Mathematical Problems of Computer Science 61, 27–49, 2024.

doi: 10.51408/1963-0113

UDC 519.237, 519.25

Comparative Analysis of Univariate SARIMA and Multivariate VAR Models for Time Series Forecasting: A Case Study of Climate Variables in Ninahvah City, Iraq

Sameera A. Othman, Hasan H. Jameel and Sadeq T. Abdulazeez

University of Duhok, College of Basic Education, Duhok, Iraq

e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Abstract

This study involves a comparison between the application of the univariate SARIMA model and the utilization of VAR methods (vector autoregressive models) for multivariate time series analysis. The analysis is conducted using three-time series variables derived from data representing the monthly average of Humidity (H), Rainfall (R), and Temperature (T) in Ninahvah City, Iraq. Both univariate and multivariate time series approaches are employed to model these series. The paper also outlines the implementation of vector autoregressive, structural vector autoregressive, and structural vector error correction models using the 'vars' package. Additionally, it provides functions for diagnostic testing, estimation of constrained models, prediction, causality analysis, impulse response analysis, and forecast error variance decomposition. Furthermore, it introduces three fundamental functions, VAR, SVAR, and SVEC, for estimating these models. The comparison between the methods is based on evaluating the mean error produced by each approach. The findings of the study indicate that univariate linear stationary methods outperform multivariate models. The analysis of the data was carried out using the R software platform. The primary objective is to assess the performance of univariate and multivariate time series models in handling the given data. The research gap lies in the need for a comparative evaluation of SARIMA and VAR methods for time series analysis in the context of monthly environmental variables. These models were chosen due to their effectiveness in capturing temporal dependencies and interactions among multiple variables in time series data, providing a comprehensive analysis of climatic patterns in Ninahvah City, Iraq. The study aims to address the research gap by comparing these models and justifying their selection based on their capabilities to analyze the specified time series data.

Keywords: Modeling, Doers, Energizers, Classifiers, Cognizing, Fundamentals, Dynamicity.

Article info: Received 22 October 2023; sent for review 10 November 2023; received in revised form 1 January 2024; accepted 15 January 2024.

1. Introduction

A multivariate time series (MTS) comprises numerous time-related variables, and it is crucial to understand that each variable's dependence is not solely influenced by its previous values but also by interactions with other variables. Future values are forecasted using this dependency. This dependency is used for forecasting future values. The goals of multivariate time series analysis are to investigate the complex establishing links among variables and enhancing forecast precision. In the early 1980s, the authors in [1] critiqued vector autoregressive models (VARs) led to vector autoregressive models becoming a standard instrument in econometrics. This strategy was immediately improved by the incorporation of non-statistical prior information since statistical tests are commonly utilized to identify connections and intricate associations among variables. In contrast to deterministic repressors, VAR models describe endogenous variables entirely through their own histories. Structured vector autoregressive models (SVARs) facilitate the explicit modeling of contemporaneous interdependencies between the variables on the left. Consequently, these models attempt to address the deficiencies associated with VAR models. Sims posed a challenge to the established multiple structural equation model paradigm initially developed by the Cowles Foundation during the 1940s and 1950s. However, Granger in [2] and later Engle and Granger in [3] introduced a powerful tool to the field of econometrics for simulating and evaluating economic relationships: the concept of co-integration.

In recent times, the study of these fields has witnessed a convergence through the application of vector error correction models (VECM) and structural vector error correction models (SVEC). A comprehensive theoretical exposition of each of these models can be found in the monographs authored by Lutkepohl [4], Hendry [5], Johansen [6], Hamilton [7], and Banerjee et al. [8]. The main aim of this study is to compare the effectiveness of the univariate SARIMA model with the utilization of VAR methods for analyzing multivariate time series data. The motivation behind this research is to understand which approach is more suitable for modeling three specific time series variables related to Humidity, Rainfall, and Temperature in Ninahvah City, Iraq. The study explores various modeling techniques, including vector autoregressive, structural vector autoregressive, and structural vector error correction models, using the 'vars' package in R. It also offers a range of functions for diagnostic testing, model estimation, prediction, causality analysis, impulse response analysis, and forecast error variance decomposition.

2. Methodology

2.1. Stationary

A time series is classified as stationary when its statistical characteristics remain consistent throughout its duration. These characteristics, such as the mean and variance, remain unchanged over time [9]. Conversely, these properties fluctuate significantly, the time series is considered non-stationary. In practical terms, one can assess the stationary of a time series by visualizing it through a plot. A time series is termed "purely stationary" when the joint distribution of $\mathrm{Z}(\mathrm{t}_1), \cdots$, $\mathrm{Z}(\mathrm{t}_\mathrm{n})$, and $\mathrm{Z}(\mathrm{t}_1+\tau), \cdots, \mathrm{Z}(\mathrm{t}_\mathrm{n}+\tau)$, where $z(t)$ represents the random variable at time t, remains constant is the same for all $\mathrm{t}_1, \cdots, \mathrm{t}_\mathrm{n}, \tau$. To put it another way, the joint distributions are mostly unaffected by changing the time origin by a specific sum; instead, they must be determined by the intervals between $t_1, \cdots, t_n$ [10]. The time series Z_t is deemed to exhibit weak stationarity when two conditions are met: (a) The expected value of $Z_t = \mu$, which is a constant vector with k −dimensions, and (b) The covariance of $Z_t = E[(Z_t - \mu)(Z_t - \mu)'] = \Sigma_z$, a constant $k \times k$ positive-definite matrix. The random vector Z's expectation and covariance matrices are indicated by the letters E(Z) and Cov(Z), respectively. To establish if the time series is stationary, the collection of autocorrelations for the time series can also be used. The univariate time series stationary is examined using the Unit Root Test, and a multivariate time series is examined using the Co-integration test [11]. Consider the following two situations:

- When each univariate time series within an MTS item exhibits stationarity, the MTS item itself is considered to be stationary.
- If any of the individual time series within a multivariate time series (MTS) exhibit non-stationarity, a cointegration test should be conducted to verify that the MTS as a whole is also non-stationary. We may make the Z_t series stationary by differencing if it has not already been done. $Z_t = Z_t - Z_{t-1} = \nabla Z_t$ denotes the differenced series. Below are the definitions of the Backward Shift Operator B:

$B^m Z_t = Z_{t-m}$. The backward difference operator ∇ is defined by $\nabla = 1 - B$. Another method for determining whether the data is stationary or not, at lag k, the autocorrelation function is defined as:

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{E[(Z_t - \mu)^2 (Z_{t+k} - \mu)^2]},$$

where z_t : stands for observation. μ: Mean of observation, ρ_k: autocorrelation function.

The cross-correlation for lag k given two time series variables X_t and Y_t is given as $r_{xy} = c_{xy}/s_x\, s_y$ where, $c_{xy} = \frac{1}{n}\sum_{t=1}^{t-k}(x_t - \bar{x})(y_{t+k} - \bar{y}),\ k = 0, 1, 2\ \ldots$; $\bar{x}$ and $\bar{y}$ are the sample means of x_t and y_t, s_x and s_y are the sample standard deviations, respectively [12], Process of White Noise. A white noise process with the formula $a_t = (a_{1t}, \cdots, a_{kt})'$ is a continuous random vector that satisfies the conditions $E(a_t) = 0$, $E(a_t a_t') = \Sigma_a$, and $E(a_t a_s') = 0$ for $s \neq t$. Unless otherwise specified, the Σ_a = covariance matrix is assumed to be non-singular as pointed in [11].

