JEAS, Vol. 31 No. 146 (2025) pp (174-196)
P-ISSN 2518-5764 - E-ISSN 2227-703X
DOI: https://doi.org/10.33095/px3b7908

Journal of Economics and Administrative Sciences

Choosing an Appropriate Wavelet for VARX Time Series Model Analysis

Intisar Ibrahim Elias *
College of Management and Economics
Zakho University, Iraq.
***Corresponding author**

Taha Hussein Ali
College of Administration and Economics
Salahaddin University- Erbil, Iraq.

Received:3/12/2024 **Accepted: 12/1/2025** **Published: 1/4/2025**

Abstract:
The paper purports to improve the accuracy of VARX (vector autoregressive with exogenous variables) models adopted for economic time series analysis through wavelet transform techniques applied for noise reduction. The research assessed various wavelet types, including Coiflets, Daubechies, Symlets, Biorthogonal, and Reverse Biorthogonal, for the most appropriate wavelet to be used for improving the performance of the models. Furthermore, it made use of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for evaluating the efficacy of each wavelet in the dimension of noise reduction and predictive accuracy enhancement.

It is conclusively deduced that preprocessing through wavelet-based techniques markedly improves the reliability of VARX models, as it removes short-term noise without affecting the long-term trends in the economy. Application of Daubechies and Reverse Biorthogonal wavelets has emerged better in all size categories in reducing AIC and BIC values. Thus, the study lauds wavelet denoising techniques in associating industries with finance to provide sound arguments for policymakers and economists to analyze more accurately complex economic relations.

This work advances development in econometric modeling, detailing the importance of wavelet transformations in improving economic forecasts. Future studies should explore combining machine learning techniques with wavelet-based VARX models for further improvements in forecasting capabilities.

Keywords: Time Series; VARX Model; Optimal Wavelet; Level and Order Wavelet; Soft Rule; Universal Threshold.

1. Introduction:

The Iraqi economy is considered one of the developing economies that depends mainly on the oil sector for export. Therefore, the relationship between the position of foreign reserves and Government Expenditures is the basis for the movement of the Iraqi economy with all its components, and the link between them is the size of oil revenues. This relationship seeks to achieve the overall economic goals represented by the internal and external balances (Al-Waili and Al-Doski, 2008, P. 2). This study presents monthly financial and monetary data of the Iraqi economy, measured in billion Iraqi dinars, spanning from January 31, 2010, to July 31, 2024, with a total of 175 observations (n=175), essential indicators used to evaluate economic performance and analyze its dynamics are covered. These primary variables include Oil Revenues (ORv), Broad Money Supply (M2), Government Expenditures (Exp), and Official Foreign Reserves (OR) (Central Bank of Iraq, 2024). Oil revenue is Iraq's most significant source of income, influencing government spending levels and the amount held in foreign reserves. Broad Money Supply (M2) reflects the money circulating in the market, providing insight into the overall health of the economy. Within this context, Government Expenditures and official reserves demonstrate tangible economic impacts primarily influenced by fluctuations in oil revenues and monetary liquidity.

A VARX Vector Autoregressive with Exogenous Variables model is an extension of the traditional VAR model. In a VAR model, the variables are modelled as linear combinations of their past values and the past values of all other variables in the system**.** The VARX model expands upon this by allowing for the inclusion of exogenous variables, which are external factors that may influence the variables in the system but are not directly determined by the system itself. These exogenous variables can provide additional explanatory power and improve the forecasting performance of the model (Pesaran, 2015, p. 563; Parra et al., 2024, p. 2).

One crucial issue is lowering noise in time series data. A novel and effective technique for financial time series analysis is wavelet analysis, which lowers data noise. On various levels, wavelets can break down a signal or a time series. Consequently, the structure of the underlying signal, as well as trends, periodicities, leaps, or singularities that are not originally visible, are revealed by this decomposition (Ali et al., 2024, p. 345; Liu& Cheng, 2024, p. 2-3).

Wave techniques play a vital role in improving the quality of analysis and the accuracy of VARX models by reducing noise in economic data. These techniques contribute to providing reliable analyses and accurate forecasts and help in understanding the relationships between economic variables and their impact on economic growth, which enables decisions to be made based on accurate and effective forecast models.

The research problem is to improve the accuracy of VARX models used in analyzing monthly economic data by reducing noise using wavelet transformation. The study aims to identify the optimal wavelet with the most efficient level and order to achieve the lowest possible values of performance criteria (AIC and BIC), which contributes to improving the accuracy of forecasts and providing clearer economic insights. The study focuses on AIC and BIC due to their widespread use and reliability in balancing model complexity and goodness-of-fit. Future research could explore alternative criteria, such as MDL (Minimum Description Length), which has shown potential advantages in specific scenarios.

This requires testing all wavelet orders, including Coiflets, Daubechies, Symlets, Biorthogonal, and Reverse Biorthogonal, to choose the most appropriate wavelet to analyze economic data more accurately and enhance the ability of models to simulate real economic patterns and predict future trends. This methodology was applied to Iraqi economic data during the period 2010-2024 to analyze the impact of external variables such as money supply and oil revenues on internal variables. These results aim to support decision-makers and researchers in better understanding economic trends and providing more accurate analysis tools.

The research focused on two main aspects, including the theoretical aspect, which dealt with the VARX model and its selection criteria, wavelet families, universal threshold, soft threshold rule, and the proposed method. The applied aspect also included a case study on real data from the Iraqi economy, using simulation data derived from the Iraqi economy to verify the algorithm's validity. Statistical tools used included Minitab V21, EViews V12, and MATLAB R2023a, along with specially designed programs in MATLAB. The study concluded with results from simulation applications.

2.1 Literature review:

Many studies discussed that Wavelet transformation techniques have advanced time series analysis, uncovering hidden patterns and trends in economic and financial data through joint time-frequency analysis:

Donoho and Johnstone (1998) presented the importance of wavelet reduction techniques in removing noise from data, which contributed to improving the accuracy and stability of models (Donoho & Johnstone, 1998, p. 879).

Percival and Walden (2000) presented the importance of choosing the appropriate wavelet type and level of analysis in improving the accuracy of estimates extracted from time series, indicating that this directly affects the performance of predictive models (Percival & Walden, 2000, p. 310).

Cazelles and colleagues (2008) pointed out that wavelet analysis provides a flexible representation of time series, which facilitates the detection of hidden dynamic structures and the improvement of economic models (Cazelles et al., 2008). Schulte (2016) used wavelet analysis to examine time series that are non-stationary and nonlinear, showing its capability to reveal hidden patterns and frequency shifts over time, offering a better understanding of complex time dynamics (Schulte, 2016).

Nicholson et al. (2017) presented the VARX-L models that use structured regularization methods to lower model complexity and boost efficiency in high-dimensional settings. These models show that they can effectively manage large and complex datasets, increasing their predictive accuracy (Nicholson et al., 2017, p.1).

In financial terms, Tilak and Krishnakumar (2018) reviewed the effectiveness of applying reverse biorthogonal spline wavelets, emphasizing their ability to improve predictive models by effectively removing noise and detecting underlying patterns (Tilak and Krishnakumar, 2018). Rhif et al. (2019) emphasized how wavelet transform helps break down signals into detailed and basic parts, which helps keep important information while efficiently eliminating noise (Rhif et al., 2019). A study by Guo et al. (2022) showed that wavelet analysis improves model efficiency through methods like Wavelet Neural Networks (WNN), which indicate substantial progress in future uses (Guo et al., 2022, p. 58869).

Arrouxet et al. (2020) showed how those wavelets provide analytical flexibility to identify time series components such as trends and volatility (Arouxet et al., 2021). Similarly, Gonçalves et al. (2022) demonstrated the utility of using wavelets in selecting dynamic structures in organic systems, noting the importance of exploring different types of wavelet families to achieve the best results (Gonçalves et al., 2022).

Ali et al. (2023) showed how using multivariate wavelet reduction techniques can enhance the performance of traditional VAR models. These techniques, which handled multidimensional economic data, resulted in a notable decrease in noise and an increase in forecast accuracy compared to unoptimized models (Ali et al., 2023, p.9).

The primary feature of VAR models is that they are based on the lagged values of the endogenous variables, whereas the ARIMAX models allow for an improvement in accuracy by incorporating exogenous variables.

Hayder et al. (2023) presented a comparative study of VAR and ARIMAX models using Iraq's general budget data (2004–2020), focusing on foreign reserves and government spending. The research that applied the mean square error (MSE) approach and MATLAB in its implementation indicated that the VAR model is suitable for multidimensional datasets with interrelated variables, whereas ARIMAX models are more effective in scenarios requiring the inclusion of exogenous variables (Haydier et al., 2023, p.249).

Ahmed and Abdulkader (2023) investigated the improvement of ARIMAX models with the help of the bivariate wavelet denoising for monthly traffic accident forecasting. In this study, A wavelet-based multiresolution analysis was applied to support the fully authenticated ARIMAX models, and that sufficiently demonstrates the importance of the wavelet transforms and the decomposition levels in the process. In doing the denoising, several wavelet families were used, two of them are Daubechies and Coiflets wavelets, and in the former, Coiflets 3 worked the best. The results showed that a bivariate wavelet enhanced the ARIMAX models (Abdulqader& Ahmed, 2023, P.15).

