Modeling via Wavelet GARCH Algorithm on Multivariate ENSO Index

Md. Nazmul Ahasan[1], Md. Abdul Khalek[2] and Md. Mesbahul Alam[2]

[1] Ph.D. Fellow, Institute of Environmental Sciences, Rajshahi University
[2] Department of Statistics, Rajshahi University

DOI: 10.29322/IJSRP.9.07.2019.p9195
http://dx.doi.org/10.29322/IJSRP.9.07.2019.p9195

Abstract- Modeling and forecasts of global oceanographic index has important implications for decision making. An effective management on climate anomalies impact depends on the performance towards better accuracy and forecasts. In this paper, an algorithm which makes use of wavelets together with a time series model, GARCH is implemented in order to improve the performance of forecasts in global climate data series. Multivariate ENSO index, MEI, for the period January, 1950 to February, 2018 is used to compare the GARCH model and a newly proposed tools with W-GARCH(1, 1) model. The goodness of performance is calculated via the Akaike information criterion, Schwarz criterion, Hannan-Quinn criterion and RMSE. The results showed that although both models fit the MEI data well, the forecast produced by the GARCH(1, 2) model underestimates the observed score while the newly proposed W-GARCH(1, 1) model generates a better accuracy of forecasts for the given data. Hence the proposed W-GARCH(1, 1) should be applied for forecasts in the fields reflected by MEI variability.

Index Terms- Multivariate ENSO index, Forecast accuracy, Wavelet-GARCH

I. Introduction

The analysis of global climate time series data provides crucial information to describe, explore and predict climate variability. It stipulates useful information about the physical, biological, or socio-economic system that affects. The purpose of global climate time series analysis is crucial as to determine some of its key properties by modeling and forecasting certain features. These properties can help to understand and forecast the system's future behavior (Ghil et al., 2002).

In this sense global climate time series index data, Multivariate ENSO index (MEI), influenced by oceanographic tele-connected factors over different time horizons that range from minutes to years, in the past, present, and future have become a topic of interests to the meteorological experts as well as data analysts for long days. It is a strong climate-dominant mode in the tropical Pacific, which has major effects on the global climate system and ecosystem as well as significant socio-economic consequences around the globe (Chen et. al., 2002; Ropelewski, 1987; Shabbar & Khandekar, 1996). Operational and strategic decision-making based on ENSO has been taken into account not just the realized effects of climate variability, but also potential effects in the decision making in different areas (Pabon, 2016). Thus, a better accuracy of performance and effective decision is vital via modeling and forecasts of MEI data series.

Following ARCH tools, GARCH takes into account excess kurtosis and volatility clustering features, two important characteristics of time series data, and provides accurate forecasts of variances and covariance's to model time-varying conditional variances. In line with this, GARCH models are useful across a wide range of applications; nevertheless it's are only part of a solution. GARCH models are parametric specifications that operate best under relatively stable conditions (Gourieroux, 2012). But GARCH often fails to capture highly irregular phenomena, especially in climatic oscillation and other highly unanticipated events that can lead to significant change in the respective arena. Furthermore, GARCH models often fail to fully capture the fat tails observed series. Heteroskedasticity explains some of the fat tail behaviors, but typically not all of it. Fat tail distributions, such as student-*t*, have been applied in GARCH modeling, but often the choice of distribution is a matter of trial and error.

Despite GARCH is unable to handle some properties in the time series data. Over the past years, GARCH model has been widely used in predicting of geophysical as well as hydrological time series (Darda, 2014; Lee at. al., 2011; Modarres & Ouarda, 2013, 2014; Romilly, 2005; Taylor & Buizza, 2004; Trombe et. al., 2012).

