Application of Elman Neural Network and SARIMA Model to Modeling Road Traffic Accident in the Kurdistan Region of Iraq

Saman Hussein Mahmood [iD] and Mryam Mohammed Ahmed [iD]

Department of Statistics and Informatics, College of Administration and Economics, Salahaddin University-Erbil, Iraq

Article information	Abstract
	Road traffic and deaths accidents are the most complex and dangerous system of all road systems that people deal with on a daily basis. In addition to the loss of life, there is also a lot of material damage to the society. Therefore, we aim to study this effective topic through the Box-Jenkins model and the Elman Neural Network which are very appropriate to choose the best and most appropriate model for the number of accidents and the number of deaths from traffic accidents in the Iraqi Kurdistan Region according to monthly during the years (2014-2021). Finally, we compared the results between both models. It was concluded that the results of the Elman model (1:2,5) are better than the SARIMA $(1,1,1)(0,1,1)_{12}$ model for the number of traffic accidents, and the results of the Elman(1:2,3) model are better than the SARIMA $(0,1,1)(1,1,2)_{12}$ model for the number of deaths from traffic accidents based on statistical measures (RMSE, MAE, MAPE), which we used it for comparison. Statistical analysis is performed using the software (Statgraphics V.19) and the program (Matlab V.18a)

Introduction

A time series data is a collection of observations made sequentially in time, or ''A sequence of observations ordered by a time parameter". When studying a phenomenon, we often encounter a dataset where observations are taken according to the order of time. These observations are called time series. [1][2][3].

The Box-Jenkins model is a technique designed based on input from a specific time series data to analyze and forecast data. It forecasts data using three principles, autoregressive, differencing, and moving averages. Each principle is used in the Box-Jenkins analysis and together they are collectively shown as an autoregressive integrated moving average, or ARIMA (p, d, q) respectively. In the time domain approach, we use time functions like the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to describe the characteristics of a time series process whose evolution is represented through various time-lag relationships. A seasonal phenomenon that recurs in many time series after a regular period of time is called the seasonal period [4].

In recent years, interest in studying time series using Artificial Neural Networks (ANN) has emerged and has been successfully applied in forecasting in different knowledge fields such as biology, finance and economics, energy consumption, medicine, etc. This offers several potential advantages with respect to alternative methods mainly ARIMA models-when it comes to dealing with problems concerning nonlinear data which normal distribution is not required. The first advantage of ANN is that it is highly diverse and

does not require formal specification of the model or the fulfillment of a certain probability distribution for the data [5].

The Elman Neural Network (ENN) was originally developed by Elman (1990) based on the Jordan network. The structure of ENN, it was a modified ANN based on the Back propagation neural network. However, unlike the BP neural network, the context layer is used to store the previous information of the hidden layer and feedback it to the next moment of the hidden layer. In this way, the context layer improves the sensitivity to historical data and makes ENN have a dynamic memory function [6].

Material and methods
Time Series:
Time series is a time-ordered sequence of observation values of a physical or financial variable made at equally spaced time. Being based on measured values and usually corrupted by noise, time series values generally contain a deterministic signal component and a stochastic component representing the noise interference that causes statistical fluctuations around the deterministic values.

The aims of time series are to analyze, describe and summarize time series data, and make forecasts. One simple method of describing a series is that of classical decomposition. The time series can be decomposed into[7]:-

- **Secular trend (T):** long-term movements of a time series (stable, increasing, or decreasing).
- **Seasonal Variations (S):** Seasonal fluctuations related to the calendar;
- **Cyclical Variations (C):** other cyclical fluctuations (such as an economic cycle or business cycle).
- **Irregular or Random Variations (I):** other random fluctuations

Types of Time Series:
Stationary Time Series:
A process is said to be stationary if the basic behavior of does not change over time. For such a process, $\mu(t)$ would not depend on time and thus could be denoted μ for all t , It is of two types [8].

A. Strictly Stationary Time Series*:*
Strict stationary involves that the mean and variance are constant in time and that the auto-covariance Cov (Z_t, Z_s) only depends on lag k = |t − s| and can be written $\gamma\_{((k))}$ [9] .

B. Weakly Stationary Time Series*:*
A time series $\{Z_t\}$ is weakly stationary if both the mean of $\{Z_t\}$ and the covariance between $\{Z_t\}$ and $\{Z_{t-k}\}$ are time-invariant, where (k) is an arbitrary integer. More specifically, $\{Z_t\}$ is weakly stationary if: $E(Z_t) = \mu$, which is a constant, and $Cov(Z_t, Z_{t-k}) = \gamma_k$ which only depends on k.

Non-Stationary Time Series:
Some time series exhibit non-stationary, which occurs in linear or nonlinear systems, as fluctuations in the system's representation across time. When there is a smoothly shifting trend component with shifts in the mean as well as fluctuations in the variance of the process, non-stationaries occur, there are two types of non-stationary [10]:

 A. Non -Stationary around mean: when the mean of a series is not constant
 B. Non Stationary around Variance: Many time series non-stationary because of their time-dependent variances and autocovariance, variance that changes through time.

Stationary Time Series Models:
Autoregressive Model (AR):
Suppose that $\{a_t\}$ is a purely random process with mean zero and variance σ_a^2 . Then a process $\{Z_t\}$ is said to be an autoregressive process of order p (abbreviated to an AR(p) process) if;[11]

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t \qquad (1)$$

This is rather like a multiple regression model, but $(Z_t)$is regressed on past values of (Z_t) rather than on separate predictor variables. This explains the prefix `auto'..First-order process, where p = 1. Then;

$$Z_t = \emptyset_1 Z_{t-1} + a_t \qquad (2)$$

Moving Average Model (MA):

A moving average (MA) process of order q is a linear combination of the current white noise term and the q most recent past white noise terms and is defined by:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \tag{3}$$

Where $\{a_t\}$ is white noise with zero mean and variance σ_a^2 .

Where θ_q is a polynomial of order q. Because MA processes consist of a finite sum of stationary white noise terms, they are stationary and hence have a time-invariant mean and auto covariance [9].

Autoregressive Moving Average (ARMA) Models:
The ARMA models is produced when MA and AR procedures are combined, order is a mixed autoregressive/moving-average process (p, q) (Chatfield, 2003). It is supplied by

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots \emptyset_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - - \theta_q a_{t-q} \tag{4}$$

Real time series data often exhibit time trend (such as slowly increasing) features that are beyond the capacity of stationary ARMA models.to remove those unstable components. Taking the difference (more than once if necessary) is a convenient and effective way to detrend and deseasonalize. We call it an autoregressive integrated moving average (ARIMA) models [12].

Seasonal Autoregressive Integrated Moving Average (SARIMA) Model:
SARIMA model is the product of seasonal and non-seasonal polynomials and is designated by SARIMA (p, d, q) x$(P, D, Q)_s$, where (p, d, q) and (P, D, Q) are non-seasonal and seasonal components, respectively with a seasonality's'. SARIMA model was defined at Equation: -

$$\Phi_P(B^s) \, \emptyset_p (B)(1\text{-}B^s)^D \, (1\text{-} B)^d \, Z_t = \Theta_Q(B^s) \, \theta_q(B) a_t \tag{5}$$

Where: (Φ and $\emptyset$) are autoregressive (AR) parameters of seasonal and non-seasonal components, correspondingly; (Θ and θ) are moving average (MA) parameters of seasonal and non-seasonal components, respectively; B = backward operator, B $(Z_t) = Z_{t-1}$; $(1\text{-}B^s)^D = D^{th}$ seasonal modification of season (s); $(1\text{-}B)^d = d^{th}$ non-seasonal difference; a_t = an individually distributed random variable; (P and p) are the orders of the AR components; (Q and q) are the orders of MA components; (D and d) are difference terms[13] .

The Stages of ARIMA Modeling:
The four following are the main steps Box–Jenkins forecasting model[11]:
1. **Model Identification**: Examine the data to select the most suitable model.
2. **Estimation:** Estimate the unknown coefficients of the model you've chosen.
3. **Diagnostic Checking**: Check the residuals from the fitted model to verify if they are sufficient.
4. **Consideration of Alternative Models if Necessary:** Alternative models may be explored until a good model is identified if the initial model appears to be unsatisfactory for some reason.
5. **Forecasting:** When such a model is discovered, calculating forecasts as conditional expectancies is usually quite simple

Artificial Neural Network (ANN):
Artificial neural networks are processing devices (algorithms) that are loosely modeled after the neuronal structure of the mammalian cerebral cortex but on much smaller scales. Computer scientists have always been inspired by the human brain. In 1943, Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician, developed the first conceptual model of an ANN. They describe the concept of a neuron, a single cell living in a network of cells that receives inputs, processes those inputs, and generates an output.[14]
ANN is a data modeling tool that depends upon various parameters and learning methods. Neural networks are typically organized in layers. Layers are made up of a number of interconnected "neurons/nodes," which contain "activation functions." ANN processes information through neurons/nodes in a parallel manner to solve specific problems. ANN acquires knowledge through learning, and this knowledge is stored within interneuron connections' strength, which is expressed by numerical values called "weights." These weights are used to compute output signal values for a new testing input signal value. Patterns are presented to the network via the "input layer," which communicates to one or more "hidden layers," where

the actual processing is done via a system of weighted "connections." The hidden layers then link to an "output layer," where the answer is output, as shown in Figure (1). [15].
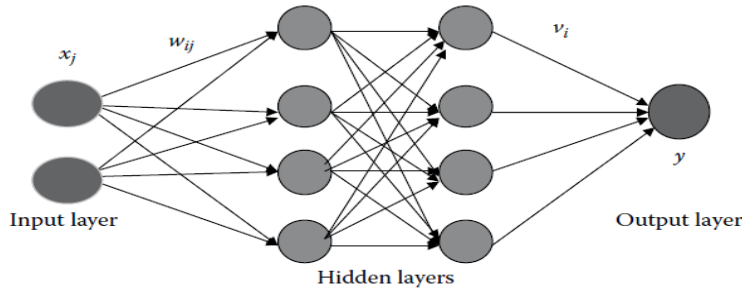

Figure 1 Structure of Artificial Neural Network

Time Series with Artificial Neural Networks [16]:-
Recently, artificial neural networks have been used in most time series forecasting applications. The following factors lead the widespread usage of artificial neural networks method in time series:
1. Analysis can be done without testing whether time series have curvilinear or linear structure.
2. When compared with conventional time series method, artificial neural network method provides much better forecasting results.
3. As conventional time series models can only be used for particular curvilinear structures, they are not flexible in general. But in the analysis using artificial neural network, no matter what the curvilinear structure of time series is.
4. Artificial neural network theory is not complex as in the conventional time series forecasting methods and it is easy to understand

Types of Neural Network Architecture:
Feed-Forward Neural Network (FFNN):
In a Feed-Forward Neural Network, neurons are organized in the form of layers. Neurons in a layer receive input from the previous layer and feed their output to the next layer. The data go from the input node to the output node in a strictly feed-forward way. There is no feedback (back loops); that is, the output of any layer does not affect the same layer [15].

Feedback Neural Network (FBN):
In a feedback neural network, the output of one layer routes back to the previous layer. This network can have signals traveling in both directions by the introduction of loops in the network. This network is very powerful and, at times, gets extremely complicated. All possible connections between neurons are allowed. Feedback neural networks are used in optimization problems, where the network looks for the best arrangement of interconnected factors. They are dynamic and their state changes continuously until they reach an equilibrium point [15].

Elman Neural Network (ENN):
Elman neural network (ENN) is a subclass of neural networks (NNs), which are constituted by a large number of neuron cell models according by a certain rules. Which feeds the output at previous moment back to the hidden–layer current input with the input data, and a network with internal delay feedback proposed by Elman in 1990, is a dynamic recurrent NNs with feedforward connections. The ENN is trained a supervised manner using a popular back-propagation algorithm, based on the inputs and targets given to the network. Meanwhile, ENN can model nonlinear dynamical systems and learn time-varying patterns, thus it has excellent ability to solve discrete time series problems. The context layer in an ENN as a self-referencing layer makes it a type of recurrent network [17].

The Structure of Elman Neural Network:
As shown in Figure (2), Elman NN consists of layers of cells (input layer, hidden layer and output layer), where each layer in the network is associated with the next layers by similar to the traditional multi-layer neural network as a feed-forward network.

There are another layer in ENN named the context layer, the inputs of this layer come from outputs of the hidden layer, the context layer is used to store the hidden layer' s output values of the previous time, which can be used in the current time step.
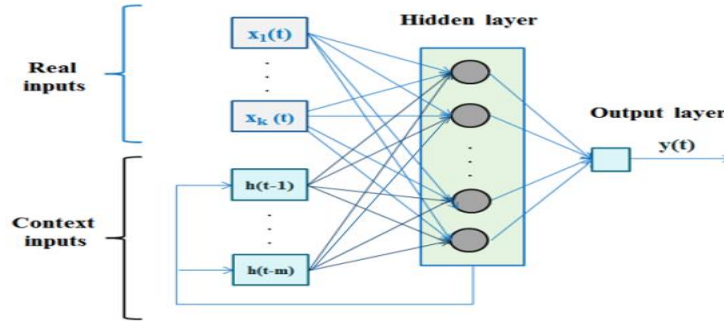

Figure 2 Elman neural networks

Because the network can store information for future reference, it is able to learn temporal patterns as well as spatial patterns. The Elman network can be trained to respond to, and to generate, both kinds of patterns [17].

2.8.2 Training an Elman NN:
Elman networks can be train following occurs at each epoch [18]:
The external input, context and output weight matrix are represented as $W^{h,i}(t)$, $W^{h,c}(t)$, $W^{o,h}(t)$ respectively. Consider the ENN structure as Figure (3), it contains a n-dimensional external input vector $x^1(t) = [x_1^1(t), x_2^1(t), \dots , x_n^1(t)]^T$ and a n-dimensional output vector $y(t) = [y_1(t), y_2(t) \dots y_n(t)]^T$,

The number of hidden neurons is m, and therefore $W^{h,i}(t) \in R^{m \times n}$, $W^{h,c}(t) \in R^{m \times m}$, and $W^{o,h}(t) \in R^{n \times m}$. The output vector of the hidden layer, $c(t-1) = [c_1(t-1), c_2(t-1), \dots , c_m(t-1)]^T$, is connected back to the hidden layer as another input vector, so $x^2(t) = [x_1^2(t), x_2^2(t), \dots , x_m^2(t)]^T = c(t-1)$ and the complete input vector is defined as:

$$x(t) = [x_1^1(t), x_2^1(t), \dots , x_n^1(t), x_{n+1}^2(t), \dots , x_k^2(t)]^T \quad = [[x^2(t)]^T [x^1(t)]^T]^T \qquad (6)$$
$$= [x_1^1(t), x_2^1(t), \dots , x_n^1(t), c_1^2(t-1), \dots, c_m^2(t-1)]^T$$

where, $k = m + n$.

The activation function of the output layer takes the sigmoid function The output vector can be computed by equations:

$$y_i(t) = f(a_i^o(t)) = \frac{1}{1+\exp(-a_i^o(t))}, i = 1, 2, \dots, n \qquad (7)$$

$$a_i^o(t) = \sum_{j=1}^m W_{ji}^{o,h}(t) * h_j(t), i = 1, 2, \dots, n \qquad (8)$$

For relationships among the input layer, the context layer and the hidden layer, define complete input weight matrix as:

$$W^h(t) = [W^{h,i}(t), W^{h,c}(t)] \in R^{m \times k} \qquad (9)$$

So the output of the complete input vector $x(t)$, the activation function of the hidden layer takes the sigmoid function.

$$h_j(t) = f(a_j^h(t)) = \frac{1}{1+\exp(-a_j^{o,h}(t))}, j = 1, 2, \dots, m \qquad (10)$$

$$a_j^{o,h}(t) = \sum_{i=1}^k W_{jl}^h(t) * x_l(t), j = 1, 2, \dots, m \qquad (11)$$

The target of ENN training algorithm is to minimize the mean-square error:

$$MSE(t) = \frac{\| e(t) \|^2}{2} \qquad (12)$$

$$e(t) = d(t) - y(t) \qquad (13)$$

Here, $d(t)$ is the desired outputs.

A standard EBP training algorithm can reduce E(t) by estimating the weight as follows:

$$W^{o,h}(t+1) = W^{o,h}(t) - \mu \frac{\partial E(t)}{\partial W^{o,h}(t)} = W^{o,h}(t) + \mu y'(t)e(t)h^T(t) \tag{14}$$

$$W^h(t+1) = W^h(t) - \mu \frac{\partial E(t)}{\partial W^h(t)} = W^h(t) + \mu h'(t)[W^{o,h}(t)]^T y'(t)e(t)x^T(t) \tag{15}$$

Here, μ is the learning rate of the EBP, and

$$y'(t) = diag[\, f'\left(a_1^o(t)\right) f'\left(a_2^o(t) \dots f'\left(a_n^o(t)\right)\right] \in R^{n \times n} \tag{16}$$

$$h'(t) = diag[\, f'\left(a_1^h(t)\right) f'\left(a_2^h(t) \dots f'\left(a_m^h(t)\right)\right] \in R^{m \times m} \tag{17}$$
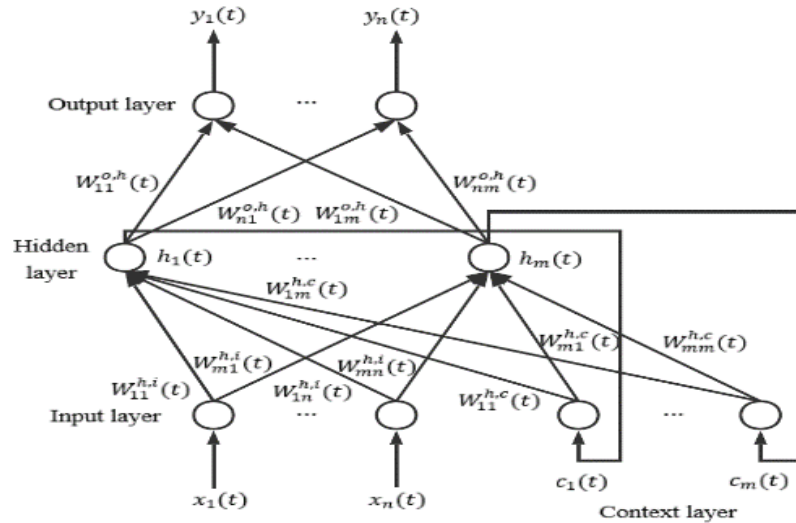

Figure 3 Training an Elman neural network with a single hidden layer

Results

We will practically apply both time series models and Elman neural networks models on road accidents data (number of accidents and the number of deaths from traffic accidents).The data starting from January 2014 to December 2021 and 96 months. The data were obtained from the General Directorate of Traffic in Erbil.

The First Series (The Number of Traffic Accidents):

The Box Jenkins methodology was applied to analyze the data. First, we analyzed the data by drawing the general trend of two series (Figure 4), and it is clear that there are oscillations, fluctuations and it is observed that these variations occur at a different rate, but on a regular basis. This pattern of variation indicating a seasonal component repeats itself every 12 months.
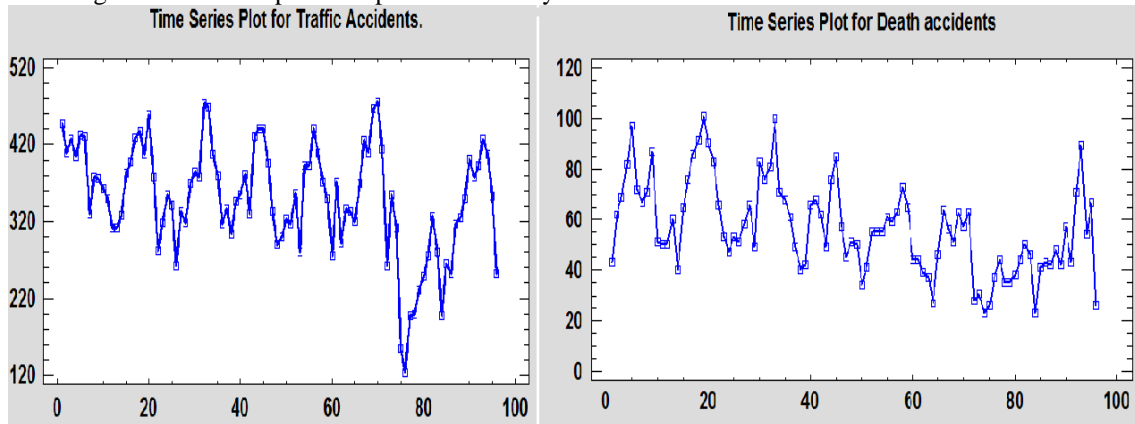

Figure 4: Time Series Plot of the Original Data for the Traffic Accidents in Iraqi Kurdistan

Figure (5), ACF and partial ACF prove that the data is non-stationary, and it can be seen that the data contains seasonal behavior, can be confirmed through the test (Box-Pierce) value is equal (133.5) and (P-value = 0.0) which rejects the null hypothesis for first series. The Box-Pierce test for (number of accident deaths) rejects the null hypothesis, according to (P-value = 0.0) for the test Box-Pierce (Q statistic = 230.054).

H0: The series is stationary.   H1: The series is not stationary.
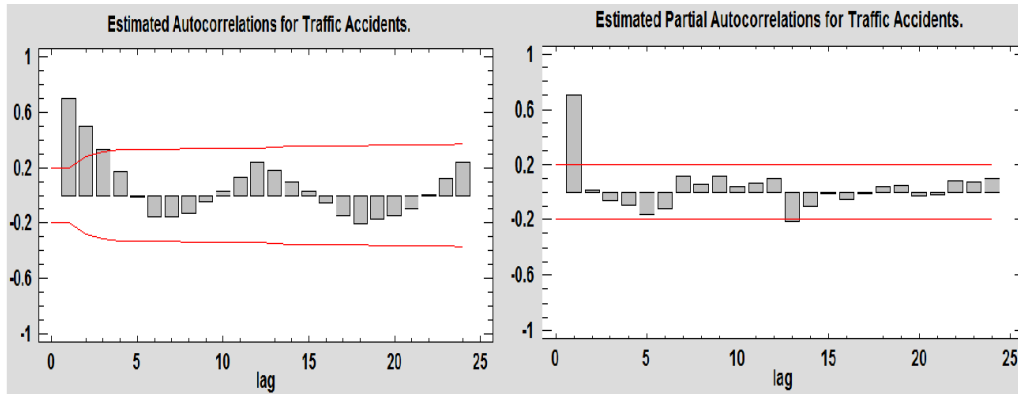

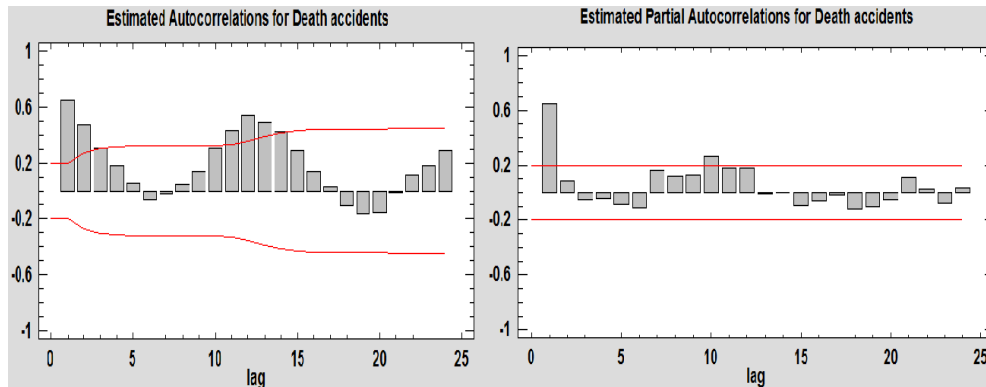Figure 5: ACF and partial ACF of the Original Data for the traffic Accidents Series

Figure 6: ACF and partial ACF of the Original Data for the accident deaths Series

Figure 7, ACF and PACF for the traffic accident series Figure 7, ACF and PACF for the traffic accident series, after natural log transformation, with first difference for non-seasonal and seasonal of degree (12) for series which transformed into stationary series. The value of the Box-Pierce statistic was(Q=32.3182 , p-value = 0.119295) confirming the stationary and the randomness of the series
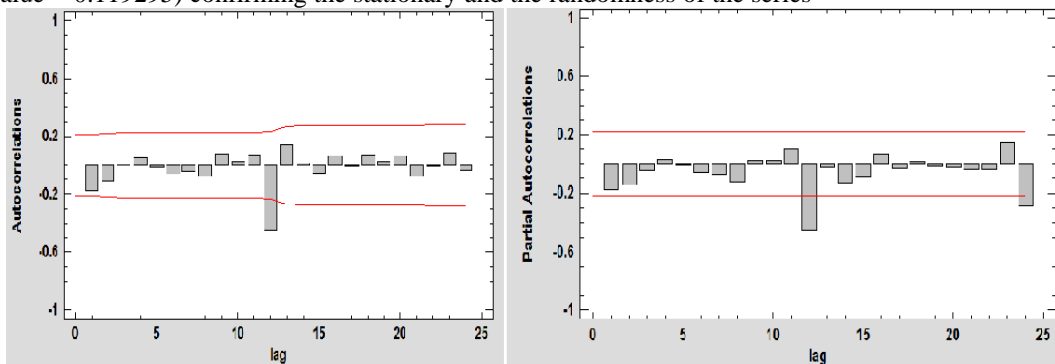

Figure 7: ACF and PACF for the traffic accident series adjusted

49

Figure 8, ACF and PACF for accident deaths series, the ideal method is to remove a non-stationary about mean (trend) taking the first difference for non-seasonal, also for the purpose of eliminating the effect of seasonality, the differences of degree (12) were taken. As presented in the value of the Box-Pierce statistic was (Q=34.3257, p-value = 0.0790443) confirming the stationary of the series.

Figure 8: ACF and PACF for the accident deaths series adjusted

Choosing Fitting Model:

Constructing an appropriate model for the modified series, depending on the ACF, PACF, and the significance of the regression coefficients. We apply the three statistical measurements (RMSE, MAE and MAPE) with AIC (Akaike Information Criteria) and SBIC(Schwarz Bayesian Information Criterion) to choose the best model,

Table (1) shows different models of SARIMA for the traffic accident series, the best and adequate model is SARIMA$(1,1,1)(0,1,1)_{12}$ having the smallest values of measurements and (AIC , SBIC) compared with the others. Random residuals for this model were confirmed by ACF and PACF with Box-Pierce (P-value = 0.772). The estimated parameters of the specified model is presented in table (2) which shows that all estimated parameters are statistically significant

Table (1) Comparison of the proposed models for the number of monthly traffic accidents

Model	RMSE	MAE	MAPE	AIC	SBIC	Sig.Coefficients
SARIMA(1,1,0)(1,1,2)12	48.30	33.40	11.22	7.84	7.95	No
SARIMA(1,1,1)(0,1,1)12	46.45	32.78	11.17	7.73	7.77	Yes
SARIMA(1,1,1)(1,1,1)12	46.86	33.04	11.28	7.78	7.88	No
SARIMA(1,1,1)(2,1,1)12	47.06	33.02	11.30	7.81	7.94	No
SARIMA(1,1,2)(1,1,1)12	47.47	33.49	11.23	7.82	7.96	No
SARIMA(0,1,1)(1,1,2)12	47.73	32.94	11.16	7.81	7.92	No

Table 2 Estimated Parameters Values of the SARIMA $(1,1,1)(0,1,1)_{12}$ Model

Parameter	Estimate	Standard Error	T	P-value
AR(1)	0.708	0.111	6.359	0.000
MA(1)	0.924	0.051	18.099	0.000
SMA(1)	0.853	0.048	17.647	0.000

Table (3) shows different models of SARIMA for the number of accidents death series which is SARIMA$(0,1,1)(1,1,2)_{12}$ having the smallest values of measurements and (AIC , SBIC) compared with the others. Random residuals for this model were confirmed by ACF and PACF with Box-Pierce (P-value = 0.396). Table (4) shows the parameters estimation of the specified model and all the parameters for the non-seasonal and seasonal are statistically significant.

Table 3 Comparison of the proposed models for the road accident deaths

Model	RMSE	MAE	MAPE	AIC	SBIC	Sig. Coefficients
SARIMA$(1,1,0)(1,1,2)_{12}$	11.31	9.02	17.70	4.94	5.04	Yes
SARIMA$(1,1,1)(0,1,2)_{12}$	11.33	8.89	18.09	4.94	5.05	No
SARIMA$(1,1,1)(1,1,1)_{12}$	11.54	9.19	19.11	4.96	5.08	No
SARIMA$(1,1,1)(2,1,2)_{12}$	10.61	8.50	17.28	4.85	5.01	Yes
SARIMA$(1,1,2)(1,1,1)_{12}$	11.40	8.91	18.06	4.97	5.11	No
SARIMA$(0,1,1)(1,1,2)_{12}$	10.53	8.42	16.45	4.79	4.89	Yes

Table 4 Estimated Parameters Values of the SARIMA $(0, 1, 1)(1, 1, 2)_{12}$ Model

Parameter	Estimate	Standard Error	T	P-value
MA(1)	0.618	0.087	7.087	0.000
SAR(1)	-0.841	0.129	-6.508	0.000
SMA(1)	-0.185	0.078	-2.382	0.020
SMA(2)	0.869	0.053	16.518	0.000

Elman Neural Network Application:
Applying Elman NNs model to study and analyze road accidents. The data used in this study is (96) observations and divided as follows (70%), which equals (68 observations) as the training set and (15%), which equals (14 observations) as the validation set, while (15%) from observations (14 observations) as a test set.
The entire input sequence is presented to the network, and its outputs are calculated and compared with the target sequence to generate an error sequence, as mentioned previously in (training an Elman NN)
The First Series (Traffic Accidents Data):
In the process of identifying the optimal ELMAN NN, we repeated the experiment several times for the number of traffic accidents, by changing the neurons number in the hidden layer from (1 to 15) in the trial, the importance of changing the number of neurons in the hidden layer is also related to the context layer because they are equal within the network. It was found that the ELMAN NN (1:2,5) model is better, The results in the table (5).

Table 5 A Comparisons of the Evaluation Indicators ENN for Traffic Accidents

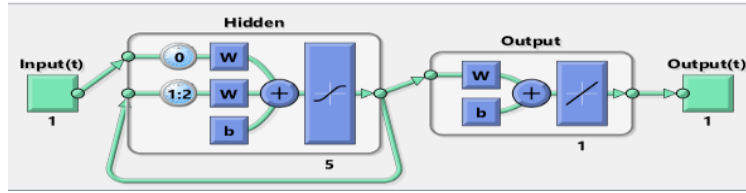| Elman net	MAE	RMSE	MAPE	R^2
(1:2,1)	9.47	12.81	3.13	96.79
(1:2,2)	91.51	124.23	28.65	14.15
(1:2,3)	11.69	14.81	3.77	95.75
(1:2,4)	53.01	66.92	17.89	61.86
(1:2,5)	**9.79**	**12.20**	**3.09**	**97.30**
(1:2,6)	15.07	19.06	4.56	95.67
(1:2,7)	52.23	67.49	18.15	10.96
(1:2,8)	53.94	68.82	18.67	6.43
(1:2,9)	50.79	65.19	17.33	28.73
(1:2,10)	72.64	91.05	23.39	2.58
(1:2,11)	25.20	34.67	8.01	76.48
(1:2,12)	17.66	24.40	5.69	89.04
(1:2,13)	11.45	17.09	3.40	94.15
(1:2,14)	67.41	86.56	22.46	4.62
(1:2,15)	63.09	88.24	21.57	2.36

Figure 9 The Configuration of the ELMAN NN (1:2,5) for the Number of Traffic Accidents

Elman NN Training Regression:
Figure (10) the regression plot is used between the targets (actual) data and the outputs (predict) data. The regression plot generally has four graphs showing for training, validation, test and combining all. It indicated our neural network structure is correct.
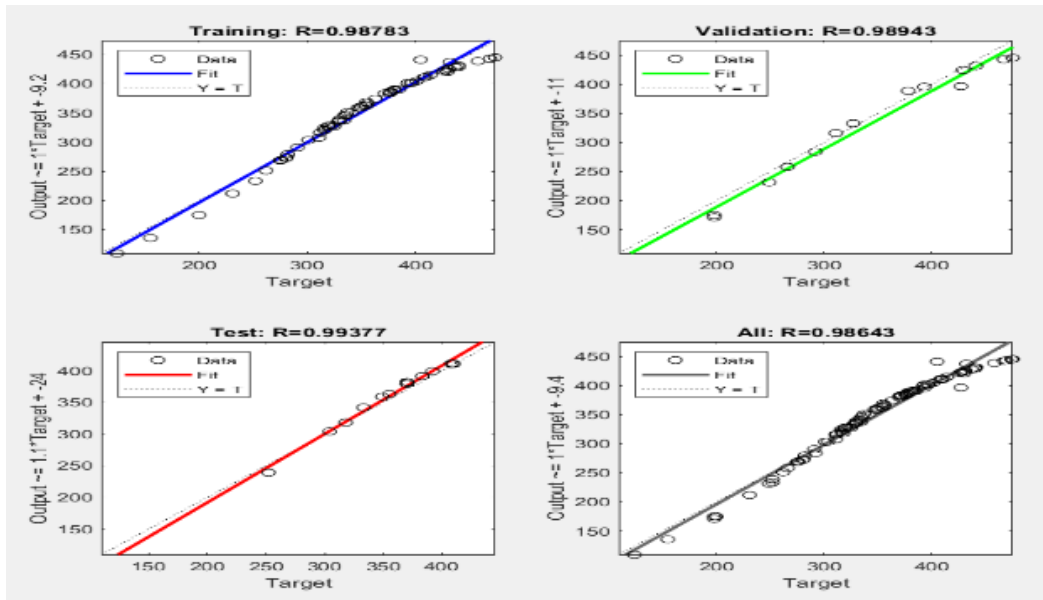

Figure 10 Regression Plots Displaying the Elman NN (1:2,5) for Traffic Accident data

Elman NN Training Time-Series Response:
Figure (11) shows the time-series response plot for confirmed cases using Elman NN. It also shows which time points are selected for training, testing and validation phases. The test result shows that Elman NN (1:2,5). It shows that the outputs were distributed evenly of the response curve and the errors were small in the training, testing, and validation subsets, indicating that the model reliably reflected the data.
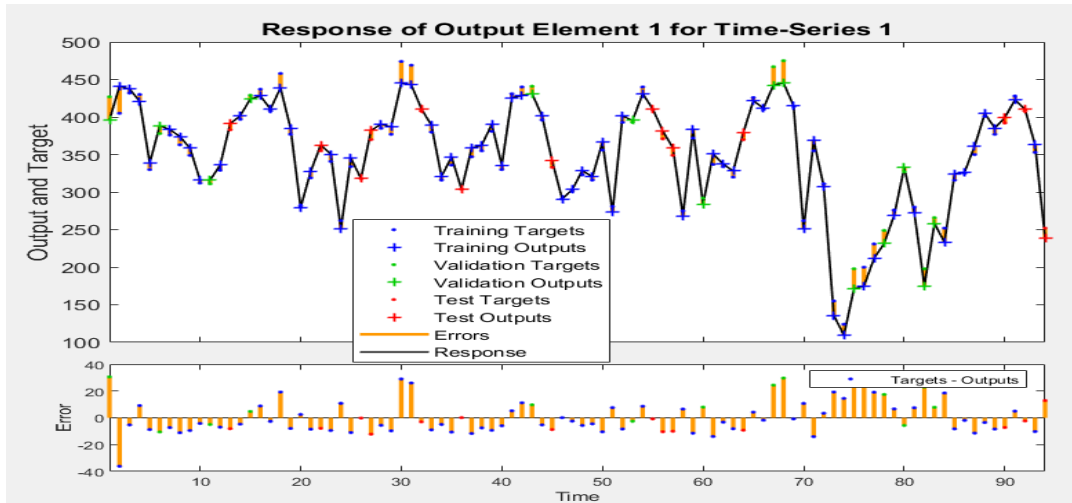
Figure 11 The Time-Series Response for the ELMAN NN (1:2,5) Model of Traffic Accident

The Second Series (Number of the traffic accident deaths):
Find the appropriate Elman model for the number of deaths by road accident, by changing the number of neurons in the hidden layer, the results in the table (6). We conclusion that the best model is Elman NN (1:2, 3).

Table 6 A Comparison of the Evaluation Indicators for Accident Deaths Occurring Monthly.

| Elman net	MAE	RMSE	MAPE	R^2
(1:2,1)	14.928	18.308	30.66%	8.24
(1:2,2)	1.8255	3.714	3.94%	95.98
(1:2,3)	**1.161**	**1.588**	**2.48%**	**99.24**
(1:2,4)	5.3493	6.947	10.19%	86.25
(1:2,5)	12.016	15.205	24.10%	41.86
(1:2,6)	3.3189	5.211	6.34%	91.95
(1:2,7)	3.1207	4.915	6.18%	92.89
(1:2,8)	10.201	12.46	20.81%	76.21
(1:2,9)	1.3129	1.6	2.76%	99.26
(1:2,10)	3.1121	4.842	6.24%	93.33
(1:2,11)	3.1958	4.516	6.08%	94.03
(1:2,12)	23.889	30.143	46.44%	58.91
(1:2,13)	13.281	17.225	24.91%	13.33
(1:2,14)	201.79	203.9	379.03%	38.40
(1:2,15)	2.3742	7.855	3.93%	83.81
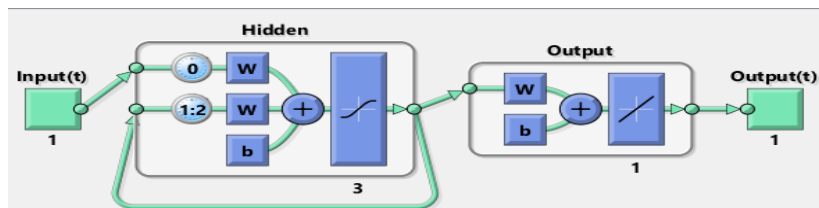
Figure 12 The Configuration of the ELMAN NN (1:2,3) Model

Elman NN Training Regression:
Figure(13) Regressions plots tell us about assess the quality of the Elman NN (1:2,3) model, displayed the plots that network outputs with respect to targets for training, validation and test sets. The quality of fit is

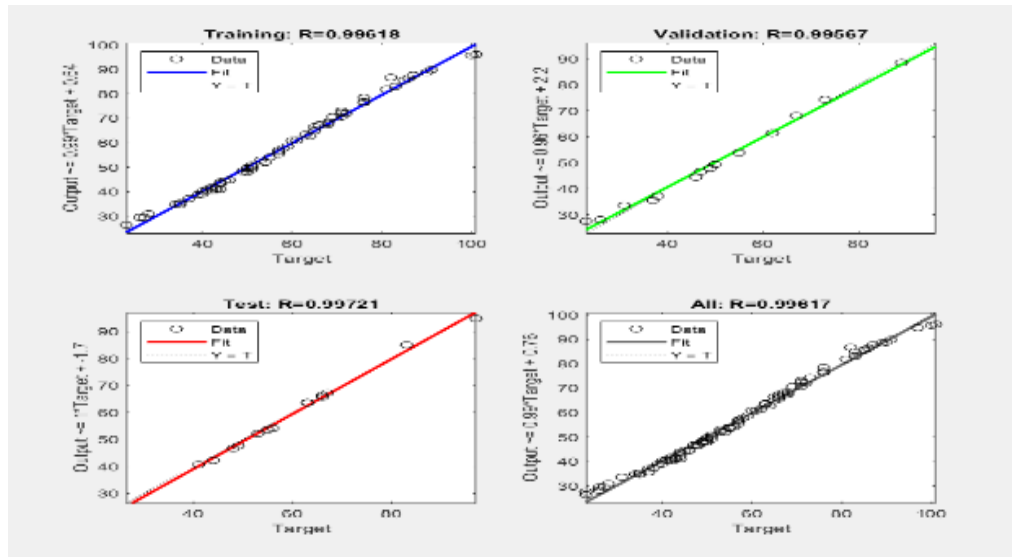reasonably good for data set, with R (correlation) values equal to 0.99618, 0.99567, and 0.99721 and 0. 99817.

Figure 13 Regression Plots Displaying the Elman NN (1:2,3) Model for Accident Deaths

Elman NN Training Time-Series Response:

The time series response in figure (14), displayed input data with predict data showing that the outputs were distributed evenly of the response curve and the errors were small in the training, testing, and validation subsets, indicating that the model reliably reflected the data.
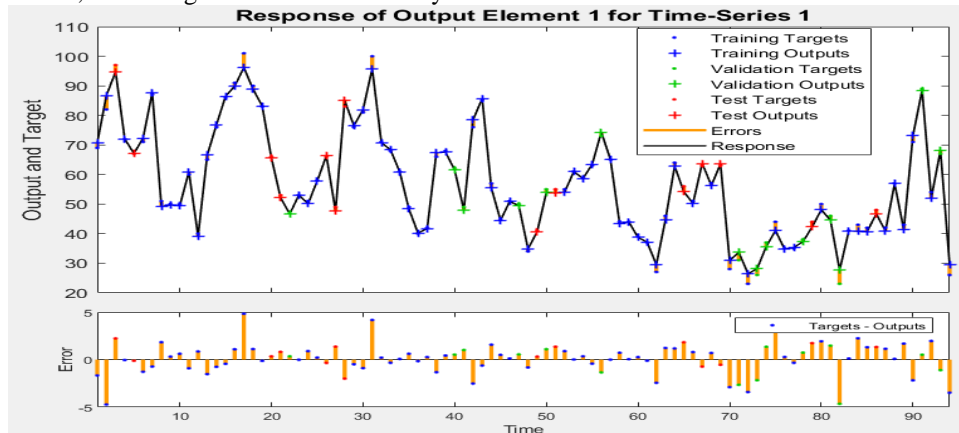

Figure 14 Time-Series Response Plot of the ELMAN NN (1:2,3) Model for Accident Deaths

Discussion

Models Comparison for the Traffic Accidents and the Number of Accident Deaths:-

Table (7) shows a comparison of the performance of two different traffic accident data models (SARIMA, Elman NN). The statistical indices (MAE, RMSE, and MAPE) used to calculate the prediction error have the lowest Elman NN value from the SARIMA model. Thus, it can be said that the Elman NN (1:2, 5) model is the best.

Table 7 Performance Comparison between Two Models (SARIMA and Elman NN) for the Number of Traffic Accidents

Models	MAE	RMSE	MAPE
SARIMA$(1,1,1)(0,1,1)_{12}$	32.7771	46.4507	11.1746
Elman NN (1:2,5)	**9.7935**	**12.20**	**3.09**

Table (8) shows that the superiority of Elman neural networks over the SARIMA model, as results were obtained with lower values of the statistical indicators (MAE, RMSE and MAPE) used to calculate the prediction error for accident deaths data. Thus, it can be said that the Elman NN (1:2,3) model is the best.

Table 8 Performance Comparison between Two Models (SARIMA and Elman NN) for Accident Deaths

Models	MAE	RMSE	MAPE
SARIMA$(0,1,1)$x$(1,1,2)_{12}$	8.41637	10.5262	16.4533
Elman NN (1:2,3)	**1.1612**	**1.588**	**2.48**

.

Conclusion

Based on the results obtained in this study, we concluded that there are seasonal patterns in the data.

1. After testing several models for time series we selected the two best models, SARIMA$(1,1,1)(0,1,1)_{12}$ for traffic accidents data and SARIMA$(0,1,1)(1,1,2)_{12}$ for the number of accident deaths, have the smallest value of (RMSE, MAE, MAPE) among all the other models.

2. In the process of selecting the best ELMAN NN It was found that the ELMAN NN (1:2,5) model is good for traffic accidents has small values for the statistical indictors, and the best model for the number of accident deaths is Elman NN (1:2, 3) has small values for the statistical indictors.

3. Comparing both models to obtain the best model, it was found that (Elman NN) is a more accurate model than the (SARIMA) model for the traffic accidents and has small values for the statistical indictors (MAE, RMSE, and MAPE).

References

1. TH, M. Using ARIMA model to forecasting with production of electrics in Australia. AL-Anbar University journal of Economic and Administration Sciences.2008; 1(2):273-282.
2. Yaffee, R.A. AND Mcgee, M. An introduction to time series analysis and forecasting: with applications of SAS and SPSS. Elsevier Science; 2000.
3. Wei, W.W. Time Series Analysis: Univariate and Multivariate Methods. 2nd Edition, Addison Wesley, New York; 2006.
4. Little, T.D. (ED). The Oxford handbook of quantitative methods in psychology. vol. 2. Oxford Library of Psychology; 2013.
5. Moreno, J.J.M. Artificial neural networks applied to forecasting time series. Psicothema. (2011); 23(2): 322-329.
6. Huang, R. and Tian, J. Wavelet-based Elman neural network with the modified differential evolution algorithm for forecasting foreign exchange rates. Journal of Systems Science and Information. 2021. 9(4): 421-439.
7. Brockwell, P.J. and Davis, R.A. Time series: theory and methods. Springer science & business media. 2009.
8. Woodward, W.A., Gray, H.L. and Elliott, A.C. Applied time series analysis with R. CRC press. 2017.
9. Cowpertwait, P.S. and Metcalfe, A.V. Introductory time series with R. Springer. 2009.
10. Kan, M.S., Tan, A.C. and Mathew, J. A review on prognostic techniques for non-stationary and non-linear rotating systems. Mechanical Systems and Signal Processing.2015; 62: 1-20.
11. Chatfield, C. The analysis of time series: an introduction. Chapman and hall/CRC.2003.
12. Fan, J. and Yao, Q. Nonlinear time series: nonparametric and parametric methods. New York: Springer. vol. 20; 2003.
13. Tadesse, K.B. and Dinka, M.O. Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa. Journal of water and land development. 2017; 35(1) :229-236.
14. Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics.1943; vol (5) :115–133
15. CHAKRAVERTY, S. AND MALL, S. Artificial neural networks for engineers and scientists: solving ordinary differential equations. CRC Press.2017.
16. Aladag, C.H. and E. Egrioglu, E. Advanced time series forecasting methods. Advances in time series forecasting. Bentham Science Publishers, Netherlands. ISBN. 2012 : ISBN-10: 1608053733.
17. Demuth, H.B. and Beale, M.H. MATLAB: neural network toolbox. MathWorks, Incorporated.1994.
18. Ren, G., Cao, Y., Wen, S., Huang, T. and Zeng, Z. A. Modified Elman neural network with a new learning rate scheme. Neurocomputing. 2018. Vol(286): pp.11-18.

تطبيق شبكة أيلمان العصبية ونموذج SARIMA لنمذجة حوادث الطرق في إقليم كردستان العراق

سامان حسين محمود و مريم محمد احمد

قسم الاحصاء والمعلوماتية، كلية الادارة والاقتصاد، جامعة صلاح الدين–اربيل، العراق

الخلاصة

تعد حوادث الطرق وحوادث الوفيات من أكثر أنظمة الطرق تعقيدًا وخطورة التي يتعامل معها الناس بشكل يومي. بالإضافة إلى الخسائر في الأرواح ، هناك أيضًا الكثير من الأضرار المادية للمجتمع. لذلك ، نهدف إلى دراسة هذا الموضوع الفعال من خلال نموذج Box–Jenkins وشبكة أيلمان العصبية والتي تعتبر مناسبة للغاية لاختيار النموذج الأفضل والأنسب لعدد الحوادث وعدد الوفيات من حوادث المرور الشهرية في إقليم كوردستان العراق خلال السنوات (2014–2021).

أخيراً، قمنا بمقارنة النتائج بين كلا النموذجين. تم الاستنتاج ان نتائج النموذج (1:2,5)Elman أفضل من نموذج $SARIMA(1,1,1)(0,1,1)_{12}$ لعدد حوادث المرور و نتائج النموذج (1:2,3)Elman أفضل من نموذج$SARIMA(0,1,1)(1,1,2)_{12}$ لعدد الوفيات من حوادث المرور استنادًا إلى المقاييس الإحصائية (RMSE, MAE, MAPE)، والتي تم إستخدامها للمقارنة. يتم إجراء التحليل الإحصائي باستخدام البرنامج (Statgraphics V.19) والبرنامج (Matlab V.18a)

الكلمات المفتاحية: الشبكة العصبية الاصطناعية ، شبكة أيلمان العصبية، السلاسل الزمنية، نماذج SARIMA ، الحوادث المرورية