

A Competitive Model to Forecast a Stock Market Index

Nhan Nguyen-Thanh ^{1,2,*}, Kun-Huang Huarng ³

¹ PhD Candidate at Department of Economics, College of Business, Feng Chia University, Taichung City, Taiwan

² Faculty of Business Administration, Ton Duc Thang University, No. 19 Nguyen Huu Tho Street, Tan Phong Ward, District 7, Ho Chi Minh City, Vietnam

³ Department of Product Innovation and Entrepreneurship, National Taipei University of Business, Taoyuan City 324, Taiwan; khhuarng@ntub.edu.tw

* Corresponding author: nguyenthahnhan@tdtu.edu.vn

Submitted: 27 August 2019, accepted: 12 April 2020, published: 14 April 2020

Abstract: This study proposes a competitive model using the Box–Jenkins approach to implement a Box–Jenkins ARIMA-GARCH model in order to improve financial forecasting. Differing from previous studies, we consider optimizing the lagged terms, which assist in capturing the relationships more properly. The competitive model is then used to forecast the stock market index in Taiwan. This study conducts out-of-sample forecasting and compares the root mean square errors (RMSEs) against previous studies. The results show that the competitive model outperformed in terms of both RMSEs and consistency.

Keywords: Box–Jenkins approach; ARIMA-GARCH; fuzzy sets; stock market index

How to cite: Nhan Nguyen-Thanh, Kun-Huang Huarng. A Competitive Model to Forecast a Stock Market Index. *J. Bus. Account. Financ. Perspect.*, 2020, 2(2): 14; doi:[10.26870/jbafp.2020.02.014](https://doi.org/10.26870/jbafp.2020.02.014).

© 2020 Copyright by the authors. Licensed as an open access article using a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.



1. Introduction

Box–Jenkins ARIMA¹ is a simulation of regressive time series models that has become widely utilized in recent years. Many kinds of Box–Jenkins ARIMA have been developed, including the ARMA approach (Rojas et al., 2008; Wang and Lu, 2006; Hwarng, 2001), the ARIMA approach (Hikichi et al., 2017; Petrevska, 2017), seasonal ARIMA (Gharbi et al., 2011; Egrioglu et al., 2009; Tsui et al., 2014), the bivariate model or ARIMA transfer (Sharma and Khare, 2001; Sun and Koch, 2001; Gröger and Rumohr, 2006), and hybrid models (Babu and Reddy, 2014; Egrioglu et al., 2009; Pong et al., 2004). Moreover, some studies have shown interest in the partitioning

¹ARIMA names as autoregressive integrated moving average that is proposed by (Asteriou and Stephen, 2011).

of the ARIMA family to improve forecasting results (Gray, 1996; Blazsek and Mendoza, 2016; Zhang, 2003; Kambouroudis et al., 2016). The Box–Jenkins ARIMA model (Box and Jenkins, 1976) has been applied to forecast linear time series for different domains, such as tourism demand (Petrevska, 2017), the number of passengers (Egrioglu et al., 2009), the exchange rate (Pong et al., 2004), CO₂ emissions (Paravantis and Georgakellos, 2007), and high-bandwidth networks (Yoo and Sim, 2016). For forecasting, many researchers have conducted new models that have outperformed their counterparts (Chen, 1996; Song, 1999; Gallant et al., 1999; Alizadeh et al., 2002; Huarng and Yu, 2006; Huarng et al., 2007; Yu and Huarng, 2008; 2010; Granger and Newbold, 1976).

The application of these Box–Jenkins ARIMA models consists of three main steps: identifying I(0) or I(1) (unit root test and differencing), Box–Jenkins optimization, and training linear and non-linear models. Many studies suggest that a hybrid model should be applied to solve both linear and non-linear approaches. However, this study uses GARCH(1,1)² as a solution for non-linear and complex relationships. Moreover, heteroskedasticity is a specific problem in both time series and statistics (Kristjanpoller and Hernández, 2017; Gray, 1996; Andersen and Bollerslev, 1998; Hakim and McAleer, 2009) and causes biased parameters, which heavily affect the forecasting results. Hence, this study proposes the Box–Jenkins ARIMA-GARCH model.