2.2. Vector Autoregressive (VAR) Model

One approach to representing the interplay among multiple time-varying variables is through the utilization of the vector autoregressive (VAR) model. This model provides a streamlined representation of dynamic interactions, wherein each internal variable is influenced by its own past values as well as the past values of all other internal variables. The simple p-lag Vector autoregressive VAR (p) method looks like this:

$$Z_t = c + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t\,; t = 0, \pm 1, \pm 2, \cdots, \quad (1)$$

where $\mathrm{Z_t} = (\mathrm{Z_{1t}}, \cdots \mathrm{Z_{kt}})'$ is a $(\mathrm{k} \times 1)$ vector of time series variable, $\emptyset_\mathrm{i}$ are fixed $(\mathrm{k} \times \mathrm{k})$ coefficient matrices, $\mathrm{c} = (\mathrm{c_1}, \cdots, \mathrm{c_k})'$ is a fixed $(\mathrm{k} \times 1)$ vector of intercept terms, $\mathrm{a_t} = (\mathrm{a_{1t}}, \cdots \mathrm{a_{kt}})'$ is a white noise procedure with $(\mathrm{k} \times 1)$. The procedure can be written clearly in matrix form:

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ \vdots \\ Z_{kt} \end{pmatrix} = \begin{pmatrix} \emptyset_{11}^1 & \emptyset_{12}^1 & \cdot & \cdot & \emptyset_{1k}^1 \\ \emptyset_{21}^1 & \emptyset_{22}^1 & \vdots & \vdots & \emptyset_{2k}^1 \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^1 & \emptyset_{k2}^1 & & & \emptyset_{kk}^1 \end{pmatrix} \begin{pmatrix} Z_{1t-1} \\ Z_{2t-1} \\ \vdots \\ Z_{kt-1} \end{pmatrix} + \begin{pmatrix} \emptyset_{11}^2 & \emptyset_{12}^2 & \cdot & \cdot & \emptyset_{1k}^2 \\ \emptyset_{21}^2 & \emptyset_{22}^2 & \vdots & \vdots & \emptyset_{2k}^2 \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^2 & \emptyset_{k2}^2 & & & \emptyset_{kk}^2 \end{pmatrix} \begin{pmatrix} Z_{1t-2} \\ Z_{2t-2} \\ \vdots \\ Z_{kt-2} \end{pmatrix} + \cdots +$$

$$\begin{pmatrix} \emptyset_{11}^p & \emptyset_{12}^p & \cdot & \cdot & \emptyset_{1k}^p \\ \emptyset_{21}^p & \emptyset_{22}^p & \vdots & \vdots & \emptyset_{2k}^p \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ \emptyset_{k1}^p & \emptyset_{k2}^p & & & \emptyset_{kk}^p \end{pmatrix} \begin{pmatrix} Z_{1t-p} \\ Z_{2t-p} \\ \vdots \\ Z_{kt-p} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{kt} \end{pmatrix}. \quad (2)$$

2.3. Stable VAR ($\boldsymbol{p}$) Processes [13]

If every root of the matrix lies within the unit circle, and the absolute values of the roots of matrix $\emptyset_\mathrm{i}$ are less than 1, then process 1 exhibits stability. That is, if $\det(\mathrm{I_n} - \emptyset_1 \mathrm{Z} - \cdots - \mathrm{Z^p}) \neq 0$ for $|\mathrm{Z}| \leq 1$, then a stationary VAR (p) process $\mathrm{Z_t}$; $\mathrm{t} = 0, \pm 1, \pm 2, \cdots$ is stable.

2.4. A Stable VAR(p) Process' Autocovariances

The result of deducting the mean from VAR (p) is

$$\mathrm{Z_t} - \mu = \emptyset_1(\mathrm{Z_{t-1}} - \mu) + \cdots + \emptyset_\mathrm{p}(\mathrm{Z_{t-p}} - \mu) + \mathrm{a_t}. \quad (3)$$

After dividing both sides by $(\mathrm{Z_{t-1}} - \mu)'$ and calculating the expectation, having at $l = 0$ by utilizing:

$$\Gamma_\mathrm{z}(\mathrm{i}) = \Gamma_\mathrm{z}(-\mathrm{i})',$$

$$\Gamma_\mathrm{z}(0) = \emptyset_1(\mathrm{Z_{t-l}} - \mu) + \cdots + \emptyset_\mathrm{p}(\mathrm{Z_{t-p}} - \mu) + \Sigma_\mathrm{a},$$

$$= \emptyset_1\Gamma_\mathrm{z}(1)' + \cdots + \emptyset_\mathrm{p}\Gamma_\mathrm{z}(\mathrm{p})' + \Sigma_\mathrm{a}. \quad (4)$$

If $\mu > 0$, then

$$\Gamma_\mathrm{z}(l) = \emptyset_1\Gamma_\mathrm{z}(l-1)' + \cdots + \emptyset_\mathrm{p}\Gamma_\mathrm{z}(l-\mathrm{p})' + \Sigma_\mathrm{a}, \quad (5)$$

If $\emptyset_1, \cdots, \emptyset_p$ and $\Gamma_z(p-1)$ are provided, the auto covariance functions $\Gamma_z(l)$ for $l \geq p$ can be derived from these equations.

2.5. A Stable VAR (p) Process's Autocorrelation

Obtaining the autocorrelations of a stable VAR (p) process is achieved by extracting information from the matrix:

$$R_z(l) = D^{-1}\Gamma_z(l)D^{-1}, \tag{6}$$

hence, D is a diagonal matrix with the Z_t component's standard deviation on the main diagonal. Consequently

$$D^{-1} = \begin{bmatrix} \frac{1}{\sqrt{\gamma_{11}(0)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\gamma_{kk}(0)}} \end{bmatrix}, \tag{7}$$

and $Z_{i,t}$ and $Z_{j,t-1}$ have the following correlation:

$$\rho_{ij}(l) = \frac{\gamma_{ij}(l)}{\sqrt{\gamma_{ii}(0)}\sqrt{\gamma_{jj}(0)}}, \tag{8}$$

which is just the $ij-$ th element of $R_z(l)$. The model's characteristic roots are, once again, the inverses of the solutions. As a result, stationarity necessitates that all characteristic roots have a modulus of less than one. The ACF satisfies the difference equation $\left(1-\emptyset_1 B-\emptyset_1 B^2-\cdots-\emptyset_p B^p\right)^p = 0,$ for $p \geq 0$ a stationary AR (p) sequence. The ACF plot of a stationary AR (p) model will display a blend of damped sinusoidal and exponential decay patterns, influenced by the unique source it originates from, leading to varying levels of similarity in the shapes observed.

2.6. Order Selection by VAR

The three selection criteria that will be utilized to evaluate the VAR process order p are as follows:

(i) Employing the Akaike Information Criterion (AIC) method [14], as introduced in [15],

$$AIC(p) = \ln\left|\widetilde{\Sigma}_\varepsilon(p)\right| + \frac{2}{N}\text{ (number of estimated parameter)}$$

$$= \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2pk^2}{N}.$$

(ii) Given Hannan and Quinn [16], the Hannan-Quinn Criterion (HQC), follows

$$\text{HQC}(p) = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2\ln k}{N} \text{ (Parameters that are freely estimated)}$$
$$= \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{2\ln(\ln N)}{N}\, pk^2.$$

(iii) Using the Bayesian Information Criterion (BIC) [17],

$$\text{BIC}(p) = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{\ln k}{N} \text{ (Parameters that are freely estimated} = \ln\left|\widetilde{\Sigma}_a(p)\right| + \frac{\ln N}{N}\, pk^2,$$

where the VAR order is p,

The estimated white noise covariance matrix Σ_ε is represented by Σ_a . In a vector time series, there are k different time series components. N is the sample size. Each estimate is selected to minimize the criterion's value in each of the aforementioned parameters.

2.7. Forecasting

If it is determined that the fitted model in 1 is sufficient, forecasts can be made being used. The following estimates are used to create forecasts:

$$\widehat{Z}_t = \hat{c} + \widehat{\emptyset}_1 Z_{t-1} + \widehat{\emptyset}_2 Z_{t-2} + \cdots + \widehat{\emptyset}_p Z_{t-p} + a_t \,;\, t = 0, \pm 1, \pm 2, \cdots. \quad (9)$$

Given the forecast origin t., the forecasts so produced are those with the smallest mean square error [4].

Using vector moving average models for forecasting (VMA). Considering that the model is recognized and serves as a source for prediction. The VMA forecast (q). Generally, for h-step forward forecast with h ≤ q, as occurs

$$Z_t(h) = \mu - \sum_{i=1}^{q} \theta_i a_{t+h-i}. \quad (10)$$

Utilizing VARMA models for prediction, we are employing the criterion of minimizing mean-squared error to delve into the future projections of a time series Z_t with a VARMA(p, q) structure, similar to the VAR models of (9). As stated below, for the VARMA (p, q) model

$$Z_t(h) = \emptyset_0 + \sum_{i=1}^{p} \emptyset_i Z_{t+h-i}. \quad (11)$$

3. Applications

Data pertaining to the monthly averages of temperature (T), rainfall (R), and humidity (H) for Ninavah, Iraq, ranging from 1976 to 2001, were examined using the R program. In this example, we'll refer to humidity as (H, Z_{1t}), precipitation as R (R, Z_{2t}), and temperature as (T, Z_{3t}). The multivariate time series can therefore be described using the random vector $Z_t = (Z_{1t}, Z_{2t}, Z_{3t}$. The time series data for these three variables are presented in Figure 1 in a variety of graphical formats. The core scientific challenge outlined in the text revolves around the thorough analysis and modeling of these multivariate time series data, specifically focusing on climate variables—humidity, precipitation, and temperature. The overarching goal is to uncover intricate relationships and discern patterns within the dataset. Additionally, the aim is to construct a robust multivariate model capable of accurate forecasting and in-depth analysis.