Karamikabir et al. (2024) reviewed and examined how wavelet reduction techniques can enhance the precision of statistical models, positioning them as excellent tools for analyzing complex financial data (Karamikabir et al., 2024).

Ali et al. (2024) studied the effectiveness of applying multivariate wavelet reduction within a VAR model to analyze expenditure and revenue data in the Kurdistan Region of Iraq. The results showed the superiority of this technique over traditional methods, which contributed to improving the accuracy of economic forecasts by removing noise from financial data and achieving accurate forecasts for the period between 2022 and 2026 (Ali et al., 2024, p.345). Finally, Parra et al. (2024) pointed out the effectiveness of advanced waves in improving the VARX models by addressing noise and preserving underlying economic patterns, allowing for a deeper understanding of dynamic economic relationships (Parra et al., 2024, p. 50).

3. Theoretical Aspect:

The theoretical aspect presented some basic concepts about research from the statistical side, as shown in the following paragraphs.

3.1 VARX Model:

In general form, the Vector Autoregressive model with Exogenous variables (VARX) is an extension of the traditional Vector Autoregressive (VAR) framework by incorporating both endogenous and exogenous variables to capture their dynamic relationships. This extension permits the model to examine the interdependence between multiple time series while considering the influence of external variables. The VARX model employs a multivariate system where past values of endogenous variables, alongside relevant exogenous variables, serve as inputs to predict multiple future outputs. This dynamic interaction allows the model to capture interdependencies within the system while accounting for the influence of external variables.

Thus, the mathematical formula representation of the VARX (s, r) model, where s indicates the number of lags for the endogenous variables and r for exogenous variables can be described by the following formula:

$$Y_t = \mathrm{C} + \sum_{i=1}^{s} \Phi_i \mathrm{Y}_{t-i} + \sum_{j=1}^{r} \theta_j \mathrm{X}_{t-j} + \epsilon_\mathrm{t} \; \ldots (1)$$

Alternatively, it can be expressed using the lag operator B, as follows:

$$\left(I - \sum_{i=1}^{s} \Phi_i B^i\right) Y_t = C + \left(\sum_{j=1}^{r} \theta_j B^j\right) X_t + \epsilon_t \ldots (2)$$

Where:
Y_t: is a $k \times 1$ vector of dependent (endogenous) variables at a time t.
Y_{t-i}: are the lagged endogenous variables for all lags ($i = 1,2, \dots, s$).
X_{t-j}: is an $m \times 1$ vector of external (exogenous) predictors at lag $t - j$.
Φ_i: $k \times k$ matrices representing the coefficients of the endogenous variables' lags.
θ_j: $k \times m$ matrices representing the coefficients of the exogenous variables' lags.
C: is a $k \times 1$ vector of intercept terms.
ϵ_t: is a $k \times 1$ vector of error terms, accounting for unexplained randomness or residual influences.
B: is the Back shift Operator, where$B^k Y_t = Y_{t-k}$. The Back shift Operator B simplifies time-series modeling by compactly representing past values of variables, enabling the representation of relationships across time lags in models like VARX.
For this formula above, use any k number of dependent variables, s lags, m external predictors, and r lags for the predictors.
This framework becomes the modelling of complex interactions between endogenous variables while simultaneously capturing the impact of external predictors. It is a commonly utilised model in time series studies, especially for grasping changing connections (Lütkepohl, 2004, p.92; Lütkepohl, 2005, p. 386; Parra et al., 2024, p. 2).

3.3 Model Selection Criteria:

Play these criteria AIC and BIC a crucial role in model selection by finding a balance between model accuracy and complexity. These criteria are extensively used throughout statistical and machine-learning fields to determine the best model from a set of different models. This study focused on AIC and BIC due to their effectiveness in achieving this balance, while dependency tests were not a primary consideration (Burnham& Anderson, 2004, p.268, 275; Acquah, 2010, p.2-3):

3.3.1 Akaike Information Criterion (AIC)

The lower AIC value represents the best model, which indicates a good fit with fewer parameters and is calculated according to the following equation:

$$AIC = 2v - 2Ln(L) \dots (3)$$

Where:
v: is the number of parameters in the model for each v = 1, 2, and 3,4.

L: is the log-likelihood function of the model and can be expressed as follows:

$$L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}(y_t - \widehat{y_t})^2$$

Where:

n: the number of observations in the time series.
σ^2: the variance of the residuals (errors).
y_t: the actual value of the time series at time$t$.
$\widehat{y_t}$: the predicted values from the model at time t.
$-\frac{1}{2\sigma^2}\sum_{t=1}^{n}(y_t - \widehat{y_t})^2$: the most significant term, representing the sum of squared errors divided by the variance.
A lower AIC value indicates a better model, as it implies a good fit with fewer parameters.

3.3.2 Schwarz Information Criterion (SIC or BIC)

The Bayesian Information Criterion (BIC) or the Schwarz information criterion called by (SIC, SBC, and SBIC) is another criterion used for model selection that also considers the goodness of fit and model complexity but supposes a stricter penalty for the number of parameters compared to AIC. It is calculated according to the following equation:

$$BIC = vLn(n) - 2Ln(L) \dots (4)$$

Where:

- n: is the number of observations.
- v: is the number of parameters in the model for each v = 1, 2, and 3, 4.
- L: is the log-likelihood function of the model.

Like AIC, a lower BIC value indicates a better model, but BIC tends to favor simpler models compared to AIC due to the $Ln(n)$ term.

3.4 Wavelets:

Wavelets are oscillatory functions characterized by an amplitude that initiates at zero, grows or shrinks, and then returns to zero in a repetitive cycle. These brief oscillations, known as wavelets, are integral to modern signal processing and time series analysis (Omer et al., 2024, p. 114).

These wavelets are advanced mathematical tools in signal processing and time series analysis, known for effectively handling non-stationary signals. Unlike Fourier transformations, which decompose signals into sine and cosine functions, wavelets provide dual localization in time and frequency domains. The mother wavelet, ψ(t), serves as the primary function from which all other wavelets are generated through scaling and translation (Mallat, 2009, pp. 57-60; Daubechies, 1992, pp. 30-35).

$$\psi^*_{a,b}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-b}{a}\right) \dots (5)$$

Thus, (a) is the scaling parameter that controls the frequency, and (b) is the translation parameter that controls the time localization.

The discrete wavelet transform (DWT**)** provides a set of coefficients in both time and frequency domains, summarizing the information of all observations with a reduced number of coefficients. DWT is widely used, especially when data contains contaminants or noise (Ali & Saleh, 2022, p. 926).

The discrete wavelet transform (DWT) denotes the coefficients utilized to represent the signal in terms of wavelets, as follows (Percival & Walden, 2000, pp. 1-19):

$$Y(t) = \sum_{a,b} W_{a,b}\, \psi_{a,b}(t) \dots (6)$$

where $W_{a,b}$ are the wavelet coefficients and $\psi_{a,b}(t)$ are the wavelet basis functions at scale a and translation b.

3.5. Wavelet Families:

The following are some of the wavelet families used in the application section of this study:

3.5.1 Coiflets Wavelets:

Coiflets wavelets are a major development in wavelet theory, especially for signal processing tasks. Ingrid Daubechies introduced them in the spring of 1989 at Ronald Coifman's request. Coifman's proposal went beyond the previous emphasis on the wavelet function (Ψ) and expanded the idea of vanishing moments to scaling and wavelet functions. Coiflets, denoted as Coif$N$ (where N is the rank), have properties tied to their rank: the candidate length is $6N$, the number of vanishing moments for the wavelet function is $L = 2N$, and for the scaling function is $L_1 = 2N - 1$. Coiflets exhibit compact support, orthogonality, and near symmetry (Ali & Mustafa, 2016, p. 436).

The mathematical form for the Coiflets wavelet ψ (t) is derived from the mother wavelet through scaling and translation, and it is calculated according to the following equation:

$$\psi_{a,b}(t) = 2^{a/2}\,\psi(\ 2^a\, t - b) \cdots (7)$$

Where (a) represents the scale parameter and (b) is the translation parameter.

The following equation explains how the denoised signal $Y_{denoised}$ is reconstructed using wavelet coefficients and wavelet basis functions after applying thresholding in the Coiflets wavelet denoising process. The equation is defined as:

$$Y_{denoised} = \sum_{a,b} \widehat{W}_{a,b}\, \psi_{a,b}(t) \cdots \; (8)$$

Thus, $\widehat{W}_{a,b}$ are the wavelet coefficients after applying thresholding to remove noise, and $\psi_{a,b}(t)$ are the wavelet basis functions used for signal reconstruction.

Equation (8) represents the wavelet noise removal equation or wavelet reconstruction equation after thresholding, which clarifies the notion of wavelet denoising, in which the signal is processed to remove unwanted noise while preserving its major components. Thus, Wavelet denoising, particularly utilizing Coiflets wavelets, is most applied in signal processing and time series analysis due to its efficiency in noise repression and keeping signal safety (Mallat, 2009, p. 74).

