STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, June 2026, pp 4757–4773.

Simulating the Interdependence Between Electricity Production and Water Release, Based on a Methodology that Combines Statistical Models and Artificial Intelligence

Hyllaa Anas Al-Omari, Najlaa Saad Ibrahim*, Alla Abdul Alsattar Hamoodat

Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Iraq

Abstract This research addresses the modeling of the dynamic relationship between water releases and electricity production in hydroelectric power systems, given the temporal complexity, instability, and long-term equilibrium structure of this relationship. Due to the limitations of traditional statistical models in representing nonlinear relationships, and the limited explanatory power of purely intelligent models, this research proposes an integrated hybrid framework combining the Vector Error Correction Model (VECM) and multilayer artificial neural networks (MLPs). The proposed framework leverages the structural properties of the VECM model to characterize cointegration and long-term equilibrium relationships, while employing the flexibility of neural networks to represent short-term nonlinear dynamics. Two interconnected hybrid models were developed, one for predicting electricity production and the other for predicting water releases, maintaining the structural and dynamic interdependence between the two variables within a unified system .The proposed model was applied to real-time daily data from the Mosul Dam hydroelectric power station in 2024, with one portion of the data allocated for estimation and another for testing predictive performance. The results clearly demonstrated the superiority of the hybrid model over the traditional model in terms of prediction accuracy, as measured by (RMSE, MAE, sMAPE and MASE) criteria, both during the training period and in future forecasting, The DM test results also showed that the accuracy of the VECM-ANN hybrid model is as good as possible. This reflects the model's ability to generalize and stabilize.

Keywords neural network, cointegration, long-term equilibrium, power production, water release, error correction model

DOI: 10.19139/soic-2310-5070-3340

1. Introduction

' The dynamic hydroelectric system at the Mosul Dam is one where the deterministic factors intertwine with hydrological and climatic elements within its operational framework. It is directly linked to electricity generation, in conjunction with the water flowing into the dam from the Tigris River, and is influenced by seasonal variations resulting from rainfall, snowmelt, and upstream flows. Daily data reveals that these water characteristics change over time, exhibiting a long-term correlation between the success and the physical outcome of converting natural energy into electrical energy.

From a statistical perspective, systems are usually analyzed using time series models. Data on water discharge and electricity generation show clear non-stationary characteristics, accompanied by random trends and high time continuity. These variables are often complementary, meaning there is a long-term equilibrium relationship despite short-term fluctuations. Cointegration theory provides a strong theoretical basis for addressing such phenomena, allowing the study of non-stationary variables within a unified system without encountering pseudo-regression problems. One of the most important practical applications of this theory is the vector error correction model

---

*Correspondence to: Najlaa Saad Ibrahim (Email: najlaa.s.a@uomosul.edu.iq). Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

(VECM), as it is able to represent long-term structural relationships, in addition to describing short-term adaptation mechanisms [15].

The research presents an integrated hybrid framework that combines a vector error correction model and multi-layer neural networks. Multi-layer neural networks are characterized by great flexibility in representing nonlinear relationships and extracting hidden patterns. The proposed model is distinguished from traditional VECM-ANN applications in that it does not limit itself to using the final outputs of the standard model as inputs to the neural network, but rather adopts a systematic integration strategy that includes the input of error correction limits, structural residuals, and optimal time delays for both variables, so that the neural network acts as a nonlinear corrector for short-term dynamics conditioned by the long-term equilibrium structure. The model is based on a two-way linking mechanism, where two interconnected hybrid models were developed, one dedicated to predicting electricity production and the other to predicting water releases, with a systematic exchange of structural information between them that maintains the reciprocal nature of the relationship instead of processing it unilaterally. The research framework consists of an introduction, followed by a literature review, then the concepts and methodology of the Vector Error Correction (VECM) model, an introduction to the concept of neural networks and the model hybridization method, data sources, applied data analysis, and finally, conclusions and recommendations related to the study of the two variables (electricity and water).

The research problem lies in constructing a joint forecasting system capable of modeling the dynamic relationship between water releases, measured in cubic meters per second, and electricity production, measured in megawatts, using daily data from a complete operating cycle. The study aims to generate short-term, single-step forecasts within a daily framework, while evaluating out-of-sample predictive power for a defined number of future observations. The model's task is not limited to improving the prediction accuracy of each individual variable, but rather aims to produce joint and conditional forecasts that reflect the dependence of each variable on the others within a single, interconnected system. This ensures the preservation of the long-term equilibrium structure while flexibly representing short-term dynamics. The correlation between the two variables is measured by testing for cointegration, analyzing the rate of adjustment towards equilibrium, examining the joint structure of variance and correlation, and evaluating the model's ability to reproduce the combined behavior of the two time series.

2. Literature Review

Previous literature has focused on time series modeling and the study of long-term relationships between economic and environmental variables, particularly in the energy and water sectors, using traditional statistical models and artificial intelligence techniques. Single-root and cointegrating models have formed the theoretical basis for non-stationary time series analysis. [9] and [15] provided a robust methodological framework for testing the existence of long-term equilibrium relationships between variables, paving the way for the use of vector error correction (VECM) models to characterize short- and long-term dynamics.

Numerous studies have shown that the VECM model has a high capacity for explaining long-term structural relationships and rapid post-shock adaptation, but its efficiency remains limited when dealing with nonlinear relationships and complex patterns in data, especially in volatile and nonlinear applications, such as energy and water resource markets [18, 19].

In contrast, artificial neural networks (ANNs) have garnered significant attention for their ability to represent nonlinear relationships and detect hidden patterns in data, proving effective in various fields such as price prediction, traffic accident analysis, and medical applications [17, 23]. However, despite their predictive power, these models suffer from weak economic interpretation and a lack of a theoretical foundation linking the results to long-term economic structures.

As a result of this gap, recent studies have turned to hybrid models that combine traditional statistical models with artificial intelligence techniques, aiming to leverage the complementary advantages of both approaches. Several studies have shown that integrating VECM with multilayer neural networks (MLPs) leads to a significant improvement in prediction accuracy by combining the representation of long-term equilibrium relationships with the ability to model short-term nonlinear dynamics [2, 22].

Within the framework of water management agreements, the scientific bodies involved in a wide range of water and electricity production projects, particularly hydroelectric power plants, establish a dynamic and integrated relationship with multiple factors and structures. This makes research into dynamic options particularly relevant for addressing this complexity. Studies indicate that using global data, along with precise evaluation criteria such as root mean square error (RMSE) and mean absolute error (MAE), yields reliable results and practical applicability for innovative future planning.

3. Unit Root Concept

The unit root idea can be used to establish the permanence of a non-stationary component in a series of data. A time series model is non-stationary and has a unit root when the regression coefficient is approximately equal to (1).