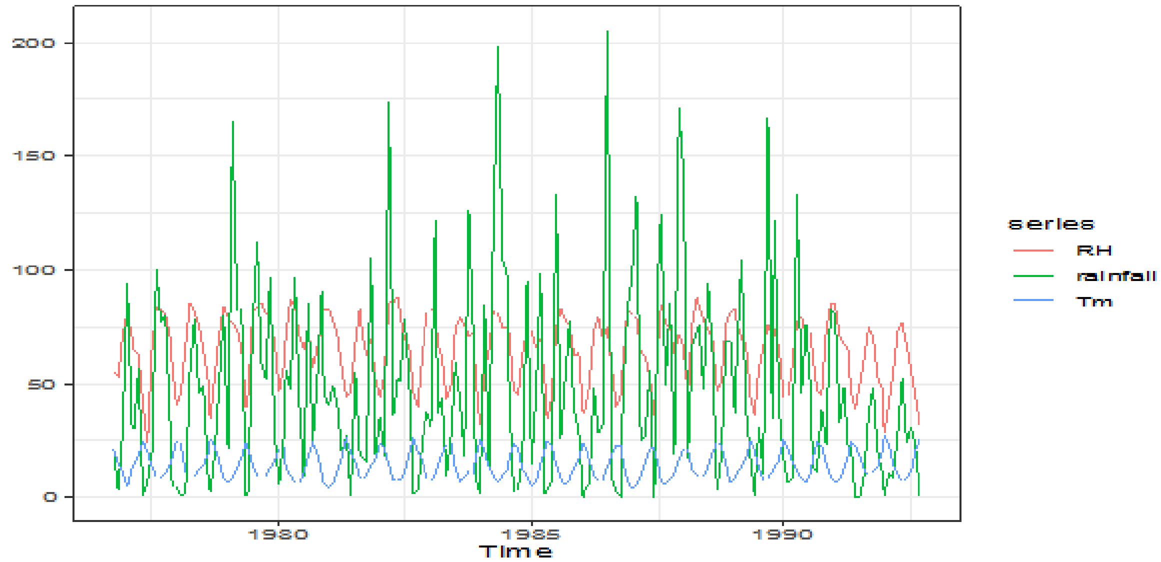

Fig. 1. The three raw series' time series plot (H, R,T).

The Unit Root test is used to determine whether univariate time series datasets are stationary. In contrast, the Co-integration test (original) is used to examine stationarity in multivariate time series datasets. The Augmented Dickey-Fuller (ADF) test can be used to determine whether a series has a unit root. This is predicated on the assumption that a trend-lined series will display a unit root and a significant p-value.

H_0: The data is non-stationary and has a unit root.

H_1: The data are static and have not yet produced the results in Table (1).

Table 1. Original and transformed data stationary testing of Nineveh time series data sets for the period 1976 – 2001

Stationary testing							
Datasets	Responses	Dickey-Fuller	p-value	Phillips-Perron Unit Root Test	p-value	KPSS Level	p-value
R	Z_t	-7.339	0.01	-70.58	0.01	0.25798	0.1
H	Z_t	-5.9331	0.01	-154.35	0.01	0.15987	0.1
T	Z_t	-8.5794	0.01	-67.175	0.01	0.03855	0.1

The alternate proposition becomes relevant when the p-value rejects the null hypothesis and exceeds the 0.05 threshold. In the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the p-value surpasses the 0.05 threshold, signifying the absence of a unit root and, consequently, stationarity within the series. To ascertain trend stationarity, researchers will assess the null hypothesis; a low p-value suggests the presence of a non-trend stationary signal with a unit root. To test the Stability, all characteristic roots should have a modulus of less than one. The results in Table (2) represent the original data.

Table 2. Roots of the stability characteristic polynomial

The characteristic polynomial's roots							
0.9018	0.9081	0.9081	0.9049	0.9049	0.8934	0.8825	0.8825
0.8803	0.8803	0.871	0.871	0.8214	0.8214	0.8148	0.8148
0.8036	0.8036	0.7866	0.7866	0.7509	0.7509	0.691	0.691
0.4724	0.2125	0.2125					
Log Likelihood -1795.702							

All of the roots k are inside the unit circle. We have no strenuous roots. Our system is generally stable.

3.1. Co-Integration Test

Co-integration testing is a method used to assess the accuracy of long-term linkages between variables because none of them now exhibit stationarity. If the variables exhibit co-integration, it implies that they have an ongoing link, even if they are not stationary at the moment [18]. They also offered the Maximum Eigen Value test and the Trace test as two more methods for counting co-integrated vectors. While the Trace test looks into the potential of r+1 co-integrating vectors, the Maximum Eigen Value test looks into the possibility of a maximum of r co-integrating vectors [19]. They claim that the Maximum Eigen Value test is the best technique for determining the number of co-integrating vectors. After d distinct differentiations, an integrated sequence of order d, designated as I(d), becomes stationary.

H_0: no co-integration of variables H_1: co- integration of variables

The results in Table 3 represent the data.

Table 3. Findings from Johansen's Co-integration Examination for H, R, T

Unrestricted Co-integration Rank Test (Trace)			
Co-integration rank(r)	Eigenvalue	Trace stat.	Critical Value 5%
r = 0*	8.766264e-02	35.55	34.91
r <= 1*	5.444259e-02	18.76	19.96
r <= 2*	4.546822e-02	8.52	9.24
Unrestricted Co-integration Rank test (Maximum Eigenvalue)			
Co-integration rank(r)	Eigenvalue	Trace stat	Critical Value 5%
r = 0*	8.766264e-02	16.79	22.00
r <= 1*	5.444259e-02	10.24	15.67
r <= 2*	4.546822e-02	8.52	9.24

The Trace test reveals the presence of three co-integrating equations with a significance level of 0.05. The asterisk (*) signifies the rejection of the hypothesis at the same 0.05 significance level.

The column of r in Table (3) represents the rank and we know that this is some indication of the number of co-integrating relationships. When r = 0, the test statistic is 35.55 > 34.91. This implies that we do not accept the null hypothesis, which proposes that r > 0. As such, there is some co-integration present. When r < 1, we do not find enough evidence to reject the null hypothesis because 18.76 < 19.9. When r < 2, this again means that we do not find enough evidence to reject the null hypothesis because 8.52 < 9.24. Therefore, we conclude that there is at most 1 co-integrating relationship that presents the Johansson test when we choose the maximal eigenvalue statistic. We are unable to dismiss the null hypothesis. None of the statistical values falls below the 5 percent threshold. It means no co-integrating relationships present the Johansson test.

3.2. The Raw Data Correlation Matrix

The three variables are highly connected, as shown by the correlation matrix ($Z_{t(corr.)}$) below. As a result, the multivariate technique will take into account the interrelation between the variables

$$\text{Corr (H,R,T)} = \begin{matrix} H \\ R \\ T \end{matrix}\begin{pmatrix} 1 & 0.636 & -0.868 \\ 0.636 & 1 & -0.461 \\ -0.868 & -0.461 & 1 \end{pmatrix},$$

$$\text{Cov (H,R,T)} = \begin{matrix} H \\ R \\ T \end{matrix}\begin{pmatrix} 229.1 & 403.2 & -79.9 \\ 403.2 & 1764 & -117.4 \\ -79.9 & -117.4 & 36.9 \end{pmatrix}.$$

3.3. The Cross Correlations

Table 4 shows the cross-correlation matrices at various lags (lags 1–12).

High values demonstrate that the variables are interdependent and a multivariate model can be successfully fitted to the data. A simple matrix $s_\ell = [s_{\ell,ij}]$ is constructed for each sample CCM $\hat{\rho}_\ell$ as follows:

$$s_{\ell,ij} = \begin{cases} + & if\ \hat{\rho}_{\ell.ij} \geq 2/\sqrt{T}, \\ - & if\ \hat{\rho}_{\ell.ij} \leq -2/\sqrt{T}, \\ . & if\ |\hat{\rho}_{\ell.ij}| < 2/\sqrt{T}, \end{cases}$$

where: $\hat{\rho}_\ell$ is a consistent estimate of ρ_ℓ, T is a total number.

The results in Table (4) represent the original data.

Table 4. displays example Cross-Correlation Matrices depicting the Monthly Simple Returns of three different Indexes in their raw form (H, R, T).