Contrary to that, wavelet transform, a nonlinear component, has been applied in many engineering, signal processing, and statistical problems (Chaovalit et. al., 2011) since the recent decade as to overcome the pitfalls of GARCH. It has also shown excellent performance in hydrological modeling (Nourani et. al., 2009; Okkan, 2012 as well as in multiple atmospheric and environmental applications (Pal & Devara, 2012; Pal et al., 2015). Wavelet transformation decomposes the main time series into subcomponents such that the decomposed data improve the model's performance by capturing useful information at various resolution levels (Karim et. al, 2013). Struzik (2001) used wavelet decomposition in forecasting time series. Hsieh et al. (2011) used Haar wavelet decomposition for removal of noise from the stock price time series data with the aim to achieve more precise forecasts. Milidiu et. al. (1999) used Haar wavelet together with a clustering algorithm to partition input data into different regions. In line with this, it is evident in the previous works that wavelet transform has also been considerably improved forecasting accuracies. Even though, we are not satisfied about simple wavelet for better accuracy of performance as model building is a continuing process of research.

In recent years, in the field of time series research, hybrid models are being proposed as an effective way to overcome the limitations of each components model as well as able to improve forecasting efficacy. More hybrid forecasting models have been proposed applying Box-Jenkins models including an ARIMA model with GARCH to time series data in various fields for their good performance. (Wang et. al., 2005) proposed an ARMA-GARCH error model to capture the ARCH effect present in daily stream flow series. Zhou et. al. (2006) applied the ARIMA-GARCH model in forecasting internet traffic, while Chen et. al. (2011) suggested ARIMA-GARCH model for short-time traffic flow prediction. Meanwhile, Liu (2013) applied ARMA-GARCH for wind speed forecasts. In accordance with that wavelet-based ARIMA model has achieved higher prediction accuracy than the conventional ARIMA and GARCH model (Kriechbaumer et. al., 2014; Rahman & Hasan, 2014; Szolgayová et. al., 2014).

To address as a key indicator of climate variability measure index in longer time domain, this paper employs time series tools, GARCH, with Wavelet Transformation (WT) techniques to develop an efficient model with applications to MEI data.

In this sense, the objective(s) of this study is to gain a better per formant accuracy in modeling and forecasts by Wavelet Transform with GARCH algorithm based on oscillatory oceanographic data series like, MEI. Therefore, the study compares the forecast error results obtained using the straight GARCH(1, 2) model with the forecast error obtained using the W-GARCH(1, 1) model about Gaussian, student-t and generalized error distribution.

Finally, as far as the authors are aware of, multivariate ENSO index (MEI) has not been yet tested on modeling and forecasting process on hybrid statistical forecasts tools like the proposed algorithm.

The paper is structured as; in section two we describe data description used in this study. In section three, four, five and six, wavelet, GARCH and data analysis method(s) is outlined. Finally, in section seven and eight presents the main discussion and conclusions of the study.

II. METHODS AND MATERIALS

A. Data Description

MEI, a more informative approach has been developed at NOAA's Climate Diagnostics Center in Boulder Colorado with multivariate ENSO index (MEI) that is derived from tropical Pacific Comprehensive Ocean–Atmosphere Data Set (COADS) integrates more information than other indices; it reflects the nature of the coupled ocean-atmosphere system better than either component (Allan et. al., 1991; Mazzarella et. al., 2010; Rasmusson & Carpenter, 1982; Ropelewski & Jones, 1987). MEI is a multivariate measure of the climatic signal calculated as the first principal component of six variables over the tropical Pacific such as sea surface temperature, sea level pressure, zonal and meridian components of the surface wind, air temperature and total cloudiness fraction of the sky. The computation details of Multivariate ENSO Index may be seen in Wolter and Timlin (1993, 1998, www.cdc.noaa.gov/people/klaus.wolter/MEI/table.html).

B. Time series forecasting model

Over the years, various time series forecasting models have been developed in literature. The random walk, autoregressive (AR), moving average (MA), and ARIMA are some widely recognized statistical forecasting models which predict future observations of a time series on the basis of some linear function of past values and white noise terms (Chaovalit et al., 2011). As such, these models impose the inherent constraint of linearity on the data generating function. To overcome this, various nonlinear models have also been developed in literature. One widely popular among them is GARCH that has many salient features. (Zhang, 2013) has rationally combined both ARIMA and ANN hybrid models in order to considerably increase the forecasting accuracy.