Often called a random walk with skew, a unit root is a random trend in a time series. It is a feature of certain random processes. One shock can have a long-lasting impact on this random process, which is a time series model. This implies that the process can be permanently impacted by the impact of a single random occurrence. Examine the regression model that follows:

$$\Delta Z_t = \sum \beta_t \Delta z_{t-1} + \gamma \mu_{t-1} + e_t \qquad \dots \; (1)$$

Where as z_t is represents the vector of variables to be selected and β is represents the short-term coefficients, ΔZ_t is the change in the variable at time t , Δz_{t-1} are lagged differences to handle autocorrelation , $\gamma \mu_{t-1}$ represents the error correction term (if analyzing cointegration) or indicates the speed of adjustment, which defines the presence of a unit root and e_t is a white noise error term.

The null hypothesis and the alternative hypothesis for the unit root are as follows:

$$H_0: \rho = 1 \qquad or \qquad H_0: \gamma = 0$$

$$H_1: \rho < 1 \qquad or \qquad H_1: \gamma < 0$$

The ADF test is a statistical test used in time series analysis to determine whether a time series contains a unit root, particularly when the results indicate an autocorrelation between errors. It is the most significant of the tests used in our study. In other words, it becomes unstable due to an accumulation of random components [1]. The primary distinction between the PP test and the ADF tests, despite their comparable scientific concepts, is that the PP test takes into account autocorrelation and heteroscedasticity in errors rather than include lag differences in the model [16].

4. Cointegration

One of the fundamental and crucial criteria for confirming whether or not two or more variables have a long-term equilibrium connection is this one. As a result, it can be described as the process of two or more time series correlating across time, with changes in one series canceling out changes in the other. This can therefore make it possible to turn an unstationary time series that has a long-term equilibrium relationship between them into a stable series when evaluated collectively. This indicates that time series are not isolated from one another, even if they move in tandem over an extended period of time. We refer to this as cointegration. The long-term correlation helps forecast the values of the dependent variable in terms of the independent variables [15, 7].

The Engle-Granger test, which was previously only valid for one pair of variables, is expanded by the Juselius-Johansen test, which is a multivariate cointegration test. There could be more than one integration vector and, hence, more than one equilibrium relationship controlling the correlations between the variables when using the Johansen test.

5. Vector Error Correction Model (VECM)

Long-term parameters must be determined using a Vector Error Correction Model (VECM). The speed at which variables return to equilibrium following a shock or disturbance is referred to as error correction in this model. The assumption that any departure from equilibrium is eventually restored is reflected in this term. Both short-term and long-term relationships between variables are guaranteed by the VECM vector. The presence of joint integration between variables is its most crucial requirement. Unlike other conventional approaches, it can be applied even with a small sample size [14].

6. Finding the lag duration, or the number of lags

The lag length is crucial for figuring out the rank of the VECM model and for use in cointegration tests like the Johansen test. The ideal lag length is determined by a number of factors, the most crucial of which are:

6.1. Akaike Information Criterion

In time series analysis and VECM models, it is one of the most crucial statistical factors used to establish the model rank, optimal number of lags, and model quality. It compares how well models match each other and is described by the following formula [3]:

$$AIC\,(p) = -2\ln\,(L)\,+2m$$

Where L is likelihood and m is represents the number of parameters.

$$AIC\,(p) \cong n\ln\left|\widehat{\Omega}\right|\,+2m \qquad \ldots(\,2\,)$$

Where as $\left|\widehat{\Omega}\right|$ is refers to a particular estimate of the covariance matrix of the residuals in the model, and the p-value that corresponds to the lowest criteria value determines the model's rank.

6.2. Schwarz Criterion (SC)

In 1978, Gideon E. Schwarz proposed the SC criterion, which is commonly referred to as the Bayesian Information Criterion (BIC). This model selection criterion serves as a substitute for both the AIC and BIC criteria. Here is what it is [3]:

$$SC\,(p) = n\ln\left|\widehat{\Omega}\right|\,+m\ln\,(n) \qquad \ldots(\,3\,)$$

6.3. Hannan-Quinn Criterion (HQ)

According to the following formula [10], Hannan and Quinn proposed this method to ascertain the model's rank:

$$HQ\,(p) = \ln\left|\widehat{\Omega}\right|\,+2mK^2\frac{\ln(\ln\,(n)\,)}{n} \qquad \ldots(\,4\,)$$

Where K represented number of variables included in the model.

6.4. Final Prediction Error (FPE)

This criterion is crucial for figuring out a model's proper rank. The FPE criterion's mathematical formula is as follows [6].

$$FPE = \frac{\left(1+\left(\frac{m}{n}\right)\right)}{\left(1-\left(\frac{m}{n}\right)\right)}\mathrm{V} \qquad \ldots(\,5\,)$$

Where V: stands for the loss function.

6.5. *Likelihood Ratio (LR)*

The degree to which the model's fit improves with the addition of new parameters is gauged by the LR criterion [12, 20]:

$$LR = 2Log\left(\frac{L(\hat{\theta})}{L(\hat{\theta}_0)}\right) \quad \ldots(\ 6\ )$$

Where $L\left(\hat{\theta}\right)$ refers to the greatest value of the probability function for the simple model (Null model), and the probability function's highest value for the more intricate model (Alternative model) is denoted by $L\left(\hat{\theta}_0\right)$.

7. Neural Networks

7.1. *Artificial neural networks (ANNs)*

A subfield of machine learning that draws inspiration from the human brain, neural networks are also known as artificial neural networks (ANNs). In order to digest information, recognize events, and balance possibilities in order to make decisions and forecast results, it employs mechanisms that are similar to how biological neurons cooperate [11, 12]. A collection of nodes known as neurons makes up a neural network. After processing inputs and one or more hidden layers using a mathematical operation, each neuron generates outputs. Each node has its own weight and boundaries and is connected to other nodes, as illustrated in Figure (1).

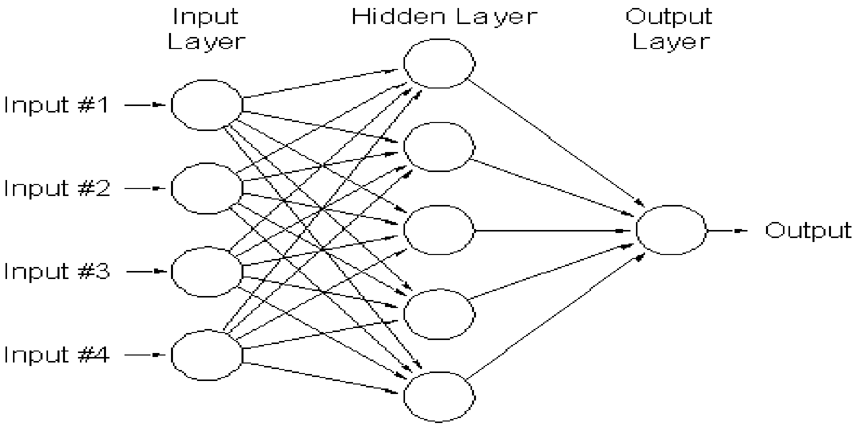

Figure 1. Original images.