Lag 1	lag 2	lag3	lag 4	lag 5	lag 6
$\begin{bmatrix} + & + & - \\ + & . & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & \cdot & + \\ - & \cdot & + \\ + & + & \cdot \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$
Lag7	**lag 8**	**lag 9**	**lag 10**	**lag11**	**lag 12**
$\begin{bmatrix} + & + & - \\ + & + & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} + & + & - \\ + & + & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} + & + & - \\ + & \cdot & - \\ - & - & + \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & - & + \\ + & + & - \end{bmatrix}$	$\begin{bmatrix} - & - & + \\ - & \cdot & + \\ + & + & - \end{bmatrix}$

Table (4) illustrates the simplified CCM for monthly data of (H, R, T). Notable cross-correlations, which are statistically significant at the estimated 5% level, are mainly observable at the lags of 8 and 9.

3.4. Selecting a Model

AIC, BIC, and HQC at various lags are shown in both Table (5), which represents the original data, and Figure (2), which depicts the data. At lag 9, a three- selection process reaches the minimal values (the bolded values). VAR (9) is, therefore, the model of choice in Table (5): Empirical Lag Selection**.**

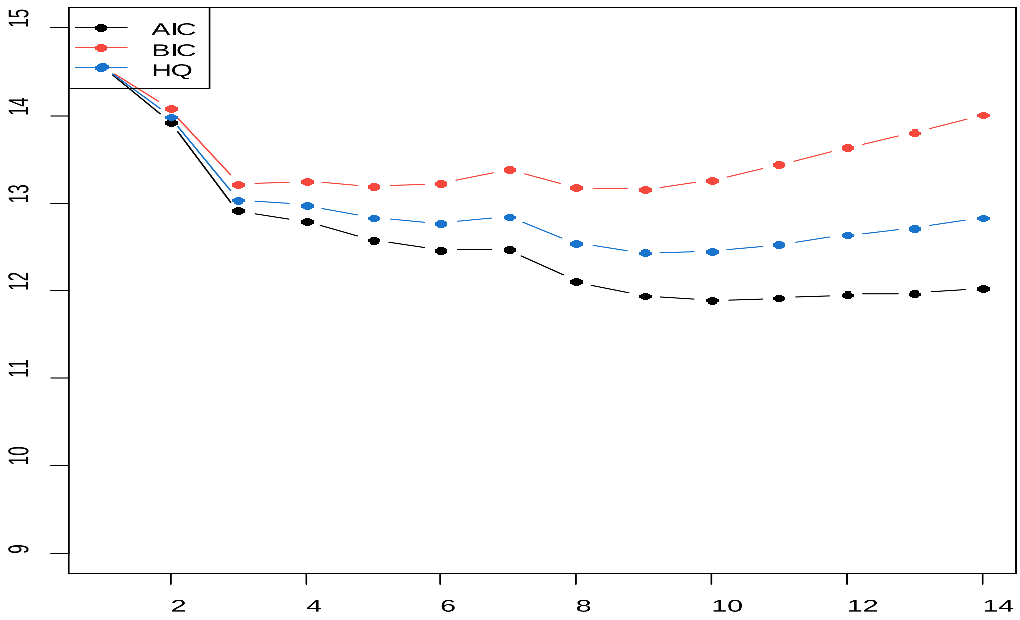

Fig. 2. Illustrates the information criteria applied to (H, R, T) data. The lines, depicted as solid, dashed, and dotted, correspond to AIC, BIC, and HQ, respectively.

Table 5. Empirical Lag Selection

	AIC(n)	HQ(n)	SC(n)	FPE(n)
selection	9	8	8	9
1	14.28293	14.34716	14.44137	15.95881
2	13.20192	13.33038	13.51880	541441.48394
3	13.04705	13.23973	13.52237	463836.94951
4	12.84215	13.09906	13.47591	378027.24396
5	12.73620	13.05735	13.52840	340211.74568
6	12.63510	13.02047	13.58574	307749.95214
7	12.24344	12.69305	13.35252	208259.68283
8	12.04285	12.55669	13.31037	170670.10872
9	11.97888	12.55694	13.40484	160410.83869
10	12.00919	12.65148	13.59359	165760.36272

3.5. Model Presentation

The VAR (9) model with significant parameters is represented in matrix form as seen in Table (6), which represents the original data, utilizing equation (2) in the approach. The optimal lag value is p = 9 according to AIC and FPE, p = 8 based on the HQ criterion, and p = 7 according to the SC criterion. They performed diagnostic analyses on the residuals after calculating a VAR with both a constant and a trend as deterministic predictors for each of the nine different lag orders.

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{pmatrix} = \begin{pmatrix} 0.341 & 0.019 & -0.187 \\ 0.567 & -0.056 & 0.393 \\ -0.011 & -0.009 & 0.466 \end{pmatrix} \begin{pmatrix} Z_{1t-1} \\ Z_{2t-1} \\ Z_{3t-1} \end{pmatrix} + \begin{pmatrix} -0.126 & 0.023 & 0.672 \\ -0.887 & 0.106 & -0.964 \\ 0.027 & -0.004 & -0.149 \end{pmatrix} \begin{pmatrix} Z_{1t-2} \\ Z_{2t-2} \\ Z_{3t-2} \end{pmatrix} +$$

$$\begin{pmatrix} 0.084 & 0.036 & 0.236 \\ -0.267 & 0.123 & -0.724 \\ -0.039 & 0.002 & -0.008 \end{pmatrix} \begin{pmatrix} Z_{1t-3} \\ Z_{2t-3} \\ Z_{3t-3} \end{pmatrix} + \begin{pmatrix} 0.007 & 0.0139 & 0.315 \\ -0.059 & 0.089 & 1.083 \\ 0.016 & 0.005 & -0.082 \end{pmatrix} \begin{pmatrix} Z_{1t-4} \\ Z_{2t-4} \\ Z_{3t-4} \end{pmatrix} +$$

$$\begin{pmatrix} 0.101 & 0.035 & 0.121 \\ -0.483 & 0.147 & -1.02 \\ -0.008 & -0.004 & -0.131 \end{pmatrix} \begin{pmatrix} Z_{1t-5} \\ Z_{2t-5} \\ Z_{3t-5} \end{pmatrix} + \begin{pmatrix} -0.084 & 0.0001 & 0.632 \\ -0.523 & 0.153 & 2.891 \\ 0.068 & -0.003 & 0.148 \end{pmatrix} \begin{pmatrix} Z_{1t-6} \\ Z_{2t-6} \\ Z_{3t-6} \end{pmatrix} +$$

$$\begin{pmatrix} 0.202 & -0.001 & -0.232 \\ 0.381 & 0.208 & -2.207 \\ -0.034 & -0.000 & 0.177 \end{pmatrix} \begin{pmatrix} Z_{1t-7} \\ Z_{2t-7} \\ Z_{3t-7} \end{pmatrix} + \begin{pmatrix} 0.277 & -0.039 & -0.497 \\ 1.002 & -0.118 & 2.337 \\ -0.069 & 0.011 & 0.291 \end{pmatrix} \begin{pmatrix} Z_{1t-8} \\ Z_{2t-8} \\ Z_{3t-8} \end{pmatrix} +$$

$$\begin{pmatrix} -0.083 & 0.006 & -0.109 \\ 0.196 & -0.075 & -0.087 \\ 0.112 & -0.010 & 0.053 \end{pmatrix} \begin{pmatrix} Z_{1t-9} \\ Z_{2t-9} \\ Z_{3t-9} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \\ a_{3t} \end{pmatrix}. \tag{12}$$

The information is presented in Table (6) along with the summarized results and the graphical representation of equation (12).

Table 6. Results for the Endogenous variables: H, R, T

Statistic	H	R	T
Multiple R-Squared	0.991	0.6873	0.9906
F-statistic	635	12.7	608.7
Adjusted R-squared	0.9894	0.6332	0.989
Residual standard error	6.893	38.87	1.643
p-value	< 2.2e-16	< 2.2e-16	< 2.2e-16

$$\text{Cov of residuals} = \begin{matrix} H \\ R \\ T \end{matrix} \begin{pmatrix} 47.514 & 169.310 & -2.705 \\ 169.310 & 1510.815 & -3.646 \\ -2.705 & -3.646 & 2.699 \end{pmatrix},$$

$$\text{Corr of residuals} = \begin{matrix} H \\ R \\ T \end{matrix} \begin{pmatrix} 1 & 0.6319 & -0.2388 \\ 0.6319 & 1 & -0.0571 \\ -0.2388 & -0.0571 & 1 \end{pmatrix}.$$

4. Diagnostic Testing

Once the multivariate model 12 has been acquired, the next step is to verify the correctness of the model fit. The following diagnostic techniques are used to this end.

4.1. Residual Autocorrelation Function

The following hypothesis is used, as described in Section (1.1.3) of the methodology**:**

$$H_0: \rho_{uv}\, i = 0 \text{ versus } H_1: \rho_{uv}\, i \neq 0$$

We had a total of n= 192 series.

