# FORECASTING WITH THE WARIMAX-GARCH-WANN HYBRID METHOD: A COMPARATIVE ANALYSIS

*By*

**Manzoor Ahmed**

**Registration No.: PIDE2018FMPHILETS02**

*Supervisor*

**Dr Ahsan ul Haq**

**Department of Economics and Econometrics**

**Pakistan Institute of Development Economics, Islamabad**

**2021**

# Pakistan Institute of Development Economics

## CERTIFICATE

This is to certify that this thesis entitled: **"Forecasting with the WARIMAX-GARCH-WANN hybrid method: A comparative analysis"** submitted by Mr. Manzoor Ahmed is accepted in its present form by the Department of Economics & Econometrics, Pakistan Institute of Development Economics (PIDE), Islamabad as satisfying the requirements for partial fulfillment of the degree of **Master of Philosophy in Econometrics**.

External Examiner:

Dr. Iftikhar Hussain Adil
Assistant Professor
School of Social Science and Humanities
NUST, Islamabad

Supervisor:

Dr. Ahsan ul Haq
Assistant Professor
PIDE, Islamabad

Head, Department of Economics & Econometrics:

Dr. Shujaat Farooq
Head
Department of Economics & Econometrics
PIDE, Islamabad

## AUTHOR'S DECLARATION

I Manzoor Ahmed hereby state that my MPhil thesis titled Forecasting with the WARIMAX-GARCH-WANN: A Comparative Approach is my own work and has not been submitted previously by me for taking any degree from Pakistan Institute of Development Economics or anywhere in the country/worldwide.

At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my MPhil degree.

Date: _______________ Signature of Student

Name of Student

# ACKNOWLEDGEMENT

All glories to the Allah, the omniscient and omnipotent, and his Benediction be upon his prophet. The savior of mankind from the darkness of ignorance a symbol to be and to do right. My deepest thanks to Allah the Almighty, who made me able to do this work. I express my heartiest obligation and appreciation to my worthy supervisor Dr. Ahsan ul Haq for their valuable guidance, valuable suggestions, sincere and sympathetic attitude, especially for their politeness during the process of writing this dissertation. Section and the Library staff for their cooperation whenever I needed them. I am also thankful to all my friends and classmates for their cooperation and encouragement. Finally, I would like to express my deepest gratitude to all my family members, especially my mother for her encouragement and endless prayers. Special words of thanks to my father for their continuous support enabled me to achieve the peak of knowledge and career.

**Manzoor Ahmed**

# ABSTRACT

Forecasting methods which incorporate appropriately chosen exogenous variables (EVs) produce enhanced forecasting performances than single variable time series methods. However, suitable exogenous variables are hardly available in practice. This study introduces a new forecasting approach, known as Wavelet Autoregressive Integrated Moving Average with WCs as EVs and Generalized Autoregressive Conditional Heteroskedasticity integrated with Wavelet Artificial Neural Network (WARIMAX-GARCH-WANN) method, to capture the data dynamics and enhance predictive power and accuracy, and, at the same time, address the challenge of non-availability of EVs. The WARIMAX-GARCH-WANN method uses Wavelet Components (WCs) extracted from the wavelet transformation of the underlying time series. These WCs are taken as conventional EVs by WARIMAX-GARCH-WANN method. Like GARCH and ARIMA-GARCH methods, the WARIMAX - GARCH-WANN method is used for high frequency time series which display nonlinear characteristics like non-constant conditional variance that hinges on lagged values of the time series. Moreover, it models frequency structure present in the data series to help achieve better performance in terms of prediction. The application of the WARIMAX-GARCH-WANN method to Wilshire 5000 Price Index commendably outperforms the WARIMAX-GARCH, WANN in terms of performance for both insample and out-of-sample forecast results.

**Keywords:** WARIMAX-GARCH-WANN; WANN; forecasting accuracy

**Table of Content**

## LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

Awareness of potential financial time series (FTS) results is important for efficient policy analysis in the financial sector. Financial sector forecasts provide the investors and fund managers in banks and insurance firms with valuable information to channel their assets correctly for higher returns. Understanding of potential financial time series (FTS) outcomes is essential for efficient policy analysis in the financial sector. For one-step ahead prediction, maintaining the data trend becomes entirely irrelevant. Nonetheless, as the need for h-step ahead forecast horizon arises, preservation of data trend becomes significant but challenging. In either case, high prediction accuracy remains indispensable. As a result, the main requirement for h-period ahead prediction models are retaining a high prediction accuracy and maintaining the data trend throughout the prediction period.

Conventional models like ARIMA, GARCH or ANN fail to acquire both high prediction accuracy and retain the data trend at the same time (Babu & Reddy, 2015). ARIMA models model the conditional mean of a time series and serve as strong forecasting tools when it comes to low frequency data. As high frequency FTS data exhibit time-dependent conditional variance, they fail to capture this volatile characteristic. GARCH models are designed to model the conditional variance. In essence, adding GARCH component to ARIMA models improves the model by taking account of the volatility in high frequency data. A further improvement in forecasting has been experienced in studies by including exogenous variables (EVs) in the ARIMA-GARCH hybrid models, known as ARIMAX-GARCH models. However, in the absence or non-availability of exogenous variables, the researchers will have to compromise on forecast accuracy (Corrêa et al., 2016).

The use of (multi-layer perception) artificial neural networks (ANNs) have remained a common practice in FTS forecasting of late. ANNs have a remarkable capability to learn nonlinear association between input and output structures (Reston Filho et al., 2014). Notwithstanding, the learning process that takes place in neural networks does not take into account the data trend; the process is developed to retain only high prediction accuracy (Babu & Reddy, 2014).

Wavelet Artificial Neural Networks (WANNs) have an edge over ANNs in part due to their fast convergence. The fast convergence takes place as a result of the low correlation between the wavelet neurons. Further, functions dilation and contraction factors improve the network's capability of approximation. In addition to the above advantages, with amazing partial characteristic and multi-resolution learning, the WANN mimics remarkably the signal, which can display functional characteristics with varying resolutions with low forecast error; that is, high accuracy (Zolfagari & Sahabi, 2019).

Correa et al. (2016) use wavelet components (WCs), using discrete wavelet decomposition (DWD) to generate them, as a solution for EVs. It is important to note that WCs obtained from DWD are stationary with a mean zero. However, their proposed Wavelet Autoregressive Integrated Moving Average with wavelet components as EVs and

Generalized Autoregressive Conditional Heteroscedasticity (WARIMAX-GARCH) model achieves enhanced prediction power as long as the stationary EVs are differenced. A serious drawback in differencing a stationary series is negative autocorrelation in the error terms.

This study intends to use wavelet ANN in WARIMAX-GARCH model instead of differencing the EVs to acquire better prediction accuracy along with keeping the data trend.

## 1.1 Research Gap

Several hybrid models have been proposed for high frequency FTS data to obtain high prediction accuracy or to capture the data dynamics or both. The hybrid model WARIMAX-GARCH-WANN, which is proposed in this study, is not seen in any other studies.

## 1.2 Motivation of the Study

All studies conducted so far have concentrated on securing accurate forecasts in financial time series. Keeping in view the volatile nature of FTS data, most of the attention has remained focused on capturing the volatility in the series. However, achieving high prediction accuracy along with keeping the data trend intact has remained the primary goal of almost all of the researchers when it comes to h-period ahead forecasting. Some of them have achieved remarkably outstanding results. Nonetheless, there is still room for further improvement. This study is motivated to bring improvements by retaining data trend and achieving high prediction accuracy simultaneously.

## 1.3 Aims and Objectives

To propose a model that retains the data dynamics, ensures high prediction accuracy and at the same time resolves the problem of exogenous variables using wavelet decomposition techniques while keeping the statistical properties intact.

## 1.4 Significance of the Study

As the proposed method used in this study captures the data dynamics and provides high prediction accuracy, it enables the investors and policy makers to arrive at informed decisions, which helps them to better channel their resources.

# CHAPTER 2

# LITERATURE REVIEW

This chapter gives a brief review of the literature on models used to generate forecasts using high frequency financial time series.

## 2.1 Empirical Review

Several techniques have been postulated for forecasting future outcomes; however, approximately all of them, one way or the other, have attracted some criticism in failing to forecast the events accurately or to fit the data perfectly. The dynamic and unstable nature of the financial time series renders the forecasting task more difficult. FTS data usually exhibit non-constant conditional error variance. Engle (1982) proposed an autoregressive conditional heteroskedastic (ARCH) model, which models the non-constant conditional variance of the innovations. Bollerslev (1986) improved upon it by adding the autoregressive component, which is called the generalized autoregressive conditional heteroskedastic (GARCH) model. The problem with GARCH model is that it can only deal with symmetric series, but several financial time series data follow an asymmetric distribution. In order to cope with this problem asymmetric GARCH models like EGARCH, APARCH, TGARCH, GJR-GARCH etc. were introduced.

The ARIMA type models popularized by Box & Jenkins (1970) with linear auto-dependent features do not serve as appropriate models if the series has volatility in conditional variance of the error term. Even if the conditional and unconditional variances are fixed, it only ensures to maintain the data trend over the forecast period at the cost of prediction accuracy (Babu & Reddy, 2015). The GARCH models, developed by Bolerslev (1986), have an edge over ARIMA models in that they capture volatility in high frequency financial time series. GARCH models in some cases are not appropriate to use as they assume Gaussian distribution. Since FTS usually exhibit fat-tail distributions, student's t-distribution better captures fat-tail features. Johnston & Scott (2000) have further explored the inappropriateness of Gaussian distribution assumption and the use of student t-distribution. GARCH models- assuming student-t distribution or skewed t-distribution – win over the linear ARIMA models.

Nonetheless, GARCH models too only retain the data dynamics at the expense of prediction accuracy. Their hybrid, like ARIMA-GARCH, models also suffer the same fate. Babu & Reddy (2015) propose a partitioning-interpolation based ARIMA-GARCH model and claim to have achieved a better forecasting technique which outperforms all other forecasting models. Notwithstanding, the partition and interpolation (PI) uses trial and error approach to determine the number of partitions. PI technique carried out with 10 partitions does not violate the Nyquist sampling theorem[1], but as the number of partitions approaches 30, PI violates Nyquist sampling rate. An appropriate estimation technique needs to be worked out to get the required number of partitions. If, Babu & Reddy (2015) assert, estimation problem of partitioning the data set is resolved and covariates are included in

[1] The Nyquist Sampling Theorem states that: A band-limited continuous-time signal can be sampled and reconstructed perfectly from the samples if the waveform is sampled over twice as fast as the highest frequency component.

the model, the PI based hybrid of ARIMA-GARCH models yield more accurate predictions while retaining the data trend.

One important development in financial time series forecasting came with the application of (multi-layer perception) ANN in time series data. ANN provides better forecasts than ARIMA and GARCH models; that is, the forecast error is smaller as compared to that of ARIMA and GARCH models. Teixeira et al (2015) proposed a hybrid of Wavelet Decomposition technique with ANN (WD-ANN), also abbreviated as WANN, and showed that the method performs better than the conventional ANN model. The critique on ANN and WANN is that they learn, using the training data, from the data set before prediction. That is, they sacrifice the data dynamics. Khandelwal et al. (2015) propose a hybrid of ARIMA and ANN model. They have shown that ANN and ARIMA are not sufficient in forecasting when applied separately. Their hybrid, ARIMA-ANN, produces remarkable forecast accuracy. Nonetheless, a model with GARCH components explicitly included in the model also captures the GARCH effect in the series and may prove to yield better forecasts.

