On the Spectral Density of the Modified-ARFIMA Model

A. Bello, M. Tasi'u[1, 2], H.G. Dikko[2] and B.B. Alhaji[3]

This study develops the Modified-ARFIMA Model and its spectral density for a recursive sequence differencing operator that can handle large data in time series that have long memory characteristics.

Keywords: Spectral density, modified-ARFIMA, long memory
DOI: https://doi.org/10.63255/03-4756.24/05

1.0 Introduction

Time series analysis uses spectral density analysis for several purposes, including data reduction, description, estimation, and hypothesis testing. Estimating spectral density functions and other spectral qualities associated with stationary multiple time series is crucial in any field where the features of the phenomenon under study can be described in terms of how it behaves in the frequency domain (Parzen, 1967). Nevertheless, Spectral analysis methods provide further tools for assessing how well different models fit the data (goodness-of-fit can be determined using sample spectra from the model's residuals) and for recommending potential models to fit (sample spectra often point out explanatory "variables" or "mechanisms" to incorporate into a time series). (Krampe & Paparoditis. 2022). The study modified the existing ARFIMA model of (Granger & Joyeux 1980; Hosking 1981) using a recursive sequence differencing operator that can handle large data sets without truncation, as observed by researchers such (Tanaka 1999) and Rahman and Jibrin 2019). The modified model is presented in Section 2. The significance of this research lies in the fact that the new model will forecast and estimate the parameters more easily than the existing model. Additionally, the primary goal of this paper is to derive the spectral density of the modified model. The spectral density of the modified model has been derived, and the summary results are presented in Section 5.

[1] Corresponding Author: M. Tasi'u; Email: bellobajogagsu@gmail.com. The views expressed in this article are those of the authors and do not necessarily reflect the position of the CISON. While efforts are made to ensure the accuracy of this paper, CISON bears no responsibility for errors, omissions, or interpretations. Authors are solely responsible for the accuracy, originality, and ethical standards of their work.
[2] Department of Statistics, Ahmadu Bello University, Zaria, Kaduna State, Nigeria
[3] Department of Mathematical Sciences, Gombe State University, Gombe, Gombe State, Nigeria

2.0 Review of Literature

Granger and Joyeux (1980) investigate the infinite filter's fractional differencing terms, which corresponds to the expansion of $(1-B)^d$. When the filter is applied to white noise, it creates a class of time series with different features, especially at low frequencies, which could be beneficial for long-memory forecasting. The aggregation of independent components is demonstrated to be a possible source of such models. The generation and estimation of these models are discussed, as well as their applicability to both simulated and real-life data. The use of infinite binomial series expansion has a limitation in that it truncates huge data into little data and does not capture the integer difference parameter d = 1.

Hosking (1981) studies the generalisation of ARIMA processes, commonly used in time series analysis, by considering fractional difference degrees. Introduce a fractional differencing operator that is based on an infinite binomial series expansion using the backwards-shift operator. This allowed for the modelling of both long-term persistence and anti-persistence in time series. Long-term persistent processes are useful in fields such as economics and hydrology, and the new family of models offers more flexibility in modelling both short and long memory behaviour of a time series at the same time. However, the limitations are the same as those previously established by Granger and Joyeux in 1980.

Beran et al. (1999) explore SEMIFARMA models through a kernel-based approach to estimate a nonparametric deterministic trend. They employ an approximate maximum likelihood method to estimate the fractional and difference parameters, as well as the autoregressive coefficients. Using the iterative plug-in concept for determining bandwidth in nonparametric regression with extended memory, a data-driven approach is proposed for estimating the entire model. However, the SEMIFARMA model has three key drawbacks: it does not account for parametric trends, relies on binomial series expansion to simplify large datasets, and fails to capture the integer differencing parameter when d =1

Meerschaert et al. (2014) present a new time series model for turbulent flow velocity data. To extend Kolmogorov's basic 5/3 spectral model, the new model uses tempered fractional calculus. To demonstrate the model's practical utility, it is applied to wind speed and water velocity in a huge lake. One limitation is that using binomial series expansion reduces large amounts of data to small amounts of data; second, the integer differencing parameter d = 1 is not captured; and third, the seasonal differencing parameter is not captured.