As a result, the residual autocorrelation function's boundary state has the form $\frac{2}{\sqrt{192}} = 0.144$, and H_0 is rejected if $|r_{uv}, i| > \frac{2}{\sqrt{N}}$ =0.144.

When examining the values of autocorrelations in the residual correlation matrices at various lags (lags 12), it was found that none of the residual autocorrelations exceeds 0.144. Figure (1) represents the original data with$|r_{uv}, i|$. This suggests that the residuals conform to a pattern consistent with white noise. To put it another way, the fitted model is sufficient.

1) Test auto correlation for serial correlation (PT) [20]

The graphs, one for each equation, demonstrate the ACF and PACF of the discrepancies, along with a discrepancy plot and a practical distribution chart. Additional justifications are provided by the plot approach for changing its design. Figures (3-5) represent the original data.

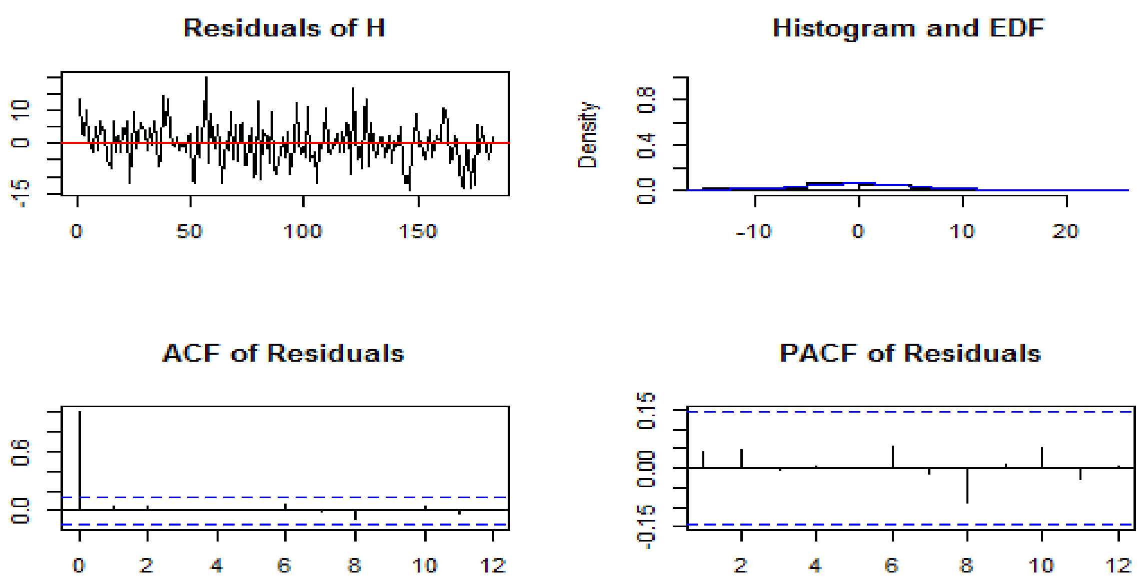

Fig. 3. Explains the Time Series Plots of Residuals (a_{1t}) for H.

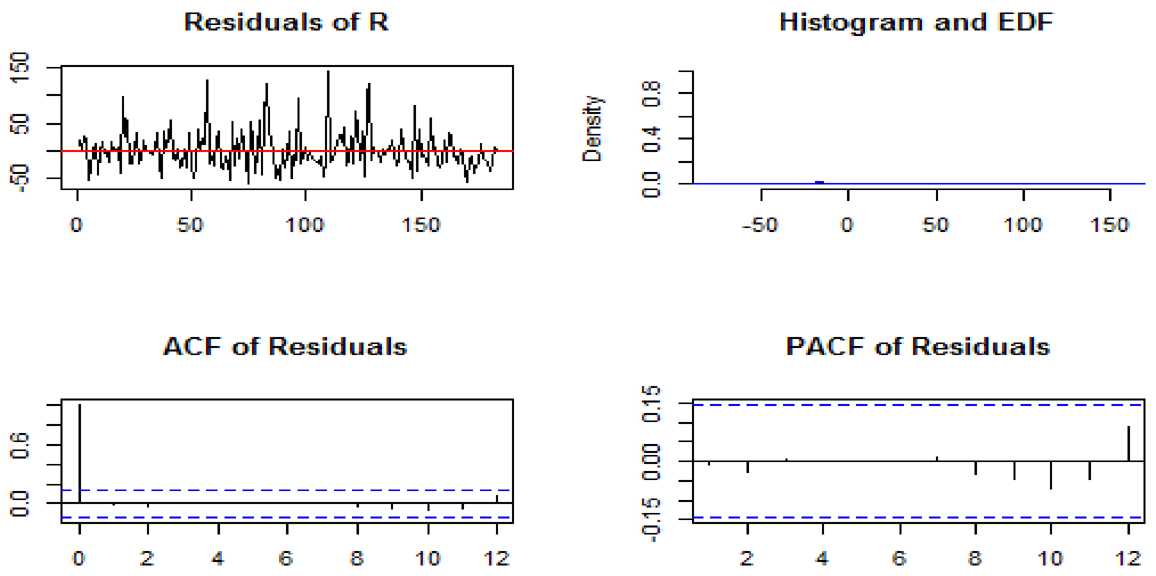

Fig. 4. Explains the Time Series Plots of Residuals (a_{2t}) for R.

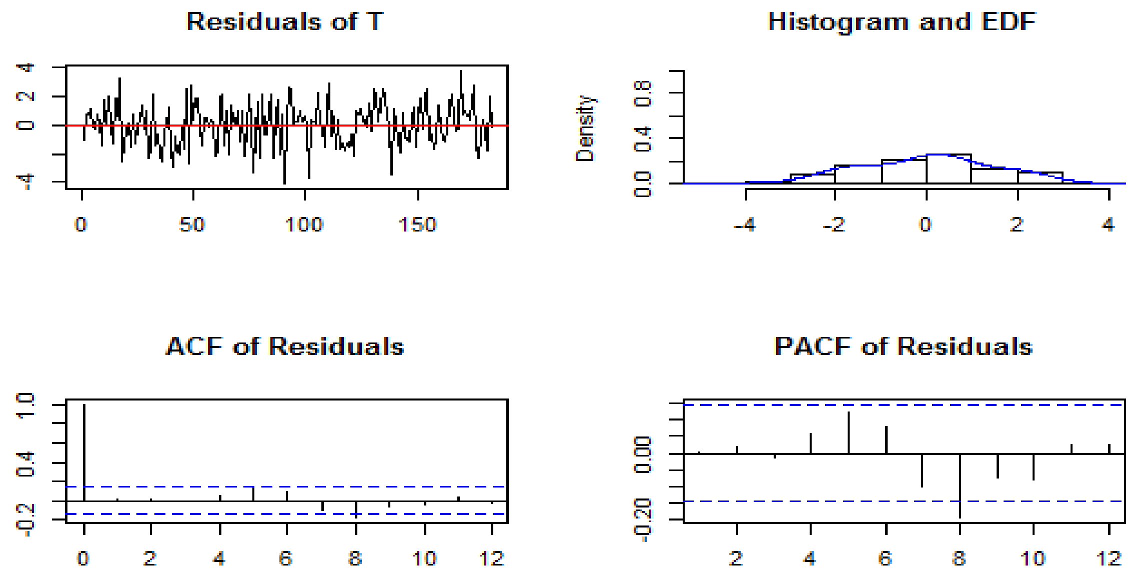

Fig. 5. Explains the Time Series Plots of Residuals (a_{3t}) for T.

Explain the Time Series Plots of Residuals (a_{3t}) for T. Heteroscedasticity: ARCH test ([21], [22]

A statistical model called autoregressive conditional heteroscedasticity (ARCH) is used to assess and forecast volatility in time series. The following regression is the foundation for the multivariate ARCH-LM test. (The univariate test is considered a specific case of the exhibit below and will be omitted):

$$vech(\hat{u}_t\,\hat{u}_t^T) = \beta_0 + B_1 vech(\hat{u}_{t-1}\hat{u}_{t-1}^T) + \ldots + B_q vech(\hat{u}_{t-q}\hat{u}_{t-q}^T) + v_t,$$

$E(u_t) = 0$ and positive time invariant unambiguous covariance matrix $E(u_t\, u_t^T) = \Sigma_u$ (white noise) define u_t as a K-dimensional process [23]. In this context, v_t represents a spherical error process, and the operator 'vech' is used to stack the columns of symmetric matrices, starting from the main diagonal and moving downward. The dimension of β_0 is $\frac{1}{2}K(K+1)$, and for the coefficient matrices B_i where $i = 1, \dots, q$, $\frac{1}{2}K(K+1) \times \frac{1}{2}K(K+1)$. The null hypothesis is: $H_0 := B_1 = B_2 = \dots = B_q = 0$ and the alternative is: $H_1: B_1 \neq 0 \cap B_2 \neq 0 \cap \dots \cap B_q \neq 0$. The test statistic is explained as: $VARCH_{\mathrm{LM}}(q) = \frac{1}{2}TK(K+1)R_m^2$ with $R_m^2 = 1 - \frac{2}{K(K+1)}\, tr(\hat{\Omega}\,\hat{\Omega}_0^{-1})$, and $\hat{\Omega}$ assigns the above-mentioned regression model's covariance matrix. $\chi^2(qK^2(K+1)^2/4)$ is the distribution of this test statistic.