Another development in time series forecasting which yields good forecasting results is a hybrid of ARIMA with exogenous variable and GARCH model known as ARIMAX-GARCH model. The exogenous variables (EVs), when chosen appropriately, enhance the forecasting performance of the model. Correa et al. (2016) claim that, in practice, exogenous variables are hard to obtain and/or are not available in many cases. In case, when EVs are not available, we have to use ARIMA-GARCH model instead, which generates poor forecasts as compared to ARIMAX-GARCH model. They generate Wavelet Components (WCs) using Wavelet Decomposition and use them as EVs to solve the issue of the non-availability of exogenous variables. Then they use a hybrid of WARIMAX and GARCH to produce h period ahead forecasts. They have shown that this model performs better than ARIMAX-GARCH, ANN and WANN. ARIMAX-GARCH models preserve the data trend and ANN and WANN maintain the prediction accuracy only but not both simultaneously. They propose that if the exogenous variable generated using wavelet components is differenced, WARIMAX-GARCH model has the ability to retain the data dynamics as well as the in both in-sample and out-of-sample forecasting accuracy.

WEVs consisting of WCs of detail are always stationary at level. The detailed components are differenced in order to obtain enhanced forecasting performance and/or get a reasonable model. As Mallat (2008) has pointed out, a wavelet function at level, which generates short duration curve images, has zero mean. From the statistical perspective, it means that the wavelet detail components are stationary around their conditional mean. Now, based on Mallat (2008), since the parameter linked with the spectral frequency of the detail WC is fixed, it shows that there is stationarity in conditional variance. As a result, a detail WC at level will always remain stationary.

Stationarity of the wavelet exogenous variable is a necessary condition that makes the WARIMAX and ARIMAX equivalent models, Correa et al. (2016). The error component of a stationary series has an autocorrelation close to zero. It is to be noted that differencing induces negative autocorrelation in the error term. When the autocorrelation is more negative than -0.5, it implies that the series has been over differenced. In order to solve this issue more lags are to be added in the series.

ANN and WANN provide good prediction accuracy. Instead of differencing the stationary series to achieve better forecasting power, we could combine ANN or WANN with WARIMAX-GARCH model, in which the stationary exogenous variables are not differenced, to capture the data dynamics as well as secure more accurate prediction power. Zolfagari & Sahabi (2019) propose ARIMAX-GARCH-AWNN and AWNN-ARIMAX-GARCH hybrids. They suggest using Mexican-hat, the second derivative of the Gaussian function, as the mother wavelet. The forecast accuracy of the hybrids looks outstanding. As most of the FT series exhibit asymmetry, Haar, Daubechies, or other asymmetric wavelet families could be used to accomplish more accurate forecast results.

## 2.2 Methodological Review

### 2.2.1 Autoregressive (AR) Model

Forecasting techniques in (financial) time series use past and current realizations to predict the future outcomes. The simplest model that relates the past values to the current value of the series to generate a forecast is an autoregressive (AR) process. The AR(p) models originated in the 1920s in the work of Udny Yule, Eugen Slutsky, and some others. The application of autoregressions is first found in the work of Yule in 1927 in which he analyzed the time-series behavior of sunspots (Klein 1997, p. 261). Mathematically,

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t \quad (2.1)$$

where, $y_t$ represents the time series, $\alpha_i$are the parameters of the autoregressive component $y_{t-i}$, $\alpha_0$ is the constant term and $\varepsilon_t$is the white noise error terms ($\varepsilon_t \sim N(0, \sigma^2)$). The lagged values, $y_{t-i}$, capture the first order autocorrelation.

### 2.2.2 Moving Average (MA) Model

It has also been noted that some series depend on the current and $q$ past random shocks, especially when the series exhibits some negative autocorrelation. The resultant model is known as MA(q) model. Symbolically the model is written as:

$$y_t = \beta_0 + \varepsilon_t - \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} \tag{2.2}$$

In the above equation, $\beta_j$are the parameters that measure the specific effect of the past error terms on $y_t$.

### 2.2.3 ARMA Model

Wold (1938) introduced the ARMA(p, q) model, the linear combination of autoregressive and moving average component, for stationary series in his PhD thesis, but he could not drive a likelihood function for estimation of the parameters. In some researchers' view (e.g., Zhang, 2003; Chaâbane, 2014), for linear (smooth) time series the ARMA models show good ability to estimate and forecast. The ARMA(p, q) model is written as:

$$y_t = \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} - \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t \tag{2.3}$$

A mixed model, containing both AR and MA components (though not preferable), sometimes provides a best fit to the data. However, it is possible for the AR and MA terms to cancel out each other's effect, even though their coefficients (as judged by the t-statistics) may appear to be significant. When the series is non-stationary, ARMA(p, q) becomes ARIMA(p, d, q). Box & Jenkins (1970) came up with the full modeling procedure for specification, estimation, diagnostics and forecasting of ARIMA(p, d, q). Symbolically, the model using lag and difference operator is as follows:

$$\Phi\ (B)\Delta^d y_t = c + \Theta\ (B)\varepsilon_t \tag{2.4}$$

Where $\Phi\ (B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\Theta\ (B) = 1 - \theta_1 B - \ldots - \theta_q B^q$ and $\Delta^d = (1 - B)^d$ indicates that the series $y_t$ has been differenced *d* times. $B$is the lag operator and $\Phi$ and $\Theta$ are the coefficients of the lags of $y_t$ and the lags of $\varepsilon_t$ , respectively. The lagged values of $y_t$ and $\varepsilon_t$ capture the remaining autocorrelation after the series has been differenced. The graphs of Autocorrelation function (ACF) and partial autocorrelation function (PACF) are used in choosing the number of AR and MA components in the model.

### 2.2.4 ARIMAX Model

Studies, including Prankratz(1991), Box & Tiao (1975), Williams (2001) and Jalalkamali et al. (2015) to name a few, have evidenced that incorporating relevant exogenous variables in ARIMA model improves the forecasting performance of the model. The specification of the ARIMAX (*p, d, q*) model is as follows:

$$\Phi\ (B)y_t = c + \Theta\ (B)\varepsilon_t + \beta x_t \tag{2.5}$$

The estimation of the above models requires covariance stationarity of $y_t$and homoscedasticity in the error terms. In financial time series, heteroskedastic conditional variances have been noticed. However, the assumption of covariance stationarity requires the unconditional variance to be constant over time.

### 2.2.5 ARCH Model

Engle (1982) noted that high frequency FTS exhibits conditional heteroskedastic variance; that is, FTS show volatility. He developed the autoregressive conditional heteroskedastic (ARCH) model to model the conditional variance. He observed that conditional (error) variance of FTS tends to depend on the past squared error terms.

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}^2 \tag{2.6}$$

The presence of ARCH effect is tested using the Breusch-Pagan test. Usual OLS estimation technique is used on the mean model to obtain the residuals. Then an auxiliary regression of the squared residuals is run on the lags of the squared residuals (Enders, 2015).

$$\hat{\varepsilon}_t^2 = \alpha + \sum_{j=1}^{q} \beta_j \hat{\varepsilon}_{t-j}^2 + u_t \tag{2.7}$$

And then compute $R^2 \times T$. Under the hypothesis of homoscedasticity, the test statistic follows a $\chi^2$ distribution. The ARCH(q) effect is present when the null hypothesis of homoscedasticity is rejected.

A serious drawback in ARCH(q) model is that it is overparameterized. We need to estimate $q+1$ parameters to obtain the required estimated values of $\sigma_t^2$.

### 2.2.6 GARCH Model

Taylor (1986) and Bolserlev(1986) independently incorporated the autoregressive component in the ARCH model to improve its fit. The resultant model is known as the generalized autoregressive conditional heteroskedastic (GARCH) model. The specification of GARCH models is as follows:

$y_t = \mu + \varepsilon_t$, where $\mu$is a constant bias term,$\varepsilon_t = \sigma_t z_t$ and$z_t \approx N(0,1)$.

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}^2 \quad (2.8)$$

Where $\omega$ is the intercept component, $\beta_j$is a measure of the reaction to conditional volatility of market shocks and $\alpha_i$is a measure of the persistence of shock in conditional volatility regardless of anything happening in the market. Black(1976) has evidenced that the returns are inversely associated with the changes in volatility of return volatility. As (conditional) variance is non-negative, it becomes a necessary condition for the coefficients to be positive. Therefore, GARCH models place nonnegativity constraints on the parameter that often get violated by the coefficient estimates and that may unnecessarily limit the dynamics of the conditional variance process. In addition, these constraints of nonnegativity can make estimation of GARCH models difficult. For instance, Engle, Lilien, and Robins (1987) had to impose a linearly decreasing structure on the $\beta_j$'s in (2.8) in order to stop some of coefficients from having negative values. Another issue in GARCH models is interpreting whether disturbances to conditional variance persist or are transitory, because the normal ways of the measurement of persistence often differ. The main question, in several studies (e.g., Poterba and Summers (1986), French, Schwert, and

Stambaugh (1987) and Engle and Bollerslev (1986a)), conducted on the time series behaviour of asset volatility, has been the length of shock persistence to volatility. If shocks to volatility persist for an indefinite time period, they may shift the entire risk premia term structure and are, therefore, more probable to have a great effect on investment in long-lived capital goods (Poterba and Summers, 1986).

### 2.2.7 GARCH-M Model

The GARCH-M model of Engle and Bollerslev (1986a) introduces another equation.

$$y_t = \mu + \lambda\sigma_t^2 + \varepsilon_t, \tag{2.9}$$

in which $\sigma_t^2$, the conditional variance of $y_t$, is used in the conditional mean equation of $y_t$. The coefficient of$\sigma_t^2$, $\lambda$, represents the feedback effect.

### 2.2.8 EGARCH Model

Nelson (1991) highlighted the presence of asymmetric effect; that is, the leverage effect. EGARCH model maps the impact of unexpected external shocks on the predicted volatility due to the exponential nature of the conditional variance. The EGARCH(*p, q*) is formulated as:

$$ln(\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \sum_{k=1}^{r} \gamma_k \frac{\varepsilon_{t-k}}{\sigma_{t-k}} + \sum_{j=1}^{q} \beta_j \, ln(\sigma_{t-j}^2) \tag{2.10}$$

Where $\alpha_i$ and $\beta_j$ show the ARCH and GARCH effects, respectively, and $\gamma_k$ denotes the asymmetric effect of disturbance on conditional variance and a negative value of this parameter is an indication of leverage effect. The logarithmic form of the conditional variance implies the existence of exponential asymmetric effect and that the forecasts from the conditional variance have nonnegative outputs (Thomas and Mitchell, 2005).

### 2.2.9 GJR-GARCH Model

Glosten et al. (1993) proposed GJR-GARCH model which also allows for leverage effect. Mathematically GJR-GARCH(p, q) model is written as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \gamma \varepsilon_{t-1}^2 I_{t-1} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \qquad (2.11)$$

Where, $\omega \geq 0,\ \alpha_i \geq 0,\ \alpha_i + \gamma \geq 0,\ \sum_{i=1}^{p} \alpha_i + \gamma + \sum_{j=1}^{q} \beta_j < 1$. $I_t = 1$ if $\varepsilon_t < 0$, and 0 otherwise. Studies like Lu et al. (2016) have shown that hybrid models formed with GJR-GARCH have high forecast performance.

### 2.2.10 ARIMA-GARCH Model

Chand et al. (2012), Hickey et al. (2012), Erdogdu (2010), Sumer et al. (2009), Nury et al. (2017) and Zolfaghari and Sahabi (2017) suggest using models of the form ARIMA-GARCH to describe the structure of the residuals' variance obtained from the best fit time series mean models. The output of the appropriate GARCH model is put in ARIMA(p, d, q) to obtain the following ARIMA-GARCH model:

$$\Phi\ (B) y_t = c + \Theta\ (B) \varepsilon_t + \varphi \sigma_t^2 \qquad (2.12)$$

### 2.2.11 ARIMAX-GARCH

Incorporating the output of GARCH model in ARIMAX gives the ARIMAX-GARCH hybrid model. Studies like Paul & Himadri (2014), Corrêa et al. (2016) and

Zulfaghari & Sahabi (2019) have shown that ARIMAX-GARCH yields better forecasts accuracy than ARIMA-GARCH. Mathematically the model is written as:

$$\Phi\ (B)y_t = c + \Theta\ (B)\varepsilon_t + \beta x_t + \varphi\sigma_t^2 \tag{2.13}$$

The prediction accuracy of the above models decreases as we employ these models for h-step ahead prediction as noted by Babu & Reddy (2014).