Pumi et al. (2019) developed a beta ARFIMA model designed for continuous random variables that take values within the range (0,1). The model incorporates explanatory variables and captures long-range dependence in the time series data. The authors also demonstrate the consistency and asymptotic normality of the estimator under specific conditions. It supports hypothesis testing, diagnostic analysis, and forecasting. To evaluate the performance of the partial likelihood estimators in finite samples, they conducted a Monte Carlo simulation. However, the model has limitations, such as the non-discreteness of the random sample and its similarity to the work of Granger and Joyeux (1980).

In their study, Monge and Infante (2022) utilised autoregressive fractionally integrated moving average (ARFIMA) models to examine the historical data on crude oil prices and determine whether shocks to the series result in lasting or short-term effects. An ARFIMA (2,d,2) model was found to be optimal, with the fractional differencing parameter estimated at 0.4. However, the large confidence interval prevents a clear rejection of either the I(0) or I(1) hypotheses. The observed uncertainty is likely a result of structural breaks or nonlinear trends present in the dataset.

Elwasify et al. (2023) examined whether long-term memory exists in the daily movements of the EGX30 stock index. Their analysis used an ARFIMA model with a mix of parametric and semi-parametric techniques. To identify long memory, they used various statistical tools such as the rescaled range (R/S) statistic, aggregated variance, and absolute moments. For estimation, they relied on semi-parametric methods like the Geweke and Porter-Hudak (GPH) method, Reisen's (SPR) method, and the local Whittle (LW) estimator, while employing parametric techniques like MLE, exact maximum likelihood (EML), modified profile likelihood (MPL), and conditional sum of squares (CSS).

In 2024, Furuoka et al. introduced a novel approach for testing fractional integration, known as the autoregressive neural network–fractional integration (ARNN–FI) test. This method is based on a multilayer perceptron neural network framework, as outlined by Yaya et al. (Oxford Bulletin of Economics and Statistics, 83(4):960–981, 2021). The study discusses the asymptotic theory and properties of the new test. Monte Carlo simulations show that as the sample size increases, the test's size and power distortions reduce. An empirical application of the test suggests that unemployment rates in three European countries do not show stationarity or mean reversion, supporting the hysteresis hypothesis.

Devianto et al. (2024) examined large-scale time series data for agricultural commodities by using an autocorrelation model that incorporates long-term trends, seasonal effects, and the impact of external variables. They chose chili prices as a case study due to their demonstration of long-term memory, seasonal variations, and sensitivity to external factors influencing production. These factors included specific periods, such as the month before the New Year and the week preceding the Eid al-Fitr celebration in Indonesia. To analyse price variations, the study applied the Seasonal Autoregressive Fractionally Integrated Moving Average (SARFIMA) model, which accounts for both seasonality and long-term memory. This model was enhanced by adding exogenous variables, resulting in the SARFIMAX model (SARFIMA with exogenous variables). A comparative analysis revealed that SARFIMAX outperformed SARFIMA in terms of accuracy, emphasising the significance of incorporating both seasonal effects and historical price data along with external factors. The enhanced SARFIMAX model provided a deeper understanding of price dynamics, offering valuable insights for sustaining chilli supply chains in the era of big data.

Swarnalatha et al. (2024) investigated the presence of long memory in time series, which is characterised by an autocorrelation function that decays slowly or follows a hyperbolic pattern. To accurately capture this behaviour, they applied the autoregressive fractionally integrated moving average (ARFIMA) model, well-suited for analysing historical financial data. The study focused on assessing the ARFIMA model's effectiveness in modelling long memory processes, using the Geweke and Porter-Hudak (GPH) method for parameter estimation. They analyzed monthly groundnut prices in Andhra Pradesh from January 2002 to December 2023. The ARFIMA (1, 0.43, 1) model provided the best fit, demonstrating strong short-term forecasting performance. It closely aligned with actual prices and outperformed the SARIMA(1,1,3)(0,1,2)12 model in terms of AIC, MSE, and RMSE. The findings indicated that the ARFIMA model offered more accurate price predictions than the SARIMA model.