C. Methods of framework

In this paper the Approximation series has been used since this series behave as the main component of the transform, while the detail series provides "small" adjustments. The procedure explained in this paper is as follows:

- Decompose through the wavelet transform by Haar wavelet transform the data.
- Use a specific GARCH model fitted to each one of the Approximation series to make the forecasting.
- The technique is compared with GARCH model used directly to forecast the data series using the standard criteria.

D. Model Selection

All of the models were fitted by the method of maximum likelihood. Some of the fitted models are not nested. Discrimination among them was performed using various criteria: The Akaike Information Criterion due to (Akaike, 1974) defined by:

$AIC=2k-2\ln L(\widehat{\Theta})$

Where k denotes the number of unknown parameters, Θ the vector of the unknown parameters and $\widehat{\Theta}$ their maximum likelihood estimates; Bayesian information criterion due to (Schwarz, 1978) defined by:

$BIC=k\ln n - 2\ln L(\widehat{\Theta})$

Where n denotes the number of observations; The Consistent Akaike Information Criterion (CAIC) due to (Bozdogan, 1987) defined by:

$CAIC=-2\ln L(\widehat{\Theta})+k(\ln n+1)$

The corrected Akaike Information Criterion (AICc) due to (Hurvich & Tsai, 1989) defined by:

$AICc=AIC+2k(k+1)(n-k-1)$;

III. RESULTS AND GRAPHS

Analytical results are explored in two ways: graphical method and estimated numerical facts. Both are supportive to each other. Below model Identification, Diagnostic and Adequacy Plots are explored and discussed

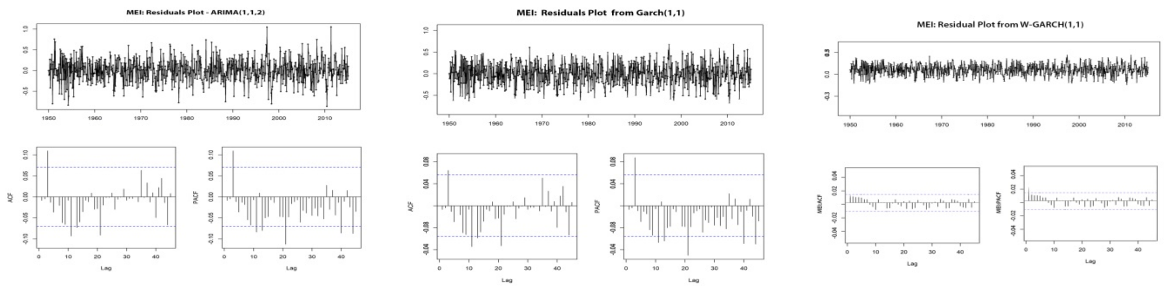
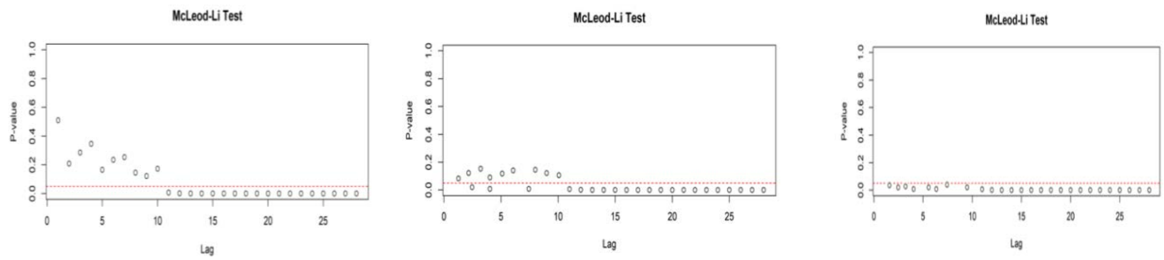


Figure 3.1: Residual plot, ACF and PACF for ARIMA, GARCH and W-GARCH

From Figure 3.1, it is evident that W-GARCH is superior to ARIMA and GARCH as in run sequence plot residuals are decreasing sequentially over the fitted models respectively. Accordingly in ACF and PACF, residuals are within the limit other than two.