3.5.2 Daubechies Wavelets:

The figure of the Daubechies wavelet was first established by Ingrid Daubechies in 1988, It is a subclass of wavelets that find use in signal processing applications and time series analysis. These wavelets are defined in terms of a certain number of vanishing moments in addition to being compactly supported such that they can be used to represent polynomials effectively. Their compact support and efficient mathematical structure have made discrete wavelet analysis practical and applicable in a wide range of fields.

These wavelets made discrete wavelet analysis practical and are denoted as D or dbN, where N represents the wavelet's order corresponding to the number of vanishing moments. For example, db2 refers to the second-order wavelet with two vanishing moments (Daubechies, 1992, p. 167; Ali et al., 2023, p.11; Jalal et al., 2024, p. 16).

The mathematical representation of Daubechies wavelets is based on the scaling function $\phi(t)$and the wavelet function $\psi(t)$, defined as follows:

$$\phi(t) = \sum_{b=0}^{N-1} a_b\, \phi(2t-b) \cdots (9)$$

$$\psi(t) = \sum_{b=0}^{N-1} c_b\, \phi(2t-b) \quad \text{where} \quad c_b = (-1)^b a_{N-1-b} \cdots (10)$$

where a_k are the scaling coefficients, c_b are the wavelet coefficients, and N denotes the number of coefficients, directly related to the vanishing moments. The wavelet coefficients c_b are derived from the scaling coefficients a_k using the relation $c_b = (-1)^b a_{N-1-b}$.

As N increases, Daubechies wavelets demonstrate improved regularity and stronger localization in the frequency domain; however, this also results in higher computational complexity.

To solve the problem of symmetry and decrease the distortion of signals during reconstruction, Daubechies suggested improved families known as Symlets and Coiflets. These families preserve the essential properties of dbN wavelets, including continuity, compact support, filter length, etc., but they manage to get better symmetry. These features make Daubechies wavelets indispensable tools for multiresolution analysis and signal decomposition (Daubechies, 1992, p. 167; GUO et al., 2022, p. 588732).

3.5.3. Symlets Wavelets:

Developed as an improvement of the original Daubechies wavelets to achieve better symmetry, Symlets wavelets were initially presented by Ingrid Daubechies in 1992. While maintaining the same number of vanishing moments and compact support, owing to these wavelets having better phase characteristics, they are said to have approximately linear phase responses. Such wavelets are particularly useful in applications where phase distortion is required to be minimal, for example, edge detection in image processing, precise signal reconstruction, and retaining details in biomedical signals such as EEG and ECG, ensuring accurate diagnostics and effective data analysis.

Thus, the mathematical framework of Symlets wavelets is based on the same formulation as Daubechies wavelets, defined by the scaling function $\phi(t)$ and the wavelet function $\psi(t)$, ensuring orthogonality and compact support. The main difference is the optimization of the scaling coefficients akin to Symlets to achieve better symmetry while maintaining the essential qualities of Daubechies wavelets, like compact support and the number of vanishing moments (Daubechies, 1992, p. 198; Mallat, 1999, p. 254).

3.5.4 **Biorthogonal Wavelets:**

Biorthogonal wavelets are a subclass of wavelets that are not strictly orthogonal but have biorthogonality. This property allows separate scaling and wavelet functions for analysis and synthesis, providing greater flexibility in signal compression and denoising applications.

The biorthogonal wavelet functions $\phi(t)$ and $\psi(t)$ are defined alongside their dual functions $\tilde{\phi}(t)$ and $\tilde{\psi}(t)$, as follows:

$$\phi(t) = \sum_b h_k\, \phi(2t-b), \qquad \tilde{\phi}(t) = \sum_b \tilde{h}_b\, \tilde{\phi}(2t-b) \cdots (11)$$

$$\psi(t) = \sum_b g_b\, \phi(2t-b),\ \tilde{\psi}(t) = \sum_b \tilde{g}_b\, \tilde{\phi}(2t-b) \cdots (12)$$

Here, h_b, g_b, $\tilde{h}_b$, and $\tilde{g}_b$ are filter coefficients satisfying the biorthogonal conditions. Structural characteristics of the biorthogonal wavelets allow greater flexibility in the filter design, providing better symmetry- vanishing moments-complexity trade-off (Strang & Nguyen, 1996, p. 208; Cohen et al., 1992, p. 488-491).

3.5.2 Reverse Biorthogonal Wavelets

Reverse biorthogonal wavelets are one of the extensions of the biorthogonal wavelet family. In comparison to the standard biorthogonal wavelets, the analysis and synthesis roles are switched for this kind of wavelet, hence making them particularly efficient for certain signal processing tasks, including edge detection and data compression.

Mathematically, the reverse biorthogonal wavelets share the same formulation as biorthogonal wavelets but with reversed filter roles for analysis and synthesis, as follows:

$$\tilde{\phi}(t) = \sum_b \tilde{h}_b\, \phi(2t-b), \psi(t) = \sum_b g_b\, \tilde{\phi}(2t-b) \cdots (13)$$

The reverse setup has special features that help maintain important properties, such as vanishing moments and compact support, while processing the symmetrical parts of signals (Cohen et al., 1992, pp. 488–491; Tilak& Krishnakumar, 2018, p.66).

3.6. **Universal Threshold Method:**

Universal thresholding is generally used in wavelet shrinkage to denoise data or images. The thresholding technique aims to select an optimal threshold value founded on the statistical characteristics of the data. The idea of universal thresholding selection is to estimate the noise level in the data and apply a threshold that adapts to this noise level. Universal thresholding is considered one common method for Stein's unbiased risk estimate (SURE), which represents an unbiased estimate of the mean squared error (MSE) of the denoised signal. The threshold value was chosen to decrease and minimize this estimated MSE. The SURE threshold is commonly called the "universal threshold" since it does well across an extensive domain of data and noise kinds without required prior knowledge of the noise properties. Then, adaptively choosing the threshold based on the same data and universal thresholding can effectively remove noise while keeping important data characteristics (Karamikabir et al., 2021, p. 1729; Jalal et al., 2024, p. 162).

The following equation calculates the universal threshold method:

$$\lambda_u = \hat{\sigma}_{MAD}\sqrt{2log(n)} \cdots (14)$$

Where (n) is the number of observations while ($\hat{\sigma}_{MAD}$) is the standard deviation estimator of the detail coefficients, and this estimate is calculated as follows:

$$\hat{\sigma}_{MAD} = \frac{MAD}{0.6745} \cdots (15)$$

The wavelet coefficients' median absolute deviation (MAD) at the delicate scale is calculated as follows:

$$MAD = median\left[\left|W_{1,0}\right|, \left|W_{1,1}\right|, \dots, \left|W_{1,\frac{N}{2}-1}\right|\right] \cdots (16)$$

Here, $W_1 = W_{1,0}\, W_{1,1}, \dots, W_{1,\frac{N}{2}-1}$ represent the discrete wavelet transformation coefficients at the first level for observations, while the constant (0.6745) is the median of the standard normal distribution (Omer et al., 2024, p. 117).

3.7. Soft Threshold Rule:

A method used in signal and image processing, especially about wavelet shrinkage, is called soft thresholding. Introduced by Donoho and Johnstone, it extends the hard thresholding method by shrinking wavelet coefficients toward zero to reduce the impact of noise. This method is mathematically defined as follows (Jalal et al., 2024, p. 163; Ali et al., 2023, p.13):

$$\widehat{W}_n = sign(W_n)\,.max(|W_n| - \lambda, 0)_+ \cdots (17)$$

or

$$\widehat{W}_n = Sign(W_n)(|W_n| - \lambda)_+ \cdots (18)$$

Where:

$$Sign\,(W_n) = \begin{Bmatrix} +1 & if\ W_n > 0 \\ 0 & if\ W_n = 0 \\ -1 & if\ W_n < 0 \end{Bmatrix} \cdots (19)$$

and

$$(|W_n| - \lambda)_+ = \begin{cases} (|W_n| - \lambda) & if\ \ (|W_n| - \lambda) \geq 0 \\ 0 & if\ \ (|W_n| - \lambda) < 0 \end{cases} \cdots (20)$$

In these equations, W_n represents the wavelet coefficients and λ is the threshold value. Soft thresholding is particularly effective at noise suppression while preserving significant signal features, although it may introduce some signal distortion. Therefore, it is the focus of the application side of wave noise removal.

The soft threshold rule works by adjusting the values below the threshold to zero, while the values above the threshold are reduced to the threshold amount. This results in a continuous mapping function. It can balance noise reduction and signal preservation, so soft thresholding is a popular choice in wavelet-based noise removal applications and studies (Li et al., 2023, p. 1244).

3.8. Proposed Method:

The use of wavelets in this study is not entirely new, as wavelets have previously been applied within the broader framework of the VARIMAX model. However, the proposed model represents a specific case of the VARIMAX model, where the moving average order is set to zero. This specific application allows for a more focused analysis and enhances the accuracy of signal processing and noise reduction within the context of the study's data.

