Analysis of Google Stock Prices from 2020 to 2023 using the GARCH Method

Khairunnisa Rifda
Sumatera Institute of Technology, Lampung, INDONESIA

M. Farhan Athaulloh
Sumatera Institute of Technology, Lampung, INDONESIA

Husni Na'fa Mubarok
Sumatera Institute of Technology, Lampung, INDONESIA

Sergii Sharov
Dmytro Motornyi Tavria State Agrotechnological University, UKRAINE

Berliyana Kesuma Hati
Sumatera Institute of Technology, Lampung, INDONESIA

Luluk Muthoharoh
Sumatera Institute of Technology, Lampung, INDONESIA

Mika Alvionita
Sumatera Institute of Technology, Lampung, INDONESIA

Article Info	Abstract
	This research focuses on Google's share price movements, considering their significant impact on the financial market, using Google's share price data from 2020 to 2023. The aim is to analyze error variance and forecast and provide valuable information to stockbrokers and investors. The ARMA model has shortcomings in dealing with volatility, so the GARCH model is used to overcome it. Research methods include financial data analysis, preprocessing, and modeling with GARCH. The rolling forecast method describes changes in price patterns over time. Evaluation using MAPE validates the prediction accuracy of the ARIMA model. The best model chosen with the most negligible AIC value criteria was the ARIMA(3,0,2)GARCH(1,1) model. The forecasting results show accurate stock price predictions with an average MAPE value of 20.7%. This research provides an essential basis for brokers and investors in making investment decisions based on a deep understanding of the dynamics of Google's share price movements in the above time frame.

INTRODUCTION

Given the company's rapid growth and significant impact on investors, Google's stock price movements have become a primary focus in financial markets [1], [2]. This phenomenon provides an essential impetus for conducting comprehensive research regarding the dynamics of stock prices in Indonesia [3], [4]. This study draws on Google stock price data from 2020 to 2023 to perform error variance analysis and forecast the provide valuable information to stock brokers and investors. Despite the coronavirus crisis, sales in 2020 increased by more than 12.7% to $182 billion. This development is mainly due to the strong growth of Google Cloud. In 2020, sales increased by 46% to $13.05 billion. However, Alphabet invests and innovates independently,

and tries to conquer more market share [5] quickly. Based on data from Kompas Daily, summarized by Yahoo Finance, Google's share price has increased by an average of 24.8 percent per year since its IPO in 2004. In 2020, Google's parent company became the third most valuable United States company with a more than 1 trillion dollars valuation. US IDR 14,422 trillion [6]. The Tokopedia shares held by Google are worth 1.1 million US dollars or IDR 16.7 billion, while the shares owned by Anderson are worth IDR 33.4 billion [7].

There is much research to predict stock prices or time series data using various approaches, including time series data [8]–[10], which can use the ARIMA [11] method or machine learning methods such as ANN [12], SVR [13], SVM [14], etc. Research by Lala Nur

Faiza and Dina Agustina uses Machine Learning Applications in Predicting Jakarta Islamic Index (JII) Stock Prices Using the Support Vector Regression Method. The conclusion of this study indicates that SVR can be implemented as a method for predicting stock prices with the smallest RMSE values for ANTM, BRIS, and BRPT shares of 0.0004 [15]. Another research conducted by Vandara Vavras Setia in 2017 using the Dividend Discount Models (DDM) method showed that all companies in 2015 were in the overlayed category [16]. Other research uses ARIMA and GARCH methods to analyze the stocks of shares listed on the LQ 45 Index. The GARCH model has the characteristic that the volatility response to a shock is the same, whether a positive or negative shock [17]. The reason for using a combination of the two models is that the ARIMA model alone cannot handle significant volatility and non-linearity data. The ARIMA(1,1,1)-GARCH (1,1) model provides the best stock price predictions for the two selected stocks [18].

