INDUSTRIAL DATA FORECASTING USING DISCRETE WAVELET TRANSFORM

S. Al Wadi*
The University of Jordan
Department of Risk Management and Insurance
Aqaba
Jordan
sadam_alwadi@yahoo.co.uk

Ahmed Atallah Alsaraireh
The University of Jordan
Department of Computer Information Systems
Aqaba
Jordan
ah_8545@yahoo.com
a.alsarairah@ju.edu.jo

Abstract. Since the industrial data plays significant element in any economic growth and these data have many factors that effect on its behavior. Therefore, in this article events of productivity of the Extractive Industry in Jordan will be forecasted using some of traditional model which is (ARIMA model) compound with Orthogonal wavelet transform (OWT) in order to improve the forecasting accuracy. First, the series of dataset will be decomposed by OWT's then the smooth's series will be predicted using ARIMA model, OWT+ ARIMA model in order to improve the forecasting accuracy. As a results the compound model (OWT+ ARIMA) is better than the ARIMA model directly in forecasting accuracy.

Keywords: operation research methods, traders satisfaction, mathematical models.

1. Introduction

Industry stock market tendencies are a challenging task. Numerous factors influence industry stock market performance, including political events, general economic conditions, and trader expectations. Though stock and futures traders rely heavily on various types of intelligent systems to make trading decisions to date their success has been limited [1]. Even financial experts find it difficult to make accurate predictions, because industrial stock market trends tend to be nonlinear, uncertain, and non-stationary. No consensus exists among experts as to the effectiveness of forecasting industrial time series. More specific, Some of the important factors influencing industrial productivity are: Technological Development, Quality of Human Resources, Availability of Finance, Managerial

∗. Corresponding author

Talent, Government Policy and Natural Factors. Consequently, Many models that has suggested in order to studying and predicting the industrial data such as [9,10] has discussed the volatility of the industrial data and provide a significant model for forecasting. In 2008, [11] investigates the interactive relationships between oil price shocks and Chinese stock market using multivariate vector auto-regression. [12] have discussed the performance of the Singular Spectrum Analysis (SSA) technique is assessed by applying it to 24 series measuring the monthly seasonally unadjusted industrial production for important sectors of the German, French and UK economies. The results are compared with those obtained using the HoltWinters and ARIMA models. Statistical selection procedures are used in a variety of applications to select the best of a finite set of alternatives. Best" is defined with respect to the (largest or smallest) mean, where the mean is inferred with statistical sampling, as in simulation optimization. Many sequential selection procedures are proposed to select a good design when the number of alternatives is large, see Alrefaei and Almomani [13], Almomani and Alrefaei [14], Almomani and Abdul Rahman [15], Almomani and Ababneh [16], Almomani and Alrefaei [17], Al-Salem et al. [18], Almomani et al. [19].

One of the most important statistical procedure is the forecasting accuracy. Therefore, many attempts have been made to forecast financial markets such as [20-25]. One of the traditional forecasted model is ARIMA model. However, the main disadvantage of ARIMA model is the enormous difficulty of interpreting the results. This study diverges from previous attempts at forecasting stock prices by proposing a method that uses the Wavelet transforms (WT) combined with ARIMA, this process create a transparent architecture. WT is a relatively new field in signal processing [2]. Wavelets are mathematical functions that decompose data into different frequency components, after which each component is studied with a resolution matched to its scale, where a scale denotes a time horizon [3]. WT is closely related to the volatile and time varying characteristics of the real-world time series and is not limited by the stationarity assumption [4]. WT decomposes a process into different scales, making it useful in distinguishing seasonality, revealing structural breaks and volatility clusters, and identifying local and global dynamic properties of a process at specific timescales [5]. WT has been shown to be particularly useful in analyzing, modeling, and predicting the behavior of financial instruments as diverse as stocks and exchange rates [6,7,26,27].

Moreover, WT has applications in forecasting in many fields such as in electronic forecasting [28], finance and economic [29, 31], also it used with short and long memory forecasting [30], and other fields [32, 33]. This study applies WT using the OWT functions which are (Haar, Daubechies, Coiflet and symmelt) to decompose the industry time series then combine the approximation coefficients with ARIMA model in order to make improve the forecasting accuracy then select the best WT function in forecasting. Then finally we compare the forecasting using the combined method with the forecasting using ARIMA model directly OWT+ ARIMA model.