3.0 Methodology

This research was intended to introduce a Modified-ARFIMA model that could be used to study the large financial and economic time series data. Given a time series $Y_1, ..., Y_m$, the observations are assumed to be trending, non-stationary and long memory. Also, they are assumed to have positive autocorrelation and long memory, denoted as d. The d is assumed to be in the range $1 < d < 1.5$. Given a series Y_m, $m \geq 0$ that is assumed or exhibit trendy (increasing and decreasing) and long memory. The fractional filter can be defined as follows: equation (1) is the differencing operator,

$$\nabla Y_m = Y_m - Y_{m-1} \qquad (1)$$

therefore,

$$d(\nabla Y_{m-1}) = d(Y_{m-1} - Y_{m-2}) \qquad (2)$$

minus equation (2) from equation (1) become

$Q_m = \nabla Ym - d(\nabla Y_{m-1})$

where Q_m is the filter, ∇ is the differencing operator and d is a sequence fractional differencing operator.

Note,

$\nabla Y_{m-1} = Y_{m-1} - Y_{m-2}$

$= Y_m - Y_{m-1} - d(Y_{m-1} - Y_{m-2})$

$$Q_m = Y_m - dY_{m-1} - Y_{m-1} + dY_{m-2} \qquad (3)$$

Using lag operators, the fractional filter in equation (3) can be represented as follows:

$= L^0Ym - dL^1Ym - L^1Ym + dL^2Ym$

$= \{1(1 - L) - dL(1 - L)\}Y_m$

$$Q_m = \{(1 - L)(1 - dL)\}Y_m \qquad (4)$$

The method for obtaining the fractional filters to induce nonstationarity is as follows:

$$Q_m = \{(1 - L)(1 - dL)\}Y_m = Y_m - Y_{m-1} - d(Y_{m-1} - Y_{m-2}) \qquad (5)$$

$$Q_{m-1} = \{(1 - L)(1 - dL)\}Y_{m-1} = Y_{m-1} - Y_{m-2} - d(Y_{m-2} - Y_{m-3}) \qquad (6)$$

$$Q_{m-n} = \{(1 - L)(1 - dL)\}Y_{m-n} = Y_{m-n} - Y_{m-n-1} - d(Y_{m-n-1} - Y_{m-n-2}) \qquad (7)$$

$$= Y_{m-n} - Y_{m-n-1} - d(Y_{m-n-1} - Y_{m-n-2}) = Y_{m-n} - Y_{m-n-1} - dY_{m-n-1} + dY_{m-n-2} \qquad (8)$$

then, by applying the lag operator in equation (8) (see Rahman and Jibrin, 2019), we can have

$$= L^n Y_m - dL^{n+1} Y_m - L^{n+1} Y_m + dL^{n+2} Y_m \qquad (9)$$

factor Y_m that is common in equation (9) we can have

$=(L^n - dL^{n+1} - L^{n+1} + dL^{n+2})Y_m$ (10)

Finally, the proposed fractional filter has the following general form:

$$Q_m = \sum_{n=0}^{N} (L^n - dL^{n+1} - L^{n+1} + dL^{n+2})Y_m \quad (11)$$

The ARMA model of Whittle (1951) $\varphi(L)Y_m = \theta(L)\varepsilon_m$ which became known to researches in time series is given as follows in the book of Box and Jenkins's (1970).

$$\phi(L)Y_m = \theta(L)\varepsilon_m \quad (12)$$

$$\varphi(L)(Q_m)Y_m = \theta(L)\varepsilon_m \quad (13)$$

where $\phi(L)$ and $\theta(L)$ are characteristic polynomials of the AR and MA process, Q_m is the sequence fractional filter, Y_m is the series, L is the lag operator, and $\varepsilon_m \sim WN(0, \sigma^2)$. The lag representation of the proposed Modified-ARFIMA model shown below in equation (14)

$$\varphi(L)\{(1 - L)(1 - dL)\}Y_m = \theta(L)\varepsilon_m \quad (14)$$

Where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$ are characteristic polynomials of AR and MA process, d is the fractional differencing filter, L is the backwards shift operator, and $\varepsilon_m \sim WN(0, \sigma^2)$ (Rahman and Jibrin, 2019).

1. The Spectral Density of the Modified-ARFIMA Model

The spectral density is a frequency domain representation of a time series that is directly related to the autocovariance time domain representation. In addition, it can be used to define a function of frequency instead of time, do regression, in which we regress, and split a time series into periodic components (Shumway and Stoffer, 2000).