The proposed algorithm includes data de-noise of the VARX Model using a variety of wavelets with the Universal thresholding estimation method and applying the soft rule through the following steps:

1.Level of wavelet decomposition, specified as a positive integer less than or equal to ($\log_2 n$) where n is the number of observations in the time series. Set L = 1, 2, …, l, where l returns the maximum level L possible for a wavelet decomposition of data using the wavelets specified by Coiflets, Daubechies, Symlets, Biorthogonal, and Reverse Biorthogonal (default value equal to 3). Each decomposition level separates the signal into approximate and detailed components, enhancing the ability to isolate noise and capture trends. The selection of wavelet families ensures flexibility to address different data types, with the default level offering a practical balance between computational efficiency and analytical precision.

2. Denoises VARX time series data for all L, as defined in equation (8).

3. Estimating the parameters of the VAR Model for each L, as described in equation (1), are estimated using statistical techniques like Ordinary Least Squares (OLS) to calculate the coefficients for each lag order. After estimating the parameters, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are computed for all models, based on equations (5) and (6), to find the best lag structure. Diagnostic tests, like Granger Causality, were not conducted in this study, as the focus was primarily on optimizing model parameters and minimizing noise to improve forecasting accuracy.

4. Determine the optimal level (OL) for wavelet decomposition that gives the least AIC value for the given data.

5. Denoises VARX time series data at OL for a variety of wavelets (Table 1) and Universal thresholding estimation method using equations (14-16) according to the Universal thresholding and applying the soft rule according to equations (17-20).

6. Determine the optimal order (OO) for wavelet decomposition that gives the least AIC value for the given data.

7. Use OL and OO for Coiflets wavelet with Universal thresholding estimation methods and apply the soft rule in data de-noise.

8. Use data de-noise from point (7) in estimating the VARX Model according to equation (1).

4. Numerical Results:

This study applies multivariate wavelet shrinkage to the VARX model, comparing its performance to both traditional approaches and models without wavelet transformations. The target is to determine the optimal wavelet and improve model accuracy that achieves the lowest AIC and BIC values, using simulated data derived from the Iraqi economy.

4.1 Simulation Study:

The simulation was carried out using projected parameters based on actual data, specifically focusing on the monthly financial and monetary statistics of the Iraqi economy, represented in billion Iraqi dinars, spanning the timeframe from January 31, 2010, to July 31, 2024, which consists of a total of 175 data points ($n = 175$). The data collection includes metrics such as Official Reserves (OR), Oil Revenues (ORv), Expenditures (Exp), and Broad Money (M2). In this research, these projected parameters acted as inputs in the simulation process, ensuring that the model accurately mirrored the real economic behaviors while enhancing its efficacy. This strategy allows for a practical assessment of the model's ability to predict outcomes and efficiency, offering important insights into the relationships between significant financial and monetary factors.

In this section, we concentrate on evaluating the performance of multivariate wavelet shrinkage using a variety of wavelets, as described below:

Table 1: Multivariate Wavelet Shrinkage with Various Wavelet Types

Wavelet Type	Order Range
Coiflets	coif1 to coif5
Daubechies	db1 to db45
Symlets	sym1 to sym30
Biorthogonal	bior1.1 to bior6.8
Reverse Biorthogonal	rbio1.1 to rbio6.8

Biorthogonal	Reverse Biorthogonal	Orders
bior1.1	rbio1.1	1
bior1.3	rbio1.3	2
bior1.5	rbio1.5	3
bior2.2	rbio2.2	4
bior2.4	rbio2.4	5
bior2.6	rbio2.6	6
bior3.1	rbio3.1	7
bior3.3	rbio3.3	8
bior3.5	rbio3.5	9
bior3.7	rbio3.7	10
bior3.9	rbio3.9	11
bior4.4	rbio4.4	12
bior5.5	rbio5.5	13
bior6.8	rbio6.8	14

In Table 1, all ranks of each of the above wavelets were tested to determine the optimal order, the optimal level, and the best wavelet type for these wavelets that achieve the lowest AIC and BIC values. The threshold level was estimated using the Universal method with soft thresholding to enhance data quality and reduce noise.

The noise level of 40 decibels (dB) was chosen to analyze the monthly financial data of the Iraqi economy for the period from January 31, 2010, to July 31, 2024, because it provides an ideal balance between reducing noise and preserving important economic signals. This level was determined through extensive testing, which involved evaluating multiple noise levels using the decibel scale, a logarithmic measure of the signal-to-noise ratio (SNR). This level allows long-term trends to be accurately extracted by reducing noise without affecting the accuracy of the data. If the noise level increases above 40 dB, this increase may cause more noise, while reducing it may lead to the loss of some important details. Therefore, 40 decibels (dB) were adopted to ensure an effective and balanced analysis of economic policies and their effects on various financial variables.

These techniques were applied using simulated data derived from the monthly financial and monetary data of the Iraqi economy, measured in 1 billion Iraqi dinars, covering the period from January 31, 2010, to July 31, 2024, with a total of 175 observations (n=175). The dataset includes indicators such as Official Reserves (OR), Oil Revenues (ORv), Expenditures (Exp), and Broad Money (M2). The VARX (s, r) model, with lag orders set to $s = 2$ and $s = 3$ for the endogenous variables (dependent variables) 'Exp' and 'OR', and $r = 1$ For the exogenous variables (independent variables), 'M2' and 'ORv' were applied to evaluate their impact on the internal economic system. The simulation was based on parameters estimated from the VARX model. The goal was to select the optimal wavelet and the most suitable model based on AIC and BIC criteria, determining the optimal level and order for denoising. This study relied on sample sizes of 150, 175,200, and 300, with 1000 iterations to evaluate performance.

The simulation experiment was conducted using different sample sizes (n) based on $n = 2^j$, where j represents the level and is a positive integer, with the maximum level denoted by (L). Four sample sizes were chosen for the simulation, as shown below:

- For $j = 7$ $(i.e, maxL = 7), n = 2^7 = 128)$, which is less than the original sample sizes of $n = 150$,$175, and$ 200
- For $j = 8$ $(i.e, maxL = 8), n = 2^8 = 256)$, which is less than the original sample sizes of $n = 300$

In summary, the simulation used sample sizes based on powers of two, providing a robust range for analysis.

Therefore, an algorithm was proposed to determine the optimal level and order for each of the previously defined wavelets, based on the minimum values of the AIC and BIC criteria in the VARX time series model. The algorithm utilized the Universal thresholding method with a soft thresholding rule to identify these optimal values. A comparison was then made between the efficiency of the proposed method, the traditional method, and the approach without wavelet transformations.

In the traditional method, the analysis level (L = 3) and order (N = 3) were applied to all wavelets, including Coiflets, Daubechies, and Symlets, while the analysis level (L = 3) and order (N = 8) were applied to two wavelets, including Biorthogonal, and Reverse Biorthogonal, which represent the default values in MATLAB23 program. The analysis was conducted on different sample sizes. This approach aims to provide a comprehensive evaluation of wavelet-based denoising techniques to improve the performance of the VARX model, contributing to more accurate and efficient analytical results.

Thus, the Multivariate Wavelet Shrinkage technique will be applied in conjunction with the VARX model for data analysis. The results of the proposed method will be compared against those of the traditional approach, and without wavelet transformation 40dB as noise level. The primary objective was to identify the optimal wavelet and model based on the AIC and BIC criteria, determining the best level and order for denoising, as shown:

4.1.1 VARX (2,1) Model:

The simulation process utilized parameters estimated from real-world economic data, representing the Iraqi economy's monthly financial and monetary dynamics. The generated data spans a range of sample sizes (150, 175, 200, and 300 observations) to assess the robustness of the VARX (2,1) model with lag orders set to $s = 2$ and $r = 1$, under varying conditions. This study examines the performance of various denoising methods applied to a two-dimensional VARX (2) model with two predictors applied to evaluate the impact of independent Variables (represents predictors), such as 'M2' (Broad Money) and 'ORv' (Oil Revenues), along with dependent Variables (represents dimensional), including 'Exp' (Expenditures) and 'OR' (Official Reserves), on the internal economic system. The analysis was conducted on sample sizes of 150, 175, 200, and 300, with 1000 iterations for robust performance evaluation. The model's performance was assessed using both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), where lower values represent optimal performance for the model. Data denoising was performed through multivariate wavelet shrinkage, employing a variety of wavelet families, including Coiflets, Daubechies, Symlets, Biorthogonal, and Reverse Biorthogonal, effectively reducing data contamination. In contrast, an alternative method without wavelet transformation is used. This approach provided a comprehensive evaluation of the effects of noise on model accuracy and the efficacy of noise removal techniques, as shown in Table 2.