2. Methodology

Table 1: shows the research framework

Industry Time Series Data
OWT (Haar, Daubechies, Symmelet, and Coiflet) functions
....
Smoothed coefficients
Improving the forecasting accuracy by using ARIMA model for the transformed data.
Selection the best model in decomposition and forecasting

2.1 Wavelet transform

WT is based on Fourier transform (FT) which shows any function as the sum of the sine and cosine functions. WT should satisfies the following condition [8]:

$$C_{\varphi} = \int_{0}^{\infty} \frac{|\varphi(f)|}{f} df < \infty, \tag{1}$$

where $\varphi(f)$ is the FT and a function of frequency f, of $\varphi(t)$. WT is a mathematical function that can be used in many fields such. WT was introduced to solve problems associated with the FT as they occur specially with non-stationary dataset or with signals that are localized in time, space, or frequency. There are two types of WT which are Father WT which describes the smooth frequency components of a signal and mother WT which describes the detailed components. Mathematically, the following equations represent the father WT and mother WT respectively, with $j = 1, ..., J$ in the J-level WT decomposition [6]:

$$\phi_{j,k} = 2^{-j/2}\phi\left(t - 2^{j}k/2^{j}\right), \ \varphi_{j,k} = 2^{-j/2}\varphi\left(t - 2^{j}k/2^{j}\right), \tag{2}$$

father WT and mother WT should satisfy the following conditions:

$$\int \phi(t)\, dt = 1 \text{ and } \int \phi(t)\, dt = 0. \tag{3}$$

Time series data (f(t)) is an input represented by WT and can be built up as a sequence of projections onto father WT and mother WT indexed by both $k, k = 0, 1, 2, ...$ and by $S = 2j, j = 1, 2, 3, ...J$. Mathematically:

$$S_{j,k} = \int \phi_j, kf(t)dt, \text{ and } d_{j,k} = \int \varphi_j, kf(t)dt. \tag{4}$$

The orthogonal wavelet series approximation to f(t) is defined by:

$$\begin{aligned} F(t) = &\sum S_j, k\phi_j, k(t) + \sum d_j, k\phi_j, k(t) \\ &+ \sum d_j - 1, k\phi_j - 1, k(t) + ... + \sum d_1, k\phi_1, k(t) \end{aligned} \tag{5}$$

$$S_j(t)) = \sum S_j, k\phi_j, k(t), \quad D_j(t) =) \sum d_j, k\phi_j, k(t). \tag{6}$$

2.2 ARIMA model

ARMA is a suitable model for the stationary time series data, although most of the software uses least square estimation which requires stationary. To overcome this problem and to allow ARMA model to handle non-stationary data, the researchers investigate a special class for the non-stationary data. This model is called Auto-regressive Integrated Moving Average (ARIMA). This idea is to separate a non-stationary series one or more times until the time series becomes stationary, and then find the fit model. ARIMA model has got very high attention in the scientific world. This model is popularized by George Box and Gwilym Jenkins in 1970s [25]. There are a huge number of ARIMA models; generally there are ARIMA (p, q, d) where: P: order of autoregressive part (AR), d: degree of first differentiation (I) and q: order of the first moving part (MA). Note that, if there is no differencing been done (d = 0), then ARMA model can be got from ARIMA model. The general mathematical ARIMA model can be defined as [25]

$$W_t = \mu + \frac{\beta(\nu)}{\varepsilon(\nu)} a_t.$$

Where:

t: Indexes time.

W_t: The response series Y_t or a difference of the response series.

μ: The mean term.

ν: The backshift operator; that is,$\nu X_t = X_{t-1}$

$\beta(\nu)$: The autoregressive operator, represented as a polynomial in the backshift operator:

$$\varepsilon(\nu) = 1 - \varepsilon_1(\nu) - ... - \varepsilon_p \nu^p$$

$\varepsilon(\nu)$: The moving-average operator, represented as a polynomial in the backshift operator:

$$\beta(\nu) = 1 - \beta_1(\nu) - ... - \beta_p \nu^p.$$

a_t: The independent disturbance, also called the random error.

2.3 Mathematical criteria

The author is used some criteria in order to make fair comparison between ARIMA and ARIMA-WT can be presented in this section. Some types of accuracy criteria have used; Root means squared error (RMSE), Percentage root mean absolute percentage error (MAPE) and mean absolute error (MAE) for the mathematical formulas refer to [9].