3) Normality: Jarque & Bera (JB), Skewness, Kurtosis

The Jarque-Bera tests for univariate and multivariate series, as well as separate tests for multivariate skewness and kurtosis (p), are performed on the VAR residuals. By performing the Jarque-Bera test on the residuals following standardization via the Choleski decomposition of the variance-covariance matrix for the centered residuals, one can create a multivariate version of this test. For the multivariate scenario, the test statistics are as follows:

$$JB = \frac{\hat{S}^2(r)}{6\,/T} + \frac{(\hat{K}\,(r) - 3)^2}{24/T},$$

where T is the sample size, $\hat{S}^2(r)$, $\hat{K}\,(r)$ are skewness and kurtosis determined the from sample data, and $\hat{K}\,(r) - 3$ is the excess kurtosis. More specifically, if $\{r_1, \dots, r_T\}$ is a variable with T observations. Beloware the definitions for sample skewness and kurtosis.

$$\hat{S}(r\,) = \frac{1}{(T-1)\hat{\sigma}_r^3}\sum_{t=1}^{T}(r_t - \bar{r})^3\,, and\ \hat{K}(r) = \frac{1}{(T-1)\hat{\sigma}_r^4}\sum_{t=1}^{T}(r_t - \bar{r})^4,$$

when $\hat{\sigma}_r^2$ considering the statistics related to sample variance, $\bar{r}$ is the sample mean of $\hat{S}(r\,)$, it is important to note that both $\hat{S}(r\,)$ and $\hat{K}(r)$ follow a normal distribution with zero mean and variances of $6/T$ and $24/T$, respectively. This is based on the assumption of normality in the data. As a result of this assumption, the JB statistic conforms to a Chi-square distribution with two degrees of freedom in the asymptotic sense.

To evaluate whether the data conforms to a normal distribution, we can use the JB statistic. If JB exceeds the critical value $JB > \chi^2_{2,1-\alpha}$, where α represents the significance level, then we have grounds to reject the null hypothesis (H_0), which posits that the data follows a normal distribution. These findings are in line with the research conducted in [24], as presented in Table 7, which showcases the original data results.

Table 7. Diagnostic tests of VAR (9) for H, R, T

Null Hypothesis Test		Statistic	p-value
no autocorrelation	PT	92.059	0.00991
no suffer from heteroscedasticity	ARCH	143.25	0.9799
not normality	JB	69.979	4.13e-13
	Kurtosis	37.419	3.751e-08
	Skewness	32.56	3.988e-07

The p-value of 0.00991 is less than the significance level of 0.05, disproving the null hypothesis that there is no autocorrelation. On the other hand, the p-value of the heteroscedasticity (ARCH) test is greater than the 0.05 level of significance, which encourages us to keep the null hypothesis in place. Practically speaking, this means that as the fitted values of the response variable increase, the variance of the residuals should not increase as well. Regarding the Portmanteau Test (PT), the p-value of the normalcy test is below the 0.05 significance level, which allows us to reject the null hypothesis.

4) Structural Stability (SVC) [25]

The stability test is used to determine if there are any structural breaks. If we are unable to test for structural breaks and one occurs, the entire estimate may be thrown off. To avoid this, we use a simple inspection technique that involves plotting the cumulative total of subsequent residuals. A structural change has occurred at that particular junction if the total sum of the data points on the chart exceeds certain essential criteria. Fig. 6, which shows the unedited dataset, serves as an illustration of this occurrence.

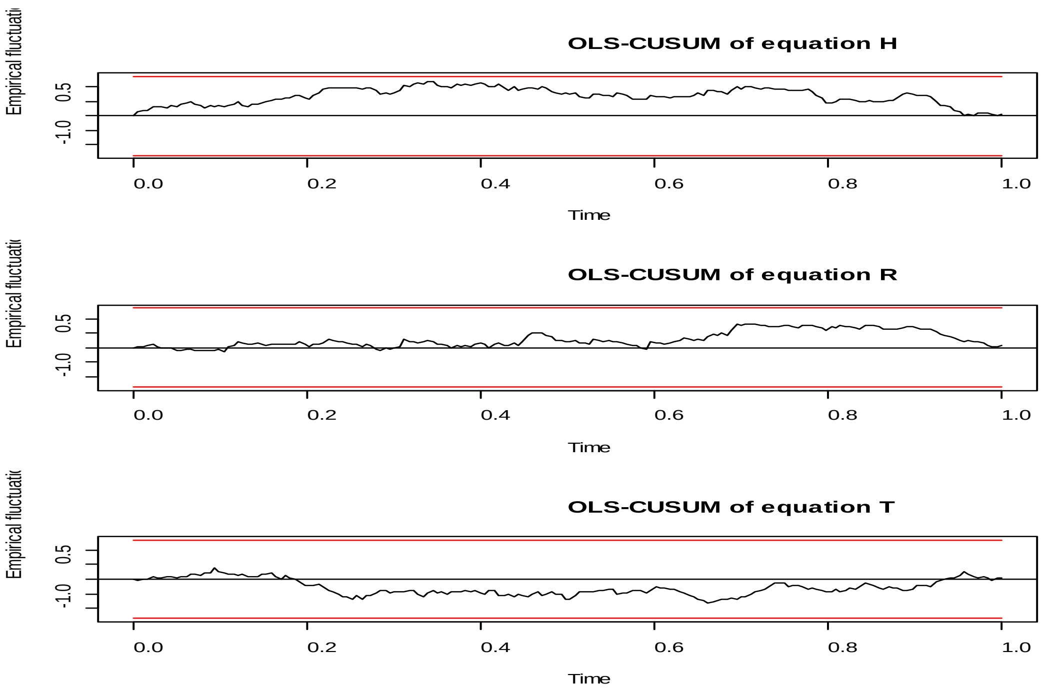

Fig. 6. CUSUM Test for H, R,T

There are no points on this graph beyond the two red lines, so the system is stable.

4.2. Granger Causality

The interdependence structure of the underlying systems of multi-variate time series was investigated. Utilizing Granger causality analysis, we can rephrase the content related to the outcomes presented in Table (8), which encapsulates the unaltered dataset.

Table 8. Causality tests for H, R, T

Null Hypothesis	Statistic(F-test)	p-value
H does not Granger-cause R ,T	3.3014	6.448e-06
R does not Granger-cause H ,T	1.9389	0.01179
No instantaneous causality between: H and R, T	55.968	7.028e-13
T does not Granger-cause H ,R	3.6086	1.0 24e-06
No instantaneous causality between: R and H ,T	53.091	2.962e-12
No instantaneous causality between: T and H,R	12.242	0.002196

We reject the null hypothesis (H_0) due to the p-value being below the significance level of 0.05.

4.3. Forecasting

The built model can be used to generate forecasts since it meets the basic assumption of the model adequacy. The MSE values produced using the program R, are shown in Table (9), which represents the data along with the multivariate model's forecasts for the period (Oct.2000 - May.2001). Table (9) represents the optimal parameters of the multivariate, univariate and MSE for the fitted ARIMA model.