From the research above, appropriate and effective investment decisions can be made based on a deeper understanding of the behavior of Google's share price movements. The increasing dynamics of share price movements in the capital market have attracted significant interest from potential investors. A deep understanding needs to be developed to provide comprehensive insight into stock prices. Therefore, this research uses Google stock data to carry out error variance modeling and forecast, which aims to provide valuable consideration and reference material for stock brokers and investors in making investment decisions. Time series data can be modeled with the Autoregressive Moving Average (ARMA) model. ARMA model requires homoscedasticity assumptions to be met. However, the financial data, such as the stock data, have high fluctuations, so the variants are heteroscedastic. One time series model to resolve such issues is Autoregressive Conditional Heteroscedasticity (ARCH). Bollerslev generalized the ARCH model into a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model in 1986. GARCH is used to overcome high order in ARCH models [19]. So, this research involves the GARCH model to handle volatility and heteroscedasticity in time series data.

METHOD

Materials
We chose Google financial data 2 from January 2020 to 31 October 2023. The data was taken from Yahoo Finance, an open-source financial data provider. The data set in the data contains open, close, high, low, volume, and adjusted Google prices. We collected about three years of Google stock price history. This research aims to find interesting facts about Google's stock price and develop a suitable model for forecasting.

Preprocessing Data
Loading Data and Initial Check
After collecting the data, we loaded the data in the next stage using the getSymbols function from the quantmod package. Then we looked at the data structure using head () and checked for missing values with colSums(is.na(GOOGL)). This process is important to ensure the data quality used in later analysis.

Data Exploratory Transformation and Analysis
After loading the data and initially checking it, we transform it. We create a data frame that contains the original price and return data and perform a log transformation process on the return data. Stock prices are created into a time series that allows visualization of price changes over time, while returns are calculated as the log of changes in stock prices. After carrying out data transformation, we explored the characteristics of the data by displaying time series plots for visualization of stock prices and return plots. With this process, we can identify and understand trends, volatility, and stationarity with the ADF test.

Modeling
Selection and Evaluation of GARCH models
In data modeling, we carry out data selection by preparing GARCH models with various configurations (ARMA(p,q) GARCH(p,q)). Then, we fit the GARCH model using the ugarchfit function from the rugarch package and perform model selection and testing based on information criteria such as AIC. Next, we conducted a persistence test on the selected model and evaluated the model convergence. Testing, comparison, and evaluation are conducted to find the best model based on the data.

Forecasting and Performance Evaluation

We also carry out forecasting tests to make predictions using the selected GARCH model. Then, it displays the forecast results using graphs to visualize the differences between the forecast and actual data. Furthermore, finally, we use an evaluation matrix such as MAPE (Mean Absolute Percentage Error) to evaluate forecast performance.
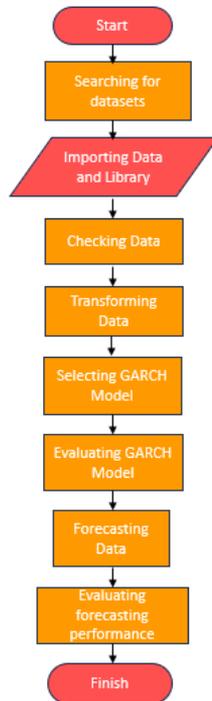

Figure 1. Material and Methods Process Flowchart

RESULTS AND DISCUSSION
Determination of ARIMA (p,d,q) and GARCH (p,d) Model Coefficients

Determining the coefficients of the ARIMA(p,d,q)GARCH(p,d) model can be done using several methods.

Iteration Method

This method is the most commonly used [20]. The procedure is as follows: 1) Determine the initial values for the ARIMA model coefficients (p,d,q) and GARCH (p,d); 2) Calculate the predicted value of data variance using the GARCH (p,d) model; 3) Calculate the residual value of the data using the ARIMA (p,d,q) model; 4) Update the coefficient values of the ARIMA (p,d,q) model using residual data value; 5) Repeat steps 2-4 until the ARIMA model coefficient values (p, d, q) and GARCH (p,d) do not change.

Maximum Likelihood Estimation Method

This method is used to determine model coefficients ARIMA (p,d,q) and GARCH (p,d), which maximize the likelihood function of the data [21].

The procedure is as follows: 1) Define the likelihood function of the data; 2) Calculate the likelihood value of the data function for each combination of values ARIMA (p,d,q) and GARCH (p,d) model coefficients; 3) Select a combination of ARIMA model coefficient values (p, d, q) and GARCH (p,d) that maximizes the likelihood function.