Computational models known as artificial neural networks are designed to mimic how biological neural networks function in the human brain. Neurons are basic, networked processing units that make up artificial neural networks. Although each neuron carries out a basic computational function, the relationships between these different parts dictate the network's overall behavior. Neural network applications include self-driving cars, picture recognition, machine translation, disease diagnosis, and stock price prediction [13].

7.2. *Multilayer neural networks (MLNs)*

This network is regarded as a computer system intended for information processing and is a form of artificial intelligence. It is made up of numerous dynamic tools and associated processing parts whose purpose is to parallelize the network's computing. The neural network is made up of several homogenous, connected processing units. Every unit is a computational tool that can be represented by a straightforward mathematical formula on the computational system made up of several interconnected processing units and distinguished by the way each one processes the data that enters it in a balanced manner. This network's structure is made up of [19]:

1. The input layer is the one that interacts with the outside world. The neural network's input layer depicts a pattern, and the output layer will generate a different pattern after a model has been shown to the input layer. Additionally, it symbolizes the circumstance for which the neural network is trained.
2. Output layer: The neural network's output layer is what genuinely shows the outside world a pattern. The kind of work that the network must perform should be closely correlated with the number of output neurons.
3. Hidden layer: The hidden layer is the layer that is placed between the input and the output layer of a neural network. When the hidden layer can be accessed, then the activation function is applied to the hidden layer. The hidden layer consists of the hidden nodes. Hidden neurons or hidden cells are cells that do not appear on either the input or the output layers.

8. The VECM-ANN Hybrid Model

The combination of Artificial Neural Networks (ANNs) and Vector Error Correction Regression (VECM) models is a complicated hybrid approach that combines the strengths of modern artificial intelligence methods and the strengths of traditional statistical models. This integration has the basis of a free philosophy which utilizes the merits of each methodology to counter the demerits. The VECM model is inferior in dealing with complex patterns in data and nonlinear interactions, although it is superior in its ability to model long-term equilibrium relationships among economic variables and also identifies error-correcting mechanisms that restore the economic variables to the equilibrium. Due to the multi-layered nature and the capacity of machine learning neural networks are famously famous in their capacity to infer complex correlations of variables and nonlinear trends. However, they lack the theoretical grounds on which to interpret long-term economic liaisons. In this way, we are able to employ structured statistical data and still maintain the flexibility of machine learning with a combination of them using the outputs of the VECM framework (as a form of neural network inputs) in the form of residuals, lagged, and equilibrium relationships. This integration provides more specific and reliable forecasts by considering short-term nonlinear dynamics as well as structural connection between the long-term structure, which forms a mixed model paradigm between the statistical accuracy and the flexibility of the computation. The following figure (2) illustrates the steps involved in hybridization by taking the output of a VECM model and using it as input to an ANN.

The process begins by examining the stability of the time series using stability tests (ADF, PP). Then, cointegration tests are performed to determine if there is a long- VECM or short-term equilibrium relationship. The extent to which a vector error correction model (VECM) is used is confirmed, and the VECM model's rank (the number of lagging values) is determined. The model is then estimated at this rank, and the diagnostic phase of the VECM model is completed at this rank. In the second phase, the model's outputs (lagging differences, extrinsic variables, and error vector) are compiled into an array. The artificial neural network then processes this array by identifying layers to extract final hybrid predictions.

9. Interdependence Simulation

The concept of correlation simulation refers to the ability of a proposed model to reproduce the shared structural properties of the real system, rather than simply reducing prediction errors. This is achieved by maintaining the long-term equilibrium relationship between the two variables electricity production and water release—according to the Johansen test:

$$\mathrm{E}_{\mathrm{t}-1} + \beta \mathrm{W}_{\mathrm{t}-1} + C \ = 0 \qquad \ldots(7)$$

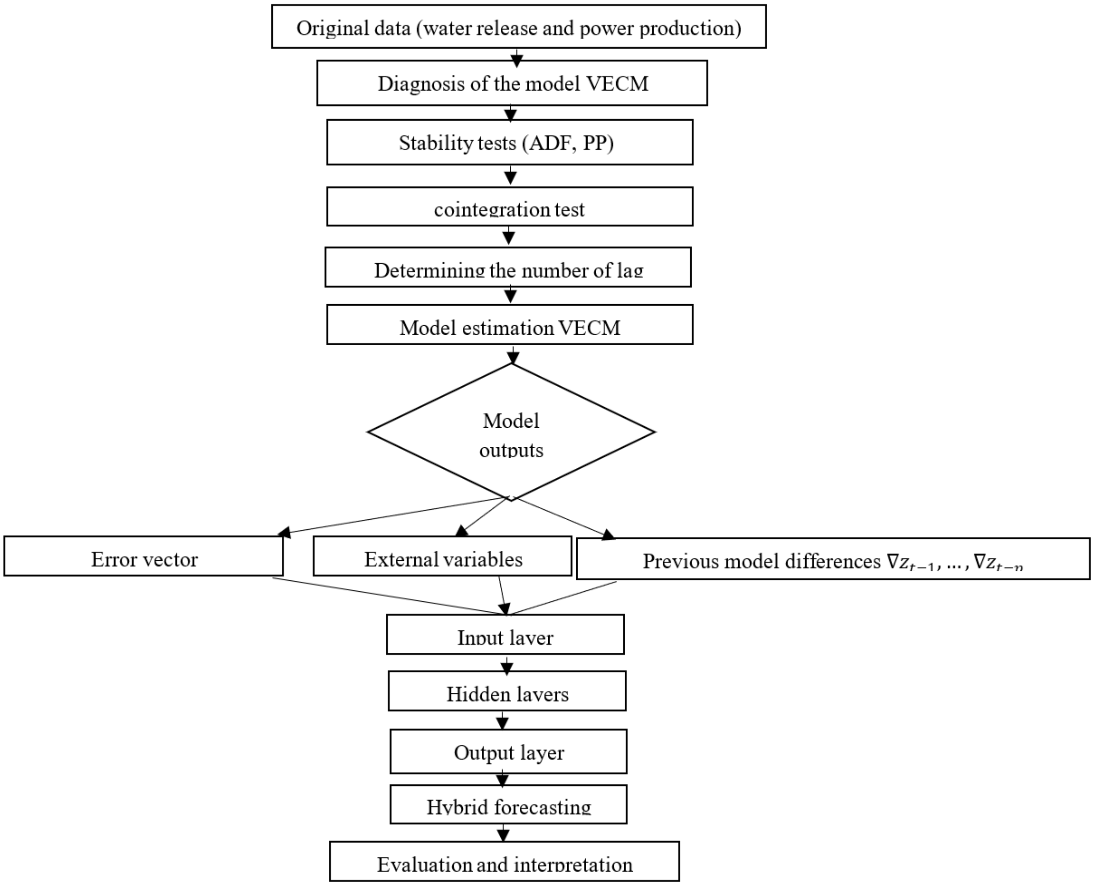

Figure 2. Flowchart of VECM-ANN Hybridization Steps.