3. Dataset, analysis and discussion

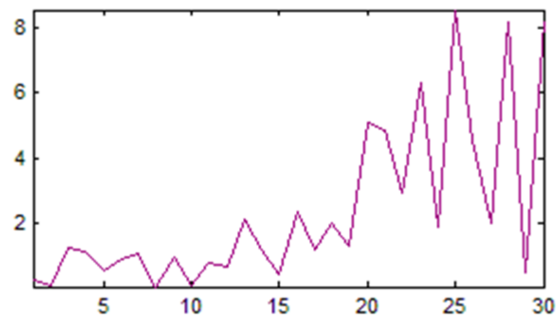

Figure 1: Original Data Set

The daily price index of ASE for a specific period of time has been selected as the statistical population, more than 4000 observations were accumulated for each variable from related databases in the mentioned period. Figure 1 shows the diagram of the dataset. Regarding the target of this paper the following table will show the results of the forecasting accuracy using ARIMA model directly and ARIMA+WT.

Table 2: forecasting results

	ARIMA directly	ARIMA+ Haar	ARIMA+ Daubechies	ARIMA + Coiflet	ARIMA+ symmelt
RMSE	1.56	0.6	0.3	0.4	0.7
MAPE	52	28.6	20	22	29
MAE	1.34	0.4	0.1	0.35	0.9

For the sake of fair comparison the same number of data set is selected. The suitable forecasted model for forecasting the sample data is the fitted Daubechies WT- ARIMA (2,0,2) with RSME equal to 0.3 as presented in Table 2. While the Fitted ARIMA model directly is ARIMA (1,0,2) with RMSE 1.56 which means that the forecasting accuracy has improved by combining OWT+ARIMA model. Also, the best forecasted model is Daubechies + ARIMA model. Moreover, to insure the results the authors have used MAPE and MAE which was the best for ARIMA + Daubechies also since these values were 20 and 0.1 respectively. While these values were more that ARIMA+ Daubechies based on other functions.

4. Conclusion

ARIMA model is the most general way of forecasting since there is no need for any assumptions and it is not limited to specific type of pattern. These models can be fitted to any set of time series data (stationary or non-stationary) by estimating the parameters p, d, and q to be suitable with the required dataset. In this study, firstly, the industry stock price is modeled using wavelet method. Secondly, we compared ARIMA+ OWT with ARIMA directly model in content of forecasting accuracy. Thirdly, we tested the accuracy of these models by using RMSE, MAPE and MSE assessing functions. Final the forecasting accuracy have improved using the suggested model. Indeed, we found ARIMA+OWT is suitable model in content of industry sector specially Daubechies WT with ARIMA model.

References

[1] Y.S. Abu-Mostafa, A.F. Atiya, *Introduction to financial forecasting*, Applied Intelligence, 1996, 205-213.

[2] A. Cohen, I. Daubechies, P. Vial, *Wavelets on the interval and fast wavelet transform*, Applied and Computational Harmonic, 1993, 5481.

[3] J.B. Ramsey, Z. Zhang, *The analysis of foreign exchange data using wave form dictionaries*, Journal of Empirical Finance, 1997, 341372.

[4] A. Popoola, K. Ahmad, *Testing the suitability of wavelet preprocessing for TSK fuzzy models*, in: Proceeding of FUZZ-IEEE: International Conference Fuzzy System Network, 2006, 13051309.

[5] R. Gencay, F. Selcuk, B. Whitcher, *Differentiating intraday seasonalities through wavelet multi-scaling*, PhysicaA, 2001, 543556.

[6] J.B. Ramsey, *The contribution of wavelets to the analysis of economic and financial data*, Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences, 1999, 25932606.

[7] K. Papagiannaki, N. Taft, Z.-L. Zhang, C.Diot, *Long-term forecasting of internet backbone traffic*, IEEE Transactions on Neural Networks, 2005, 11101124.

[8] R. Gencay, F. Selcuk, B. Whitcher, *An introduction to wavelets and other filtering methods in finance and economics*, Academic Press, NewYork, 2002.

[9] Firas Muhammad Al-Rawashdi, S. Alwadi and Mohammad Hasan Saleh, *Wavelet methods in forecasting for insurance companies listed in amman stock exchange*, European Journal of Economics, finance and administrative sciences, 82 (2015), 54-60.

[10] S. Al Wadi, *Improving volatility risk forecasting accuracy in industry sector*, International Journal of Mathematics and Mathematical Sciences, 2017.

[11] R.G. Cong, Y.M. Wei, J.L. Jiao, and Y. Fan, *Relationships between oil price shocks and stock market: An empirical analysis from China*, Energy Policy, 36 (2008), 3544-3553.