Table 9. Multivariate VAR (9) model's and univariate and MSE for the fitted ARIMA (H, R,T) time series

Time Series	VAR(9)	SARIMA $(1,0,0)(1,1,1)_8$
	MSE	MSE
H	49.2073	15.6396
R	497.190	366.388
T	3.2624	2.2405

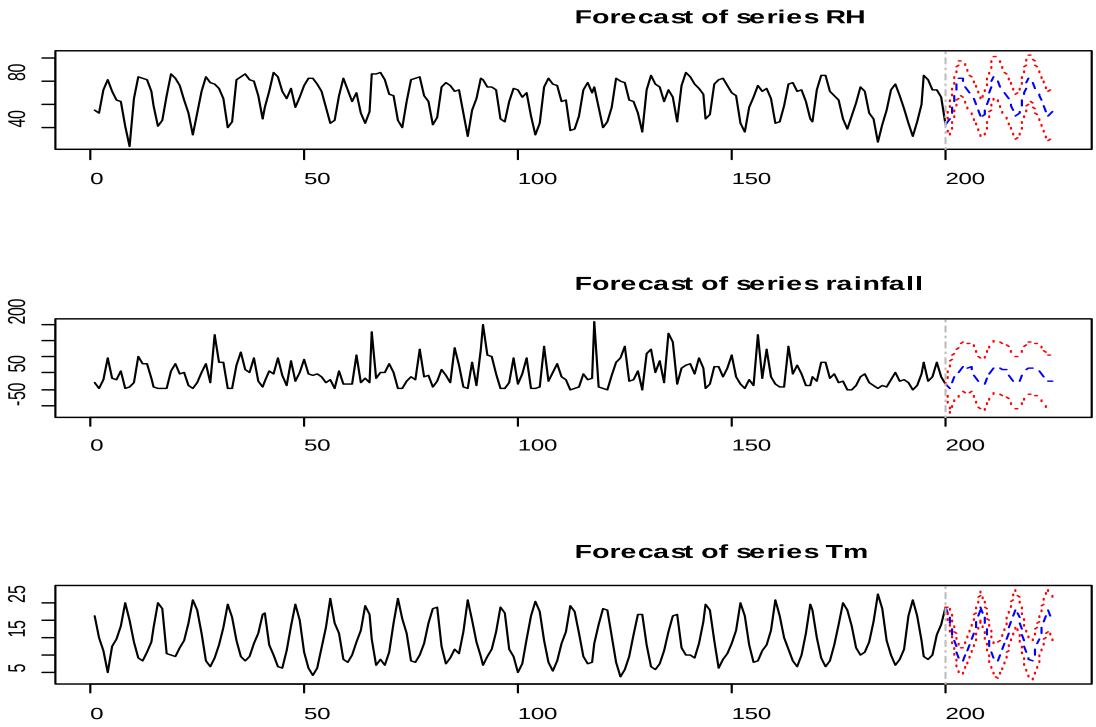

Fig. 7. Forecasts of the Multivariate Model VAR (9)

4.4. Forecast Error Variance Decomposition (FEVD)

A Forecast Error Variance Decomposition assesses the mutual influence of variables through the utilization of the VAR model. To determine the FEVD, we analyze the forecast errors from each equation within the fitted VAR model. Subsequently, the prepared VAR model quantifies the proportion of each error manifestation attributed to unanticipated fluctuations in the counterpart variable (forecast errors). The variance decomposition method aids in the interpretation of the VAR model. The amount of variance in the dependent variable described by each independent variable can be determined. FEVD describes how a potential shock in a one-time series affects the future uncertainty in the other time series of the system. Since this process progresses over time, a shock to a time series can be insignificant in the short run but critical in the long run. When a vector autoregression (VAR) model is used, FEVD, a crucial technique in econometrics and many multivariate time series analytic contexts, helps to comprehend its consequences. The degree to which one variable in the autoregression influences the others is revealed by this decomposition of variance. It evaluates the percentage of forecast error variation for each variable that may be attributable to external shocks affecting the other variables in the context of the data shown in Fig.7.

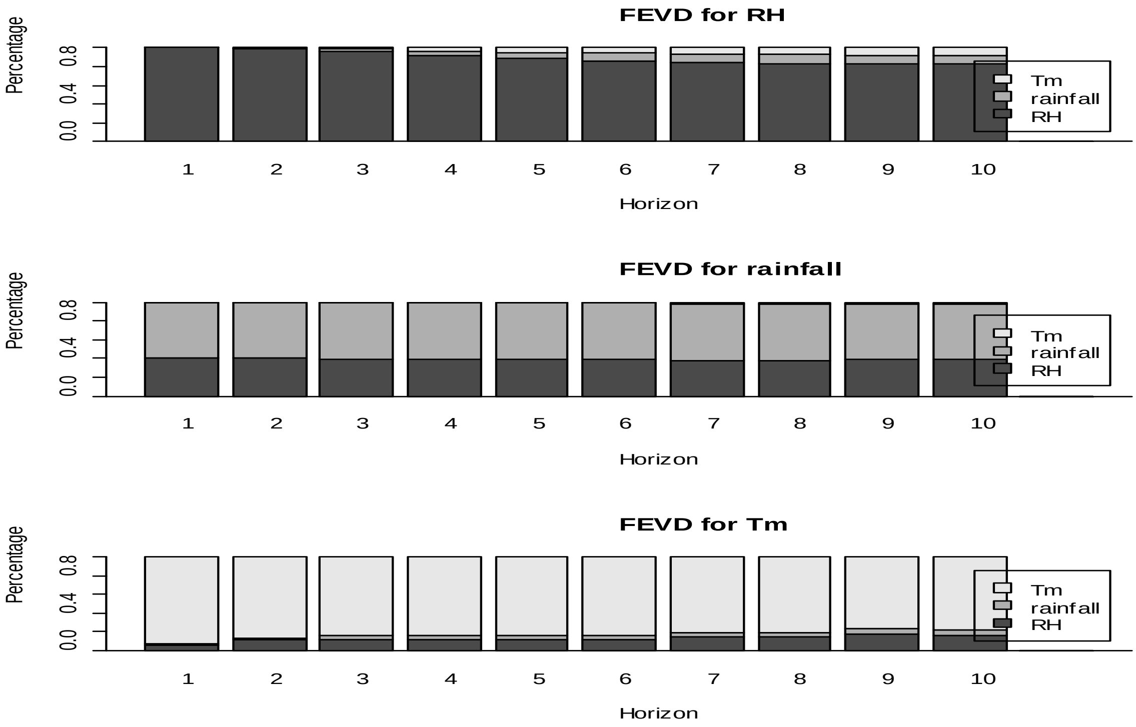

Fig. 8. Forecast Error Variance Decomposition from VAR (9) model fit.

These Fig. 8 graphs show percentages of the shook. The first plot depicts the FEVD for RH starts. It appears that although we were borderline on whether to conclude that Granger causes RH starts, the FEVD reveals that the magnitude of the causality is tiny anyway, while that of RH is greater on rainfall and Tm starts. The second plot shows the FEVD for rainfall. It appears that although we were borderline on whether to conclude that RH starts rate Granger cause rainfall and Tm.

5. Conclusion

The evolution of numerous vars package functions and strategies is described in this article. These improvements give researchers an easy-to-use setting for conducting VAR, SVAR, and SVEC analyses. This is primarily accomplished by putting impulse response function implementation approaches into practice, breaking down forecast error variance, making forecasts, and offering diagnostic testing tools. It also provides tools for determining the model's ideal lag duration, evaluating stability and causation, and performing further diagnostic tests. The article also covers how to determine the co-integrating rank using VECM, which can easily be changed into its level-VAR equivalent. The data was not stationary, as we observed. However, an effective method of transforming a non-stationary series is stationary. To ascertain the model's order, compute the differences and build a correlogram. SARIMA (1,0,0)(1,1,1)8 was chosen for univariate, and the VAR model was then used. Then, using MSE, we assess the forecasting precision. After examining each forecasting accuracy, we concluded that SARIMA would produce better results than VAR in the presence of low-correlated variables and the absence of numerous co-integrations among variables because of its higher forecasting accuracy. They should be aware that there is a correlation. When variables exhibit a strong correlation, the VAR model can be utilized to yield highly favorable outcomes. Limitations of this study include the focus on monthly environmental variables in Ninahvah City, Iraq. Future research could explore other regions, incorporate additional variables, and assess model performance under diverse climatic conditions.

Reference

[1] C.A. Sims, "Macroeconomics and reality", Econometrica: *Journal of the Econometric Society*, pp. 1-48, 1980.

[2] C.W. Granger, "Some properties of time series data and their use in econometric model specification", *Journal of Econometrics*, vol. 16, no. 1, pp. 121-130, 1981.

[3] R.F. Engle, & C.W.J. Granger, "Co-integration and error correction: representation, estimation, and testing", Econometrica: *Journal of the Econometric Society*, pp. 251-276, 1987.

[4] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media, 2005.

[5] D. F. Hendry, *Dynamic econometrics*, Oxford University Press, 1995.

[6] S. Johansen, "Likelihood-Based Inference in Cointegrated Vector Autoregressive Models", *Oxford University Press*, New York, 1995.

[7] J. D. Hamilton, *Time series analysis*, Princeton University Press, Princeton, N.J., 1994.

[8] A. Banerjee, J.J. Dolado, J.W. Galbraith, & D. Hendry, *Co-integration, error correction, and the econometric analysis of non-stationary data*, Oxford University Press, 1993.