reproducing the dynamic adjustment mechanism after shocks, as captured by the VECM model.

$$\nabla \mathrm{E_t} = \alpha_1 ECT_{t-1} + \sum_{j=1}^{p} \mathcal{L}_{1j} \nabla \mathrm{E_{t-j}} + \sum_{j=1}^{p} \ell_{1j} \nabla W_{\mathrm{t-j}} + e_{1t} \qquad \dots (8)$$

$$\nabla \mathrm{W_t} = \alpha_2 ECT_{t-1} + \sum_{j=1}^{p} \mathcal{L}_{2j} \nabla \mathrm{E_{t-j}} + \sum_{j=1}^{p} \ell_{2j} \nabla W_{\mathrm{t-j}} + e_{2t} \qquad \dots (9)$$

The following table (1) shows the symbols used in the equations above.
and replicating the shared structure of fluctuations and correlation between the two series. Correlation simulation is evaluated by comparing the characteristics of the predicted series with those of the actual data, including the stability of the shared relationship, the convergence of the correlation structure, and the model's ability to maintain consistency between the two time paths.

Table 1. Symbols used in equations (7,8 and 9).

symbols	Description
$\mathrm{E_t}$	Electricity production over time t.
$\mathrm{W_t}$	Water release over time t.
$\mathrm{E_{t-1}}$	Electricity production lagged.
$\mathrm{W_{t-1}}$	Water release lagged.
$\nabla \mathrm{W_t}$	First difference in water release
$\nabla \mathrm{E_t}$	First difference in electricity production
$\nabla \mathrm{E_{t-j}}$	lagged first difference of electricity
$\nabla \mathrm{W_{t-j}}$	lagged first difference of water release
e_{1t}	Random error in the electricity equation
e_{2t}	Random error in the water equation
α_1	Electricity adjustment speed
α_2	Water adjustment speed
$\mathrm{ECT_{t-1}}$	Error-Correction
β	Water variable coefficient
C	Equilibrium equation constant
$\mathcal{L}$	lagged difference coefficient in electricity
ℓ	lagged difference coefficient in water

10. Predictive criteria

1. One statistical measure that is identical to MSE is the Root Mean Square Error (RMSE) [18, 2]. This is its formula.

$$\text{RMSE } = \sqrt{\text{MSE }} = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_t\right)^2}{n}} \qquad \ldots(10)$$

   Where y_t represented the value determined at time t, $\hat{y}_t$ The estimated value from the model is represented at the same time, n represents the number of observation.

2. The Mean Absolute Error (MAE) is a measure that transforms errors in positive values as it involves the summation of the absolute error and then the average of the differences between the actual values. The average error is determined by use of the following equation [5, 8]:

$$\text{MAE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{t}=1}^{\mathrm{n}} |\mathrm{e_t}| \qquad \ldots(11)$$

Where $|\mathrm{e_t}| = |y_t - \hat{y}_t|$ represents the prediction error at time t .

11. Practical Aspect

Accurate predictive models that take into account the dynamic interaction between water release and power production are necessary given the growing difficulties in managing energy and water resources. Because of its nonlinearity and the way variables interact, this relationship poses a methodological problem, particularly in hydropower plants. a hybrid model was created that combines artificial neural networks (ANN) to capture the non-linear patterns and the vector error correction model (VECM) to reflect the long-term structural link. The VECM model was estimated using Eviews10 software, and the hybrid model was programmed and compared using Matlab software.

11.1. Study Data

The data for this study were obtained from the Mosul Dam Power Stations Directorate and comprised 366 daily observations dating back to 2024, reflecting a complete annual cycle of hydrological and operational changes in the hydroelectric system. The data included two main variables: $(\mathrm{W_t})$ water releases measured in cubic meters per second $\left(m^3/S\right)$and ($\mathrm{E_t}$) electricity production measured in megawatts (MW). These represent the primary inputs and outputs of the dam's hydroelectric power generation process. Water releases reflect the inflows into the dam from the Tigris River, which are influenced by several natural and climatic factors. These include the volume of water flowing from upstream sources within Turkish territory, as well as rainfall and snowmelt levels during the spring. These factors lead to significant seasonal variations in available water quantities, which directly impact electricity generation levels. Given the nature of daily data and the non-linear interactions it contains between variables, it represents a suitable testing environment for evaluating the efficiency of standard models and proposed hybrid models, particularly with regard to their ability to represent long-term equilibrium relationships and short-term dynamics simultaneously. The data were recorded regularly on a daily basis, and no values were missing. A box plot was used to check for outliers, and no outliers were found, meaning all data fell within the dam's operational range and no values were excluded. Table (2) below shows the descriptive statistics for both variables ($\mathrm{W_t}$, $\mathrm{E_t}$) and reveals a high degree of variability between them. Water discharge exhibits a higher mean and greater dispersion than electricity generation. The difference between the minimum and maximum values indicates significant fluctuations, rendering traditional models unreliable for prediction. This necessitates the adoption of hybrid modeling, which can detect complex nonlinear patterns.

Table 2. descriptive statistics for both variables ($\mathrm{W_t}$, $\mathrm{E_t}$).

$\mathrm{W_t}$	$\mathrm{E_t}$	statistics
352.02	198.13	mean
145.08	94.28	StDev
21049.60	8887.86	Variance
165	0	Minimum
350	197	Maximum
750	414	Median

The data was divided into (336) observations for the estimation (training) process to enable the VECM model to be estimated, as it requires a large dataset. (30) observations were left for testing.
The following figure (3) shows the plot of both original series ($\mathrm{E_t}$, $\mathrm{W_t}$), which are found to be non-stationary.

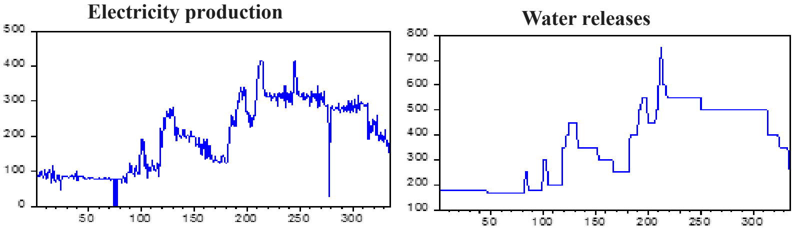

Figure 3. Data diagram for the water release and electricity production chain.