To calculate or derive the Modified-ARFIMA model spectral density function can be recall from equation (14)

$$\varphi(L)\{(1 - L)(1 - dL)\}Y_m = \theta(L)\varepsilon \quad (15)$$

the spectral density for the above equation (15) (see Cryer and Kung, 2015), can be represented as this

$$S(\lambda) = \frac{\sigma^2}{2\pi}\left|\frac{\theta(e^{-i\lambda})}{\varphi(e^{-i\lambda})(e^{-i\lambda})d(e^{-i\lambda})}\right|^2 \quad (16)$$

$$= \frac{\sigma^2 \left|\theta(e^{-i\lambda})\right|^2}{2\pi \left|\varphi(e^{-i\lambda})\right|^2 \left|(e^{-i\lambda})\right|^2 \left|d(e^{-i\lambda})\right|^2} \tag{17}$$

then, let consider the numerator of the above equation (17) and prove the spectral.

$$\left|\theta(e^{-i\lambda})\right|^2 = \left|(1-\theta e^{-i\lambda})\right|^2 = (1-\theta e^{-i\lambda})(1-\theta e^{i\lambda}) \tag{18}$$

$= (1 - \theta e^{-i\lambda})(1 - \theta e^{i\lambda})$
$= (1 - \theta^{i\lambda} - \theta e^{-i\lambda} + \theta 2 e^{-i\lambda} \times e^{i\lambda})$
$= 1 - {}_\theta e^{i\lambda} - {}_\theta e^{-i\lambda} + {}_\theta 2 e^{-i\lambda + i\lambda}$
$= 1 - \theta e^{i\lambda} - \theta e^{-i\lambda} + \theta 2 e^{0}$
$= 1 - \theta[e^{i\lambda} + e^{-i\lambda}] + \theta 2 \times 1$
$= 1 + \theta 2 - \theta[e^{i\lambda} + e^{-i\lambda}]$

note that, $e^{i\lambda}=\cos(i_\lambda) + i\sin(i_\lambda)$, also $e^{-i\lambda} = \cos(-i_\lambda) + i\sin(-i_\lambda)$

$= 1 + \theta^2 - \theta[\cos(i\lambda) + i\sin(i\lambda) + \cos(-i\lambda) + i\sin(-i\lambda)]$

note that, $\cos(-i\lambda)=\cos(i\lambda)$ but $i\sin(-i\lambda)=-i\sin(i\lambda)$

$= 1 + \theta^2 - \theta[\cos(i\lambda) + i\sin(i\lambda) + \cos(i\lambda) - i\sin(i\lambda)]$
$= 1 + \theta^2 - \theta[\cos(i\lambda) + \cos(i\lambda)]$
$= 1 + \theta^2 - \theta[2\cos(i\lambda)]$

$$= 1 + \theta^2 - 2\theta\cos(i\lambda) \tag{19}$$

secondly let consider the denominator

$$\left|\varphi(e^{-i\lambda})(e^{-i\lambda})d(e^{-i\lambda})\right|^2 = \left|\varphi(e^{-i\lambda})\right|^2 \left|(e^{-i\lambda})\right|^2 \left|d(e^{-i\lambda})\right|^2 \tag{20}$$

$$\left|\varphi(e^{-i\lambda})\right|^2 = \left|(1-\varphi e^{-i\lambda})\right|^2 = (1-\varphi e^{-i\lambda})(1-\varphi e^{i\lambda})$$

$(1 - \phi e^{-i\lambda})(1 - \phi e^{i\lambda}) = 1 - \phi e^{i\lambda} - \phi e^{-i\lambda} + \phi 2 e^{-i\lambda} \times \phi e^{i\lambda}$
$= 1 - \phi e^{i\lambda} - \phi e^{-i\lambda} + \phi 2 e^{-i\lambda + i\lambda}$
$= 1 - \phi e^{i\lambda} - \phi e^{-i\lambda} + \phi 2 e^{0}$
$= 1 - \phi e^{i\lambda} - \phi e^{-i\lambda} + \phi 2 \times 1$
$= 1 + \phi 2 - \phi[e^{i\lambda} + e^{-i\lambda}]$
$= 1 + \phi^2 - \phi[\cos(i\lambda) + i\sin(i\lambda) + \cos(-i\lambda) + i\sin(-i\lambda)]$