Table 2: Average AIC and BIC criteria for a Two-dimensional VARX (2) model with two predictors

Sample size	Method	Optimal Wavelet Selection	Level (L)	Order (N)	AIC	BIC
150	Classical	coif3	3	3	1737.8525	1791.8023
	Proposed	coif4	*4	*4	*219.5024	*273.4522
	Classical	db3	3	3	948.4846	1002.4345
	Proposed	db27	*6	*27	* -1253.1773	* -1199.2275
	Classical	sym3	3	3	948.4846	1002.4345
	Proposed	**sym16	**3	**16	** -1621.5881	**-1567.6382
	Classical	bior3.3	3	8	1970.4478	2024.3976
	Proposed	Bior4.4	*3	*14	*-826.2192	*-772.2694
	Classical	rbio3.3	3	8	-33.2619	20.6879
	Proposed	rbio3.3	*5	*8	*-1504.7243	*-1450.7745
	Without wavelet		-	-	15217.5245	15271.4743
175	Classical	coif3	3	3	-51.7009	5.0584
	Proposed	coif4	*4	*4	*-631.1704	*-574.4111
	Classical	db3	3	3	891.6282	948.3874
	Proposed	**db33	**7	**33	**-2964.0153	**-2907.2561
	Classical	sym3	3	3	891.6282	948.3874
	Proposed	Sym21	*4	*21	*-1801.2615	*-1744.5022
	Classical	bior3.3	3	8	557.9473	614.7065
	Proposed	bior5.5	*3	*13	*-1899.3099	*-1842.5506
	Classical	rbio3.3	3	8	-1486.9191	-1430.1598
	Proposed	rbio3.5	*6	*9	*-2199.8322	*-2143.0729
	Without wavelet		-	-	16280.8476	16337.6068
200	Classical	Coif3	3	3	-4050.3018	-3991.113
	Proposed	coif4	*6	*4	*-4298.6937	*-4239.5049
	Classical	db3	3	3	-3770.9408	-3711.752
	Proposed	**db27	**6	**27	**-4899.3125	**-4840.1237
	Classical	sym3	3	3	-3770.9408	-3711.752
	Proposed	Sym19	*6	*19	*-4384.991	*-4325.8022
	Classical	Bior3.3	3	8	-3337.088	-3277.8992
	Proposed	Bior5.5	*6	*13	* -4583.605	* -4524.4164
	Classical	rbio3.3	3	8	-2570.4253	-2511.2365
	Proposed	rbio3.5	*5	*9	*-4536.769	*-4477.5809
	Without wavelet		-	-	-1982.462	-1923.2731
300	Classical	coif3	3	3	-6263.2594	-6196.7117
	Proposed	coif4	*7	*4	*-6474.048	*-6407.5003
	Classical	Db3	3	3	-5942.9143	-5876.3666
	Proposed	db25	*7	*25	*-7289.878	*-7223.3311
	Classical	sym3	3	3	-5942.9143	-5876.3666
	Proposed	sym20	*7	*20	* -6721.0114	* -6654.4637
	Classical	bior3.3	3	8	-5093.2203	-5026.6726
	Proposed	Bior5.5	*7	*13	*-7332.2785	*-7265.7308
	Classical	rbio3.3	3	8	-3967.3531	-3900.8054
	Proposed	**rbio3.7	**5	**10	**-7530.8911	**-7464.3434
	Without wavelet		-	-	-2975.1493	-2908.6016

An asterisk (*) denotes the optimal value for all the wavelets corresponding to the minimum AIC and BIC values.

An asterisk (**) denotes the optimal wavelet value corresponding to the minimum AIC and BIC values.

4.1.1.1 Discussion:

Table (2) presents the results, which consistently show that, for each sample size (150, 175,200, and 300), the proposed method using universal thresholding with the soft rule achieved the lowest AIC and BIC values, as indicated by the asterisks (*). This demonstrates that the method, when applied with optimal levels (L) and orders (N), was the most efficient at denoising the data

The results presented in Table (2) provide a clear indication of the optimal wavelet performance across various sample sizes for a two-dimensional VARX (2) model with two predictors. The wavelets were evaluated based on the lowest AIC and BIC values, as indicated by the asterisks (**) and highlighted in yellow, which serve as indicators of model efficiency in removing noise from the data.

This indicates that the optimal wavelet was sym16 at a decomposition level (L) of 3 (where the maximum level is 7, depending on the sample size (150)) and an order (N) of 16 for the sample size of 150. For the sample size of 175, the optimal wavelet was the db33 wavelet, which showed superior performance at a level (L) of 7 (where the maximum level is 7, depending on the sample size of 175) and an order (N) of 33. For the sample size of 200, the db27 wavelet was optimal at level (L) 6 (where the maximum level is (7), depending on the sample size (200)) and order (N) 27. Lastly, for the sample size of 300, the rbio3.7 wavelet was optimal at level (L) 5 (where the maximum level is (8), depending on the sample size (300)) and order (N) 10 delivered the best results. All the optimal wavelets mentioned earlier are as follows: sym16 (Symlets), db33, db27 (Daubechies wavelet), and rbio3.7 (Reverse Biorthogonal wavelet), the most efficient for denoising the data for each sample size (150, 175,200, and 300). For Table 2, in the context of a two-dimensional VARX (2) model with two predictors. The results in table (3) are as follows:

Table 3: Optimal Wavelet based on lowest AIC and BIC criteria for a two-dimensional VARX (2) model with Two Predictors.

Sample size	Method	Wavelet	Level (L)	Order (N)	AIC	BIC
150	Proposed	**sym16	**3	**16	** -1621.5881	**-1567.6382**
175	Proposed	**db33	**7	**33	**-2964.0153	**-2907.2561**
200	Proposed	**db27	**6	**27	**-4899.3125	**-4840.1237**
300	**Proposed**	**rbio3.7**	**5**	**10**	**-7530.8911**	**-7464.3434**

The results presented in Table (3) demonstrate in the context of a two-dimensional VARX (2) model with two Predictors, that Daubechies wavelets, specifically db27 and db33, exhibited outstanding performance across various sample sizes (n=175, 200). Similarly, Reverse Biorthogonal wavelets (Rbio3.7) performed exceptionally well, particularly with larger sample sizes (n=300). Additionally, Symlets (sym16) were found to be highly effective for smaller sample sizes (n=150), delivering optimal results in noise reduction. Consequently, the wavelets sym16, db33, db27, and rbio3.7 were identified as the most efficient across different sample sizes, with their optimal levels and orders adjusting based on the specific sample size.

4.1.2 VARX (3,1) model:

Building upon the methodology used in the VARX (2) model, the VARX (3) model was evaluated using wavelet shrinkage techniques to enhance performance across varying sample sizes and noise levels. The simulation process utilized parameters estimated from real-world economic data, representing the Iraqi economy's monthly financial and monetary dynamics. The generated data spans a range of sample sizes (150, 175, 200, and 300 observations) to assess the robustness of the VARX (3,1) model with lag orders set to $s = 3$ and $r = 1$, under varying conditions. Performance was evaluated using AIC and BIC criteria, with wavelet families, levels and orders tested to determine the optimal wavelets for noise reduction and model accuracy.

The results of these analyses, highlighting the efficacy of wavelet-based denoising techniques, are shown in Table 4.

Table 4: Average AIC and BIC criteria for a Two-dimensional VARX (3) model with two predictors

Sample size	Method	Optimal Wavelet Selection	Level (L)	Order (N)	AIC	BIC
150	Classical	Coif3	3	3	1783.2407	1861.1682
	Proposed	Coif4	*4	*4	*334.4387	*412.3662
	Classical	db3	3	3	1429.5726	1507.5002
	Proposed	**db35	**7	**35	**-3277.8955	**-3199.968
	Classical	sym3	3	3	1429.5726	1507.5002
	Proposed	sym17	3	17	-1774.3687	-1696.4412
	Classical	bior3.3	3	8	1511.3589	1589.2865
	Proposed	bior3.9	*3	*11	*-637.4929	*-559.5654
	Classical	rbio3.3	3	8	-108.6466	-30.719
	Proposed	rbio3.5	*5	*9	*-1421.0344	*-1343.1069
	Without wavelet		-	-	16527.0374	16604.9649
175	Classical	coif3	3	3	215.9443	297.9299
	Proposed	coif4	*4	*4	*-622.1887	*-540.2031
	Classical	db3	3	3	1305.6826	1387.6682
	Proposed	**db33	**7	**33	**-3835.4062	**-3753.4206
	Classical	sym3	3	3	1305.6826	1387.6682
	Proposed	Sym24	*4	*24	*-2545.4321	*-2463.4465
	Classical	bior3.3	3	8	1157.474	1239.4596
	Proposed	bior5.5	*3	*13	*-1638.4979	*-1556.5124
	Classical	rbio3.3	3	8	-1261.3547	-1179.3691
	Proposed	rbio3.3	*5	*8	*-2071.981	*-1989.9954
	Without wavelet		-	-	17693.2802	17775.2658
200	Classical	Coif3	3	3	-4175.9077	-4090.4127
	Proposed	coif5	*7	*5	*-4556.6753	*-4471.1803
	Classical	db3	3	3	-3669.0354	-3583.5404
	Proposed	**db31	**6	**31	**-5307.1394	**-5221.6445
	Classical	sym3	3	3	-3669.0354	-3583.5404
	Proposed	sym22	*6	*22	*-4725.7761	*-4640.2812
	Classical	Bior3.3	3	8	-3499.3787	-3413.8837
	Proposed	bior5.5	*6	*13	*-4826.8713	*-4741.3763
	Classical	rbio3.3	3	8	-2612.0086	-2526.5136
	Proposed	rbio3.5	*5	*9	*-4649.9835	*-4564.4886
	Without wavelet		-	-	-1989.4083	-1903.9133
300	Classical	Coif3	3	3	-6397.7891	-6301.6647
	Proposed	Coif5	*8	*5	*-6809.118	*-6712.9936
	Proposed	**db32	**7	**32	**-7818.039	**-7721.9146
	Classical	sym3	3	3	-5656.095	-5559.9706
	Proposed	sym23	*7	*23	*-7214.6237	*-7118.4993
	Classical	bior3.3	3	8	-5356.7852	-5260.6608
	Proposed	Bior5.5	*7	*13	*-7525.7819	*-7429.6575
	Classical	rbio3.3	3	8	-4098.5746	-4002.4502
	Proposed	rbio3.5	*6	*9	*-7447.9478	*-7351.8234
	Without wavelet		-	-	-3027.8542	-2931.7298

An asterisk (*) denotes the optimal value for all the wavelets corresponding to the minimum AIC and BIC values.