[9] J.M. Liu, R. Chen, L.M. Liu, & J.L. Harris, "A semi-parametric time series approach in modeling hourly electricity loads", *Journal of Forecasting*, vol. 25, no. 8, pp. 537-559, 2006.

[10] G.S. Maddala and I.M. Kim, *Unit roots, cointegration, and structural change*, 1998.

[11] W.W. Wei, *Multivariate time series analysis and applications*, John Wiley & Sons, 2018.

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

[13] A.M. Sathe, & N.S. Upadhye, "Estimation of the Parameters of Vector Autoregressive (VAR) Time Series Model with Symmetric Stable Noise", *arXiv preprint* arXiv:2104.07262, 2021.

[14] K. Yamaoka, T. Nakagawa, & T. Uno, "Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations", *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 6, no. 2, pp. 165-175, 1978.

[15] H. Akaikei, "Information theory and an extension of maximum likelihood principle", *In Proc. 2nd Int. Symp. on Information Theory*, pp. 267-281, 1973.

[16] E.J. Hannan, & B.G. Quinn, "The determination of the order of an autoregression", *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 41, no. 2, pp. 190-195, 1979.

[17] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.

[18] S. Johansen, "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, vol. 12, no. 2-3, pp. 231-254, 1988.

[19] G.S. Maddala and I.M. Kim, *Unit roots*, *cointegration, and structural change*, 1998.

[20] V. Patilea, & H. Raïssi, "Corrected portmanteau tests for VAR models with time-varying variance", *Journal of Multivariate Analysis*, vol. 116, pp. 190-207, 2013.

[21] H. Bouakez, & M. Normandin, "Fluctuations in the foreign exchange market: How important are monetary policy shocks?", *Journal of International Economics*, vol. 81, no. 1, pp. 139-153, 2010.

[22] O.C. Onifade, & S.O. Olanrewaju, "Investigating performances of some statistical tests for heteroscedasticity assumption in generalized linear model: A Monte Carlo simulations study", *Open Journal of Statistics*, vol. 10, no. 3, pp. 453-493, 2020.

[23] H. Lütkepohl, "Structural vector autoregressive analysis for cointegrated variables", *Allgemeines Statistisches Archiv*, vol. 90, pp. 75-88, 2006.

[24] C.M. Jarque, & A.K. Bera, "A test for normality of observations and regression residuals", *International Statistical Review/Revue Internationale de Statistique*, pp. 163-172, 1987.

[25] H. Herwartz, "Hodges–Lehmann detection of structural shocks–an analysis of macroeconomic dynamics in the euro area", *Oxford Bulletin of Economics and Statistics*, vol. 80, no. 4, pp. 736-754, 2018.

Միաչափ SARIMA և բազմաչափ VAR մոդելների համեմատական վերլուծություն ժամանակային շարքերի կանխատեսման համար. կլիմայի փոփոխականների դեպքի ուսումնասիրություն Իրաքի Նինահվա քաղաքում

Սամիրա Ա. Օթմա, Հասան Հ. Ջամիլ և Սադեք Թ. Աբդուլազիզ

Դուհոկի համալսարան, Հիմնական կրթության քոլեջ, Դուհոկ, Իրաք
e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Ամփոփում

Այս ուսումնասիրությունը ներառում է համեմատություն SARIMA միաչափ մոդելի կիրառման և VAR մեթոդների (վեկտորային ավտոռեգրեսիվ մոդելներ) օգտագործման միջև բազմաչափ ժամանակային շարքերի վերլուծության համար: Վերլուծությունն իրականացվում է եռաժամանակյա շարքի փոփոխականների միջոցով, որոնք ստացվել են Իրաքի Նինահվա քաղաքում խոնավության (H), տեղումների (R) և ջերմաստիճանի (T) ամսական միջինը ներկայացնող տվյալներից: Այս շարքերը մոդելավորելու համար օգտագործվում են և՛ միաչափ, և՛ բազմաչափ ժամանակային շարքերի մոտեցումները: Հոդվածը ուրվագծում է նաև վեկտորային ավտոռեգրեսիայի, կառուցվածքային վեկտորի ավտոռեգրեսիայի և կառուցվածքային վեկտորի սխալի ուղղման մոդելների իրականացումը vars փաթեթի միջոցով: Բացի այդ, այն ապահովում է ախտորոշիչ թեստավորման, սահմանափակ մոդելների գնահատման, կանխատեսման, պատճառահետևանքային վերլուծության, իմպուլսային արձագանքի վերլուծության և կանխագուշակման սխալի շեղումների տարրալուծման գործառույթներ: Բացի այդ, այս մոդելները գնահատելու համար ներդրվում են երեք հիմնարար գործառույթներ՝ VAR, SVAR և SVEC: Մեթոդների համեմատությունը հիմնված է յուրաքանչյուր մոտեցման արդյունքում առաջացած միջին սխալի գնահատման վրա: Հետազոտության արդյունքները ցույց են տալիս, որ միաչափ գծային ստացիոնար մեթոդները գերազանցում են բազմաչափ մոդելներին: Տվյալների վերլուծությունը կատարվել է R ծրագրային հարթակի միջոցով: Հիմնական նպատակը տվյալների մշակման մեջ միաչափ և բազմաչափ ժամանակային շարքերի մոդելների կատարողականի գնահատումն է: Հետազոտության բացը կայանում է ամսական բնապահպանական փոփոխականների համատեքստում ժամանակային շարքերի վերլուծության SARIMA և VAR մեթոդների համեմատական գնահատման անհրաժեշտության մեջ: Այս մոդելներն ընտրվել են ժամանակային սերիաների տվյալների մեջ բազմաթիվ փոփոխականների միջև ժամանակային կախվածությունների և փոխազդեցությունների գրանցման արդյունավետության շնորհիվ՝ ապահովելով Իրաքի Նինահվա քաղաքի կլիմայական օրինաչափությունների համապարփակ վերլուծություն: Ուսումնասիրության նպատակն է լրացնել հետազոտության բացը՝ համեմատելով այս մոդելները և հիմնավորելով դրանց ընտրությունը՝ հիմնվելով նշված ժամանակային շարքի տվյալները վերլուծելու նրանց կարողությունների վրա:

Բանալի բառեր՝ Միաչափ ժամանակային շարքեր, Բազմաչափ գործընթաց, Խաչաձև հարաբերակցություն և VAR, Կանխատեսում, ARCH-LM թեստ, Կառուցվածքային կայունություն (SVC)

Сравнительный анализ одномерных SARIMA и многомерных VAR моделей для прогнозирования временных рядов: тематическое исследование климатических переменных в городе Нинахва, Ирак

Самира А. Отман, Асан А. Джамил и Садек Т. Абдулазиз

Университет Духока, Колледж базового образования, Духок, Ирак
e-mail: Sameera.Othman@uod.ac, hasan.hazim@uod.ac, sadiq.taha@uod.ac

Аннотация

Данное исследование включает сравнение применения одномерной модели SARIMA и использования методов VAR (векторных авторегрессионных моделей) для многомерного анализа временных рядов. Анализ проводится с использованием переменных трех временных рядов, полученных на основе данных, представляющих среднемесячные значения влажности (H), количества осадков (R) и температуры (T) в городе Нинахва, Ирак. Для моделирования этих рядов используются как одномерные, так и многомерные подходы к временным рядам. В статье также описывается реализация моделей векторной авторегрессии, структурной векторной авторегрессии и структурной векторной коррекции ошибок с использованием пакета vars. Кроме того, он предоставляет функции для диагностического тестирования, оценки моделей с ограничениями, прогнозирования, анализа причинно-следственных связей, анализа импульсных характеристик и разложения дисперсии ошибок прогноза. Кроме того, для оценки этих моделей вводятся три фундаментальные функции: VAR, SVAR и SVEC. Сравнение методов основано на оценке средней ошибки, создаваемой каждым подходом. Результаты исследования показывают, что одномерные линейные стационарные методы превосходят многомерные модели. Анализ данных проводился с использованием программной платформы R. Основная цель — оценить эффективность одномерных и многомерных моделей временных рядов при обработке данных. Пробел в исследованиях заключается в необходимости сравнительной оценки методов SARIMA и VAR для анализа временных рядов в контексте ежемесячных переменных окружающей среды. Эти модели были выбраны из-за их эффективности в определении временных зависимостей и взаимодействий между множеством переменных в данных временных рядов, обеспечивая всесторонний анализ климатических моделей в городе Нинахва, Ирак. Исследование направлено на устранение пробелов в исследованиях путем сравнения этих моделей и обоснования их выбора на основе их возможностей анализировать указанные данные временных рядов.

Ключевые слова: Одномерный временной ряд, Многомерный процесс, Взаимная корреляция и VAR, Прогноз, Тест ARCH-LM, Структурная устойчивость (SVC)