$= 1 + \phi^2 - \phi[\cos(i\lambda) + i\sin(i\lambda) + \cos(i\lambda) - i\sin(i\lambda)]$
$= 1 + \phi^2 - \phi[\cos(i\lambda) + \cos(i\lambda)]$
$= 1 + \phi^2 - \phi[2\cos(i\lambda)]$
$= 1 + \phi^2 - 2\phi\cos(i\lambda)$

the second prove for spectral denominator

(21)

$$\left|e^{-i\lambda}\right|^2 = \left|1 - e^{-i\lambda}\right|^2 = (1-e^{-i\lambda})(1-e^{i\lambda}) \tag{22}$$

$(1 - e^{-i\lambda})(1 - e^{i\lambda}) = (1 - e^{-i\lambda} - e^{i\lambda} + e^{-i\lambda} \times e^{i\lambda})$
$= 1 - e^{-i\lambda} - e^{i\lambda} + e^{-i\lambda+i\lambda}$
$= 1 - e^{-i\lambda} - e^{i\lambda} + e^{0}$
$= 1 - e^{-i\lambda} - e^{i\lambda} + 1$
$= 2 - [e^{i\lambda} + e^{-i\lambda}]$
$= 2 - [\cos(i\lambda) + i \sin(i\lambda) + \cos(-i\lambda) + i \sin(-i\lambda)]$
$= 2 - [\cos(i\lambda) + i \sin(i\lambda) + \cos(i\lambda) - i \sin(i\lambda)]$
$= 2 - [\cos(i\lambda) + \cos(i\lambda)]$
$= 2 - 2 \cos(i\lambda)$

(23)

thirdly prove of spectral denominator

$$\left|de^{-i\lambda}\right|^2 = \left|(1-de^{-i\lambda})\right|^2 = (1-de^{-i\lambda})(1-de^{i\lambda}) \tag{24}$$

$= (1 - de^{-i\lambda})(1 - de^{i\lambda)}) = (1 - de^{-i\lambda} - de^{i\lambda} + d2e^{-i\lambda} \times e^{i\lambda})$
$= 1 - de^{-i\lambda} - de^{i\lambda} + d2e^{-i\lambda+i\lambda}$
$= 1 - de^{-i\lambda} - de^{i\lambda} + d2e^{0}$
$= 1 - de^{-i\lambda} - de^{i\lambda} + d^2 \times 1$
$= 1 + d2 - d[e^{i\lambda} + e^{-i\lambda}]$
$= 1 + d^2 - d[\cos(i_\lambda) + i \sin(i_\lambda) + \cos(-i_\lambda) + i \sin(-i_\lambda)]$
$= 1 + d^2 - d[\cos(i\lambda) + i \sin(i\lambda) + \cos(i\lambda) - i \sin(i\lambda)]$
$= 1 + d^2 - d[\cos(i\lambda) + \cos(i\lambda)]$
$= 1 + d^2 - d[2 \cos(i_\lambda)]$
$= 1 + d^2 - 2d \cos(i_\lambda)$ (25)

Therefore, the spectral density for the Modified-ARFIMA model is presented by the combination of the following equations (19), (21), (23) and (25) (Cryer and Kung, 2015) below

$$S(\lambda) = \frac{\sigma^2[1+\theta^2 - 2\theta\cos(i\lambda)]}{2\pi[1+\varphi^2 - 2\varphi\cos(i\lambda)][2-2\cos(i\lambda)][1+d^2 - 2d\cos(i\lambda)]} \tag{26}$$

5.0 Summary

The spectral density for the Modified Autoregressive Fractional Integrated Moving Average MARFIMA (p, d, q) model has three components. These components are the Moving Average component denoted by θ_j, the Autoregressive component denoted by ϕ_j, and the Fractional filter (1-L)(1-dL) component. The spectral density component for modified ARFIMA model as derived in section 3 is summarised and for crystal clear are given as follows.

i. MA(q) $= 1 + \theta^2 - 2\theta \cos(i\lambda)$
ii. AR(p) $= 1 + \phi^2 - 2\phi \cos(i\lambda)$

iii. Fractional Filter $= [2 - 2\cos(i\lambda)]\,[1 + d^2 - 2d\cos(\lambda_j)]$

The identified optimal model was ARFIMA (2, d, 2), with the fractional differencing parameter estimated at 0.4. However, the wide confidence interval hinders the ability to decisively reject the I(0) or I(1) hypotheses. This uncertainty is probably due to the influence of structural breaks or nonlinear trends within the data.