(a) ARIMA (b) GARCH (c) W-GARCH

Figure 3.2: Plots of test data about McLeod-Li test for ARIMA, GARCH and W-GARCH model

In Figure 3.2, according to McLeod-Li test it also suggests an improvement is observed for W-GARCH.

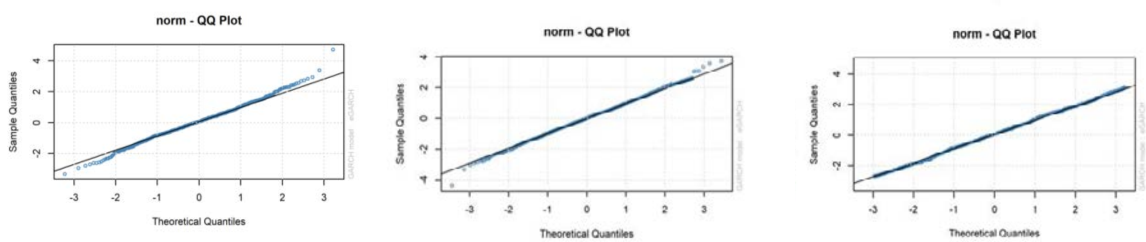


Figure 3.3: QQ Plots of for ARIMA, GARCH and W-GARCH model

In Figure 3.3 the standardized residuals from the three models are presented. The QQ plot for the W-GARCH is the only plot showing residuals that can be considered to follow a normal distribution, while for the ARIMA and G this assumption is not achieved.

IV. PERFORMANCE OF MODELS

Table 1: Model selection about Normal, Student-t and Generalized Error Distributions (GED)

Model	Normal		Student's t		GED	
	GARCH(1,2)	W-GARCH(1,1)	GARCH(1,2)	W-GARCH(1,1)	GARCH(1,2)	W-GARCH(1,1)
R^2	0.919	0.991781	0.879206	0.903889	0.919193	0.913856
AIC	0.265	-4.303453	0.209454	-1.426408	0.259436	---
SC	0.307	-4.267684	0.307145	-1.384678	0.307128	---
HQ	0.281	-4.289697	0.277794	-1.41036	0.277777	---
DW	1.944	0.6123	1.627257	1.789733	1.933189	1.79673

In this paper, the minimum value of R-squared, AIC (Akaike information Criteria), SC, DW and HQ is considered to select the best model for the experiment. All choices of GARCH models for data are included in this test between (0, 1) and (2, 2). Some of the combination measures based on degree of parameters are complex roots that should not be satisfied the estimation procedure.

The GARCH model for the original index data is considered as GARCH (1,2) with AIC (Akaike Information Criteria) about Gaussian distribution to 0.259454 as presented in table-1, while the GARCH model for the transform data by using wavelet transform is selected as Wavelet-GARCH (1, 1) with AIC equal to -4.303453 as presented in table-1as well. Although the fit GARCH model for the transform data using Haar wavelet transform is selected as Wavelet-GARCH(1, 1) based on AIC about Gaussian distribution equal to -4.303453, table-1shows some other criteria of various distributions about the result. All of these criteria explain that the Wavelet GARCH model is better than the GARCH model.

Moreover, the Haar wavelet transform gives more sufficient result and better than Daubechies wavelet transform in the forecasting that is not cited here. However, in some statistical literature, Daubechies wavelet transform is better than Haar wavelet in the decomposition, but in this paper we found a different result, the reason is related to the data set since just the approximation series have used in the comparison.

Furthermore, the standard errors of model parameters which measure the variation between index data after wavelet transforms are also small. All of these criteria explain that the Haar wavelet transform gives more sufficient result and better than GARCH in the forecasting.