Bayesian Estimation Method

This method is used to determine model coefficients ARIMA (p,d,q) and GARCH (p,d) are based on a priori information [22], [23]. The procedure

are as follows: 1) Define a priori the probability distribution for the coefficients ARIMA (p,d,q) and GARCH(p,d) models; 2) Calculate the posterior value of the probability distribution for the coefficients ARIMA (p,d,q) and GARCH (p,d) models; 3) Select a combination of ARIMA model coefficient values (p, d, q) and GARCH (p,d) with the highest posterior probability.

Determining the ARIMA (p,d,q) and GARCH (p,d) model
Data complexity

More complex data requires more complex methods to determine model coefficients.

Data availability

The amount of data available also influences the choice of method. More straightforward methods are better if the amount of data available is limited.

Selection of the Best Model

Selection of the best model using AIC (Akaike Information Criterion) is based on the assumption that the best model is the model that has the smallest AIC value. The AIC value is calculated using the following formula:

$$AIC = 2k - 2\ln(L)$$

Where: 1) k is the number of parameters in the model; 2) L is the likelihood function of the data.

AIC provides a trade-off between model complexity and capability model to explain data. More complex models will have a smaller AIC value and more parameters than is known. These unknown parameters can lead to overfitting, that is, a model that fits the training data too well and cannot be used to predict new data. Choosing the best model using AIC can be done as follows: 1) Define several candidate models with different parameters; 2) Calculate the AIC value for each candidate model; 3) Choose the model with the smallest AIC value. Here are some candidate models we can try (see table 1).

Table 1. Determination of ARIMA (p,d,q) and GARCH (p,d) coefficients

Model	p	d	q	p	d
ARIMA(0,0,0)GARCH(1,1)	0	0	0	1	1
ARIMA(1,0,1)GARCH(1,1)	1	0	1	1	1
ARIMA(2,0,2)GARCH(1,1)	2	0	2	1	1
ARIMA(1,0,2)GARCH(1,1)	1	0	2	1	1
ARIMA(2,0,1)GARCH(1,1)	2	0	1	1	1
ARIMA(3,0,1)GARCH(1,1)	3	0	1	1	1
ARIMA(3,0,2)GARCH(1,1)	3	0	2	1	1
ARIMA(1,0,3)GARCH(1,1)	1	0	3	1	1

We get the following results after calculating the AIC value for each candidate model (see table 2).

Table 2. Model with AIC value

Model	AIC
ARIMA(0,0,0)GARCH(1,1)	-5.0702
ARIMA(1,0,1)GARCH(1,1)	-5.0724
ARIMA(2,0,2)GARCH(1,1)	-5.0664
ARIMA(1,0,2)GARCH(1,1)	-5.0703
ARIMA(2,0,1)GARCH(1,1	-5.0702
ARIMA(3,0,1)GARCH(1,1)	-5.0710
ARIMA(3,0,2)GARCH(1,1)	-5.0753
ARIMA(1,0,3)GARCH(1,1)	-5.0709

Based on these results, the ARIMA (3,0,2) GARCH (1,1) model has the smallest AIC value -5.0753. Therefore, the ARIMA (3,0,2) GARCH (1,1) model is the best model to predict the future price of Google shares.

Forecasting using the Best Model
This study used two approaches: forecasting with unconditioned Sigma and rolling forecast with conditioned Sigma. In time series, forecast rolling is a method that uses recent historical data to predict future values. This method is often used to predict data that fluctuates or has patterns that cannot be predicted well.

The latest historical data is used to train the prediction model in the rolling forecasting method. This model is then used to predict future value. Once future values are predicted, the latest historical data is added to the training data. The model is then retrained using this new training data. This process continuously repeats itself to produce a prediction of future value.

The following are the steps for implementing the forecast method rolling: 1) Determine the prediction model that will be used; 2) Select a period of historical data to use as training data; 3) Train the prediction model using the training data; 4) Predict future values using pre-trained models; 5) Add the latest historical data to the training data; 6) Retrain the prediction model using the new training data; 7) Repeat steps 4-6 until the desired future value is predicted.

The rolling forecast method has several advantages, those are: 1) This method is relatively simple to implement; 2) This method can be used to predict data that fluctuates or has patterns that cannot be predicted well.