To confirm the stationarity of the two series ($\mathrm{E_t}$, $\mathrm{W_t}$), the ADF and PP statistical tests were used to examine the unit root, as shown in the following table (3).
It is clear from the table above that both original series ($\mathrm{E_t}$, $\mathrm{W_t}$), are non-stationary. The values (P-value) in the brackets below the calculated values for the ADF and PP tests indicate the non-stationarity of the two series because they are greater than the significance level of 0.05. After taking the first difference for both series, indicated by the symbol ($\nabla \mathrm{E_t}$, $\nabla \mathrm{W_t}$), and re-evaluating the ADF and PP tests, the two series were found to be stationary because the values ($\text{P-value} < 0.05$) indicate acceptance of the alternative hypothesis, which states that the two series are stationary.

Table 3. Unit root test results for the study variables.

PP	ADF	Variables
-2.38377 (0.1471)	-1.84493 (0.3584)	E_t
-32.00408 (0.0000)	-17.8892 (0.0000)	∇E_t
-1.66327 (0.4491)	-1.63913 (0.4614)	W_t
-17.77625 (0.0000)	-17.7747 (0.0000)	∇W_t

11.2. The test for Juselius-Johansen cointegration

As shown in the following two tables (4 ; 5) , the long-run equilibrium connection was tested using the Juselius-Johansen cointegration test both in the absence of the trend component and the constant component, as well as again in the absence of the trend and the presence of the constant component on the data I (1), revealing an integral relationship. We observe that the alternative hypothesis, which asserts the existence of simultaneous integration, is accepted since the computed values of the two tests (Maximum Eigenvalue and Trace Test) in the two joint integration test cases are higher than the table values at a significance level of 0.05. This suggests that the two variables have a long-term equilibrium relationship, thus we may utilize the VECM model to describe the long-term link between the variables. The first differences were applied to the short-term inputs of the hybrid model.

Table 4. Results Trace Test of the cointegration.

Prob.**	0.05 Critical Value	Trace Statistic	Eigenvalue	Hypothesized No. of CE(s)	Cases
0.0001	15.49471	413.7641	0.578787	$r = 0$	Absence of both the general direction and the fixed direction.
0.0000	3.841466	125.8468	0.314714	$r \leq 1$	
0.0000	25.87211	414.9212	414.9212	$r = 0$	Absence of the general direction and the presence of the fixed direction.
0.0000	12.51798	127.0016	127.0016	$r \leq 1$	

Table 5. Results Maximum Eigenvalue of the cointegration.

Prob.**	0.05 Critical Value	Max-Eigen Statistic	Eigenvalue	Hypothesized No. of CE(s)	Cases
0.0001	14.26460	287.9172	0.578787	$r = 0$	Absence of both the general direction and the fixed direction.
0.0000	3.841466	125.8468	0.314714	$r \leq 1$	
0.0001	19.38704	287.9196	414.9212	$r = 0$	Absence of the general direction and the presence of the fixed direction.
0.0000	12.51798	127.0016	127.0016	$r \leq 1$	

11.3. Calculating the Model Rank and Estimating It

The following table (6) displays the rank of the VECM model based on the criteria (LR, FPE, AIC, SC, HQ).

We observe from the table (6) above that the appropriate rank for the model is at the second gap according to the criteria (LR, FPE, AIC), as it had the lowest values at this gap.

The electricity production variable was normalized ($\beta_1 = 1$), and the water discharge variable was left to be estimated in order to explain the effect of water Release on electricity generation. The following equation illustrates

Table 6. Criteria for selecting the rank of the VECM model.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-3634.897	NA	14657305	22.17620	22.19933	22.18543
1	-3026.113	1206.433	366861.3	18.48849	18.55788*	18.51617*
2	-3020.118	11.80613*	362428.2*	18.47633*	18.59197	18.52247
3	-3018.651	2.872020	368071.7	18.49177	18.65367	18.55637
4	-3013.799	9.438020	366169.5	18.48658	18.69473	18.56963
5	-3011.829	3.807414	370739.1	18.49896	18.75337	18.60046
6	-3010.196	3.135775	376142.0	18.51339	18.81406	18.63335
7	-3008.042	4.112800	380414.6	18.52464	18.87157	18.66306
8	-3007.434	1.152830	388387.4	18.54533	18.93850	18.70219

the long-term equilibrium relationship:

$$ECT_{t-1} = E_{t-1} + \beta_2 W_{t-1} + C \qquad \ldots (12)$$

The model VECM (2) was estimated, and the cointegration equation was as follows, and the model was estimated at order two:

$$\mathrm{E_t} = 0.636323\mathrm{W_t} - 25.759421 \qquad \ldots (\ 13\ )$$

According to the preceding equation, the amount of electricity production increases by 0.636 units for every unit increase in water release. The baseline level of power production is represented by the constant value (25.759421). The following are the equations for producing power and releasing water:

$$\nabla \mathrm{E_t} = 0.15834 - \ 0.5505\mathrm{E_{t-1}} - \ 0.63632\mathrm{W_{t-1}} \ + \ 25.75942 \ - \ 0.16244\nabla \mathrm{E_{t-1}}$$

$$- \ 0.06626\nabla \mathrm{E_{t-2}} \ - \ 0.01993\nabla \mathrm{W_{t-1}} \ + \ 0.10099\nabla \mathrm{W_{t-2}} \ + e_{1t} \qquad \ldots (14)$$

$$\nabla \mathrm{W_t} = 0.20742 + 0.20334\mathrm{E_{t-1}} \ - \ 0.63632\mathrm{W_{t-1}} + \ 25.75942 \ + \ 0.04468\nabla \mathrm{E_{t-1}}$$

$$- \ 0.01296\nabla \mathrm{E_{t-2}} \ + \ 0.02671\nabla \mathrm{W_{t-1}} \ + \ 0.02034\nabla \mathrm{W_{t-2}} + e_{2t} \qquad \ldots (\ 15\ )$$

The two equations above demonstrate good adaptive behavior for the variables of water release and energy generation, as indicated by the adjustment speed coefficients, which look like this:

1. The electricity production variable responds rapidly to long-term equilibrium shocks, as indicated by the negative electricity adjustment rate coefficient (($\alpha_1 = -0.55050$). This implies that changes in power production are used to swiftly and efficiently repair any departure from the equilibrium relationship with water outflow.
2. The water outflow variable responds to imbalances very slowly, as indicated by the positive water adjustment rate coefficient ($\alpha_2 = 0.20334$). This illustrates how water release systems are more static and steadier than those that production electricity.
3. The model concludes that electricity production adapts quickly to balance the system by a rate of (55%) per day, while water discharge adapts slowly by a rate of (20%) per day, with a conversion efficiency of 0.64 MW per unit of water.