[12] H. Hassani, S. Heravi, A. Zhigljavsky, *Forecasting european industrial production with singular spectrum analysis*, International journal of forecasting, 25 (2009), 103-118.

[13] M.H. Alrefaei, M.H. Almomani, *Subset selection of best simulated systems*, Journal of the Franklin Institute, 344 (2007), 495-506.

[14] M.H. Almomani, M.H. Alrefaei, *A three-stage procedure for selecting a good enough simulated system*, Journal of Applied Probability and Statistics, 6 (2012) (2012), 13-27.

[15] M.H. Almomani, R. Abdul Rahman, *Selecting a good stochastic system for the large number of alternatives*, Communications in Statistics-Simulation and Computation, 41 (2012), 222-237.

[16] M.H. Almomani, F. Ababneh, *The expected opportunity cost and selecting the optimal subset*, Applied Mathematical Sciences, 9 (2015), 6507-6519.

[17] M.H. Almomani, M.H. Alrefaei, *Ordinal optimization with computing budget allocation for selecting an optimal subset*, Asia-Pacific Journal of Operational Research, 33 (2016), DOI: 10.1142/S0217595916500093.

[18] M. Al-Salem, M.H. Almomani, M.H. Alrefaei, Ali Diabat, *On the optimal computing budget allocation problem for large scale simulation optimization*, Simulation Modelling Practice and Theory, 71 (2017), 149-159.

[19] M.H. Almomani, M.H. Alrefaei, S. Al Mansour, *A combined statistical selection procedure measured by expected opportunity cost*, Arabian Journal for Science and Engineering, 2017.

[20] S. Al Wadi, *Improving volatility risk forecasting accuracy in industry sector*, International Journal of Mathematics and Mathematical Sciences, 2017.

[21] A.M. Awajan, M.T. Ismail, S. Al Wadi, *A hybrid EMD-MA for forecasting stock market index*, Italian Journal of Pure and Applied Mathematics, 38 (2017), 1-20.

[22] A.M. Awajan, M.T. Ismail, S.A. Wadi, *Forecasting time series using EMD-HW bagging*, International Journal of Statistics and Economics, 18 (2017), 9-21.

[23] J.J. Jaber, N. Ismail, S. Al Wadi, M.H. Saleh, *Forecasting of volatility risk for Jordanian banking sector*, Far East Journal of Mathematical Sciences, 101 (2017), 1491-1507.

[24] S. Al Wadi, F. Ababneh, H. Alwadi, M.T. Ismail, *Maximum overlap discrete wavelet methods in modeling banking data*, Far East Journal of Applied Mathematics, 84 (2013).

[25] M.T.I.S. Al Wadia, M. Tahir Ismail, *Selecting wavelet transforms model in forecasting financial time series data based on ARIMA model*, Applied Mathematical Sciences, 5 (2011), 315-326.

[26] S. Al Wadi, M.T. Ismail, S.A.A. Karim, *A comparison between the Daubechies wavelet transformation and the fast Fourier transformation in analyzing insurance time series data*, Far East J. Appl. Math, 45 (2010), 53-63.

[27] S. Al Wadi, M.T. Ismail, S.A. Abdul Kari, *Discovering Structure breaks in Amman stocks market*, Journal of Applied Sciences, 11 (2011), 1273-1278.

[28] S. Li, P. Wang, L. Goel, *Electric load forecasting using wavelet transform and extreme learning machine*, In ESANN, 2014.

[29] L. Bai, S. Yan, X. Zheng, B.M. Chen, *Market turning points forecasting using wavelet analysis*, Physica A: Statistical Mechanics and its Applications, 437 (2015), 184-197.

[30] O. Renaud, F. Murtagh, J.L. Starck, *Wavelet-based forecasting of short and long memory time series*, Université de Genève/Faculté des sciences économiques et sociales, 2002.

[31] J.B. Ramsey, *The contribution of wavelets to the analysis of economic and financial data*, Wavelets: The Key to Intermittent Information, Volume Wavelets: the key to intermittent information, 221-236.

[32] C.A.G. Santos, P.K.M.M. Freire, G.B.L. Silva, R.M. Silva, *Discrete wavelet transform coupled with ANN for daily discharge forecasting into Três Marias reservoir*, Proceedings of the International Association of Hydrological Sciences, 364 (2014), 100-105.

[33] S. Schlüter, C. Deuschle, *Using wavelets for time series forecasting: Does it pay off?*, IWQW discussion paper series, 2010.

Accepted: 14.03.2018