However, the rolling forecast method also has several weaknesses, namely: 1) This method can produce less accurate predictions if the data is historically insufficient; 2) This method can produce less accurate predictions if data patterns change over time.
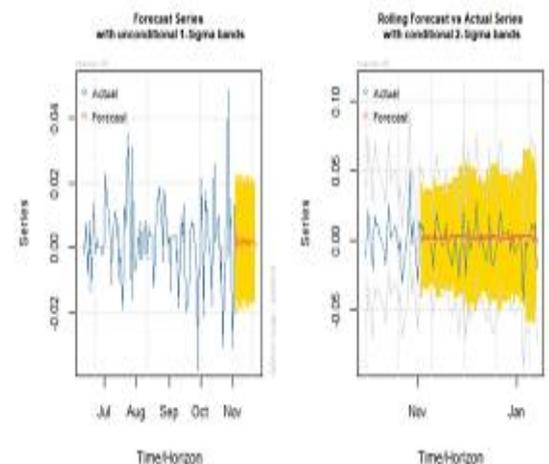
Figure 2. Forecasting results of 20 new data; Left: Forecasting with unconditioned Sigma Right: Rolling forecasting with conditioned Sigma

Comparative Evaluation of Predicted Results with Actual Data

Mean Absolute Percentage Error (MAPE) is a measure of prediction accuracy commonly used in time series data [24]. MAPE is calculated by dividing the absolute amount of the difference between the predicted and actual values by the actual value, then multiplying by 100. The MAPE formula is as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| x100 \qquad (1)$$

Where:
- y_t Is the actual value in period t
- $\widehat{y_t}$ is the predicted value in period t
- ⬚ is the number of periods

MAPE has several advantages, namely:
- Easy to calculate
- It can be used for time series data with varying patterns

However, MAPE also has several weaknesses, namely:
- Can be affected by large actual values
- It cannot be used for time series data with actual values. Negative.

To evaluate the comparison of predicted results with actual series data when using MAPE, the steps taken are as follows: 1) Calculate the MAPE value for each period t; 2) Average the MAPE values for all periods.

A smaller MAPE value indicates the prediction is more accurate [25], [26]. MAPE values below 10% indicate that predictions are very accurate. MAPE values between 10% and 20% indicate that predictions are pretty accurate. MAPE values above 20% indicate that predictions are less accurate [27]. The following are the results of stock price predictions using the ARIMA model.

```
MAPE(Comparison$x2023.10.30,Comparison$Return)
```

Figure 3. Results of comparison of actual data with model predictions using the MAPE method

Based on these results, the average MAPE value is 20.7%. Therefore, the ARIMA model's stock price predictions are entirely accurate [28].

The lower the MAPE value, the ability of the forecasting model used can be said to be good, and for MAPE, there is a range of values that can be used as measurement material regarding the ability of a forecasting model, the range of values can be seen in table 1 [29].

Table 3. The signification of MAPE

MAPE	Signification
<10%	Excellent forecasting ability
10-20%	Good forecasting ability
20-50%	Reasonable forecasting ability
>50%	Bad forecasting ability

CONCLUSION

The best model was selected based on the AIC (Akaike Information Criterion) criteria, and the ARIMA (3,0,2) GARCH (1,1) model was selected as the best model for predicting Google's stock price in the future. Forecasting evaluation uses the rolling forecasting method and MAPE (Mean Absolute Percentage Error) to broadcast forecasting performance. The results of this research provide comprehensive insight and can be a valuable reference for stock brokers and investors in making investment decisions.

REFERENCES

[1]   Ratih Kusumawardhani, Risal Rinofah, and Dandi Rukmana, "Google trend and stock market: Does it matter?," *J. Innov. Bus. Econ.*, vol. 5, no. 02, pp. 85–92, 2022, doi: 10.22219/jibe.v5i02.18678.

[2]   L. Bijl, K. Glenn, and M. Peter, "Google searches and stock returns," *Int. Rev. Financ. Anal.*, vol. 45, no. 1, pp. 150–156, 2016, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S105752191630045X

[3]   R. Robiyanto, M. A. Santoso, A. D. R. Atahau, and H. Harijono, "The Indonesia stock exchange and its dynamics: An analysis of the effect of macroeconomic variables," *Montenegrin J. Econ.*, vol. 15, no. 4, pp. 59–73, 2019, doi: 10.14254/1800-5845/2019.15-4.5.

[4]   Z. A. Rafsanjani, D. Nurtiyasari, and A. Syahputra, "The Dynamics of Stock Price Change Motion Effected by Covid-19 Pandemic and the Stock Price Prediction Using Multi-layered Neural Network," *Int. J. Comput. Sci. Appl. Math.*, vol. 7, no. 1, p. 8, 2021, doi:

10.12962/j24775401.v7i1.7023.

[5] J. Kane, "Google stock forecast: 2023 and Beyond," *Lite Finance*, p. 1, 2021.

[6] W. K. Pertiwi and R. Wahyudi, "17 Tahun Lalu, Google Perdana Lepas Saham 85 Dollar Per Lembar Artikel ini telah tayang di Kompas.com dengan judul '17 Tahun Lalu, Google Perdana Lepas Saham 85 Dollar Per Lembar', Klik untuk baca: https://tekno.kompas.com/read/2021/ 08/19/16270007/17-tahu," *Kompas*, Jakarta, 2021. [Online]. Available: https://tekno.kompas.com/read/2021/ 08/19/16270007/17-tahun-lalu-google-perdana-lepas-saham-85-dollar-per-lembar

[7] E. Ulya, Fika Nurul Djumena, "Sah, Google dan Temasek Kini Jadi Pemegang Saham Baru Tokopedia," *Kompas*, Jakarta, 2020. [Online]. Available: https://money.kompas.com/read/2020 /11/17/060300226/sah-google-dan-temasek-kini-jadi-pemegang-saham-baru-tokopedia

[8] F. Aslam, M. A. Yaqub, and B. Bashir, "International Journal of Management Research and Emerging Sciences Predicting Stock Market Direction using Machine Learning Models," vol. 83, no. 1, p. 2019, 2019, [Online]. Available: http://ijmres.pk/

[9] S. Maddodi and K. G. N. Kumar, "Stock Market Forecasting: A Review of Literature," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. Vol 5, No, p. 11, 2021, [Online]. Available: https://ojs.stmikpringsewu.ac.id/index. php/ijiscs/article/download/1064/pdf

[10] S. S. Ishak *et al.*, "Indonesian Consumer Price Index Forecasting Using Autoregressive Integrated Moving Average," *Int. J. Electron. Commun. Syst.*, vol. 3, no. 1, p. 33, 2023, doi: 10.24042/ijecs.v3i1.18252.

[11] S. Khandelwal and D. Mohanty, "Stock Price Prediction Using Arima Model," *Int. J. Mark. Hum. Resour. Res.*, vol. 2, no. 2, pp. 2746–4040, 2021.

[12] Z. Haider Khan, T. Sharmin Alin, and A. Hussain, "Price Prediction of Share Market Using Artificial Neural Network 'ANN,'" *Int. J. Comput. Appl.*, vol. 22, no. 2, pp. 42–47, 2011, doi: 10.5120/2552-

3497.

[13] P. Meesad and R. I. Rasel, "Predicting stock market price using support vector regression," *2013 Int. Conf. Informatics, Electron. Vision, ICIEV 2013*, no. May, 2013, doi: 10.1109/ICIEV.2013.6572570.

[14] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, "Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications," *Int. J. Financ. Stud.*, vol. 11, no. 3, 2023, doi: 10.3390/ijfs11030094.

[15] L. N. Faiza and D. Agustina, "Aplikasi Machine Learning dalam Prediksi Harga Saham Jakarta Islamic Index (JII) Menggunakan Metode Support Vector Regression," vol. 8, no. 3, pp. 79–88, 2023.

[16] V. V. Setia, "Analisis Penilaian Harga Saham Menggunakan Metode Dividend Discount Model (DDM) sebagai Dasar Pengambilan Keputusan Investasi (Studi pada Perusahaan yang termasuk dalam Indeks LQ-45 di Bursa Efek Indonesia Tahun 2013-2015).," *Apl. Adm.*, vol. 20, no. 1, pp. 26–38, 2018, doi: 10.30649/aamama.v20i1.90.