With water preserving structural stability and electricity serving as the quick adjuster, the disparity in adjustment speeds produces a dynamic equilibrium in the system.

11.4. The VECM Model Diagnostic Tests

The LM test was used to determine whether autocorrelation exists in the VECM model residuals up to the tenth gap. The LM test results, listed in Table (7), show that all p-values are greater than the significance level of 0.05, confirming the absence of autocorrelation in the model residuals.

Table 7. LM test results.

Lag	LRE* stat	df	Prob.	Rao F-stat	Df	Prob.
1	8.122757	4	0.0872	2.040302	(4, 648.0)	0.0872
2	6.304352	4	0.1775	1.581328	(4, 648.0)	0.1775
3	9.598185	4	0.0578	2.413653	(4, 648.0)	0.0578
4	3.071596	4	0.5459	0.768533	(4, 648.0)	0.5459
5	3.995304	4	0.4066	1.000363	(4, 648.0)	0.4066
6	4.852462	4	0.3028	1.215786	(4, 648.0)	0.3028
7	1.951661	4	0.7446	0.487897	(4, 648.0)	0.7446
8	1.429778	4	0.8390	0.357287	(4, 648.0)	0.8390
9	0.975503	4	0.9135	0.243683	(4, 648.0)	0.9135
10	2.726461	4	0.6046	0.681997	(4, 648.0)	0.6046

Table (8) below shows the results of the Portmanteau test, which examines whether the model remnants are white noise and the suitability of the chosen model to the data. Note that all probability values for the Q-Stat test after the second gap are greater than 0.05, indicating acceptance of the null hypothesis that the model remnants are white noise. The first and second gaps of the test were neglected because the model rank is second, and correlation values included within the model rank are not used to examine model accuracy.

Table 8. Portmanteau test results.

Lags	Q-Stat	Prob.*	Adj Q-Stat	Prob.*	Df
1	0.061754	—	0.061940	—	—
2	0.243649	—	0.244934	—	—
3	6.636855	0.3557	6.696260	0.3499	6
4	9.615605	0.4748	9.711226	0.4662	10
5	13.55605	0.4833	13.71174	0.4714	14
6	18.33206	0.4340	18.57539	0.4184	18
7	20.24593	0.5677	20.53035	0.5499	22
8	21.66286	0.7070	21.98215	0.6897	26
9	22.63650	0.8299	22.98284	0.8160	30
10	25.37001	0.8573	25.80098	0.8426	34

11.5. The Hybrid Model of VECM-ANN

In order to handle the distinct features of each variable while preserving their interdependence, two distinct yet complimentary hybrid models were created. To predict power production, a hybrid electricity model was developed, and to predict water release, a hybrid water model. By exchanging knowledge and inputs, the two models were combined. To manage intricate nonlinear interactions, both models rely on a multi-layer perceptron (MLP) neural network design. The dual hybrid model's structure is depicted in the following diagram (4).

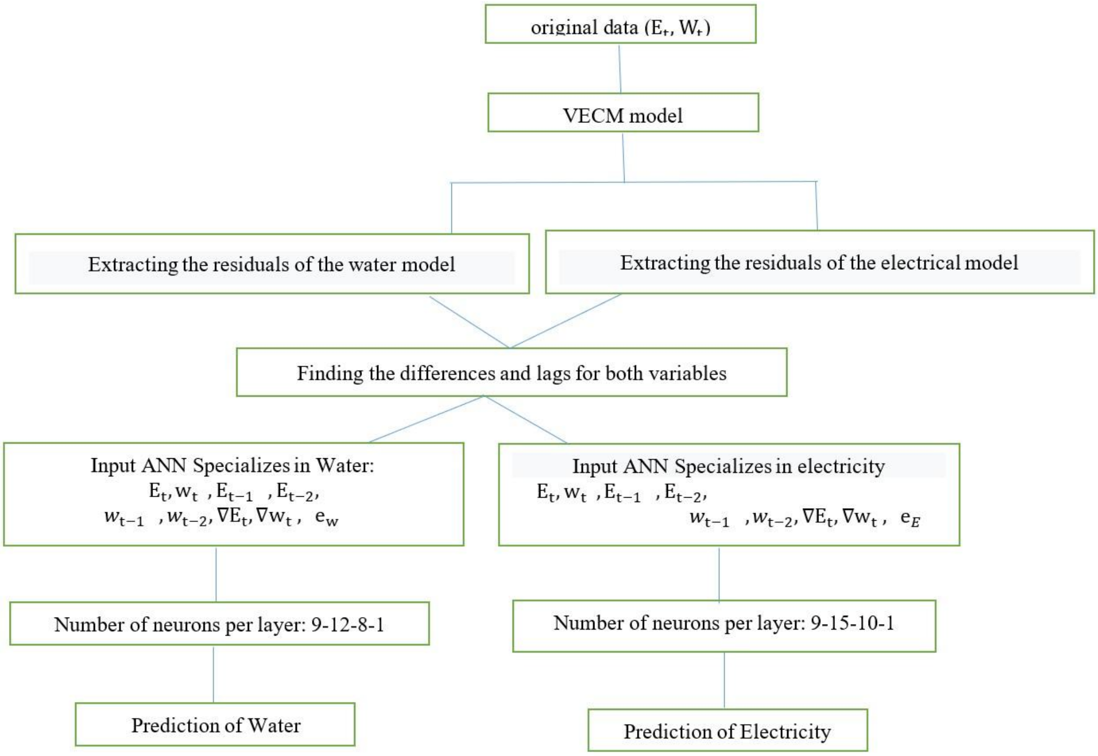

Figure 4. Structure of the dual hybrid model.

The hybridization process was conducted in two stages. In the first stage, the VECM model was estimated, and residuals were extracted for both the electricity and water models, which contain nonlinear patterns. In the second stage, hybridization was performed by taking the VECM model outputs and using them as inputs for the neural network. The following table (9) shows the network structure for both hybrid electricity and water models.

Table 9. Information contained in the hybrid electricity and water models.

Components of a neural network	Hybrid Electricity Model	Hybrid Water Model
The input layer (9) nodes.	The original data for electricity and time transformations (first difference), the lagged values (lag 2, lag 1) for electricity, the original and transformed data (first difference) for water release, the lagged values (lag 2, lag 1) for water release, and the typical residuals for electricity	The initial water and electricity data with their transformations, the lagged values of both variables, and the typical residuals of the water release.
The first hidden layer	15 neurons, the activation function was linear.	12 neurons, the activation function was linear.
The second hidden layer	10 neurons, the activation function was linear.	8 neurons, the activation function was linear.
The output layer	A single node and the activation function was linear(purelin).	A single node and the activation function was linear(purelin).
The training function	Levenberg-Marquardt	Levenberg-Marquardt
Performance evaluation standard	MSE ($1e^{-6}$)	MSE ($1e^{-6}$)
Training epochs	500	500
Implementing the algorithm	MATLAB 2017	MATLAB 2017

Two overriding controls were implemented: First, the training data was randomly split into (85%) training and (15%) validation sets early stopping, which is activated by monitoring validation errors during training, stopping the training if the validation error increases for six consecutive cycles. second the complexity of the neural network, using several hidden layers [15, 10] for electricity and [12, 8] for water. The hybrid and traditional models' performance evaluations during the training period are shown in the table (10) below using (RMSE, MAE, sMAPE and MASE) criteria. The results' consistency and robustness were confirmed by the hybrid model's notable superiority on both criteria.

Table 10. Scale of prediction accuracy criteria for the training period.

Criteria	VECM		VECM-ANN	
	$\mathrm{E_t}$	$\mathrm{w_t}$	$\mathrm{E_t}$	$\mathrm{w_t}$
RMSE	38.3388	18.2055	7.5786	6.9522
MAE	19.4813	7.9378	3.8647	2.2153
sMAPE	8.79	1.95	2.71	0.54
MASE	0.7347	2.0956	0.1458	0.5848

As can be seen from the above table, the hybrid model's RMSE , MAE, sMAPE and MASE values are lower than those of the traditional model. The hybrid model's improvement rate in the RMSE criterion for water release was 61.81%, while it was 80.23% for electricity generation. This notable advantage of the hybrid model shows how well the multi-layer perceptron network can represent intricate nonlinear interactions between system variables. It also demonstrates how well statistical modeling and machine learning techniques work together to greatly increase predicting accuracy. The models' performance on the 30-observation future data (test period) for predictive comparison is displayed in the table (11) below.

Table 11. Scale of forecast accuracy criteria for the test period.

Criteria	VECM		VECM-ANN	
	$\mathrm{E_t}$	$\mathrm{w_t}$	$\mathrm{E_t}$	$\mathrm{w_t}$
RMSE	75.5949	219.8516	12.7206	33.2955
MAE	75.1081	216.3964	7.4876	13.9882
sMAPE	47.51	89.29	4.02	3.78
MASE	11.4039	41.8366	1.1369	2.7044

With an RMSE of 83.17% for electricity production and 84.86% for water release, the hybrid model significantly improved the accuracy of predictions for the 30 future data observations, as shown in the above table. The model's results on the training data, which showed improvements of 80.23% for electricity and 61.81% for water, are consistent with this exceptional performance. This is an indicator of the ability to generalize the hybrid model to new data and maintain a high level of predictive accuracy in later usage. These outcomes are a significant breakthrough towards precision in water release planning and energy production as the hybrid model is capable of relying on the future to make assumptions. Figure (5) below shows the graphs comparing the predictions and errors of the hybrid and standard models for the training and testing period.

The above diagram, which shows the plotting of the original values versus the predicted values for both the hybrid model and the traditional model (VECM), shows that the blue line indicates the actual values and the red line indicates the predicted values according to the two models for the training period, while the green line indicates the predicted values for the testing period. The plotting of the hybrid model for both variables, water and electricity, shows that there is a great convergence to the actual values, which indicates the accuracy of the hybrid model in capturing complex patterns and dynamic relationships between the data. This convergence confirms the importance of integrating the VECM model for modeling long-term equilibrium relationships with multi-layered neural networks, as the hybridization resulted in a highly accurate and efficient predictive model.

To differentiate between the accuracy of the VECM and VECM-ANN models, the Diebold-Mariano (DM) test, proposed by Diebold and Mariano and based on the mean squared error (MSE), was used to determine whether one prediction model predicts more accurately than another (Chen et al. 2014). The null hypothesis states that the

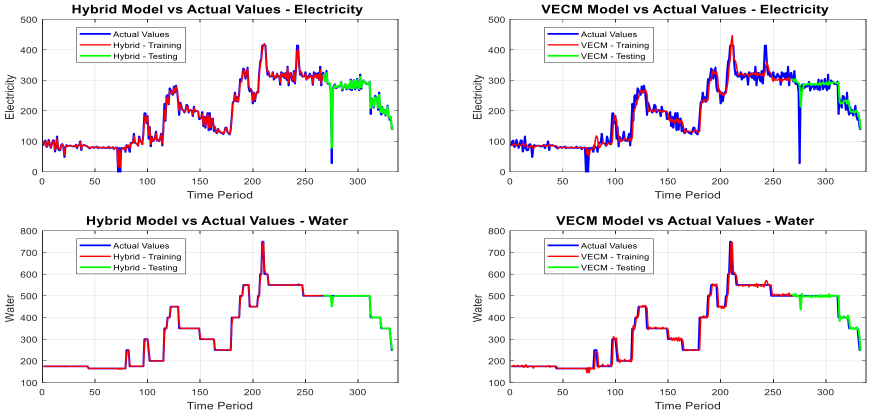

Figure 5. Comparison chart of predictions for the traditional and hybrid models.

predictive accuracy of the two models is equal, while the alternative hypothesis states that the predictive accuracy of the two models is unequal, with one being superior. The following table (12) shows the results of the DM test comparing the VECM and VECM-ANN models.

Table 12. Results of the DM test comparing the VECM and VECM-ANN models.

Variable	MD statistical	p-value
E_t	20.37	0.0000
w_t	16.33	0.0000

The DM test results shown in the table above indicate that the p-values for both variables are less than the 0.01 significance level. This supports the alternative hypothesis, suggesting that the prediction accuracy of the VECM-ANN hybrid model is significantly better than that of the traditional model VECM.

11.6. Sensitivity Analysis of ANN Architectures

Table (13) below shows the results of the RMSE test for four hybrid neural network architectures applied to the test data to demonstrate the superiority of the proposed hybrid model.

Table 13. Sensitivity Analysis Results for Different Artificial Neural Network Architectures

Architectures		RMSE	
E_t	w_t	E_t	w_t
(15,10)	(12,8)	12.7206	33.2955
(20,15)	(18,12)	24.9095	33.5708
(10,5)	(8,5)	46.1667	40.7764
(25,20)	(20,15)	19.4091	56.8226

As shown in table (12) above, the proposed hybrid model (15,10) for electricity and (12,8) for water is the best due to its lowest RMSE value compared to the other architectures.