V. CONCLUSION

An effective management on climate anomalies impact depends on the performance towards GARCH and Wavelet are two widely used forecasting models captured better accuracy. Although wavelet and GARCH both are suitable for an irregular low and high frequency signals in global climatic time series but their hybrid effect gives a better results of forecasts. As such, in this paper, we have proposed a hybrid forecast method that applies GARCH and Wavelet in conjugate.

In concluding remarks of the study, if the Wavelet transform is used for forecasts in MEI data series, then the result of the W-GARCH(1, 1) model attained better forecast and forecasting accuracy is more stable than the original index data.

ACKNOWLEDGMENT

We would like to express our gratitude to University Grants Commission (UGC), Bangladesh, for their financial support to this study. We are also grateful to the anonymous reviewers, for reading our manuscript and giving us valuable suggestions and comments.

REFERENCES

[1] Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.

[2] Allan, R. J., Nicholls, N., Jones, P. D., & Butterworth, I. J. (1991). A further extension of the Tahiti-Darwin SOI, early ENSO events and Darwin pressure. Journal of Climate, 4(7), 743-749.

[3] Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52(3), 345-370.

[4] Chaovalit, P., Gangopadhyay, A., Karabatis, G., & Chen, Z. (2011). Discrete wavelet transform-based time series analysis and mining. ACM Computing Surveys (CSUR), 43(2), 6.

[5] Chen, C., Hu, J., Meng, Q., & Zhang, Y. (2011). Short-time traffic flow prediction with ARIMA-GARCH model. Paper presented at the Intelligent Vehicles Symposium (IV), 2011 IEEE.

[6] Chen, C., McCarl, B., & Hill, H. (2002). Agricultural value of ENSO information under alternative phase definition. Climatic Change, 54(3), 305-325.

[7] Darda, M. A. (2014). On Modelling Temperature Volatility Dynamics. Jahangirnagar university journal of information technology (jujit) vol. 3, issue. 1, september 2014, vol. 3,(issue. 1,september 2014).

[8] Ghil, M., Allen, M., Dettinger, M., Ide, K., Kondrashov, D., Mann, M.,Varadi, F. (2002). Advanced spectral methods for climatic time series. Reviews of geophysics, 40(1).

[9] Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. Biometrika, 76(2), 297-307.

[10] Karim, S. A. A., Ismail, M. T., Hasan, M. K., & Sulaiman, J. (2013). Denoising the temperature data using wavelet transform. Applied Mathematical Sciences, 7(117), 5821-5830.

[11] Kriechbaumer, T., Angus, A., Parsons, D., & Casado, M. R. (2014). An improved wavelet–ARIMA approach for forecasting metal prices. Resources Policy, 39, 32-41.

[12] Lee, T., Ouarda, T., & Ousmane, S. (2011). Predictability of climate indices with time series models. Stochastic Hydrology of the Great Lakes–A Systemic Analysis, 45.

[13] Mazzarella, A., Giuliacci, A., & Liritzis, I. (2010). On the 60-month cycle of multivariate ENSO index. Theoretical and applied climatology, 100(1-2), 23-27.

[14] Milidiu, R. L., Machado, R. J., & Rentería, R. P. (1999). Time-series forecasting through wavelets transformation and a mixture of expert models. Neurocomputing, 28(1), 145-156.

[15] Modarres, R., & Ouarda, T. (2014). Modeling the relationship between climate oscillations and drought by a multivariate GARCH model. Water Resources Research, 50(1), 601-618.

[16] Modarres, R., & Ouarda, T. B. (2013). Testing and modelling the volatility change in ENSO. Atmosphere-Ocean, 51(5), 561-570.

[17] Nourani, V., Alami, M. T., & Aminfar, M. H. (2009). A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. Engineering Applications of Artificial Intelligence, 22(3), 466-472.

[18] Okkan, U. (2012). Wavelet neural network model for reservoir inflow prediction. Scientia Iranica, 19(6), 1445-1455.