### 2.2.12 Artificial Neural Networks (ANNs)

The ANN, due to its potential to learn nonlinear relationship between input and output patterns, is employed in FTS data to forecast the volatilities (Reston Filho et al. 2014).

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g(\beta_{0j} + \sum_{i=1}^{L} \beta_{ij} y_{t-i}) + \varepsilon_t \tag{2.14}$$

Where, $\alpha_j (j = 0,1,\ldots,q)$ and $\beta_{ij} (i = 0,1,\ldots,L; j = 0,1,\ldots,q)$ are known as connection weights; k shows the count of input nodes; q denotes the number of hidden nodes. $\varepsilon_t$ represents the approximate error term. Since g(.) is a nonlinear transfer function, the ANN model above, it performs a nonlinear mapping of the lagged data points $y_{t-i}$ to generate forecasts for $y_t$. The transfer function used as the hidden layer in ANN is a logistic function and is mathematically defined by

$$g(\beta_{0j} + \sum_{i=1}^{L} \beta_{ij} y_{t-i}) = \frac{exp(\beta_{0j} + \sum_{i=1}^{L} \beta_{ij} y_{t-i})}{1 + exp(\beta_{0j} + \sum_{i=1}^{L} \beta_{ij} y_{t-i})} \tag{2.15}$$

Generally, ANN is written as

$$y_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-L}, w) + \varepsilon_t \tag{2.16}$$

Where $w$represents the vector of parameters in ANN model. In practice $w$is unknown. The numerical estimation requires an objective function that applies an algorithm to optimize the weights of training data. The Levenberg-Marquardt's algorithm[2] is used as the numerical criteria to reduce the in-sample sum of squared errors.

$$y_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-L}, \widehat{w}) + \hat{\varepsilon}_t, \tag{2.17}$$

with $\hat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-L}, \widehat{w})$ containing the optimal ANN end results at time $t$, which produces forecasts of $y_t$.

### 2.2.13 Wavelets

Wavelet analysis, in recent times, has found great use in non-stationary time series. Wavelets offer an alternative to the window-based techniques for examining localized frequency behaviour of the data whose characteristics vary across time. It can be viewed as an extention of the classical Fourier analysis in which short waves, called wavelets, are used in place of sines and cosines. Discrete wavelet transformation (DWT) separates the signal into a time domain and a frequency domain at the same time. On the contrary, Fourier Transform decomposes the raw signal (time domain) into processed signal (frequency domain); time information is complete last (Ortega & Khashanah, 2014).

Any function $f(t)$ can be split into component parts by a series of projections onto the wavelet basis:

[2] The Levenberg – Marquardt algorithm, also known as the Damped least-squares method, is used in mathematics and computing to solve problems with non-linear minus squares.

$$S_{m,n} = \int f(t)\Phi_{m,n}(t)dt \quad (2.18)$$

$$d_{m,n} = \int f(t)\psi_{m,n}(t)dt \quad (2.19)$$

where $m$ denotes the count of multiresolution; $\Phi$ represents the father wavelet and $\psi$ is the mother wavelet. $S_{m,n}$ and $d_{m,n}$ represent the smooth and the detailed coefficients, respectively. $\Phi_{m,n}$ and $\psi_{m,n}$, the scaling and translation of $\Phi$ and $\psi$, are defined by

$$\Phi_{m,n}(t) = 2^{-m/2}\Phi(2^{-m}t - n) \quad (2.20)$$

$$\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n) \quad (2.21)$$

A number of wavelets like Haar, Daubechies, Morlet and Mexican Hat, to name a few, are available. Morlet and Daubechies wavelets have useful applications in image processing but are not without problems. Mexican Hat wavelets are hard to compute. Hair wavelets are useful for time series analysis because they capture variations between adjacent observations (Ortega & Khashanah, 2014; Lahmiri, 2014; Murtagh, Starck, & Renaud, 2004; Li, Li, Zhu, & Ogihara, 2002).

### 2.2.13 Wavelet Decomposition (WD)

The proposed hybrid model uses discrete wavelet decomposition to split the series $y_t$ into an approximate component $\tilde{y}_{A_{m_0},t}$ and several detail components$\sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t}$. Type of the time series' variations determines the number of detail components (Zulfaghari & Sahabi 2019).

$$y_t = \tilde{y}_{A_{m_0},t} + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t} + \varepsilon_t \qquad (2.22)$$

Where, $\tilde{y}_{A_{m_0},t} = \sum_{n=0}^{2^{(M-m_0)}-1} a_{m_0,n}\, \phi_{m_0,n}(t)$ and $\tilde{y}_{D_m,t} = \sum_{n=0}^{2^{(M-m_0)}-1} d_{m,n}\omega_{m,n}(t)$, with $t = 1,\dots,T$. Diagrammatically, wavelet decomposition can be depicted as below:

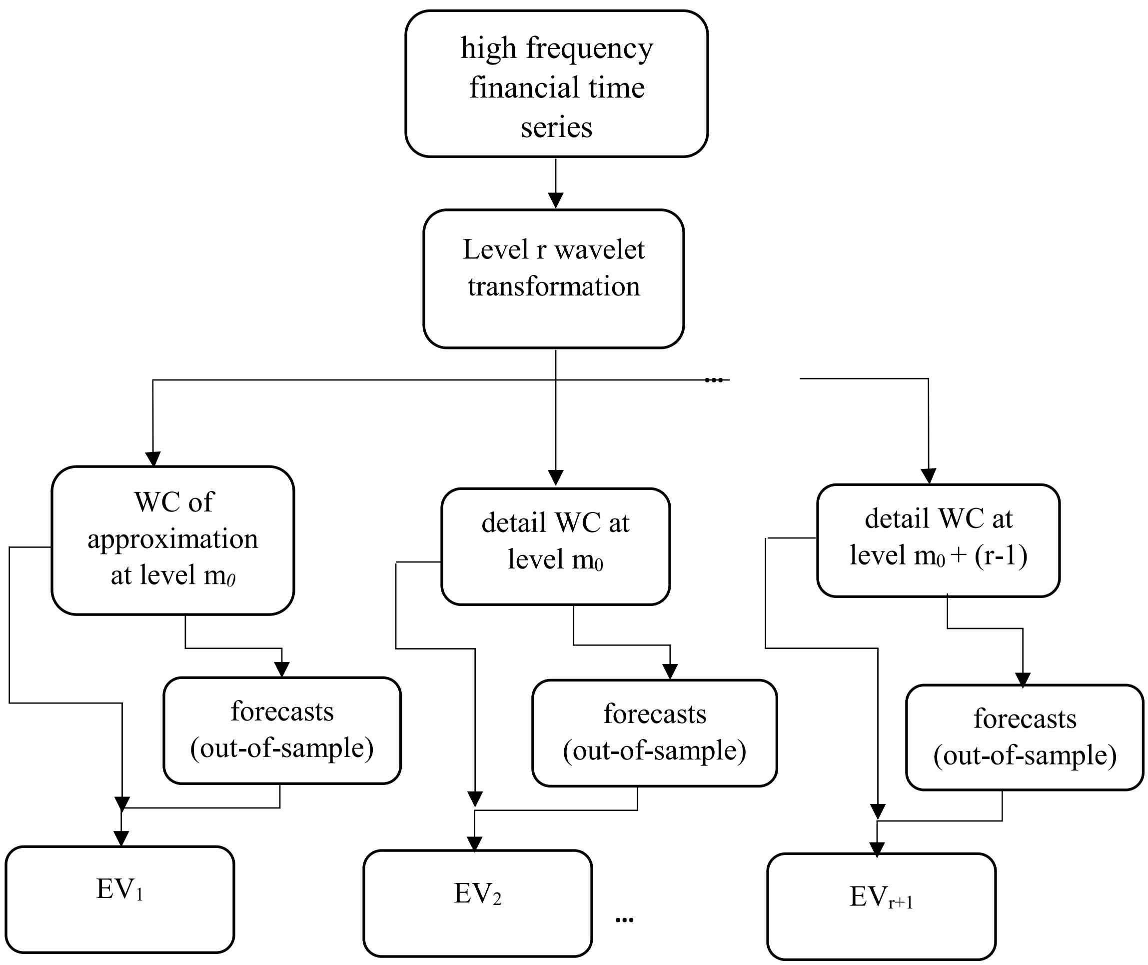



**Figure 2.1: The Flowchart of the Discrete Wavelet Decomposition**

DWT comes up with a couple of limitations. Firstly, DWT requires that the of size of the dataset be dyadic; that is, a power of 2. Secondly, the output produced by DWT remain highly dependent on origin of the signal being analyzed. A slight change in origin has effects on the produced outputs, and this issue is referred to as Circular Shift. Owing to the circular shift the transformed signals are difficult to match with time. An alteration of DWT called Maximal Overlap Discrete Wavelet Transformation (MODWT) is used to resolve the above two limitations. MODWT is invariant to circular shift and is not constrained by the restriction of the dyadic length; hence, the signal interpretation becomes easier for time series analysis. In this study, DWT and MODWT will be used interchangeably. Gençay, Selçuk, & Whitcher (2002), Ortega & Khashanah (2014) and Lahmiri (2014) practically apply wavelets in Finance and Economics.

### 2.2.14 Wavelet Artificial Neural Network (WANN)

Wavelet decomposition combined with ANN, called wavelet ANN, as shown in several studies (e.g., Teixeira Júnior, 2015), achieves remarkable predictive accuracy in forecasting in time series. For the series $y_t$ $(t = 1, \dots, T)$, WANN is constructed in two steps:

Step 1: A wavelet decomposition of $y_t$ at level r is carried out, creating one approximate WC at level $m'$, denoted by $y_{A_{m'},t}$ $(t = 1, \dots, T)$, and r detail components of wavelet transform at levels $m', m' + 1, \dots, m' + (r - 1)$, denoted by $\tilde{y}_{D_m,t}$ $(t = 1, \dots, T)$, where r, $m' \in \mathbb{Z}$; and

Step 2: The WCs, both of approximation and detail, then are modelled using ANN. The general form of wavelet ANN model is given by

$$y = f(y_{A_{m',t,L}}; y_{D_{m',t,L}}; \ldots; y_{D_{m'+m'+(r-1)',t,L}}; \widehat{w}) + \hat{\varepsilon}_t \tag{2.23}$$

Where $y_{A_{m',t,L}} = (y_{A_{m',t}}, \ldots, y_{A_{m',t-K}})$ and $y_{D_{m',t,L}} = (y_{D_{m,t}}, \ldots, y_{D_{m,t-L}})$ with the input data $m = m', m' + 1, \ldots, m' + (r - 1)$. The Levenberg-Marquardt's algorithm is used to obtain the optimal solution for $\widehat{w}$; such that $min_w \sum_{t=1}^{T} \varepsilon_t{}^2$is obtained. The forecasts of $y$are given as

$\hat{y} = f(y_{A_{m',t,L}}; y_{D_{m',t,L}}; \ldots; y_{D_{m'+m'+(r-1)',t,L}}; \widehat{w})$, and $\hat{\varepsilon}_t$represents the forecast error of$\hat{y}$. The time subscript *t* is removed from the output,$y$, to differentiate it from the variable used dependent variable in WARIMAX-GARCH-WANN. Following is a general diagram of the WANN:

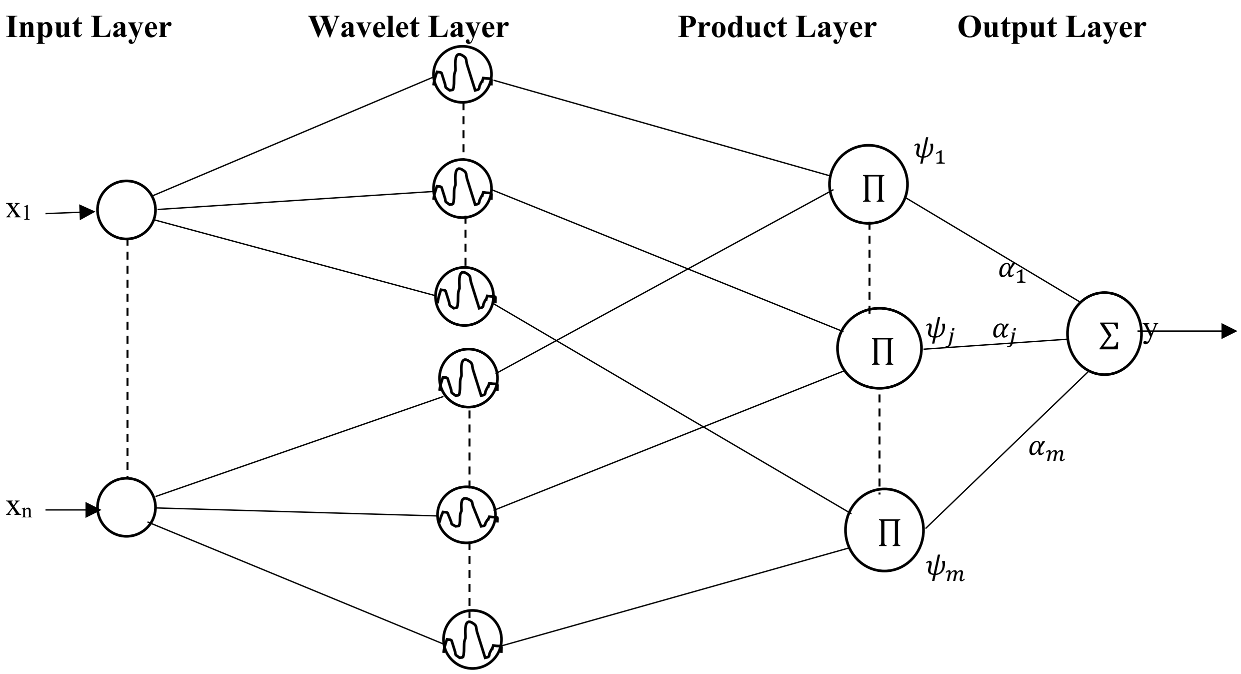



**Figure 2.2: The General Architecture of a WANN Model**

### 2.2.15 ARIMAX-GARCH-AWNN

Zulfaghari & Sahabi (2019) propose the following ARIMAX-GARCH-AWNN hybrid model and achieve enhanced prediction power. Where AWNN stands for adaptive wavelet neural networks.

$$\Phi\ (B)y_t = c + \Theta\ (B)\varepsilon_t + \beta x_t + \varphi\sigma_t^2 + \delta y \tag{2.24}$$

Where $y = \sum_{j=1}^{m} \omega_j \phi_j + \sum_{i=1}^{n} \nu_i x_i + g$, $\omega_j$ is the weight of the layer between the *j*th node of the product layer and the output node and $\nu_i$ is the weight of input between the *i*th input node and output node.

### 2.2.16 WARIMAX-GARCH

The composite model proposed by Correa et al. (2016), known as WARIMAX-GARCH (p, d, q)x(P, D, Q) model, is given by

$$\phi(B)\Phi(B^S)\nabla^d\nabla_S^D y_t = \varphi_{A_{m_0}}\nabla^k g(\tilde{y}_{A_{m_0},t,C}) + \sum_{m=m_0}^{m_0+(r-1)} \nabla^{k_m}\tilde{y}_{D_m,t,C}\varphi_{D_m,C} + \theta(B)\Theta(B^S)e_t + \varphi\sigma_t^2 \tag{2.25}$$

Where, $\phi(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \ldots - \theta_q B^q)$, $\nabla^d = (1 - B)^d$, $\nabla_S^D = (1 - B^S)^D$, $\Theta(B^S) = (1 - \theta_1 B^S - \ldots - \theta_Q B^{SQ})$, and $k \in \mathbb{Z}$.

WARIMAX-GARCH with exogenous variables at level becomes:

$$\phi(B)\Phi(B^S)\nabla^d\nabla_S^D y_t = \varphi_{A_{m_0}}\nabla^k g(\tilde{y}_{A_{m_0},t,C}) + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t,C}\varphi_{D_m,C} + \theta(B)\Theta(B^S)e_t + \varphi\sigma_t^2 \tag{2.26}$$

Each WC obtained in section (2.2.15) is modeled separately using a specific ARIMA-GARCH to produce out-of-sample forecasts. These forecasts produce the completed WCs (CWCs). Algebraically,

$\tilde{y}_{A_{m_0},t,C}$ or $x_{1,t}(t = 1, \ldots T, T + 1, \ldots, T + h)$, which is the CWC of approximations at level $m_0$of $y_t(t = 1, \ldots, T)$. And $\tilde{y}_{D_m,t,C}$ or $x_{i,t}(t, \ldots, T, T + 1, \ldots, T + h; i = 1, \ldots, r + 1)$ is composed of the CWCs of detail at level $m$ of $y_t(t = 1, \ldots, T)$, where $m_0 \leq m \leq m_0 + (r - 1)$. The above $r + 1$ CWCs are used as wavelet exogenous variables in equations (2.25) and (2.26).

# CHAPTER 3

# DATA & METHODOLOGY

High frequency FTS like Wilshire 5000 price index is used to compare the forecasting performance of the proposed model with the other existing models. The output of equation (2.23) is put in equation (2.26) to form the WARIMAX-GARCH-WANN model.

$$\phi(B)\Phi(B^S)\nabla^d\nabla_S^D y_t = \varphi_{A_{m_0}} g(\tilde{y}_{A_{m_0},t,C}) + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t,C}\varphi_{D_m,C} + \theta(B)\Theta(B^S)e_t + \varphi\sigma_t^2 + \delta y \quad (3.1)$$

When the series does not exhibit seasonal patterns, the model becomes:

$$\phi(B)\nabla^d y_t = \varphi_{A_{m_0}} g(\tilde{y}_{A_{m_0},t,C}) + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t,C}\varphi_{D_m,C} + \theta(B)e_t + \varphi\sigma_t^2 + \delta y \quad (3.2)$$

The stock prices are affected by numerous shocks to the economy like policy changes, disasters, pandemics, economic recessions, etc. Such shocks may persist indefinitely unless the government implements a policy change to counter the shock. The 2008 recession and the Covid-19 pandemic appear to have the effects on stock prices and the economy of the countries at large. Therefore, Bai-Perron test was used to find out the number of structural breaks in the series which might have been caused by different shocks to the economy. Fourier ADF and Fourier LM tests, developed by Enders & Lee (2012a,b), were used to confirm whether the Wilshire 5000 price index contained any unit roots. As high frequency FTS exhibit linear and nonlinear characteristics, DW and BDS tests were applied to test for the presence of linear and nonlinear autocorrelation. The lag order will be chosen on the basis of Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) graphs. The selection of best

WARIMAX-GARCH model is based on the forecasting performance that is measured by their Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE); and model based on Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) also known as Schwardz Information Criteria (SIC). Model with lowest RMSE, MAPE, and MAE value and with the greatest AIC or BIC is considered better than others WARIMAX-GARCH models are.

# CHAPTER 4

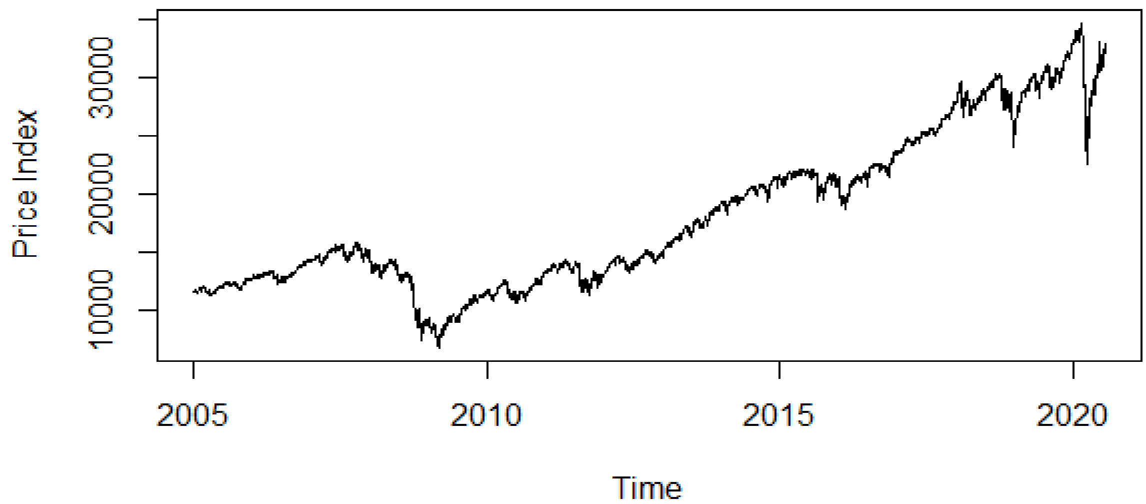
# EMPIRICAL RESULTS AND DISCUSSIONS

This chapter presents the obtained results of the WARIMAX- GARCH-WANN model applied to the time series of daily Wilshire 5000 price index. For comparison, WARIMAX-GARCH, and WANN methods have also been used on the price index data. In section 4.1, for the justification of the selected model, a statistical analysis of the time series is carried out. In Section 4.2, an appropriate WARIMAX-GARCH model was chosen to model the price index series. Section 4.3 contains the application of WARIMAX-GARCH-WANN method, and its key statistical tests that determine the validation of the model. Section 4.4 compares the forecasting performance of the proposed method with the WARIMAX-GARCH and WANN methods.

## 4.1 The daily time series of Wilshire 5000 price index

The daily observations for the price index, price returns, are the daily index value computed at market close. A daily time series of Wilshire 5000 price index from 6$^{th}$ of January 2005 to 17$^{th}$ of July 2020 was used for the purpose of testing the forecasting power of the proposed method as compared with other leading models as used in Correa et al. (2016). Figure 4.1 depicts the time series plot of the 3907 daily data points. The first 3877 data points were allocated as in-sample training set and the remaining 30 data points were allocated for out-of-sample model testing. In this study no exogenous variables were used.

**Figure 4.1: Wilshire 5000 Daily Price Index from 2005 to 2020**




It can be observed from Figure 4.1 that the series exhibits sharp drops around periods 2008, 2020 and in between, which are indicative of volatility in the series. An upward trend is evident from the graph which shows that the series evolves around a nonconstant mean. And more importantly, the series clearly shows high frequency characteristics. An ARIMA model integrated with wavelet decomposition and a GARCH component can capture the high-frequency traits along with the volatile nature of the series.

**Table 4.1a: Bai-Perron multiple breakpoint test of Wilshire 5000 price index**

| Breaks | No. of Coefficients | Sum of Sq. Residuals | Log-L | Schwarz Criterion | LWZ Criterion |
|---|---|---|---|---|---|
| 0 | 1 | 1.71E+11 | -39903.96 | 17.59637 | 17.60098 |
| 1 | 3 | 3.91E+10 | -37023.83 | 16.12589 | 16.13971 |
| 2 | 5 | 1.51E+10 | -35161.14 | 15.17637 | 15.19940 |
| 3 | 7 | 1.32E+10 | -34900.56 | 15.04717 | 15.07942 |
| 4 | 9 | 1.04E+10 | -34445.89 | 14.81860 | 14.86006 |
| 5 | 11 | 9.81E+09 | -34321.63 | 14.75921 | 14.80989 |

Bai-Perron multiple breakpoint test, using the global information criteria specification, identified five breakpoints in series. Both the Schwartz criterion and LWZ criterion selected five structural breaks on the basis of minimum information criterion values as displayed in shading. Though the graph of the series displays more than five breaks in the series, Bai-Perron test selects only five structural breaks. Nonetheless, the test results and the graph suggest that the series has multiple breakpoints. Therefore, ordinary unit root tests - like ADF, KPSS, PP, Narayan and Popp, to name a few – do not offer a solid testing ground for non-stationarity in the series.

**Table 4.1b: Unit root tests under multiple unknown structural breaks at level**

| Model / Test | Break in Level | | Break in Level and Trend | |
|---|---|---|---|---|
| | Test statistic | Critical Value (1%) | Test statistic | Critical Value (1%) |
| Fourier ADF | -1.751 | -4.310 | -4.528 | -4.800 |
| Fourier LM | -4.328 | -4.560 | -- | -- |

**Table 4.1c: Unit root tests under multiple structural breaks at first difference**

| Model / Test | Break in Level | | Break in Level and Trend | |
|---|---|---|---|---|
| | Test statistic | Critical Value (1%) | Test statistic | Critical Value (1%) |
| Fourier ADF | -16.532 | -3.610 | -16.583 | -4.240 |
| Fourier LM | -9.997 | -4.560 | -- | -- |

Fourier ADF and Fourier LM tests for unit root with multiple breakpoints were carried out on the series at level. The null and alternative hypotheses of the two tests are H0: Unit root series with unknown number of level breaks and Ha: Stationary process with unknown number of level breaks, respectively. Both the tests did not reject the presence of unit root at 1% level of significance. However, the tests applied to the first difference gave highly significant results at 1% level as given in Table 4.2. The Autocorrelation Function (ACF) in figure 4.2a exhibits spikes which do not die out, showing the non-stationarity of the series at level; however, the ACF of the first difference in figure 4.2c tapers off after the first lag, indicating stationarity. PACF in figure 4.2b shows a significant spikes at lag one and two which indicate that the appropriate model is ARIMA. The LM test for ARCH effect gave a chi-square value of 486.9426 with a p-value of 0.000 rejecting the null hypothesis of no arch effect at 1% level of significance, which is evident that GARCH would be a suitable choice to capture the volatility. Hence, as the study deals with a single series with its WCs, a WARIMAX-GARCH model was deemed reasonable to model the process.

The conducted investigation produced multi-step forecast values, of both the out-of-sample and in-sample, from WARIMAX-GARCH model (used for comparison purposes) and from the WARIMAX-GARCH-WANN model. For statistical validation, Fourier ADF and Fourier LM tests for unit root, Ljung-Box, DW and BDS tests were conducted for randomness in the error terms and first order autocorrelation, respectively. The ACF and PACF plots were constructed, and BDS test was carried out to check for the presence of nonlinear serial auto-dependent residuals and ARCH test was conducted to see if conditional heteroskedasticity existed. R, GAUSS and Eviews softwares were used to obtain the ACF and PACF graphs and to perform the above tests.

Figure 4.2a: ACF of the Wilshire 5000 price index at level

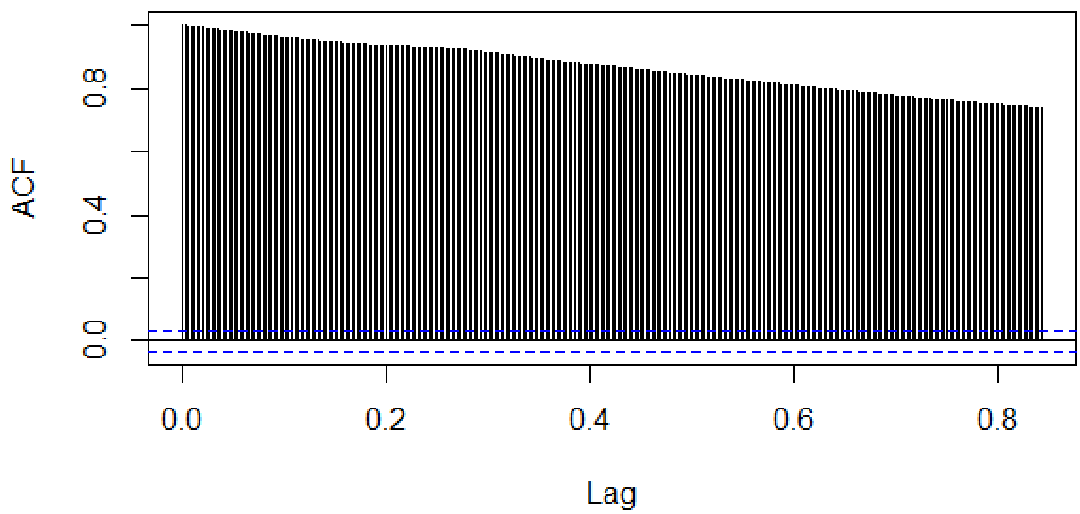



Figure 4.2b: PACF of the Wilshire 5000 price index at level

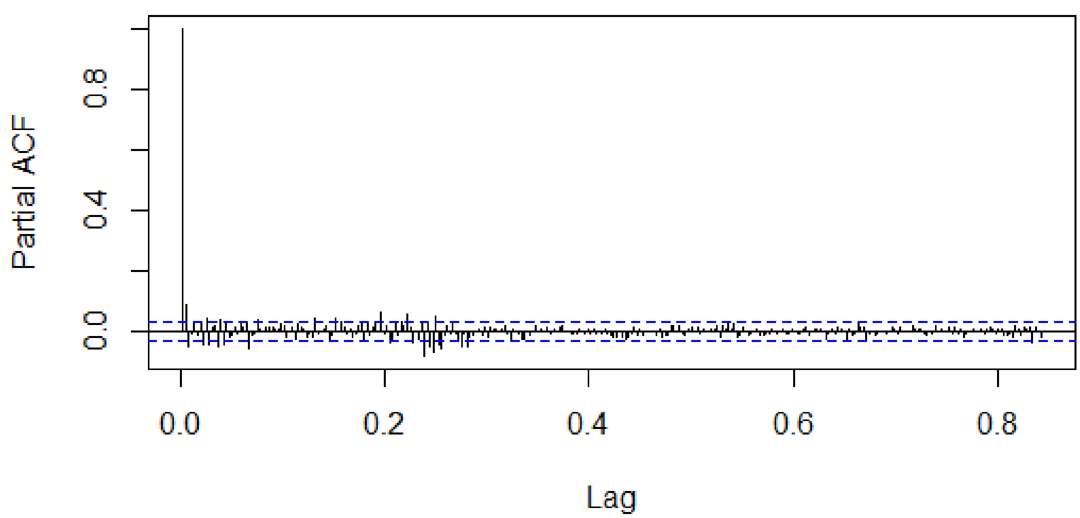


**Figure 4.2c: ACF of the Wilshire 5000 price index at first difference**

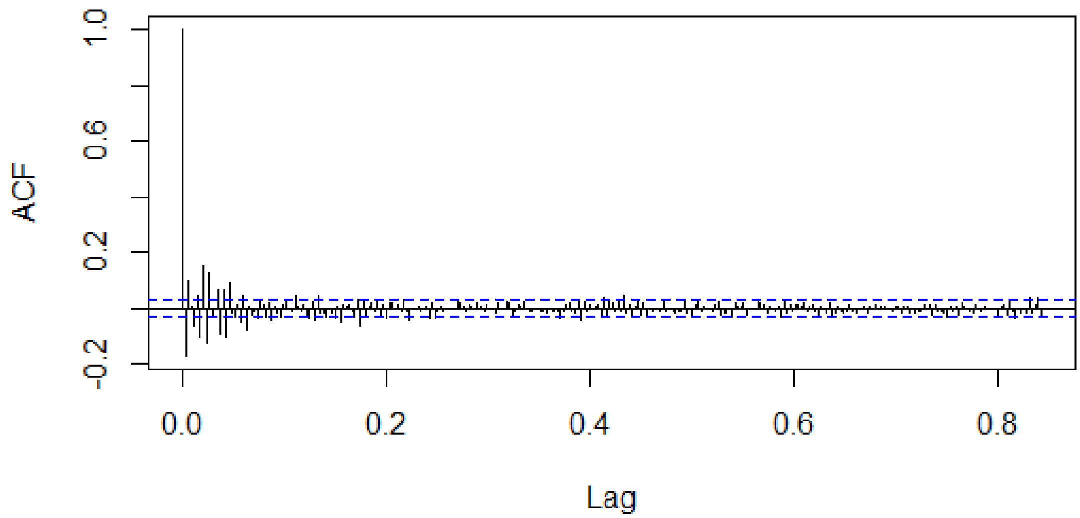
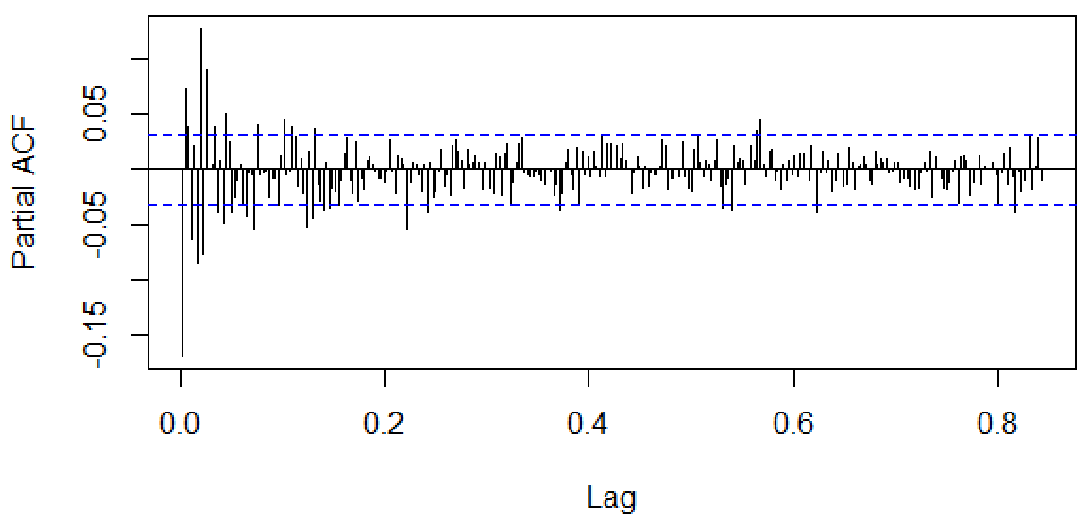



**Figure 4.2d: PACF of the Wilshire 5000 price index at first difference**



### 4.2 The WARIMAX-GARCH Method

A level-2 wavelet transformation was used to split the time series into an approximate and two detail components in MATLAB (version 2020a) software. A WARIMAX- GARCH (2, 1, 1) x (1, 1) model using WCs as exogenous variables, with student's t-distribution, came out to be the best fit to the first difference of Wilshire 5000 price index, $y'_t$ where $t$ = 1, …, 3903. The MLE gave statistically significant results of the parameter estimates at 1% significance level except the conditional variance coefficient, which is significant at 5% significance level.

**Table 4.2a: the WARIMAX-GARCH(2,1,1)x(1,1) estimation output**

| Variable | Coefficient | Standard Error | t-statistic | P-value |
|---|---|---|---|---|
| $\triangle\, y_{t-2}$ | 0.222645 | 0.012456 | 17.87383 | 0.0000 |
| $e_{t-1}$ | 0.454774 | 0.019129 | 23.77443 | 0.0000 |
| $\triangle\, y_{A_2,t-1}$ | -0.343622 | 0.022390 | -15.34706 | 0.0000 |
| $y_{D_1,t}$ | 0.646617 | 0.023025 | 28.08385 | 0.0000 |
| $y_{D_1,t-1}$ | -1.392607 | 0.035096 | -39.67936 | 0.0000 |
| $y_{D_2,t}$ | 0.233658 | 0.021610 | 10.81236 | 0.0000 |
| $y_{D_2,t-1}$ | -0.836907 | 0.022907 | -36.53539 | 0.0000 |
| $ln\sigma_t^2$ | -407.5250 | 174.6999 | -2.332715 | 0.0197 |

**Table 4.2b: BDS test outcomes from the ordinary standard residuals of WARIMAX-EGARCH(2,1,1)(1,1) model.**

| Dimension | BDS Statistic | Std. Error | z-Statistic | p-values |
|---|---|---|---|---|
| 2 | -0.013065 | 0.002036 | -6.416787 | 0.0000 |
| 3 | 0.007444 | 0.003248 | 2.292251 | 0.0219 |
| 4 | 0.007998 | 0.003884 | 2.059235 | 0.0395 |
| 5 | 0.026923 | 0.004068 | 6.619040 | 0.0000 |
| 6 | 0.030745 | 0.003942 | 7.799209 | 0.0000 |

Among all possible WARIMAX-GARCH models, the WARIMAX-EGARCH (2, 1, 1) x (1, 1) produced more accurate forecasts in terms of in-sample RMSE, MAE and MAPE. The outcomes of the BDS test, shown in Table 4.2b, applied to the residuals of WARIMAX-EGARCH (2, 1, 1) x (1, 1) reject the null hypothesis of linear auto-dependence in the ordinary least squares residuals. In addition, the computed DW statistic of 2.162100 indicates no significant first order autocorrelation in the residuals.

The BDS test was used to determine if there existed linear auto-dependence in the residuals of WARIMAX-EGARCH (2, 1, 1) x (1, 1) model. The p-value of dimensions 2, 5 and 6 suggest the existence of no linear dependence on its own lagged values at 1% significance level and the p-values of dimensions 3 and 4 suggest no linear aut-dependence in the errors at 5% significance level. Which suggests that the series better mapped by WARIMAX-EGARCH (2,1,1) x (1,1) as compared with the other WARIMAX-GARCH models. The model selection was based on the lowest MAPE, MAE and RMSE values, and largest $R^2$.

### 4.3 The WARIMAX-GARCH-WANN Model

MATLAB (version 2020a) was used to obtain a wavelet transformation of level 2. The plots of the approximate and detail WCs are shown in Figures 4.3a, 4.3b and 4.3c below.

**Figure 4.3a: Wavelet Component of Approximation at level 2**

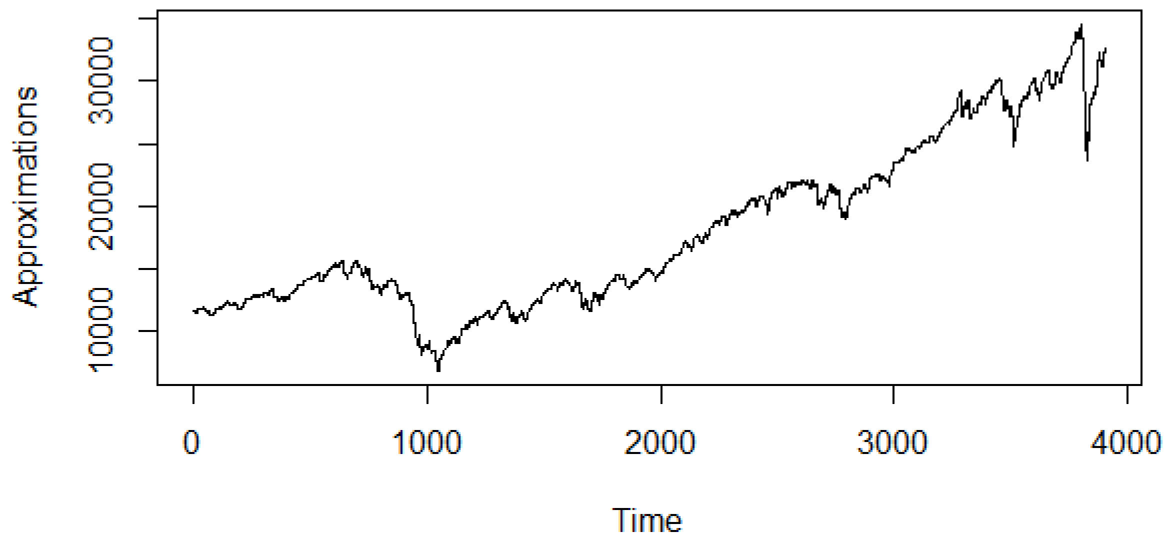



**Table 4.3a: Bai-Perron Multiple Breakpoint Test of Wavelet Component of Approximation at level 2**

| Breaks | No. of Coefficients | Sum of Sq. Residuals | Log-L | Schwarz Criterion | LWZ Criterion |
|---|---|---|---|---|---|
| 0 | 1 | 1.71E+11 | -39902.71 | 17.59573 | 17.60034 |
| 1 | 3 | 3.90E+10 | -37018.35 | 16.12308 | 16.13690 |
| 2 | 5 | 1.50E+10 | -35147.16 | 15.16921 | 15.19224 |
| 3 | 7 | 1.31E+10 | -34884.52 | 15.03896 | 15.07121 |
| 4 | 9 | 1.03E+10 | -34424.71 | 14.80776 | 14.84921 |
| 5 | 11 | 9.69E+09 | -34299.05 | 14.74765 | 14.79832 |

The graph of the WC of approximation at level 2 clearly depicts several breakpoints. Therefore, to formally test for the breaks, Bai-Perron multiple breakpoint test, using the global information criteria specification, was carried out and it identified five breakpoints in the WC of approximation series. Both the criteria, Schwartz and LWZ, selected five breakpoints on the basis of minimum information criterion values as shown in shading. Although the graph shows more than five breakpoints, Bai-Perron test, due to its limitations, identifies five breaks from 2008 to 2017. As a result, Fourier ADF and Fourier LM were considered as the most appropriate unit root test with multiple structural breaks. The tests were applied to the series at level and at first difference. The results are shown in the tables below.

**Table 4.3b: Unit root tests under multiple unknown structural breaks at level**

| Model / Test | Break in Level | | Break in Level and Trend | |
|---|---|---|---|---|
| | Test statistic | Critical Value (1%) | Test statistic | Critical Value (1%) |
| Fourier ADF | -1.617 | -4.310 | -4.326 | -4.800 |
| Fourier LM | -4.133 | -4.560 | -- | -- |

**Table 4.3c: Unit root tests under multiple structural breaks at first difference**

| Model / Test | Break in Level | | Break in Level and Trend | |
|---|---|---|---|---|
| | Test statistic | Critical Value (1%) | Test statistic | Critical Value (1%) |
| Fourier ADF | -16.255 | -3.610 | -16.310 | -4.240 |
| Fourier LM | -8.477 | -4.560 | -- | -- |

Both the test did not reject the null hypothesis of unit root with multiple unknown structural breaks at level. Nevertheless, the two unit root tests gave significant results at first difference, indicating stationarity of the series at first difference.

Figure 4.3b: Wavelet Component of Detail at level 1

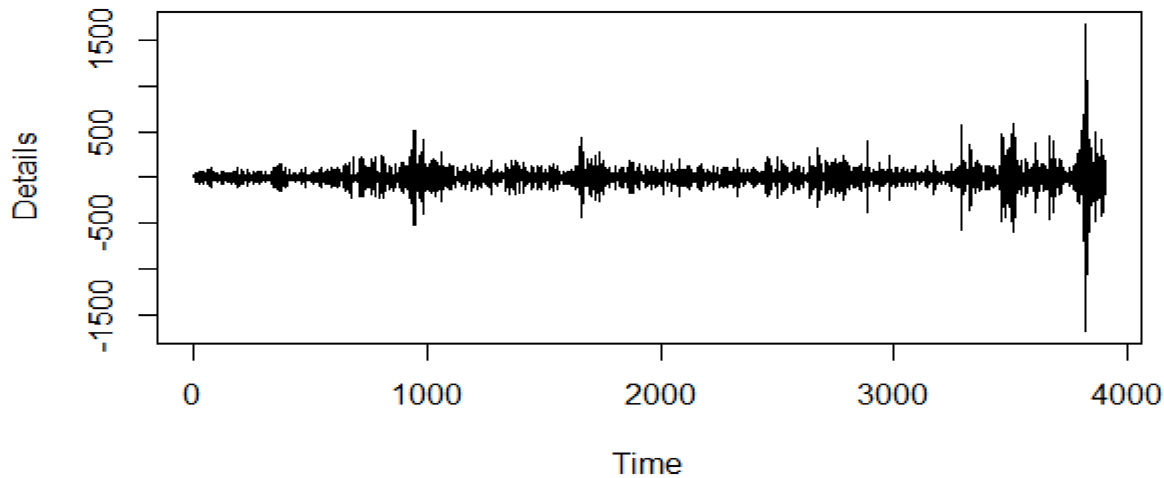
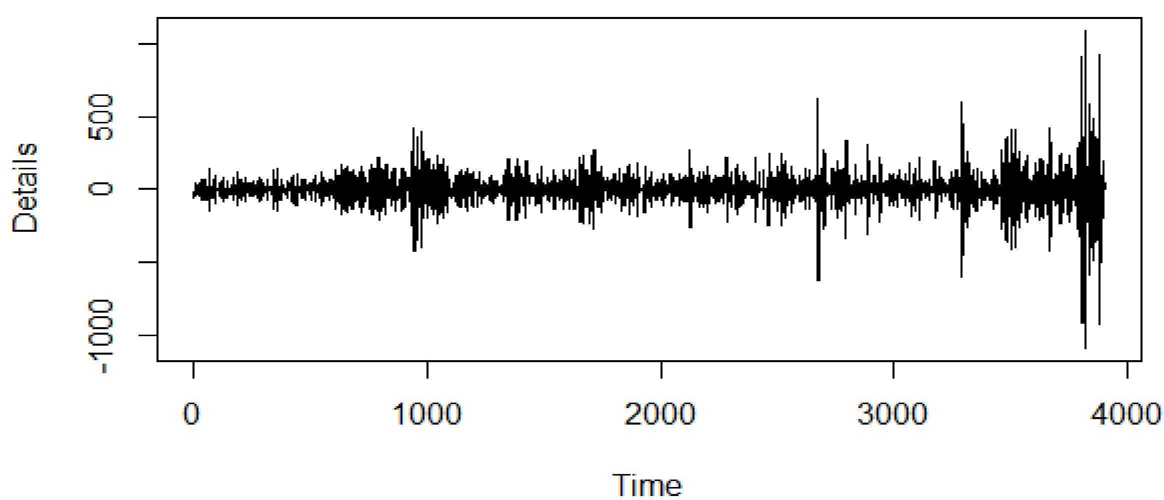



Figure 4.3c: Wavelet Component of Detail at level 2




The series $y_t$ was decomposed into an approximate component $\tilde{y}_{A_2,t}$ and two detail components $\tilde{y}_{D_1,t}$ and $\tilde{y}_{D_2,t}$ where $t$ =1, …, 3907. In the second step, theses WCs were used as exogenous variables in ARIMA-GARCH model, written as WARIMAX-GARCH. Different lagged values of the WEVs were tried for the selection of the best WARIMAX-GARCH model. And the model was used to produce 30-step-ahead forecasts from the level of the series. The best formation of WARIMAX-GARCH method on the basis of its forecasting performance is algebraically given as

$$(1-\phi_1 B-\phi_2\,)y_t \;=\; (1-\theta_1 B)\,e_t \;+\triangle\;\tilde{y}_{A_2,t-1}\varphi_{A_2,t-1}\;+\sum\nolimits_{j=0}^{1}\tilde{y}_{D_1,t-j}\,\varphi_{D_1,t-j}\;+$$
$$\sum\nolimits_{j=0}^{1}\tilde{y}_{D_2,t-j}\,\varphi_{D_2,t-j}\;+\varphi\;ln\sigma_t^2 \qquad (4.1)$$

The three WEVs, each lagged once, were required in WARIMAX-GARCH to produce better forecasts. In addition to the WEVs, the log of conditional variance is used to map the non-linear effect in the series. The best GARCH model which helped achieve better forecast accuracy was an EGARCH (1, 1) with student's t distribution. Algebraically,

$$ln(\sigma_t^2)=\omega+\alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|+\gamma\,\frac{\varepsilon_{t-1}}{\sigma_{t-1}}+\beta ln(\sigma_{t-1}^2) \qquad (4.2)$$

The parameter estimates of the WARIMAX-GARCH model were obtained using MLE method. All of them were significant at 1% level of significance.

The WEVs were used as input variables in ANN, known as WANN, and obtained the required output of WANN. The outputs were incorporated in WARIMAX-GARCH. Mathematically,

$$(1-\phi_1 B-\phi_2\,)y_t = (1-\theta_1 B)\,e_t \,+\triangle\,\tilde{y}_{A_2,t-1}\varphi_{A_2,t-1} \,+\sum\nolimits_{j=0}^{1}\tilde{y}_{D_1,t-j}\,\varphi_{D_1,t-j} \,+ \sum\nolimits_{j=0}^{1}\tilde{y}_{D_2,t-j}\,\varphi_{D_2,t-j} \,+\varphi\, ln\sigma_t^2+\delta y \tag{4.3}$$

**Table 4.3d: the WARIMAX-GARCH-WANN estimation output**

| Variable | Coefficient | Standard Error | t-statistic | P-value |
|---|---|---|---|---|
| $\triangle\, y_{t-2}$ | 0.140345 | 0.011089 | 12.65602 | 0.0000 |
| $e_{t-1}$ | 0.451566 | 0.019032 | 23.72722 | 0.0000 |
| $\triangle\, y_{A_2,t-1}$ | -0.352847 | 0.022312 | -15.81448 | 0.0000 |
| $y_{D_1,t}$ | 0.640105 | 0.022920 | 27.92798 | 0.0000 |
| $y_{D_1,t-1}$ | -1.395568 | 0.034909 | -39.97770 | 0.0000 |
| $y_{D_2,t}$ | 0.225232 | 0.021530 | 10.46133 | 0.0000 |
| $y_{D_2,t-1}$ | -0.840035 | 0.022787 | -36.86488 | 0.0000 |
| $ln\sigma_t^2$ | -1763.914 | 268.4902 | -6.569753 | 0.0000 |
| $y$ | 0.001188 | 0.000176 | 6.762612 | 0.0000 |

**Figure 4.4a: ACF of ordinary residuals obtained from WARIMAX-GARCH-WANN (2,1,1) x (1,1)**

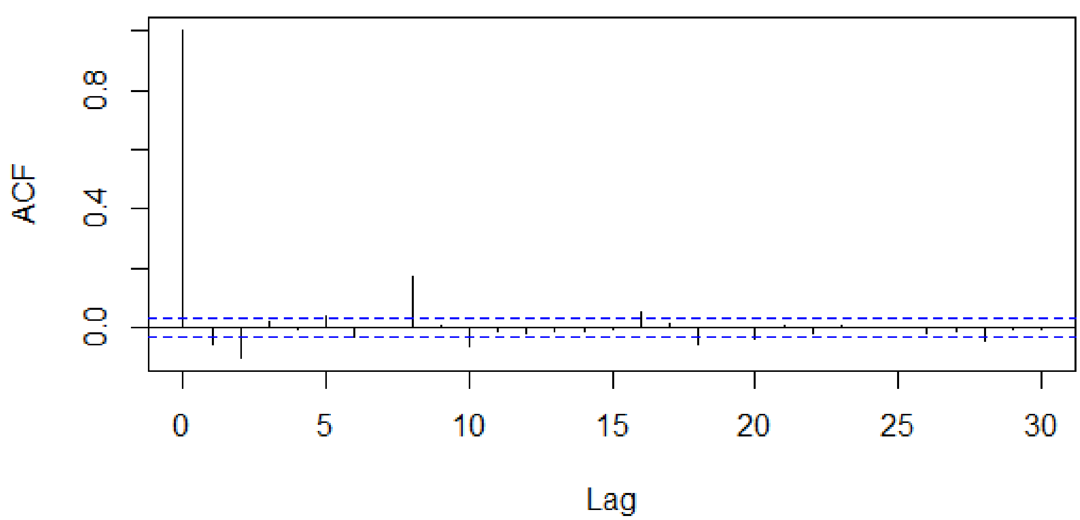


**Figure 4.4b: PACF of ordinary residuals obtained from WARIMAX-GARCH-WANN (2,1,1) x (1,1)**

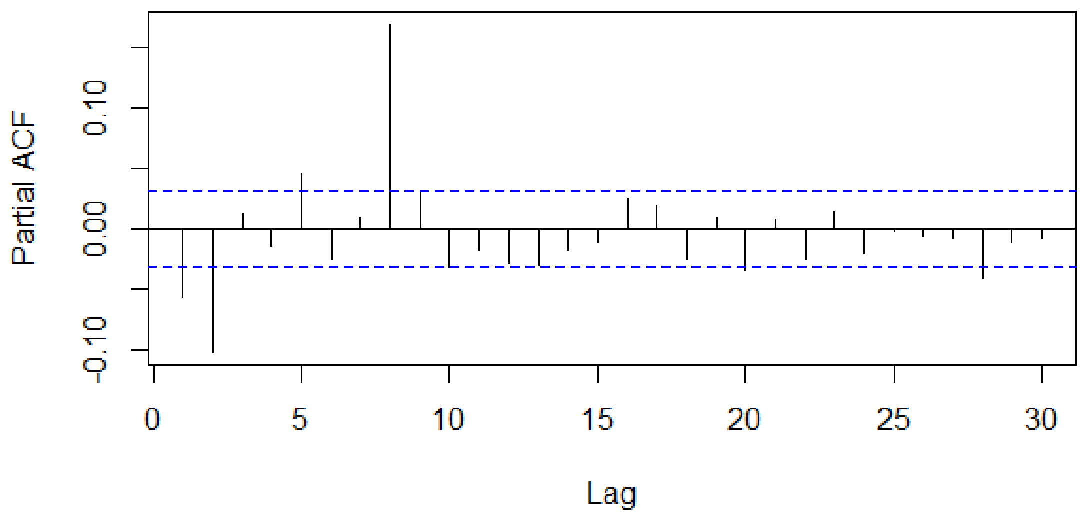


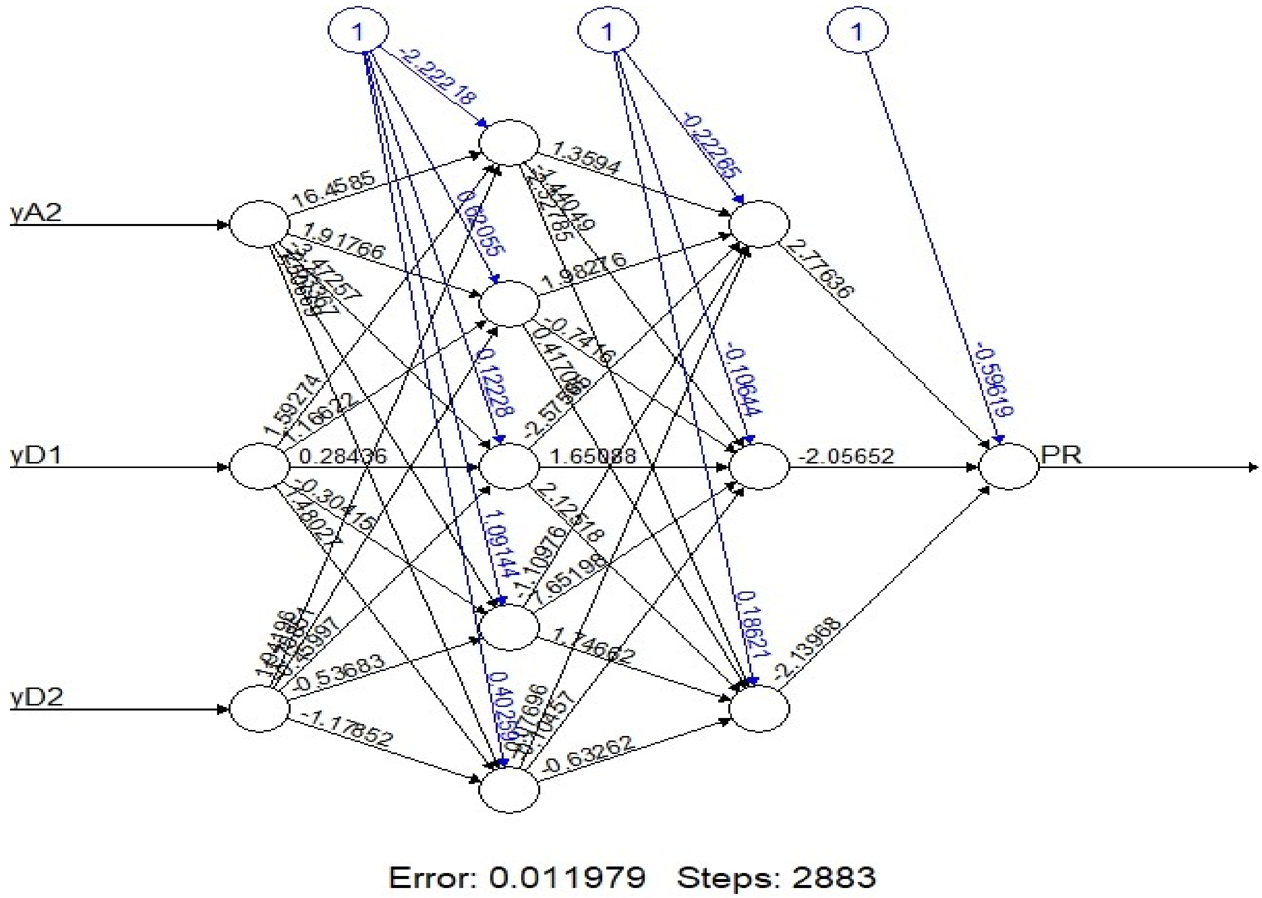

Figure 4.4a and Figure 4.4b display the ACF and the PACF, respectively, up to lag 30 of the ordinary least residuals from the estimated WARIMAX-GARCH-WANN(2,1,1)x(1,1) model. It is to be noted that estimated ACF tapers off after the first lag suggesting stationarity in the error terms. The PACF of residuals show that lags, other than lags five and eight, are within 99% CI indicating no sign of linear auto-dependence in the residuals obtained from WARIMAX-GARCH-WANN model. The ARCH test gave insignificant results suggesting no ARCH effect in the residuals obtained from the above model. Which indicates that the model has captured the volatility. The DW statistic was 2.108644 showing no significant first order autocorrelation in the error term.

### 4.4. WANN Model

WCs were used as input variables in ANN. Daubechies functions were used for each WC. Total three hidden layers were used with 5 and 3 neurons, respectively.

Leverberg-Marquardt training algorithm was applied at each hidden layer, respectively, in order to obtain optimal weights. Figure 4.5 displays the obtained optimal weights WANN methods.

**Figure 4.4: Wavelet ANN plot with Optimal Weights**




### 4.5 Comparison of the Forecasting Performance

Table 4.2 contains the RMSE, MAE and MAPE values for both the in-sample and the out-of-sample forecasts of WANN, WARIMAX-GARCH and the WARIMAX-

GARCH-WANN methods. The optimal WARIMAX-GARCH model given in Section 4.2 is a WARIMAX-EGARCH (2, 1, 1) x (1, 1). It is evident from the inputs in Table 4.2 that the WARIMAX – GARCH-WANN method has lower RMSE, MAE and MAPE values. Therefore, it has a better performance than WARIMAX-GARCH and WANN models in terms of both the in sample and out-of-sample horizons. Particularly, the WARIMAX-GARCH-WANN method generates improved in and out-of-sample forecasts, having a RMSE of 0.00097912, MAPE of 0.069494% and MAE of 0.00040330 against RMSE of 0.007569, MAPE of 103.512% and MAE 0.00430041 of WARIMAX-GARCH model and RMSE of 0.00305133, MAE of 0.00139755 and MAPE of 0.228% of WANN model. From the results, it appears that WCs - which are used as exogenous variables in the model- and the WANN component in ARIMA-GARCH model improved its forecasting performance significantly. Put differently, the WARIMAX-GARCH-WANN captured the dynamics of the underlying series remarkably better as compared to the benchmark methods and give out better forecast results than the WARIMAX-GARCH and WANN methods.

**Table 4.4: The in-sample and out-of-sample Forecast Results Comparison**

| | **Out-of-sample** | | | **In-sample** | | |
|---|---|---|---|---|---|---|
| | RMSE | MAE | MAPE | RMSE | MAE | MAPE |
| WANN | 0.00305133 | 0.00139755 | 0.227746 | 0.007341 | 0.002527 | 0.124051 |
| WARIMAX-GARCH | 0.00756972 | 0.00430041 | 103.5186 | 0.009187 | 0.005443 | 185.3334 |
| WARIMAX-GARCH-WANN | 0.00097912 | 0.00040330 | 0.069494 | 0.000482 | 0.001537 | 0.087541 |

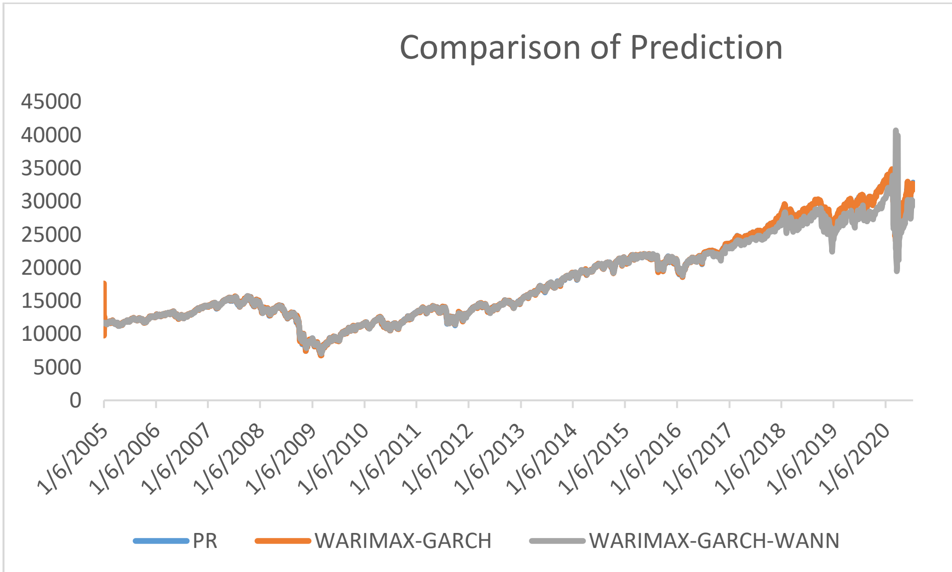



**Figure 4.6: The Comparison of Predictions from WARIMAX-GARCH-WANN and WARIMAX-GARCH**

Figure 4.5 above gives the actual values of the price index, PR, WARIMAX-GARCH-WANN and WARIMAX-GARCH models. The outputs from WARIMAX-GARCH-WANN almost mimic the series perfectly. However, the WARIMAX-GARCH values deviate from the actual values at the end period from 2017 onwards.

# CHAPTER 5

# CONCLUSION AND RECOMMENDATIONS

In this study, a new approach to forecasting high frequency time series, WARIMAX-GARCH-WANN, is proposed. The proposed method uses wavelet components derived from the wavelet transformation of the series and treats them as exogenous variables and produces outstanding forecasting performances improvements over conventional WARIMAX-GARCH and WANN models. The wavelet components incorporated in this model exhibit good statistical properties which make them suitable for use as EVs by both the WARIMAX-GARCH as well as WARIMAX-GARCH-WANN methods. For example, the components of detail show stationarity at level and display conditional volatility- as present in a number of high frequency financial time series– that helps account for nonlinear effects in the final model. Moreover, it can easily be verified that WCs have strong association with the regressand they come from. The method proposed in this study was applied to the daily time series of Wilshire 5000 price index. The results, as compared with the WARIMAX-GARCH and WANN model predominantly produce enhanced forecasting performance. It will be interesting to note if the proposed method can produce similar results when applied to other high frequency time series.

# REFERENCES


Babu, C. N., & Reddy, B. E. (2014, October). Selected Indian stock predictions using a hybrid ARIMA-GARCH model. *In 2014 International Conference on Advances in Electronics Computers and Communications (pp. 1-6)*. IEEE.

Babu, C. N., & Reddy, B. E. (2015). Prediction of selected Indian stock using a partitioning interpolation based ARIMA–GARCH model. *Applied Computing and Informatics*, *11(2)*, 130-143.

Black, F. (1976). Studies of stock market volatility changes. *1976 Proceedings of the American Statistical Association Bisiness and Economic Statistics Section*, 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Box, G. E. P., & Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*-Hoi den- Day. San Francisco.

Box, G. E., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical association*, *70(349)*, 70-79.

Chaâbane, N. (2014). A hybrid ARFIMA and neural network model for electricity price prediction. *International journal of electrical power & energy systems*, *55*, 187-194.

Chand, S., Kamal, S., & Ali, I. (2012). Modelling and volatility analysis of share prices using ARCH and GARCH models. *World Applied Sciences Journal*, *19*(1), 77-82.

Corrêa, J. M., Neto, A. C., Júnior, L. T., Franco, E. M. C., & Faria Jr, A. E. (2016). Time series forecasting with the WARIMAX-GARCH method. *Neurocomputing*, *216*, 805-815.

Enders, W., & Lee, J. (2012). The flexible Fourier form and Dickey–Fuller type unit root tests. Economics Letters, 117(1), 196-199.

Enders, W., & Lee, J. (2012). A unit root test using a Fourier series to approximate smooth breaks. Oxford bulletin of Economics and Statistics, 74(4), 574-599.

Enders, W. (2015). *Applied econometric time series*. John Wiley & Sons.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987- 1007.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, *5*(1), 1-50.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica: journal of the Econometric Society*, 391-407.

Erdogdu, E. (2010). Natural gas demand in Turkey. *Applied Energy*, *87*(1), 211-219.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, *19*(1), 3.

Gençay, R., & Selçuk, F. B. Whitcher (2002): An Introduction to Wavelets and Other Filtering Methods in Finance and Economics.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.

Hickey, E., Loomis, D. G., & Mohammadi, H. (2012). Forecasting hourly electricity prices using ARMAX–GARCH models: An application to MISO hubs. *Energy Economics*, *34*(1), 307-315.

Jalalkamali, A., Moradi, M., & Moradi, N. (2015). Application of several artificial intelligence models and ARIMAX model for forecasting drought using the Standardized Precipitation Index. *International journal of environmental science and technology*, *12*(4), 1201-1210.

Johnston, K., & Scott, E. (2000). GARCH models and the stochastic process underlying exchange rate price changes. *Journal of Financial and Strategic Decisions*, *13*(2), 13-24.

Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. *Procedia Computer Science*, *48*, 173-179.

Klein, J. L., & Klein, D. (1997). *Statistical visions in time: a history of time series analysis, 1662-1938*. Cambridge University Press.

Lahmiri, S. (2014). Wavelet low-and high-frequency components as features for predicting stock prices with backpropagation neural networks. *Journal of King Saud University-Computer and Information Sciences*, *26*(2), 218-227.

Li, T., Li, Q., Zhu, S., & Ogihara, M. (2002). A survey on wavelet applications in data mining. *ACM SIGKDD Explorations Newsletter*, *4*(2), 49-68.

Lu, X., Que, D., & Cao, G. (2016). Volatility forecast based on the hybrid artificial neural network and GARCH-type models. *Procedia Computer Science*, *91*, 1044-1049.

Mallat, S. (2008). A Wavelet Tour of Signal Processing: The Sparse Way, *832 pp. Academic, Burlington, Mass*.

Murtagh, F., Starck, J. L., & Renaud, O. (2004). On neuro-wavelet modeling. *Decision Support Systems*, *37*(4), 475-484.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Nury, A. H., Hasan, K., & Alam, M. J. B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University-Science*, *29*(1), 47-61.

Ortega, L., & Khashanah, K. (2014). A neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns. *Journal of Forecasting*, *33*(2), 134-146.

Pankratz, A. (1991). Forecasting with dynamic regression models. *Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York: Wiley, 1991*.

Paul, R. K., & Himadri, G. (2014). Development of out-of-sample forecasts formulae for ARIMAX-GARCH model and their application. *Journal of the Indian Society of Agricultural Statistics*, *68*(1), 85-92.

Poterba, J. M., & Summers, L. H. (1986). Reporting errors and labor market dynamics. *Econometrica: Journal of the Econometric Society*, 1319-1338.

Reston Filho, J. C., Affonso, C. D. M., & de Oliveira, R. C. (2014). Energy price prediction multi-step ahead using hybrid model in the Brazilian market. *Electric power systems research*, *117*, 115-122.

Sumer, K. K., Goktas, O., & Hepsag, A. (2009). The application of seasonal latent variable in forecasting electricity demand as an alternative method. *Energy policy*, *37*(4), 1317-1322.

Taylor, S. J. (1986). Modelling Financial Time Series. *Wiley, New York*.

Teixeira Júnior, L. A., Souza, R. M. D., Menezes, M. L. D., Cassiano, K. M., Pessanha, J. F. M., & Souza, R. C. (2015). Artificial neural network and wavelet decomposition in the forecast of global horizontal solar radiation. *Pesquisa Operacional, 35(1)*, 73-90.

Thomas, S., & Mitchell, H. (2005). GARCH modeling of high-frequency volatility in Australia's National Electricity Market. *System*, 1-39.

Williams, B. M. (2001). Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling. *Transportation Research Record*, *1776*(1), 194-200.

Wold, H. (1938). *A study in the analysis of stationary time series* (Doctoral dissertation, Almqvist & Wiksell).

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175.

Zolfaghari, M., & Sahabi, B. (2017). Impact of foreign exchange rate on oil companies' risk in stock market: A Markov-switching approach. *Journal of Computational and Applied Mathematics*, *317*, 274-289.

Zolfaghari, M., & Sahabi, B. (2019). A hybrid approach to model and forecast the electricity consumption by Neuro Wavelet and ARIMAX-GARCH models. *Energy Efficiency*, 1-24.