References

Beran, J. (1999). SEMIFAR models, a semiparametric fractional frame work for modelling trends, long-range dependence and nonstationarity. *Preprint*, University of Konstanz.

Box, G. E. P. & Jenkins, G. M. (1970).*Time series analysis, forecasting and control*, San Francisco: Holden-Day.

Devianto, D., Wahyuni, E., Maiyastri, M., & Yollanda, M. (2024). The seasonal model of chili price movement with the effect of long memory and exogenous variables for improving time series model accuracy. *Frontiers in Applied Mathematics and Statistics*, *10*, 1408381.

Elwasify, Alshaimaa & Isa, zaidi, Estimation of the Difference Parameter in ARFIMA Models Using Parametric and Semi-Parametric Methods (August 31, 2023). International Journal of Science, Mathematics and Technology Learning 31(2), 2023, Available at SSRN: https://ssrn.com/abstract=5004200

Furuoka, F., Gil-Alana, L. A., Yaya, O. S., Aruchunan, E., & Ogbonna, A. E. (2024). A new fractional integration approach based on neural network nonlinearity with an application to testing unemployment hysteresis. *Empirical Economics*, *66*(6), 2471-2499.

Granger C.W.J & Joyeux .R, (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time series Analysis*, 1, 15-29.

Hosking, J. R. M., (1981). Fractional differencing, *Biometrika*, 68, 165-176.

Jonathan D. Cryer & Kung – Sick Chen (2015); Time series Analysis with Application in R: Springer Texts in Statistics Second Edition 55- 130: e-ISBN: 978-0-387-75959-3

Krampe, J., & Paparoditis, E. (2022). Frequency domain statistical inference for high-dimensional time series. *arXiv preprint arXiv:2206.02250*.

Meerschaert, M.M., Sabzikar, F., Phanikumar, M.S. & Zeleke, A. (2014). Tempered fractional time series model for turbulence in geophysical flows. *Journal of Statistical Mechanics: Theory and experiment*, 1-13.

Monge, M., & Infante, J. (2023). A fractional ARIMA (ARFIMA) model in the analysis of historical crude oil prices. *Energy Research Letters*, *4*(1).

Parzen, E. (1967). The role of spectral analysis in time series analysis. *Revue de l'Institut International de Statistique*, 125-141.

Pumi, G., Valk, M., Bisognin, C., Bayer, F. M. & Prass, T. S. (2019). Beta autoregressive fractionally integrated moving average models. *Journal of Statistical Planning and Inference*, 200: 196-212.

Rahman, R.A. & Jibrin, S.A. (2019). Modeling and forecasting Tapis crude oil price: A long memory approach, {*Proceedings of the International Conference on Mathematical Sciences and Technology*} (MathTech2018), Pulau-Penang, Malaysia, AIP Conf. Proc. 2184, https://doi.org/10.1063/1.5136403. R

Shumway and Stoffer (2000). Time Series Analysis and Its and Its Applications Shumway and Stoffer, Springer, 2000

Swarnalatha , P., V. Srinivasa Rao, G. Raghunadha Reddy, Santosha Rathod, D. Ramesh, & K. Uma Devi. 2024. "Modelling Long Memory Time Series for Groundnut Prices in Andhra Pradesh With Autoregressive Fractionally Integrated Moving Average for Forecasting". *Journal of Scientific Research and Reports* 30 (7):289-302. https://doi.org/10.9734/jsrr/2024/v30i72145.

Tanaka, K. (1999). The nonstationary fractional unit root. Econometric theory, 15(4), 549-582.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Thesis, Uppsala University, Almqvist and Wiksell, Uppsala.

Yaya, O. S., Ogbonna, A. E., Furuoka, F., & Gil-Alana, L. A. (2021). A new unit root test for unemployment hysteresis based on the autoregressive neural network. *Oxford Bulletin of Economics and Statistics*, *83*(4), 960-981.