[17] I. S. Mubarokah, A. Fitrianto, and F. M. Affendi, "Perbandingan Model Garch Simetris dan Asimetris Pada Data Kurs Harian," *Indones. J. Stat. Its Appl.*, vol. 4, no. 4, pp. 627–637, 2020, doi: 10.29244/ijsa.v4i4.709.

[18] A. H. A. Zili, D. Hendri, and S. A. A. Kharis, "Peramalan Harga Saham dengan Model Hybrid Arima-Garch dan Metode Walk Forward," *J. Stat. dan Apl.*, vol. 6, no. 2, pp. 341–354, 2022, doi: 10.21009/jsa.06218.

[19] V. O. Nur Laily, B. Warsito, and D. A. I Maruddani, "Comparison of ARCH / GARCH model and Elman Recurrent Neural Network on data return of closing price stock," *J. Phys. Conf. Ser.*, vol. 1025, no. 1, 2018, doi: 10.1088/1742-6596/1025/1/012103.

[20] A. J. M. Tarigan, M. Mardiningsih, and S. Suwilo, "The search for alternative algorithms of the iteration method on a system of linear equation," *Sinkron*, vol. 7, no. 4, pp. 2124–2424, 2022, doi:

10.33395/sinkron.v7i4.11817.

[21] Bernadhita Herindri Samodera Utami, D. Herinanto, and M. Gumanti, "Aplikasi Metode Maximum Likelihood Estimation pada Data Binomial Interval-Tersensor," *Transform. J. Pendidik. Mat. dan Mat.*, vol. 6, no. 1, pp. 75–85, 2022, doi: 10.36526/tr.v6i1.1934.

[22] K. R. Katianda, R. Goejantoro, and A. M. Ade Satriya, "Estimasi Parameter Model Regresi Linier dengan Pendekatan Bayes (Studi Kasus: Kemiskinan di Provinsi Kalimantan Timur pada Tahun 2017) Estimation Parameter of Linear Regression Model with Bayes Approach (Case Study: Poverty of East Kalimantan Province in ," *J. EKSPONENSIAL*, vol. 11, no. 2, pp. 127–132, 2020.

[23] F. Yanuar, R. Febriyuni, and I. R. HG, "Bayesian Generalized Self Method to Estimate Scale Parameter of Invers Rayleigh Distribution," *CAUCHY J. Mat. Murni dan Apl.*, vol. 6, no. 4, pp. 270–278, 2021, doi: 10.18860/ca.v6i4.11482.

[24] M. B. S. Junianto, "Fuzzy Inference System Mamdani dan the Mean Absolute Percentage Error (MAPE) untuk Prediksi Permintaan Dompet Pulsa pada XL Axiata Depok," *J. Inform. Univ. Pamulang*, vol. 2, no. 2, p. 97, 2017, doi: 10.32493/informatika.v2i2.1511.

[25] I Putu Susila Handika and I Kadek Susila Satwika, "Enhancing Sales Forecasting Accuracy Through Optimized Holt-Winters Exponential Smoothing with Modified Improved Particle Swarm Optimization," *J. Nas. Pendidik. Tek. Inform.*, vol. 12, no. 2, pp. 203–212, 2023, doi: 10.23887/janapati.v12i2.65462.

[26] D. Alita, A. D. Putra, and D. Darwis, "Analysis of Classic Assumption Test and Multiple Linear Regression Coefficient Test for Employee Structural Office Recommendation," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 3, p. 295, 2021, doi: 10.22146/ijccs.65586.

[27] P. Bagus, P. Putra, C. Wiedyaningsih, and E. Yuniarti, "Forecasting Drug Demand Using The Single Moving Average At Prof . dr . I . G . N . G . Ngoerah Hospital," vol. 19, no. 3, pp. 394–402, 2023.

[28] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Stock Price Prediction using the ARIMA Model," *Proc. - UKSim-AMSS 16th Int. Conf. Comput. Model. Simulation, UKSim 2014*, no. June, pp. 106–112, 2014, doi: 10.1109/UKSim.2014.67.

[29] P. C. Chang, Y. W. Wang, and C. H. Liu, "The Development of a Weighted Evolving Fuzzy Neural Network for PCB sales forecasting," *Expert Syst. Appl.*, vol. 32, no. 1, pp. 86–96, 2007, doi: 10.1016/j.eswa.2005.11.021.