12. Conclusion

In the context of the in-depth research conducted in this study we can come to the following conclusions:

1. This study proposed a dual hybrid approach combining a vector error correction model (VECM) with multilayer neural networks (MLPs) to simulate the correlation between water release and electricity production in a hydroelectric power system. The results showed that combining the structural equilibrium representation of the VECM model with the nonlinear approximation capabilities of neural networks significantly improves prediction compared to the conventional econometric model. The hybrid model outperformed both the long-term equilibrium relationship and the short-term nonlinear adjustments, reflecting the operational reality of hydroelectric power plants where water availability is constrained by production, while production decisions simultaneously influence water release strategies.
2. The estimated adjustment parameters revealed an asymmetric adaptation: electricity production acts as a rapidly adjusting variable that corrects deviations from equilibrium, while water releases act as a slowly adjusting operational control variable. This interpretation suggests that operational decisions at the station level respond quickly to maintain a balance in electricity demand, while hydrological policies remain constrained by safety, storage targets, and downstream commitments. Therefore, the equilibrium factor obtained in this study is interpreted as an operational transformation relationship, not simply a physical efficiency factor for the turbines.

13. Limitations and Future Research Directions

It is crucial to understand the primary limitation of the VECM-ANN hybrid model, which was constructed, trained, and validated using actual data from the Mosul Dam (hydroelectric power plant data), despite the fact that it performed remarkably well in this investigation. As a result, it is unclear if it can be applied to other power plants with different operating traits or climates. The model's low parameters (RMSE, MAE, sMAPE, and MASE) amply illustrate its great accuracy and good connection with the patterns and relationships present in the training data. External validation should be given top priority in future studies by using the hybrid model on separate datasets from various regions. For assessing the hybrid model across a broader variety of predictive properties, like demand projections and climatic indicators, this would be essential. The model's applicability as a general planning tool for integrated water-energy nexus management would be strengthened by this assessment, which would also increase the model's adaptability to new situations.

Reproducibility Statement

MATLAB R2017b was used to perform statistical analyses of the study data using the packages available in the program for both the VECM model and the artificial neural network. Figure (2) shows the complete pseudocode for the hybridization algorithm, and Table (9) shows the configurational architecture of the hybrid model.

REFERENCES

1. R. Aakash, *Testing for Time Series Stationarity: A Practical Guide to the Dickey-Fuller and Augmented Dickey-Fuller Tests*, SSRN, 2025.
2. R. T. Ahmed, and O. S. Salim, *Modeling and analyzing banana prices in the city of Mosul using the ARFIMA model Predictive Market Study*, Iraqi Journal of Statistical Sciences, vol. 21, no. 2, pp. 92–108, 2024.
3. H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*, In Second International Symposium on Information Theory, 1973.
4. H. Akaike, *A new Look at the Bayes Procedure*, Biometrika, vol. 65, no. 1, pp. 53–59, 1978.
5. I. M. Alwan, *Using time series models to study stock price fluctuations in the Iraq Stock Exchange for the period 2017-2018*, Department of Statistics, University of Baghdad, 2017.
6. Z. Assem, and H. A. Hayawi, *Identification State Space models and some Time Series models*, Iraqi Journal of Statistical Sciences, vol. 18, no. 1, pp. 30–37, 2021.
7. H. Chen, Q. Wan, and Y. Wang, *Refined Diebold-Mariano Test Methods for the Evaluation of Wind Power Forecasting Models*, Energies, vol. 7, pp. 4185–4198, 2014.
8. M. Eid, A. Alhussan, and E. M. El-kenawy, *Potato Consumption Forecasting Based on a Hybrid Stacked Deep Learning Model*, Potato Research (European Potato Journal), 2024.

9. R. Engle, and C. Granger, *Cointegration and Error Correction: Representation, Estimation and Testing*, Econometrica, vol. 55, pp. 251–276, 1987.
10. N. S. Fadhil, and Z. M. Albazzaz, *Comparing SVR and Random Forest Forecasting based on Autoregressive Time Series with Application*, Iraqi Journal of Statistical Sciences, vol. 20, no. 2, pp. 155–165, 2023.
11. A. Fahim, A. M. Osman, Z. Tarek, and A. M. Elshewey, *Electricity Bill Prediction Based on a Particle Swarm Optimized Multilayer Perceptron Model*, Engineering, Technology & Applied Science Research, vol. 15, no. 6, pp. 29515–29522, 2025.
12. K. F. Hamad, B. Celik, and R. M. Ahmed, *Classification of Circular Mass of Breast Cancer Using Artificial Neural Network vs. Discriminant Analysis in Medical Image Processing*, Iraqi Journal of Statistical Sciences, vol. 21, no. 1, pp. 46–58, 2024.
13. E. Hannan, and B. Quinn, *The Determination of the Order of an Autoregression*, Journal of the Royal Statistical Society Series B, vol. 41, no. 2, pp. 190–195, 1979.
14. H. A. A. Hayawi, N. S. Ibrahim, and L. J. Mohammed, *Using the fuzzy technique to identification stochastic linear dynamic systems*, Journal of Statistics and Management Systems, vol. 24, no. 1, pp. 167–181, 2021.
15. S. Johansen, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Clarendon Press, 1995.
16. T. Grindheim, *Time Series: Forecasting and Evaluation Methods With Concentration On Evaluation Methods for Density Forecasting*, Master's thesis, Department of Mathematics, University of Bergen, 2018.
17. S. H. Mahmood, and M. M. Ahmed, *Application of Elman Neural Network and SARIMA Model to Modeling Road Traffic Accident in the Kurdistan Region of Iraq*, Iraqi Journal of Statistical Sciences, vol. 20, no. 2, pp. 43–56, 2023.
18. U. Mustofa, F. Dhia, S. Yusuf, A. Faiz, and Widiarti, *Application of Vector Error Correction Model (VECM) and Impulse Response Function for Analysis Data Index of Farmers' Terms of Trade*, Indian Journal of Science and Technology, vol. 10, no. 19, 2017.
19. A. E. Obayelu, and A. S. Salau, *Agricultural response to prices and exchange rate in Nigeria: Application of co-integration and Vector Error Correction Model (VECM)*, Journal of Agricultural Sciences, vol. 1, no. 2, pp. 73–81, 2010.
20. P. C. Phillips, and P. Perron, *Testing for a Unit Root in Time Series Regression*, Biometrika, vol. 75, no. 2, pp. 335–346, 1988.
21. S. E. Said, and D. A. Dickey, *Testing for unit roots in autoregressive-moving average models of unknown order*, Biometrika, vol. 71, no. 3, pp. 599–607, 1984.
22. R. M. Saliwa, and D. R. Matar, *Comparison of prediction performance between some artificial neural networks and the Box-Jenkins methodology with application*, Iraqi Journal of Statistical Sciences, vol. 16, no. 28, pp. 51–67, 2019.
23. A. Sahed, and H. Kahoui, *Comparison between Neural Network Model of Radial Basis Function, multilayer perceptron and ARIMA Model to Predict Traffic Accidents in Tlemcen Region During the Period 2011-2019*, Journal of Strategy and Development, vol. 11, no. 4, pp. 295–311, 2021.
24. L. Wasserman, *Statistics/Data Mining Dictionary*, Pittsburgh, Pennsylvania, 2003.