An asterisk (**) denotes the optimal wavelet value corresponding to the minimum AIC and BIC values.

4.1.2.1 Discussion

Table (4) presents that the results consistently shown, for each sample size (150, 175,200, and 300), the proposed method using universal thresholding with the soft rule achieved the lowest AIC and BIC values, as indicated by the asterisks (*). This demonstrates that the method, when applied with optimal levels (L) and orders (N), was the most efficient at denoising the data.

The results presented in Table 4 provide a clear indication of the optimal wavelet performance across various sample sizes for a two-dimensional VARX (3) model with two predictors. The wavelets were evaluated based on the lowest AIC and BIC values, asterisk (**) and highlighted in yellow, which serve as indicators of model efficiency in removing noise from the data.

This indicates that the optimal wavelet was db35 at a decomposition level (L) of 7 (where the maximum level is 7), depending on the sample size (150) and an order (N) of 35 for a sample size of 150. For the sample size of 175, the optimal wavelet was the db33 wavelet, which showed superior performance at a level (L) of 7 (where the maximum level is 7), depending on the sample size (175) and order (N) of 33. For the sample size of 200, the db31 wavelet was optimal at level (L) 6 (where the maximum level is (7), depending on the sample size (200)) and order (N) 31. Lastly, for the sample size of 300, the db32 wavelet was optimal at level (L) 7 (where the maximum level is (8), depending on the sample size (300)) and order (N) 32 delivered the best results. All the optimal wavelets mentioned earlier are as follows: db35, db33, db31, and db32 (Daubechies wavelet), the most efficient for denoising the data for each sample size (150, 175,200, and 300). For Table (4), in the context of a two-dimensional VARX (3) model with two Predictors, the results in Table (5) are as follows:

Table 5: Optimal Wavelet based on lowest AIC and BIC criteria for a Two-dimensional VARX (3) model with Two Predictors.

Sample size	Method	Wavelet	Level (L)	Order (N)	AIC	BIC
150	Proposed	**db35	**7	**35	**-3277.8955	**-3199.968**
175	Proposed	**db33	****7**	******33	**-3835.4062	**-3753.4206**
200	Proposed	**db31	**6	**31	**-5307.1394	**-5221.6445**
300	**Proposed**	**db32**	**7**	**32**	**-7818.039**	**-7721.9146**

An asterisk (**) denotes the optimal wavelet value corresponding to the minimum AIC and BIC values.

The results presented in Table (5) demonstrate in the context of a two-dimensional VARX (3) model with two Predictors that Daubechies wavelets, specifically db35, db33, db31, and db32, exhibited outstanding performance across various sample sizes (n=150, 175, 200, 300). These wavelets were highly efficient in noise reduction, with optimal analysis levels (L) and orders (N) adjusted according to the sample size. Specifically, the db35 wavelet was optimal for a sample size of 150, while db33 was ideal for 175. For a sample size of 200, db31 showed the best performance, and db32 proved effective for larger sample sizes (n=300).

Across all tables (2, 3, 4, and 5), the proposed method using the universal threshold with the soft rule consistently outperformed traditional methods and those that did not incorporate wavelet transformations across various sample sizes. This superiority is reflected in the significantly lower AIC and BIC values when the optimal wavelets were used, demonstrating the method’s efficiency in noise reduction and improving model accuracy. While the classical method showed some effectiveness, it did not achieve the same level of performance as the proposed approach.

In contrast, methods that did not apply wavelet transformations produced higher AIC and BIC values, indicating lower efficiency and a weaker model fit compared to wavelet-based techniques.

Without wavelet transformations, the model struggled to remove noise from the data, leading to inaccurate parameter estimates and weaker predictive outcomes.

Finally, for all tables (4 and 5), the results demonstrate that for each sample size (150,175, 200, and 300), the proposed method consistently achieved the lowest AIC and BIC values, denoted with an asterisk (**). Specifically, the 2-dimensional VARX (3) model with 2 Predictors demonstrates more efficient performance, indicating an improved model fit, with lower AIC and BIC values compared to the 2-dimensional VARX (2) model with 2 Predictors, as shown in tables 2 and 3.

4.2 Proposed VARX (3) Model for Simulation Study:

Based on the results previously reviewed, the proposed VARX (3) model was selected as the best model for the simulation study, outperforming the two-dimensional VARX (2) model in terms of efficiency and estimation accuracy. The Daubechies wavelets, specifically db35, db33, db31, and db32, demonstrated outstanding performance across various sample sizes (n=150,175,200, and300), achieving the lowest AIC and BIC values. Therefore, we will select the proposed method for the simulation study in the case of n = 175 as an example due to the abundance of cases.

As described in the theoretical framework, a two-dimensional VARX (3) model with two - external Predictors was applied for the simulation of financial and economic data, using the Proposed Method, demonstrates superior performance, particularly when applying the optimal db33 wavelet at the optimal level (L=7), order (N=33), and universal thresholding with the soft rule. Consequently, the VARX (3) model with a constant term was selected based on achieving the lowest AIC and BIC values. Table 6 presents the estimation results for this model With the estimated number of parameters N = 18 parameters, and the mathematical form of the 2-dimensional VARX ($s = 3, r = 1$) model with 2 external Predictors, including both the dependent variables (Expenditures (Exp) and Official Reserves (OR)) and the external predictors (Broad Money (M2) and Oil Revenue (ORv)), according to equations (2), is as follows:

$$\left(\mathbf{1} - \mathbf{\Phi_1 B} - \mathbf{\Phi_2 B^2} - \mathbf{\Phi_3 B^3}\right)\mathbf{Yt} = \mathbf{C} + (\mathbf{\theta_1 B})\begin{bmatrix} \mathbf{x}_{1t} \\ \mathbf{x}_{2t} \end{bmatrix} + \boldsymbol{\epsilon}_\mathbf{t}$$

$$\left(\mathbf{1} - \mathbf{\Phi_1 B} - \mathbf{\Phi_2 B^2} - \mathbf{\Phi_3 B^3}\right)\begin{bmatrix} \mathbf{Exp}_t \\ \mathbf{OR_t} \end{bmatrix} = \begin{bmatrix} \mathbf{C_1} \\ \mathbf{C_2} \end{bmatrix} + (\mathbf{\theta_1})\begin{bmatrix} \mathbf{M2}_{t-1} \\ \mathbf{ORv_{t-1}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_{\mathbf{1,t}} \\ \boldsymbol{\epsilon}_{\mathbf{2,t}} \end{bmatrix}$$

$$\begin{aligned}\Bigg(\mathbf{1} - \begin{bmatrix} 2.596 & 0.55151 \\ -0.0026436 & 2.7598 \end{bmatrix}\mathbf{B} - \begin{bmatrix} -2.5669 & -1.2347 \\ 0.010397 & -2.5825 \end{bmatrix}\mathbf{B^2} \\ - \begin{bmatrix} 0.9546 & 0.68159 \\ -0.0074847 & 0.8198 \end{bmatrix}\mathbf{B^3}\Bigg)\begin{bmatrix} \mathbf{Exp}_t \\ \mathbf{OR_t} \end{bmatrix} \\ = \begin{bmatrix} 0.00028523 \\ 0.00009365 \end{bmatrix} + \begin{bmatrix} 0.0096401 & 0.0083217 \\ -0.00095031 & 0.0048387 \end{bmatrix}\begin{bmatrix} \mathbf{M2}_{t-1} \\ \mathbf{ORv_{t-1}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_{\mathbf{1,t}} \\ \boldsymbol{\epsilon}_{\mathbf{2,t}} \end{bmatrix}\end{aligned}$$

Table 6: Estimation Results for the Proposed AR-Stationary 2-Dimensional VARX (3) Model with 2 Predictors

Parameter	Value	Standard Error	t Statistic	P-Value
Constant (1)	0.00028523	0.00062823	0.45402	0.64981
Constant (2)	9.365e-05	9.6176e-05	0.97373	0.33019
AR {1} (1,1)	2.596	0.028908	89.805	0
AR {1} (2,1)	-0.0026436	0.0044255	-0.59735	0.55027
AR {1} (1,2)	0.55151	0.24758	2.2276	0.025906
AR {1} (2,2)	2.7598	0.037902	72.814	0
AR {2} (1,1)	-2.5669	0.051385	-49.954	0
AR {2} (2,1)	0.010397	0.0078667	1.3217	0.18628
AR {2} (1,2)	-1.2347	0.47501	-2.5993	0.0093417
AR {2} (2,2)	-2.5825	0.07272	-35.513	3.0482e-276
AR {3} (1,1)	0.9546	0.028397	33.616	9.8731e-248
AR {3} (2,1)	-0.0074847	0.0043474	-1.7216	0.085134
AR {3} (1,2)	0.68159	0.23246	2.9321	0.0033666
AR {3} (2,2)	0.8198	0.035587	23.036	2.0175e-117
Beta (1,1)	0.0096401	0.0035208	2.7381	0.0061802
Beta (2,1)	-0.00095031	0.000539	-1.7631	0.077885
Beta (1,2)	0.0083217	0.0074386	1.1187	0.26326
Beta (2,2)	0.0048387	0.0011388	4.249	2.1472e-05

Table (6) underscores the statistical significance of key estimated parameters, highlighting the model's robustness. Figure 1 illustrates the effectiveness of the VARX (3) model in capturing the data dynamics for the Exp and OR variables. The first figure compares the observed values (blue lines) with the estimated values (red lines) for both Exp and OR. The figure shows a strong alignment between the observed and estimated values for Exp, indicating the model's ability to effectively capture seasonal patterns and fluctuations. For the OR variable, the model also demonstrates reasonable alignment, although some minor deviations are noticeable, particularly in the middle of the series. These deviations suggest that the model could be improved to handle such sudden changes more accurately.

