimation For the selected data values has overcome the ARIMA (1, 1, 1) Model. Actual variable do not have the stationary conditions so First order difference is calculate and it satisfies the stationary, so we have used d=1. Then the p, d and q values (without changing the d=1) have found the different BIC values. Among the BIC values ARIMA (1, 1, 1) having the minimum value. Based on ARIMA (1, 1, 1) values are forecasted. Our data has trend and no seasonality so we choose Holt's trend DES method. Different α and β values (0 to 1) were obtained for forecasted. Comparing with other α and β values forecasted value is much closer to the actual value α= 0.9,β = 0.2 values. The universe of discourse S, using this formula S= [Dmax+D1, D min -D2] where D1 and D2 are two suitable positive numbers. So we get the range S= [79-70.5], partition the universe of discourse into ten equal length of intervals s1,s2,…..s10, where S1=(70.5-71.35) S2=(71.35-72.2) S3(72.2-73.05) S4=(73.05-73.09) S5=(73.09-74.75) S6=(74.5-75.6) S7=(75.6-76.45) S8=(76.45-77.3) S9=(77.3-78.15) S10=(78.15-79).Let's say Fuzzy no M1 is the interval of S1, Fuzzy no M2 is the interval of S2Fuzzy set are defined on the universal set U. "the deviation of petrol price" is a linguistic variable that assumes the following linguistic values:M1=(enormously low level), M2 =(very very low level), M3 =(very low level), M4=(low level), M5=(normal low level), M6= (normal high level), M7= (high level), M8=(very high level),M9=(very very high level), M10=(enormously high level), All the linguistic value has a corresponding fuzzy variable which, according to a certain rule is assigned aligned with a corresponding fuzzy set formative the meaning of this variable. If the value of variable U in method is established as the middle point of the corresponding interval of fuzzy set is as follows adopted from Song and Chissom (1993), Mi(i=1...10) will be defined as follows: Mi=(μMi (si)/si)si ∈S μMi(ui)∈[0,1]is a fuzzy set. The accuracy of the fitted model is measure by MAE, MSE, RMSE, which is a measure of accuracy of the fitted model.
Based on the above three models the predicted values are compared numerically and graphically. For all measures, smaller values generally indicate a better fitting model for forecasting errors shows in table (1).

$$MAE = \sum_{t=1}^{n} \frac{(Forecasted\ Value - Actual\ Value)}{n} \quad (12)$$

$$MSE = \sum_{t=1}^{n} \frac{(Forecasted\ Value - Actual\ Value)^2}{n} \quad (13)$$

$$RMSE = \sqrt{MSE} \quad (14)$$

We evaluate the forecasted values of DES and ARIMA (1, 1, 1), and Fuzzy Time series. Fuzzy Time series values are very closer to actual values.

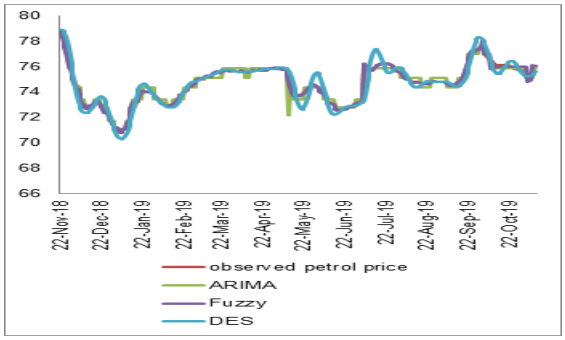


FIGURE 5 *SHOWS THE LINE CHART OF ACTUAL AND FORECASTED VALUES,*

which shows that Actual values are closely associated with the Fuzzy time series model.

The following Table 1 shows, MSE, MAE, and RMSE.

TABLE 1

FORECASTED ERROR VALUES

Errors	*DES*	*ARIMA(1,1,1)*	*Fuzzy time series*
MSE	*0.0333*	*0.0310*	*0.0073*
MAE	*0.0956*	*0.0843*	*0.0798*
RMSE	*0.1826*	*0.1761*	*0.0856*

4.CONCLUSIONS

Forecasting of daily petrol price prediction is a very deadly job. A lot of Forecasting techniques have tried and used for forecasting over the years. In this paper forecasted daily petrol price using DES and ARIMA (1, 1, 1) and Fuzzy Time series .The ARIMA models present some difficulties in estimating and validating the model because of following some condition (stationary, at least fifty observation .. etc) DES is very flexible to use the non-linear model we can adjust the α and β values, we can reduce the error values. Classical time series model depend large quantity of past data but Fuzzy time series model, small amount of data we can calculate the error values such as MSE, RMSE, and MAE are compared for three models. In this case, Fuzzy time series model error values are less than the other two models. It concludes that fuzzy time series model is more appropriate for Forecasting of daily petrol price prediction.

5 REFERENCES

[1] R. G. Brown, Statistical Forecasting for Inventory Control, New York: McGraw-Hill ,1959.

[2] L. Broze and G. Mélard. Exponential smoothing: Estimation by maximum likelihood, Journal of Forecasting, Vol 9, pp 445-455, 1990.

[3] C. Chatfield, and M. Yar. Prediction intervals for multiplicative Holt-Winters, International Journal of Forecasting, vol 7, pp 31-37, 1991.

[4] Song, Q., and B. S. Chissom ,Fuzzy time series and its models,Fuzzy Sets andSystems, 54(3), 269-277(1993).

[5] Song, Q., and B. S. Chissom , Forecasting enrollments with fuzzy time series- Part II,Fuzzy Sets and Systems, 62(I), 1-8(1994).

[6] Chen, S. M., Forecasting enrollments based on fuzzy time series,Fuzzy Sets and Systems, 81(3), 311 – 319(1996).

[7] T. M. J. A. Cooray, Applied Time Series Analysis and Forecasting, Narosa Publishing House, 2008.

[8] Wang, C. CA Comparison study between fuzzy time series model and ARIMA Model for Forecasting Taiwan Export, Expert Systems with Applications, 38, 9296–9304,2011.

[9] Yun-sheng Hsu et.al ,A Comparison of ARIMA fore casting and heuristic modeling, Applied Financial Economics, 21, 1095-1102,2011.

[10] Niyimbanira .F.,Fuel price and exchange rate dynamics in south Africa:A Time series Analysis ,Corporate Ownership and Control ,12(4)2015.

[11] Hansun S. A New Approach of Brown's Double Exponential Smoothing Method in Time Series Analysis, Balkan Journal of Electrical &Computer Engineering,4(2)2016.

[12] Edward.A. and Manoj.J.,One Forecastmodel using ARIMA for stock price of Automobile sector International Journal of Research in Marketing.6(4)2016.

[13] Tularam, G.A. and Saeed, T. Oil-Price Forecasting Based on Various Univariate Time-Series Models. American Journal of Operations Research, 6, 226-235, 2016.

[14] Tsai M-C, Cheng C-H, Tsai M-I, Shiu H-Y, Forecastingleading industry stock prices based on a hybrid time-series forecast model. PLoS 13(12),2018.

[15] K. Senthamarai Kannan , M. Sulaiga Beevi , and S. Syed Ali Fathima ,Comparison of Fuzzy Time Series and ARIMA Model,INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 08, AUGUST 2019