[19] Pabón, J. D., Martínez, R. (2016). Preface: Third ENSO Conference: findings and key messages. Advances in Geosciences, 42, 91.

[20] Pal, S., & Devara, P. (2012). A wavelet-based spectral analysis of long-term time series of optical properties of aerosols obtained by lidar and radiometer measurements over an urban station in Western India. Journal of Atmospheric and Solar-Terrestrial Physics, 84, 75-87.

[21] Pal, S., Lopez, M., Schmidt, M., Ramonet, M., Gibert, F., Xueref-Remy, I., & Ciais, P. (2015). Investigation of the atmospheric boundary layer depth variability and its impact on the 222Rn concentration at a rural site in France. Journal of Geophysical Research: Atmospheres, 120(2), 623-643.

[22] Rahman, M. J., & Hasan, M. A. M. (2014). Performance of wavelet transform on models in forecasting climatic variables Computational Intelligence Techniques in Earth and Environmental Sciences (pp. 141-154): Springer.

[23] Rasmusson, E. M., & Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. Monthly weather review, 110(5), 354-384.

[24] Romilly, P. (2005). Time series modelling of global mean temperature for managerial decision-making. Journal of environmental management, 76(1), 61-70.

[25] Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. Monthly weather review, 115(8), 1606-1626.

[26] Ropelewski, C. F., & Jones, P. D. (1987). An extension of the Tahiti–Darwin southern oscillation index. Monthly weather review, 115(9), 2161-2165.

[27] Shabbar, A., & Khandekar, M. (1996). The impact of el Nino-Southern oscillation on the temperature field over Canada: Research note. Atmosphere-Ocean, 34(2), 401-416.

[28] Struzik, Z. R. (2001). Wavelet methods in (financial) time-series processing. Physica A: Statistical Mechanics and its Applications, 296(1), 307-319.

[29] Szolgayová, E., Arlt, J., Blöschl, G., & Szolgay, J. (2014). Wavelet based deseasonalization for modelling and forecasting of daily discharge series considering long range dependence. Journal of Hydrology and Hydromechanics, 62(1), 24-32.

[30] Taylor, J. W., & Buizza, R. (2004). A comparison of temperature density forecasts from GARCH and atmospheric models. Journal of Forecasting, 23(5), 337-355.

[31] Trombe, P.-J., Pinson, P., & Madsen, H. (2012). A general probabilistic forecasting framework for offshore wind power fluctuations. Energies, 5(3), 621-657.

[32] Wang, W., Van Gelder, P., Vrijling, J., & Ma, J. (2005). Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes. Nonlinear processes in Geophysics, 12(1), 55-66.

[33] Wolter, K., & Timlin, M. S. (1993). Monitoring ENSO in COADS with a seasonally adjusted principal component index. Paper presented at the Proc. of the 17th Climate Diagnostics Workshop.

[34] Ye, L., Yang, G., Van Ranst, E., & Tang, H. (2013). Time-series modeling and prediction of global monthly absolute temperature for environmental decision making. Advances in Atmospheric Sciences, 30(2), 382.

[35] Zhou, B., He, D., & Sun, Z. (2006). Traffic predictability based on ARIMA/GARCH model. Paper presented at the Next Generation Internet Design and Engineering, 2006. NGI'06. 2006 2nd Conference on.

AUTHORS

First Author: Md. Nazmul Ahasan, M. Sc. in Statistics, Ph.D. Fellow, Institute of Environmental Sciences, Rajshahi University, e-mail: nazmula30@gmail.com.
Second Author: Md. Abdul Khalek, Ph. D.; Associate Professor, Department of Statistics, Rajshahi University, e-mail: mak.stat09@gmail.com
Third Author: Md. Mesbahul Alam, Ph. D.; Professor, Department of Statistics, Rajshahi University, e-mail: mesbahru@gmail.com.

Corresponding Author: Md. Nazmul Ahasan, Ph.D. Fellow, e-mail: nazmula30@gmail.com.