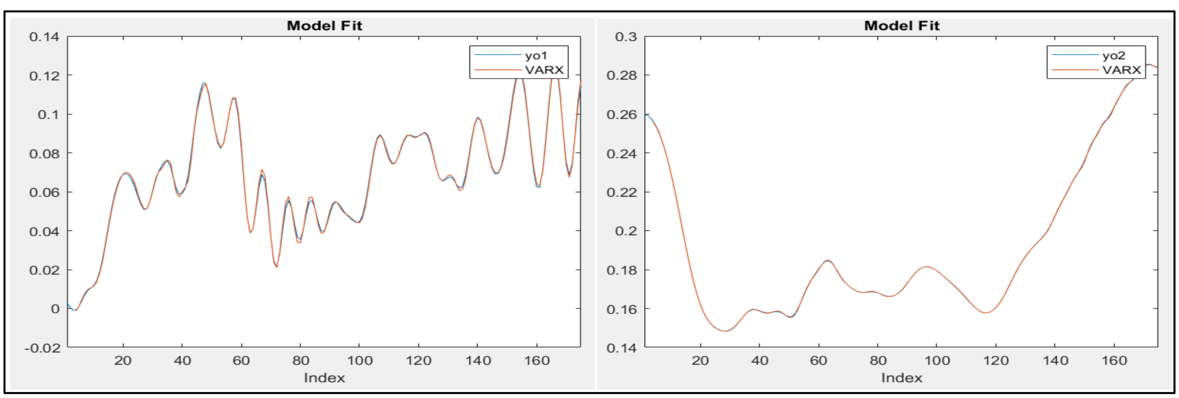

Figure 1: Fit of the Proposed VARX (3) Model for Dependent Variables Exp and OR

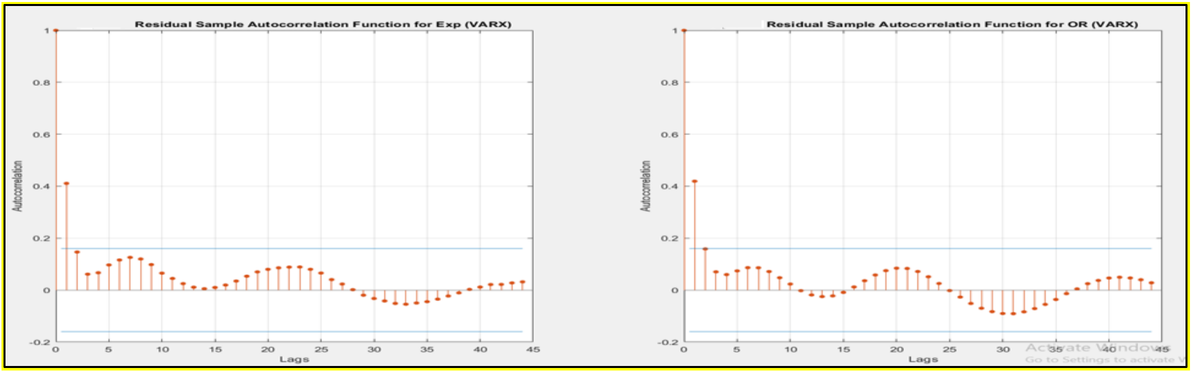

Figure 2: Residual Sample Autocorrelation Function the Proposed VARX (3) Model

Figure 2 illustrates the residual autocorrelation functions for the dependent variables Exp and OR. The residuals for Exp mostly fall within confidence intervals, indicating that the model effectively captured its core dynamics. In contrast, OR residuals show some deviations beyond the confidence bounds, suggesting potential unexplained patterns. Based on the Box-Ljung Test results for both Exp and OR residuals, the test statistics (Q-Stat) for both variables are below the critical value $\chi^2(0.05{,}44) = 60.46$. For Exp, $Q - \text{Stat} = 44.83 < 60.46$, with a p-value of 0.21 ($p > 0.05$), indicating no significant autocorrelation at a 5% significance level. Similarly, for OR, $Q - \text{Stat} = 39.561 < 60.46$, with a p-value of 0.662 ($p > 0.05$), confirming the absence of significant autocorrelation. These findings align with the visual evidence in Figure 2, where most residual autocorrelation values for both variables fall within the confidence intervals, validating the residual's independence. This outcome underscores the model's robustness in capturing the core dynamics of Exp and OR while minimizing unexplained patterns.

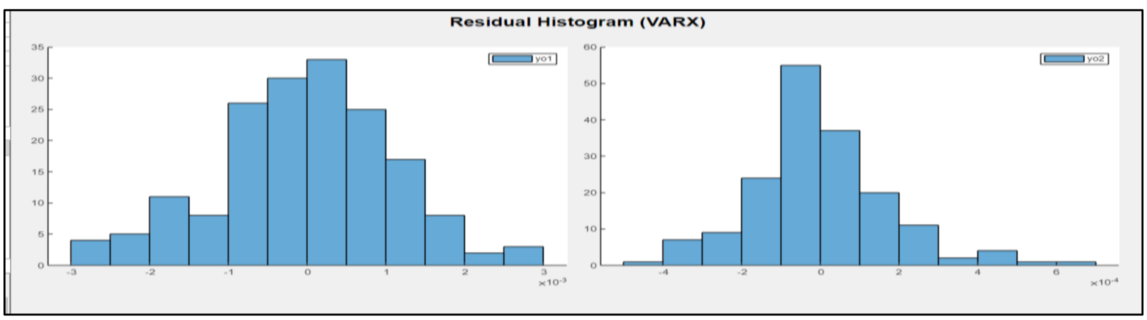

Figure 3: Histogram of the residuals FOR the Proposed VARX (3) Model for Dependent Variables Exp and OR

Figure 3 presents the residual histogram for the Exp and OR variables, revealing the error distribution generated by the VARX (3) model. For Exp, the residuals closely approximate a normal distribution, concentrated around zero, indicating a well-fitting model. The Jarque-Bera p-value for Exp is 0.5 (greater than 0.05), confirming that the residuals follow a normal distribution . For OR, slight deviations from zero suggest the influence of unmodeled fluctuations, though the overall distribution remains acceptable without significant outliers. The Jarque-Bera p-value for OR is 0.06 (greater than 0.05), further supporting the conclusion that the residuals are normally distributed. Combined with the residual sample autocorrelation function, these diagnostics confirm the model’s effectiveness in capturing data dynamics with small, unbiased residuals. The VARX (3) model is deemed suitable for forecasting Exp and OR, with the potential for minor enhancements to improve accuracy in cases of sudden variations.

5. Conclusion:

The research led to the following conclusions from the Simulation Study Analysis:

● The proposed method outperforms using universal thresholding with the soft rule consistently the classical and without variety wavelet transformation (coif3, db3, sym3, bior3.3, and rbio3.3) methods, which achieve the lowest AIC and BIC values for all sample sizes (150, 200, and 300) in VAR (2) and VAR (3) models.

● Although the classical method is effective, it underperforms the proposed approach. The higher AIC and BIC values in the models without wavelet transformation (coif3, db3, sym3, bior3.3, and rbio3.3) variety highlight the importance of wavelet-based noise removal to enhance model accuracy in parameter estimation and prediction.

● For each sample size (150,175, 200, and 300), both AIC and BIC values decrease as the sample size of the data increases. These findings, presented in Tables 2 through 5, support this observed trend. Thus, the inverse relationship between the sample size and the AIC and BIC values is underscored across all Tables.

● The optimal wavelet is selected with the determination of optimal wave levels and arrangements that achieved the lowest AIC and BIC values, where the proposed method achieved the best noise reduction performance across all sample sizes (150, 200, and 300) in the VAR (2) and VAR (3) models, as shown in tables 2-5.

● Higher order VARX models (p=2, 3) consistently achieved lower AIC and BIC values across all sample sizes (150, 200, and 300), demonstrating greater efficiency to the data compared to the lower-order models.