For comparison purposes, this study chooses various fuzzy time series models, including first-order models (Chen, 1996), bivariate models (Yu and Huarng, 2008), multivariate models (Huarng et al., 2007), and hybrid models (Huarng and Yu, 2006; Yu and Huarng, 2010). In these studies, the fuzzy models do not handle the lagged terms. Therefore, the competitive model is expected to outperform these models.

The contributions of this study are as follows. First, the competitive model successfully builds a time series approach to solve both the linearity and non-linearity of the data with optimizing lagged terms, and presents good forecasting results. Second, the application of the Box–Jenkins approach can better capture the lagged term relationships and thus provide a better forecast. Third, optimizing the lagged terms allows the competitive model (a univariate model) to compete with a bivariate model.

This study proposes a competitive model using the Box–Jenkins approach to implement a Box–Jenkins ARIMA-GARCH model to improve forecasting. Toward that end, the remainder of this paper consists of the following sections. Section 2 reviews the concepts of Box–Jenkins ARIMA, fuzzy time series, and neural network (NN) models. Section 3 describes the data and explains the competitive model. Section 4 uses an example to demonstrate the forecasting analysis. Section 5 compares the performance of the empirical models. Section 6 concludes the paper.

2. Literature Review

2.1. Box–Jenkins ARIMA

ARIMA(p, d, q) is a well-known linear approach that has been applied in many studies in the forecast literature. Before using ARIMA(p, d, q), the stationarity of the data series (order d) and the order (p, q) should be determined. The best-suited ARIMA can be validated by Akaike information criterion (AIC). For stock index forecasting, many researchers suggest that ARIMA be combined with a non-linear approach, because the asymmetric volatility of the stock index is typically researchers' target interest (Wang et al., 2012; Mostafa, 2010; Kang and Yoon, 2013). The Box–Jenkins can optimize p and q choices in the ARIMA equation as:

$$FD(t, t-1) = AR_p + MA_q = A_o + \sum_{p-1}^n A_p D(t-p, t-p-1) + e_{pt} + e_{qt} + \sum_{q-1}^n W_q e_{t-q}. \quad (1)$$

In ARIMA, the autoregressive series $\left(\sum_{p-1}^n A_p D(t-p, t-p-1)\right)$ captures the linear trend of data, and the moving average series $\left(\sum_{q-1}^n W_q e_{t-q}\right)$ captures the linear terms in error. Building a hybrid model, this study combines the GARCH model with ARIMA to better predict the results of a stock market index, called GARCH(1,1).

²GARCH is the well-known approach named as generalized autoregressive conditional heteroskedasticity.

This study also uses the applications of the GARCH model, which is provided to deal with sensitiveness, non-stationarity, and asymmetric volatility series (Engle, 2002). However, various hybrid models have been employed to solve both linear and non-linear characteristics of the stock index problem (Blazsek and Mendoza, 2016; Kazem et al., 2013). GARCH(1,1) is one of the most popular methods (Sbrana and Silvestrini, 2013; Kömm and Küsters, 2015). The square error of ARIMA u_{t-1}^2 can be used exogenously, and σ_{t-1}^2 can also be adjusted exogenously in the GARCH function, as in Equation (2). Hence, GARCH(1,1) can exclude the condition of heteroskedasticity in the volatility of a stock market index.

$$\sigma_t^2 = \alpha_o + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (2)$$

2.2. Fuzzy Set Time Series

Fuzzy time series models have been utilized for decades by many researchers. Song and Chissom (1993) proposed the foundation for fuzzy time series models, including (1) to define the universe of discourse and intervals; (2) to fuzzify; (3) to establish fuzzy relationships; and (4) to forecast. We use Chen's (1996) model (referred to as Model 1) as an example of a first-order model and conduct similar forecasts. The heuristic model integrates the heuristic to improve a fuzzy time series (Huarng, 2001) and also is conducted as a multivariate model (Huarng et al., 2007) (referred to as Model 2).

2.3. NN Models

The neural network (NN) is a non-linear technique that is similar to the human brain architecture and is applied widely in forecasting. The first NN-fuzzy time series in forecasting was suggested by Huarng and Yu (2006) (Model 3). The basic framework uses the most significant degrees of membership for each observation both for in-sample and out-of-sample forecast when the other ones are ignored, which may affect the outcome. The univariate NN-fuzzy time series model (Yu and Huarng, 2010) (Model 4), which uses all the degrees of membership to establish a fuzzy relationship, is a more innovative and complicated model. The bivariate NN-fuzzy time series model (Yu and Huarng, 2008) (Model 5) performs better than a univariate model by using D(TAIFEX³) rather than D(TAIEEX⁴) as the input to generate a forecast series.

3. Research Method

3.1. Data

This study employed data for the daily closing stock market index of Taiwan, Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEEX). To facilitate comparisons, the sample size was set to be the same as in Yu and Huarng (2008) from years 2000 to 2004. To achieve forecasting, a previous study stated the importance of out-of-sample observations (Martin and Witt, 1989). Observations from January to October were considered as in-sample data (training sample). This study also used out-of-sample observations for each year, with November to December as the out-of-sample data (testing sample). Hence, the ratios were consistently $\frac{10}{12} : \frac{2}{12}$ every year.

3.2. The Competitive Model

We conducted the Box–Jenkins ARIMA-GARCH(1,1) as follows, named as Model 6.

Step 1. Unit root test

Checking for data stationarity is an important step. In the case of stationarity, the TAIEEX at time $(t - 1)$ was directly used to forecast the TAIEEX at time (t) by Box–Jenkins ARIMA($p;0;q$)-GARCH(1,1). The indirect

³TAIFEX denotes Taiwan Future Exchange, <https://www.taifex.com.tw>.

⁴TAIEEX denotes Taiwan Capitalization Weighted Stock Index, <https://wn.com/taieex/news>.

case implies that Box–Jenkins ARIMA($p;1;q$)-GARCH(1,1) was applied to use the TAIEX in first difference at time ($t - 1$) and to forecast the difference of the TAIEX at time (t) if the data series exhibited non-stationarity. We used the augmented Dickey–Fuller (ADF) unit root test to detect data stationarity (Dickey and Fuller, 1979).

Step 2. Difference

Many previous researchers used the differences of a stock market index for prediction by the ARIMA family (Babu and Reddy, 2014). Hence, the order d is usually set to be 1. Following the unit root results, if first differences are used as the input series, then the first difference series are calculated by Equation (3):

$$D(t, t - 1) = stockindex_t - stockindex_{t-1}. \tag{3}$$

Step 3. Box–Jenkins optimization

The orders p and q were determined by the Box–Jenkins method (Zhang, 2003; Hosking, 1981). Each year’s data were used to conduct the Box–Jenkins ARIMA ($p;d;q$) individually. The orders p and q can be observed by the trend and correlation analysis of the series. The order p can be picked up from ACF (autocorrelation), and the order q can be picked up from PACF (partial autocorrelation). The orders p and q were substituted into ARIMA ($p;d;q$) in order to optimize the order by considering the smallest Schwarz information criterion (SIC) and Akaike information criterion (AIC).

Step 4. Building the competitive model

Heteroskedasticity is a statistical problem that causes a bias parameter. Hence, GARCH(1,1) equation is usually used to provide the solution for homoskedasticity and to enhance the robustness for the ARIMA family (Brooks, 2014). The hybrid model is a combination of the optimized ARIMA($p;d;q$) model with GARCH(1,1), due to its capabilities in handling non-linear relationships. The hybrid simultaneous model is listed as Equation (4) below, which combines Equations (1) and (2):

$$\left\{ \begin{array}{l} FD(t, t - 1) = AR_p + MA_q = A_0 + \sum_{p-1}^n A_p D(t - p, t - p - 1) + e_{pt} + e_{qt} + \sum_{q-1}^n W_q e_{t-q} \\ u_t = e_{pt} + e_{qt} \\ \sigma_t^2 = \alpha_o + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \end{array} \right. \tag{4}$$

Here, $FD(t, t - 1)$ is the forecasted values of the stock market index in differencing at time (t); $D(t - p, t - p - 1)$ is the autoregressive series at time ($t - p$); e_{t-q} is the moving average series at time ($t - q$), and e_{pt} and e_{qt} are error terms in the autoregressive series and moving average series, respectively. Thus, u_t is the total error term, and σ_t^2 is the variance series of heteroskedasticity estimated by the error term of the autoregressive moving average series.

Step 5. Forecasting

Similar to previous studies, this research takes the backward induction of the difference in forecast values to result in the forecasted stock market index. The output of the Box–Jenkins ARIMA-GARCH(1,1) model still forecasts the index ($FStockindex_t$). Therefore, the index at time ($t - 1$) can be calculated as the input of Equation (5):

$$FStockindex_t = Stockindex_{t-1} + FD(t, t - 1) \tag{5}$$

Step 6. Performance evaluation

Following a previous study (Yu and Huarng, 2010), this research also uses root mean square errors (RMSEs) to compare the performance, as in Equation (6):

$$RMSE = \sqrt{\left(\sum_{i=k+1}^n (ActualTAIEX_t - ForecastTAIEX_t)^2 \right) / (n - k)} \tag{6}$$

where there are n observations, including k in-sample and $n - k$ out-of-sample observations.

4. Forecasting Analysis

We took TAIEX data in the year 2000 as an example to demonstrate the empirical analysis.

Step 1. Unit root test

The null hypothesis of the ADF tests is that TAIEX is non-stationary. We performed this test and rejected the null hypothesis. The results are shown in Table 1. Here, ARIMA($p;d;q$) was applied by using the differences of the TAIEX at time $(t - 1)$ as the input to forecast the TAIEX at time (t) . The integrated order d was fixed at 1. The ARIMA ($p;1;q$) model was thus executed.

	2000	2001	2002	2003	2004
$TAIEX_t$	0.2622	-0.8788	-0.8427	-0.7568	-1.6883
$D(TAIEX_t)$	-15.75 ***	-14.41 ***	-15.26 ***	-15.05 ***	-14.85 ***

*** denotes significance at 99%, ** significance at 95%, and * significance at 90%. Note: The null hypothesis of the Augmented Dickey-Fuller (ADF) test is that the TAIEX has a unit root.

Table 1 Results of ADF unit root tests.

Step 2. Difference

The stock market index on January 5 was 8849.87, and it was 8756.55 on January 4. Hence, $D(1/5/2000,1/4/2000) = 8849.87 - 8756.55 = 93.32$.

Step 3. Box-Jenkins optimization

By using Box-Jenkins methods, the value of order p can be considered as suitable when the autocorrelation value in the ACF column is observed from the highest -0.178 (lag $p = 18$) to the smallest value 0.084 (lag $p = 28$) and larger than 0.05 . We employed a similar approach with the order q (Table 2). We considered ARIMA(18;1;4,15) as the most optimized, in which SIC = 12.9664 and AIC = 12.9126.

Model	Available order to be chosen as p and q lagged terms (see the full table in Appendix A)						
AR(p)	4	13	14	15	18	21	28
	-0.163	0.09	0.119	-0.123	-0.178 **	0.088	0.084
MA(q)	4	13	14	15	18	21	28
	-0.168 **	0.101	0.104	-0.136 **	-0.139	0.085	0.126
Model selection by information criteria							
ARIMA family	SIC			AIC		Rank	
ARIMA(18;1;4)	12.9674			12.9254		3	
ARIMA(18;1;15)	12.9879			12.9458		4	
ARIMA(18;1;4,15)	12.9664			12.9126		1	
ARIMA(0;1;4,15,18)	12.9686			12.9131		2	

*** denotes significance at 99%, ** significance at 95%, and * significance at 90%.

Table 2 Box-Jenkins method optimization and results for the year 2000.

Step 4. Building the competitive model

By using Equation (4), we optimized the error terms u_t of both autoregressive series e_{pt} and moving average series (e_{pt}) in GARCH(1,1). The new hybrid model, which can solve for heteroskedasticity, is expected to generate a

better output. We followed Box–Jenkins ARIMA-GARCH(1,1) methods to optimize the forecast model for each year. Table 3 lists all the results.

Year	Best model
2000	ARIMA(18;1;4,15)-GARCH(1,1)
2001	ARIMA(4;1;4,15)-GARCH(1,1)
2002	ARIMA(9;1;4)-GARCH(1,1)
2003	ARIMA(8;1;8)-GARCH(1,1)
2004	ARIMA(29;1;2,15)-GARCH(1,1)

Table 3 Optimized model for each year.

Step 5: Forecasting

As in Equation (4), the closing index on November 1 was 5425.02 and on November 2 it was 5625.08. Employing the model, the forecast of stock difference ($FD(t, t - 1)$) was 14.5274 on November 2. Hence, the output of the model ($FStockindex_t$) was computed as 5439.55 for November 2. Table 4 presents the forecasts.

Date(t)	Actual TAIEX(t)	Inputs		Outputs	
		First difference series ($D(t, t - 1)$)	Actual TAIEX ($t - 1$) ($Stockindex_{t-1}$)	Forecast value of stock market index difference ($FD(t, t - 1)$)	Forecast TAIEX (t) ($FStockindex_t$)
11/1/2000	5425.02	-119.16	5544.18	-0.33156	5543.85
11/2/2000	5626.08	201.06	5425.02	14.5274	5439.55
11/3/2000	5796.08	170.00	5626.08	59.2410	5685.32
11/4/2000	5677.30	-118.78	5796.08	13.9798	5810.06
11/6/2000	5657.48	-19.82	5677.30	45.7822	5723.08
11/7/2000	5877.77	220.29	5657.48	0.21743	5657.70
11/8/2000	6067.94	190.17	5877.77	-89.554	5788.22
11/9/2000	6089.55	21.61	6067.94	20.0731	6088.01
11/10/2000	6088.74	-0.81	6089.55	34.4934	6124.04
11/13/2000	5793.52	-295.22	6088.74	-129.906	5958.83
11/14/2000	5772.51	-21.01	5793.52	-136.168	5657.35
11/15/2000	5737.02	-35.49	5772.51	-33.7512	5738.76
11/16/2000	5454.13	-282.89	5737.02	-19.9644	5717.06
11/17/2000	5351.36	-102.77	5454.13	25.2855	5479.42
11/18/2000	5167.35	-184.01	5351.36	-35.5636	5315.80
11/20/2000	4845.21	-322.14	5167.35	11.3685	5178.72
11/21/2000	5103.00	257.79	4845.21	38.9101	4884.12
11/22/2000	5130.61	27.61	5103.00	10.8685	5113.87
11/23/2000	5146.92	16.31	5130.61	48.9828	5179.59
11/24/2000	5419.99	273.07	5146.92	52.0762	5199.00
11/27/2000	5433.78	13.79	5419.99	-124.454	5295.54
11/28/2000	5362.26	-71.52	5433.78	-48.3960	5385.38
11/29/2000	5319.46	-42.80	5362.26	-9.66358	5352.60
11/30/2000	5256.93	-62.53	5319.46	-96.5995	5222.86

Table 4 Forecast from the Box–Jenkins ARIMA-GARCH(1,1) model.

Step 6. Performance evaluation

For the year 2000, the RMSE of Model 6 was 122.68.

5. Empirical Analysis

We repeated the forecasting for all years and compared the performance of six models in terms of RMSE, as in Table 5. In the table, the competitive model (Model 6) performed best among all the models, because it had more years with smaller RMSEs than the other models. To show consistency, the competitive model also outperformed the other models, because it had more total number of years with smaller RMSEs than each model.

Model	2000	2001	2002	2003	2004
(1) First-order model (Chen, 1996)	176.32	147.84	101.18	74.46	84.28
(2) Multivariate model (Huarng et al., 2007)	154.42	124.02	95.73	70.76	72.35
(3) NN model (Huarng and Yu, 2006)	152	130	84	56	N/A
(4) NN based fuzzy time series (Yu and Huarng, 2010)	149.59	98.31	78.71	58.78	55.91
(5) Bivariate fuzzy time series (Yu and Huarng, 2008)	67	120	69	52	60
(6) The competitive model	122.68	108.62	66.11	55.56	52.99

Table 5 Performance evaluation by root mean square errors (RMSEs). NN: neural network.

The competitive model could capture both the linearity and non-linearity of the data, which is the same as any hybrid models in previous studies. In the model, the Box–Jenkins approach allows the researcher to optimize the lagged terms of both autoregressive series and moving average series by minimizing the white noise, which means that more correct information is taken into consideration. Hence, the forecasting results are expected to improve. Model 5 (bivariate NN-fuzzy approach), using all the degrees of membership, exhibited good results. The drawback of taking all the degrees of membership for training and forecasting is that there can be too many fuzzy sets or inputs for the NN.

6. Conclusion

This study proposes a Box–Jenkins ARIMA-GARCH model as a competitive model to improve forecasting, as it optimizes the lagged terms of both autoregressive series and moving average series by minimizing the white noise. Due to the coverage of most optimizing information, the model performed better than many previous studies. Another advantage of taking the optimized lagged terms is that the competitive model can improve a univariate model to compete with a bivariate model.

The competitive model can easily expand its function and also calculate fuzzy relationships. Following the empirical results in this study, the competitive model can solve problems of both linear and non-linear data. For future work, if other suitable inputs can be observed, then the competitive model can be easily expanded to bivariate models.

Acknowledgements

Author named Nhan, Nguyen-Thanh is very appreciation to Prof. Yung-Chang, Wang, and Prof. Kuo-Hsuan, Chin on their supporting comments.

Appendix A. ACF and PACF of the First Difference of TAIEX in the Year 2000

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. .	. .	1	0.037	0.037	0.3763	0.540
. .	. .	2	0.043	0.042	0.8847	0.643
. .	. .	3	0.045	0.042	1.4412	0.696
* .	* .	4	-0.163	-0.168	8.7337	0.068
. .	. .	5	0.016	0.026	8.8085	0.117
. .	. .	6	-0.049	-0.039	9.4727	0.149
. .	. .	7	-0.029	-0.012	9.7123	0.205
. .	. .	8	-0.037	-0.063	10.094	0.259
. .	. .	9	0.009	0.027	10.115	0.341
. .	. .	10	-0.005	-0.017	10.124	0.430
. .	. .	11	0.029	0.030	10.359	0.498
. .	. .	12	-0.004	-0.026	10.364	0.584
. *	. *	13	0.090	0.101	12.654	0.475
. *	. *	14	0.119	0.104	16.722	0.271
* .	* .	15	-0.123	-0.136	21.087	0.134
. .	. .	16	-0.031	-0.050	21.364	0.165
. .	. .	17	-0.045	-0.005	21.964	0.186
* .	* .	18	-0.178	-0.139	31.245	0.027
. .	. .	19	0.016	-0.005	31.324	0.037
. .	. .	20	-0.014	0.008	31.379	0.050
* .	* .	21	-0.088	-0.085	33.673	0.039
. .	. .	22	0.011	-0.035	33.710	0.053
. .	. .	23	-0.045	-0.047	34.319	0.061
. .	. .	24	0.046	0.047	34.951	0.069
. .	. .	25	0.022	-0.012	35.093	0.087
. .	. .	26	-0.007	-0.026	35.108	0.109
. .	. .	27	-0.011	-0.055	35.143	0.135
. *	. *	28	0.084	0.126	37.277	0.113
. .	. .	29	-0.006	0.009	37.289	0.139
. .	. .	30	0.031	0.016	37.574	0.161

References

- Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3), 1047–1091. [[CrossRef](#)]
- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885–905. [[CrossRef](#)]
- Asteriou, D., & Stephen, G. H. (2011). ARIMA models and the Box–Jenkins methodology. *Applied Econometrics*, 2(2), 265–286.

- Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. *Applied Soft Computing*, 23(10), 27–38. [[CrossRef](#)]
- Blazsek, S., & Mendoza, V. (2016). QARMA-Beta-t-EGARCH versus ARMA-GARCH: an application to S&P 500. *Applied Economics*, 48(12), 1119–1129.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco, CA: Holden-Day.
- Brooks, C. (2014). *Introductory econometrics for finance*. Cambridge: Cambridge University Press.
- Chen, S. M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, 81(3), 311–319. [[CrossRef](#)]
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427–431.
- Egrioglu, E., Aladag, C. H., Yolcu, U., Basaran, M. A., & Uslu, V. R. (2009). A new hybrid approach based on SARIMA and partial high order bivariate fuzzy time series forecasting model. *Expert Systems with Applications*, 36(4), 7424–7434. [[CrossRef](#)]
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*, 20(3), 339–350. [[CrossRef](#)]
- Gallant, A. R., Hsu, C. T., & Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *The Review of Economics and Statistics*, 81(4), 617–631. [[CrossRef](#)]
- Gharbi, M., Quenel, P., Gustave, J., Cassadou, S., La Ruche, G., Girdary, L., & Marrama, L. (2011). Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. *BMC Infectious Diseases*, 11, 166. [[CrossRef](#)]
- Granger, C. W. J., & Newbold, P. (1976). *Forecasting economic time series*. New York, NY: Academic Press.
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62. [[CrossRef](#)]
- Gröger, J., & Rumohr, H. (2006). Modelling and forecasting long-term dynamics of Western Baltic macrobenthic fauna in relation to climate signals and environmental change. *Journal of Sea Research*, 55(4), 266–277.
- Hakim, A., & McAleer, M. (2009). Forecasting conditional correlations in stock, bond and foreign exchange markets. *Mathematics and Computers in Simulation*, 79(9), 2830–2846. [[CrossRef](#)]
- Hikichi, S. E., Salgado, E. G., & Beijo, L. A. (2017). Forecasting number of ISO 14001 certifications in the Americas using ARIMA models. *Journal of Cleaner Production*, 147(23), 242–253. [[CrossRef](#)]
- Hosking, J. R. (1981). Fractional differencing. *Biometrika*, 68(1), 165–176. [[CrossRef](#)]
- Huang, K., & Yu, T. H.-K. (2006). The application of neural networks to forecast fuzzy time series. *Physica A: Statistical Mechanics and Its Applications*, 363(2), 481–491. [[CrossRef](#)]
- Huang, K.-H., Yu, T. H.-K., & Hsu, Y. W. (2007). A multivariate heuristic model for fuzzy time-series forecasting. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(4), 836–846. [[CrossRef](#)]
- Huang, K. (2001). Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems*, 123(3), 369–386. [[CrossRef](#)]

- Hwang, H. B. (2001). Insights into neural-network forecasting of time series corresponding to ARMA (p, q) structures. *Omega*, 29(3), 273–289. [[CrossRef](#)]
- Kambouroudis, D. S., McMillan, D. G., & Tsakou, K. (2016). Forecasting Stock Return Volatility: A Comparison of GARCH, Implied Volatility, and Realized Volatility Models. *Journal of Futures Markets*, 36(12), 1127–1163. [[CrossRef](#)]
- Kang, S. H., & Yoon, S. M. (2013). Modeling and forecasting the volatility of petroleum futures prices. *Energy Economics*, 36(2), 354–362. [[CrossRef](#)]
- Kazem, A., Sharifi, E., Hussain, F. K., Saberi, M., & Hussain, O. K. (2013). Support vector regression with chaos-based firefly algorithm for stock market index forecasting. *Applied Soft Computing*, 13(2), 947–958. [[CrossRef](#)]
- Kömm, H., & Küsters, U. (2015). Forecasting zero-inflated price changes with a Markov switching mixture model for autoregressive and heteroscedastic time series. *International Journal of Forecasting*, 31(3), 598–608.
- Kristjanpoller, R. W., & Hernández, P. E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, 84, 290–300. [[CrossRef](#)]
- Martin, C. A., & Witt, S. F. (1989). Accuracy of econometric forecasts of tourism. *Annals of Tourism Research*, 16(3), 407–428. [[CrossRef](#)]
- Mostafa, M. M. (2010). Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait. *Expert Systems with Applications*, 37(9), 6302–6309. [[CrossRef](#)]
- Paravantis, J., & Georgakellos, D. (2007). Trends in energy consumption and carbon dioxide emissions of passenger cars and buses. *Technological Forecasting and Social Change*, 74(5), 682–707. [[CrossRef](#)]
- Petrevska, B. (2017). Predicting tourism demand by ARIMA models. *Economic Research-Ekonomska Istraživanja*, 30(1), 939–950. [[CrossRef](#)]
- Pong, S., Shackleton, M. B., Taylor, S. J., & Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR (FI) MA models. *Journal of Banking and Finance*, 28(10), 2541–2563. [[CrossRef](#)]
- Rojas, I., Valenzuela, O., Rojas, F., Guillén, A., Herrera, L. J., Pomares, H., Marquez, L., & Pasadas, M. (2008). Soft-computing techniques and ARMA model for time series prediction. *Neurocomputing*, 71(4), 519–537. [[CrossRef](#)]
- Sbrana, G., & Silvestrini, A. (2013). Aggregation of exponential smoothing processes with an application to portfolio risk evaluation. *Journal of Banking and Finance*, 37(5), 1437–1450. [[CrossRef](#)]
- Sharma, P., & Khare, M. (2001). Short-term, real-time prediction of the extreme ambient carbon monoxide concentrations due to vehicular exhaust emissions using transfer function-noise model. *Transportation Research Part D: Transport and Environment*, 6(2), 141–146. [[CrossRef](#)]
- Song, Q., & Chissom, B. S. (1993). Forecasting enrollments with fuzzy time series—Part I. *Fuzzy Sets and Systems*, 54(1), 1–9. [[CrossRef](#)]
- Song, Q. (1999). Seasonal forecasting in fuzzy time series. *Fuzzy Sets and Systems*, 107(2), 235–236. [[CrossRef](#)]
- Sun, H., & Koch, M. (2001). Case study: analysis and forecasting of salinity in Apalachicola Bay, Florida, using Box-Jenkins ARIMA models. *Journal of Hydraulic Engineering*, 127(9), 718–727. [[CrossRef](#)]
- Tsui, W. H. K., Balli, H. O., Gilbey, A., & Gow, H. (2014). Forecasting of Hong Kong airport's passenger throughput. *Tourism Management*, 42, 62–76. [[CrossRef](#)]

- Wang, X. K., & Lu, W. Z. (2006). Seasonal variation of air pollution index: Hong Kong case study. *Chemosphere*, *63*(8), 1261–1272. [[CrossRef](#)]
- Wang, J. J., Wang, J. Z., Zhang, Z. G., & Guo, S. P. (2012). Stock index forecasting based on a hybrid model. *Omega*, *40*(6), 758–766. [[CrossRef](#)]
- Yoo, W., & Sim, A. (2016). Time-series forecast modeling on high-bandwidth network measurements. *Journal of Grid Computing*, *14*(3), 463–476. [[CrossRef](#)]
- Yu, T. H.-K., & Huarng, K.-H. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications*, *34*(4), 2945–2952. [[CrossRef](#)]
- Yu, T. H.-K., & Huarng, K.-H. (2010). A neural network-based fuzzy time series model to improve forecasting. *Expert Systems with Applications*, *37*(4), 3366–3372. [[CrossRef](#)]
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*(8), 159–175. [[CrossRef](#)]