● For all tables (4, and 5), the results demonstrate that for each sample size (150,175, 200, and 300), the proposed method uses universal thresholding with the soft rule to choose the optimal wavelets with determine the optimal level and optimal order that consistently achieved the lowest AIC and BIC values for the 2-dimensional VARX (3) model with 2 Predictors, demonstrates more efficient performance, indicating an improved model fit compared to the 2-dimensional VARX (2) model with 2 Predictors, as shown in Tables (2 and 3).

● The Daubechies wavelets, db27 and db33, consistently showed excellent performance across sample sizes of 175 and 200. For larger samples, such as 300, Reverse Biorthogonal wavelets like Rbio3.7 were highly effective, while Symlets (specifically sym16) were especially suitable for smaller samples, such as 150, in terms of reducing noise in a two-dimensional VARX (2) model with two predictors, as shown in Table 3.

● The Daubechies wavelets (db35, db33, db31, and db32) displayed high efficiency in noise reduction across different sample sizes. The db35 wavelet was optimal for a sample size of 150, db33 was best suited for 175, db31 performed ideally at 200, and db32 was most effective with the largest sample size of 300, in the two-dimensional VARX (3) model, as detailed in Table 5.

● Through diagnostic tests, the strength of the model has been validated. The analysis of the residuals, using the autocorrelation function (ACF) of the residuals and the residuals' histogram, showed minimal autocorrelation and a near-normal error distribution. These findings enhance the model's reliability and accuracy, supporting its suitability for future predictions.

● The noise removal wavelet-based is necessary to significantly improve the model performance, as the VAR (3) model shows the best fit and accurately captures the time series dynamics. Diagnostic tests validated the model's robustness, confirming an absence of autocorrelation and substantial heterogeneity, which enhances the reliability of the predictions.

Authors Declaration:
Conflicts of Interest: None
-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted for Republication and Attached to The Manuscript.
- Ethical Clearance: The Research Was Approved by The Local Ethical Committee at The University.

References:

Abdulqader, Q. M., & Ahmed, N. M. (2023). Enhancing the ARIMAX Model by Using the Bivariate Wavelet Denoising: Application on Road Traffic Accidents Data. *Iraoi Journal of Statistical Sciences*, *20*(2).

Acquah, H. D. (2010). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics, 2*(1), 1–6.

Al-Doski, A. S., & Al-Waili, S. F. (2008). The reality of the Iraqi economy concerning the parallel sector. *Prepared by the consulting office at Duhok University College with the support of the Center for International Private Enterprise (CIPE), Iraq - Duhok.*

Ali, S., Murshed, S. M., & Papyrakis, E. (2023). Oil, export diversification and economic growth in Sudan: evidence from a VAR model. *Mineral Economics*, *36*(1), 77-96. https://doi.org/10.1007/s13563-022-00310-w

Ali, T. H., Albarwari, N. H. S., & Ramadhan, D. L. (2023). Using the hybrid proposed method for quantile regression and multivariate wavelet in estimating the linear model parameters. *Iraqi Journal of Statistical Sciences, 20*(1), 9-24.

Ali, T. H., Mustafa, Q., & Raza, M. S. (2016). Reducing the orders of the mixed model (ARMA) before and after the wavelet de-noising with application. *Journal of Humanity Sciences, 20*, 433-442.

Ali, T. H., Raza, M. S., & Abdulqader, Q. M. (2024). VAR time series analysis using wavelet shrinkage with application. *Science Journal of University of Zakho*, *12*(3), 345-355. https://doi.org/10.25271/sjuoz.2024.12.3.1304

Ali, T. H., & Saleh, D. M. (2022). Proposed hybrid method for wavelet shrinkage with robust multiple linear regression model: With simulation study. *Qalaai Zanist Journal*, *7*(1), 920-937.

Arouxet, M. B., Pastor, V. E., & Vampa, V. (2021). Using the wavelet transform for time series analysis. In *Applications of Wavelet Multiresolution Analysis* (pp. 59-74). Springer International Publishing.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261-304. https://doi.org/10.1177/0049124104268644

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J. O., Jenouvrier, S., & Stenseth, N. C. (2008). Wavelet analysis of ecological time series. *Oecologia, 156*, 287-304.

Central Bank of Iraq. (2024). Economic and statistical data. Retrieved from https://cbi.iq

Cohen, A., Daubechies, I., & Feauveau, J. C. (1992). Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics, 45*(5), 485-560.

Dastgerdi, A. K., & Mercorelli, P. (2022). Investigating the effect of noise elimination on LSTM models for financial markets prediction using Kalman Filter and Wavelet Transform. *WSEAS Transactions on Business and Economics*, *19*, 432-441. https://doi.org/10.37394/23207.2022.19.39

Daubechies, I. (1992). *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics (SIAM). https://doi.org/10.1137/1.9781611970104

Donoho, D. L., & Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *The Annals of Statistics*, *26*(3), 879-921. https://doi.org/10.1214/aos/1024691081

Gonçalves, M. A., Gonçalves, A. S., Franca, T. C., Santana, M. S., da Cunha, E. F., & Ramalho, T. C. (2022). Improved protocol for the selection of structures from molecular dynamics of organic systems in solution: The value of investigating different wavelet families. *Journal of Chemical Theory and Computation, 18*(10), 5810-5818.

Guo, T., Zhang, T., Lim, E., Lopez-Benitez, M., Ma, F., & Yu, L. (2022). A review of wavelet analysis and its applications: Challenges and opportunities. *IEEE Access*, *10,* 58869-58903.

Haydier, E. A., Albarwari, N. H. S., & Ali, T. H. (2023). The Comparison Between VAR and ARIMAX Time Series Models in Forecasting. *Iraqi Journal of Statistical Sciences*, *20*(2), 249-262.

Jalal, S. A., Saleh, D. M., Sedeeq, B. S., & Ali, T. H. (2024). Construction of the Daubechies wavelet chart for quality control of the single value. *Iraqi Journal of Statistical Sciences*, *21*(1), 160-169. https://doi.org/10.33899/iqjoss.2024.183257

Karamikabir, H., Afshari, M., & Lak, F. (2021). Wavelet threshold based on Stein's unbiased risk estimators of restricted location parameter in multivariate normal. *Journal of Applied Statistics, 48*(10), 1712-1729.

Kirillov, O. E. (2024). Application of the Wavelet Transform in the Analysis of Non-stationary Processes in Aerodynamic Experiments. *Lobachevskii Journal of Mathematics, 45*(5), 2058-2066. https://doi.org/10.1134/S1995080224602248

Li X, Liao K, He G, Zhao J. Research on improved wavelet threshold denoising method for non-contact force and magnetic signals. *Electronics.* 2023;12(5):1244. https://doi.org/10.3390/electronics12051244.

Liu, B., & Cheng, H. (2024). De-noising classification method for financial time series based on ICEEMDAN and wavelet threshold, and its application. *EURASIP Journal on Advances in Signal Processing*, *2024*(1), 19.

Liu, B., & Cheng, H. (2024). De-noising classification method for financial time series based on ICEEMDAN and wavelet threshold, and its application. *EURASIP Journal on Advances in Signal Processing*, *2024*(1), 19.

Lütkepohl, H. (2004). *Applied time series econometrics*. Cambridge University Press.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media. p. 1-558

Mallat, S. (2009). The sparse way. In *A wavelet tour of signal processing* (3rd ed., pp. 1-795). Academic Press.

Mallat, S. (1999). Introduction to wavelets. In *A wavelet tour of signal processing* (pp. 1-532). Academic Press.

Omer, A. W., Sedeeq, B. S., & Ali, T. H. (2024). A proposed hybrid method for multivariate linear regression model and multivariate wavelets (Simulation study). *Polytechnic Journal of Humanities and Social Sciences, 5*(1), 112-124.

Parra, L. C., Ortubay, A. S., Nentwich, M., Madsen, J., & Babadi, B. (2024). VARX Granger analysis: Modeling, inference, and applications. *arXiv preprint arXiv:2404.10834.*

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis.* Cambridge University Press. https://doi.org/10.1017/CBO9780511841040

Pesaran, M. H. (2015). Introduction to econometric models. In *Time series and panel data econometrics* (pp. 1-1024). Oxford University Press.

Rhif, M., Ben Abbes, A., Farah, I. R., Martínez, B., & Sang, Y. (2019). Wavelet transform application for/in non-stationary time-series analysis: A review. *Applied Sciences*, *9*(7), 1345. http://dx.doi.org/10.3390/app9071345

Schulte, J. A. (2016). Wavelet analysis for non-stationary, nonlinear time series. *Nonlinear Processes in Geophysics*, *23*(4), 257-267. https://doi.org/10.5194/npg-23-257-2016

Strang, G., & Nguyen, T. (1996). Introduction to wavelets and filter banks. In *Wavelets and Filter Banks* (pp. 1-515). Wellesley-Cambridge Press.

Tilak, T. N., & Krishnakumar, S. (2018). Reverse biorthogonal spline wavelets in undecimated transform for image denoising. *Int. J. Comput. Sci. Eng*, *6*, 66-72.

Tsay, R. S. (2013). *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons.