Abstract

Objectives: This paper gives a brief survey analysis on time series forecasting models developed so far and how soft computing techniques are becoming popular for time series forecasting. Statistical Analysis: This paper gives an introductory study on time series models developed in literature so far. First all traditional models were developed among them ARIMA became popular but only for linear data. Then after Soft Computing models came into picture like ANN and it was proved to be best for nonlinear forecasting. But none of the individual models are capable to handle all types of datasets as real data is a mixture of linear and nonlinear both. So there arises a need for hybridization. Findings: Many traditional methods have been developed since last few decades for time series forecasting, however their performance is not up to the mark till today. Recent trends have proven that soft computing techniques like neural network, support vector machine are good alternatives to conventional methods. The accuracy of time series model is very important and difficult task to achieve, and individual models are not able to perform well for all types of time series data. So, hybridization of models is better to achieve good accurate forecasting results. Application/Improvements: Implementation of a hybrid model using soft computing techniques which can give better accurate results as real data are made linear and nonlinear both. This hybrid model can be used in applications like weather forecasting, exchange rate forecasting, etc.

Keywords: ARIMA, Artificial Neural Network, Hybrid Model, Soft Computing Techniques, Time Series, Time Series Forecasting

1. Introduction

Time series forecasting is a remarkable and difficult active analysis issue with growing attention to numerous areas. The inspiration of forecasting as from is to explore out potential useful relationships and discover connected applied math regularities for the observations of an underlying time series in order that helpful deciding is created prior to. Nonetheless, correct and reliable forecasting for the long run trend is ordinarily intense and backbreaking in true issues.

Forecasting future is best exploitation time series forecasting. With the time series prediction, past knowledge assortment of constant variable are analyzed to develop a model which might predict the longer term availableness of constant. Then established model is employed so as to extrapolate the time series into the future. This modeling approach is especially helpful once very little information is offered concerning past variables and no alternative things is known.

Our purpose in this paper is to do a comparative study of all the models of Time series forecasting and how soft computing is used for time series. The detail paper is organized as follows. Section 2 describes traditional model ARIMA of time series. Section 3 describes regarding soft computing approach models for time series like ANN and SVM. Section 4 tells regarding how hybrid models are used today to boost accuracy of your time series prediction then concludes this work by Section 5.
1.1 Time Series Forecasting

Time Series Definition “A time series is a gathering of past observations, measured as a rule over consecutive interim”\(^3\).

It is typically created of four components: Trend, Cyclical, Seasonal and Irregular elements. Statistic pattern are often describe by these four elements as shown in\(^3\). There are varied totally different approaches for statistic prediction and a number of the popular models are justify in next section.

2. Auto Regression Integrated Moving Average (ARIMA) Model

For statistic prediction initially of all ancient applied mathematics models were developed like moving average, exponential smoothing and autoregressive integrated moving average. Among all ARIMA became well-liked for linear forecasting\(^8\).

In this strategy, the given measurement data are introductory checked for stationarity and if are not then a differing operation is performed. On the off chance that the data are still non-stationary, differencing is afresh performed till the data are at last made stationary\(^9\). On the off chance that the differencing is performed d times, the mix request of the ARIMA philosophy is asserted to be d. The resultant data are sculptural as an Auto-Regressive Moving Average (ARMA) measurement as takes after\(^5,9\).

The information value at any given time \(t\), say \(y_t\), is taken into account as a function of the previous p information values, say \(y_{t-1}, y_{t-2}, \ldots, y_{t-p}\) and therefore the errors at times \(t, t-1, \ldots, t-q\) say \(n_t, n_{t-1}, \ldots, n_{t-q}\). The corresponding ARMA equation is shown in eq. 1.\(^9\) In eq. 1, \(a_1\) to \(a_p\) are the Autoregressive (AR) coefficients and \(b\) to \(b_q\) are the MA coefficients. So the statistic model is denoted as ARIMA (p, d, q) because it is a combination of AR and MA. The ARMA model accept that the blunder arrangement \(n_t\) is commotion and is Gaussian circulated, that the difference of this mistake is furthermore a model parameter\(^9\). The ARIMA displaying technique has 3 stages: 1. recognizing the model request, i.e., recognizing p and q; 2. assessing the model coefficients; and 3. forecast the data\(^5,6,8\).

The model coefficients are assessed utilizing the Box–Jenkins strategy. At long last the model coefficients are measurable, the future estimations of the measurement are anticipated utilizing the accessible past data values and hence the model coefficients. ARIMA models anticipate straight measurement data with amazing accuracy\(^2,5\).

\[ y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + n_t + b_1 n_{t-1} + \cdots + b_q n_{t-q} \] (1)

In an autoregressive integrated moving average model, the longer term value of a variable is thought to be a linear operate of numerous past observations and random errors\(^3,12\). ARIMA consolidates an assumption that the real information is regularly straight this suspicion is a reward and disadvantage each for ARIMA because it offers smart forecast result for linear statistic and it cannot deal with nonlinear data\(^5,7,12\).

3. Soft Computing Approach Models

Soft Computing is becoming extremely popular and it’s getting used in all the fields for obtaining sensible results. And as from literature survey it’s prove that ancient models don’t seem to be capable to handle nonlinear statistic knowledge properly. Thus soft computing and intelligence system are smart alternative for statistic forecasting\(^2,4,13,14\). These soft computing models are explained very well here during this section.

3.1 Artificial Neural Network (ANN) Model

The neural-system configuration is closeness to the neurons inside of the mind, therefore the name “artificial neural system.” And there is likewise two or a considerable measure of layers\(^2\). As a sample, a three-layer ANN has three layers, in particular an info layer, a concealed layer, and a yield layer. The inputs might be of any extent. Additionally, the measure of neurons inside of the shrouded layer is adaptable. A commonplace three-layer ANN, is appeared in Figure 1.To model measurement data misuse such a network, the arrangement \(y_i\) is considered as a nonlinear work of \(y_{i-1}, \ldots, y_{i-n}\). The

![Figure 1. Three layer architecture of ANN.](image)
relating comparison is appeared in eq. 2. In eq. 2, the work \(g\) could be a nonlinear work, and \(v\) could be a clamor or blunder term. The ANN model yield is portrayed as far as information and shrouded layer weight parameters. The exchange work of the shrouded layer is typically a sigmoid work, appeared in eq. 3, which of the yield layer is straight function. The ANN model yield is portrayed as far as information and shrouded layer weight parameters. The exchange work of the shrouded layer is typically a sigmoid work, appeared in eq. 3, which of the yield layer is straight function. The ANN model yield is portrayed as far as information and shrouded layer weight parameters. The exchange work of the shrouded layer is typically a sigmoid work, appeared in eq. 3, which of the yield layer is straight function.

\[
y_t = g(y_{t-1}, y_{t-2}, \ldots, y_{t-n}) + v_t
\]  
(2)

\[
Sigmoid(\alpha) = \frac{1}{1 + e^{-\alpha}}
\]  
(3)

The results of ANN are better for many time series forecasting compared to ARIMA as ANN can handle nonlinear data. And it is very flexible and universal approximator.

### 3.2 Support Vector Machine (SVM) Model

Second Soft Computing Approach is Support Vector Machine. After introduction of an alternative called loss function SVM has been used for time series forecasting also. The ability of SVM to resolve nonlinear estimation problem makes it capable to use for statistic forecasting.

The basic plan of SVM as shown in Figure 2 is mapping of data \(x\) into a high dimensional feature house by a nonlinear mapping space so do linear regression.

Consider a given training set of \(n\) data points \(G = \{x_i, d_i\}^n_{i=1}\) with input data \(x_i \in \mathbb{R}^p\), \(p\) is total number of data patterns and \(d_i \in \mathbb{R}\) is the output.

The SVM regression as shown in Figure 2 approximated by the following function eq. 4.

\[
f(x) = w \phi(x) + b, \quad \phi: \mathbb{R} \rightarrow F, w \in \mathbb{F}
\]  
(4)

Where \(b\) is a scalar threshold; \(\phi\) is that the high dimensional feature space that is nonlinearly mapped from the input space \(x\). Thus, the regression within the high-dimensional feature house responds to nonlinear regression in low dimension input house, that disregards the real number computation between \(w\) and \(\phi\) in the high-dimensional feature space. The coefficients \(w\) and \(b\) calculable by minimizing the regularized perform:

\[
R(C) = \frac{1}{2} \|w\|^2 + \frac{C}{n} \sum_{i=1}^{n} L_\varepsilon(d_i, y_i)
\]  
(5)

\[
L_\varepsilon(d_i, y_i) = \begin{cases} 
|d_i - y_i| - \varepsilon, & \text{if } |d_i - y_i| \geq \varepsilon \\
0, & \text{otherwise}
\end{cases}
\]  
(6)

To obtain the estimation of \(w\) and \(b\) eq. 5 is converted to the primal function given by eq. 7 by introducing the positive slack variable \(\xi\) and \(\xi^*\) as follows:

\[
\begin{align*}
\text{Minimize} & \quad R(w, \xi^*) = \frac{1}{2} \|w\|^2 + \frac{C}{n} \sum_{i=1}^{n} (\xi_i + \xi_i^*) \\
\text{subject to} & \quad w \phi(x_i) + b_i - y_i \leq \varepsilon + \xi_i \\
& \quad \xi_i, \xi_i^* \geq 0
\end{align*}
\]  
(7)

The term \(\frac{1}{2} \|w\|^2\) is the weights vector norm, \(d_i\) the desired output value and \(C\) is the referred as regularized constant. Finally, by introducing Lagrange multipliers as given and exploiting the optimality constraints, the choice operate given by equivalent weight of eq. 4 has the subsequent express form:

\[
f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x_i) + b
\]  
(8)

In eq. 8, \(\alpha_i\) and \(\alpha_i^*\) are the so-called Lagrange multipliers. They satisfy \(\alpha_i x_i = 0\) and \(\alpha_i \geq 0\) and \(\alpha_i^* \geq 0\) where \(i=1,2,\ldots,n\) and are obtained by maximizing the dual function of eq. 7, and the maximal dual function in eq. 7 which has the following form:

\[
\begin{align*}
\hat{\xi}(\alpha, \alpha^*) &= \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) - \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) - \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) (\alpha_i^* - \alpha_i) K(x_i, x_j)
\end{align*}
\]  
(9)

With constraints:

\[
\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0, \quad 0 \leq \alpha_i \leq C \quad i = 1,2,\ldots,n \\
0 \leq \alpha_i^* \leq C \quad i = 1,2,\ldots,n
\]

\(K(x_i, x_j)\) is the kernel function. The value of kernel function is equal to inner product of two vectors \(x_i\) and \(x_j\) in the feature space \(\phi(x_i)\) and \(\phi(x_j)\) i.e. \(K(x_i, x_j) = \phi(x_i) \phi(x_j)\). Generally Gaussian function is used as kernel function and its equation is as follows:

\[
K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right)
\]  
(10)

![Figure 2. Architecture of SVM.](image-url)
4. Hybrid Model

In this Section we will take about hybrid models.

4.1 ARIMA and ANN Hybrid Model

As from analysis of the literature it’s found that none of individual model is nice or has been accepted as commonplace for forecasting every type of knowledge. Although these models offer predictions with higher accuracy in some cases, there’s scope for additional improvement within the accuracy if the nature of the given statistic is taken under consideration before applying the models. In real world data cannot be only linear or nonlinear it’s continuously a mix of each. therefore it’s recommended to develop an acceptable hybrid model which might work on each linear and nonlinear statistic data.

In 2016 Zhang planned his 1st hybrid model using ARIMA and ANN that was a straightforward combination of each. Then in 2017 Khashei planned his model as an update to Zhang’s model that gave higher accuracy than Zhang’s model. Here he assumes that real information is that the addition of linear and nonlinear information. In 2018 once more Khashei planned his hybrid model of ARIMA and ANN that assumes that real information could be a function of linear and nonlinear each. In 2019 proposed a hybrid model of ARIMA and ANN with a new class of classifier that first separate linear and nonlinear information then ARIMA and ANN does its work. All the categories of hybrid ARIMA and ANN models were able to offer higher accuracy than individual and were conjointly able to solve some limitations of individual models too.

Figure 3 shows the working of hybrid model proposed by Khashei.

4.2 Adaptive Neuro-Fuzzy Inference System (ANFIS) Model

An Adaptive Neuro-Fuzzy Inference System (ANFIS) is created of Neural Network (NN) and Fuzzy Inference System (FIS). Here NN develops acceptable if-then rules and membership functions for FIS from the given input-output information pairs. The standard ANFIS structure given in Figure 4 has six layers including input (layer 0), membership (layer 1), rules (layer 2), normalization (layer 3), function (layer 4) and output (layer 5) layers.

Layer 0: x and y present within the layer 0 are the inputs and also the corresponding membership functions are $A_1, A_2$ and $B_1, B_2$ respectively.

Layer 1: The nodes of layer 1 work out on the membership values related to the inputs as given in equations.

Output $O_{1,i}$ for node $i = 1, 2$ (bell shape):

$$O_{1,i} = \mu_{A_i} = \frac{1}{1 + \left(\frac{x - c_i}{b_i}\right)^2}$$

Similarly Output $O_{1,i}$ for node $i = 3, 4$:

$$O_{1,i} = \mu_{B_{i-2}}(y)$$

Where $a_i$ and $b_i$ are the membership functions and $c_i$ is center of curve.

Layer 2: Every node in during layer multiplies the incoming signals from previous layer nodes and finds the firing strength of a rule by eq. 13.

$$O_{2,i} = w_i = \mu_{A_i}(x) . \mu_{B_i}(y)$$

Layer 3: Subsequently, nodes present throughout the layer 3 calculate normalized firing strength as shown in eq. 14.

$$O_{3,i} = \frac{w_i}{w_1 + w_2}$$

Layer 4: Node $i$ during this layer computes the contribution of the $i^{th}$ rule towards the model output, with the subsequent node function.
Where $\mathbf{w}_i$ is the output of layer 3 and $p_i, q_i$ and $r_i$ are the parameter set.

**Layer 5:** The single node during this layer computes the overall output of the ANFIS as

$$O_{h,t} = \sum \mathbf{w}_i f_i$$

But there is still scope to enhance hybridizing techniques also. As we all know that SVM is an alternate to ANN thus if some models using good features of SVM and ARIMA are developed that they’ll once more provide sensible and higher results than ARIMA and ANN hybrid. Same approach using ANFIS additionally can also be used to produce good results for forecasting.

In Table 1 a comparative analysis of accessible models is shown.

### Table 1. Comparison of models

| Model          | Advantages                                                                 | Disadvantages                                                                 |
|----------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| ARIMA          | It is very popular for linear modeling                                      | It cannot handle nonlinear data.                                               |
|                | It is very popular for linear modeling                                      | It assumes that real data is always linear in nature.                         |
|                |                                                                              | It presumes that the data is always linear in nature which is not true.        |
| Neural Network | Very well works with nonlinear data.                                        | Has problem of over fitting                                                   |
|                | Data driven, universal approximator                                         | During linear modeling yields mixed results.                                  |
| Support Vector Machine | It gives good regression results                                            | Result depends on parameters values.                                          |
|                | Good alternative to ANN.                                                    | Kernel function parameters should be selected properly.                      |
|                | It does not have any problem like over fitting.                            |                                                                              |
| Hybrid [ARIMA+ANN] | Can work well for both linear and nonlinear data.                           | If the model is not fitted properly the result are poor then the individuals. |
|                | Advantages of ARIMA and ANN both are there.                                |                                                                              |

### 5. Conclusion and Future Work

The time series forecasting and its models are studied here during this paper. Conjointly studied regarding the utilization of soft Computing Approach for time series prediction like ANN and SVM. From the study we found that time series information is combination of linear and nonlinear each therefore individual model does not seem to be capable to handle it properly. Therefore hybridizing of models by exploitation the strength of individual models could be a new analysis space in time series. Hybrid model of ARIMA and ANN are planned however there are still scope of improvement for hybrid model as we all know that SVM could be a sensible alternative of ANN. Therefore there’s a quest gap in hybridization which might be solved exploitation SVM.

### 6. Reference

1. Li CS, Chiang TW. Complex Neurofuzzy ARIMA Forecasting-A New Approach using Complex Fuzzy Sets, IEEE Trans Fuzzy Syst. 2013; 21(3):567−84. DOI: 10.1109/tfuzz.2012.2226890.
2. Bautu E, Barbulescu A. Forecasting Meteorological Time Series using Soft Computing Methods: An Empirical Study, Appl. Math. 2013; 7(4):1297−306.
3. Adhikari R, Agrawal RK. An Introductory Study on Time Series Modeling and Forecasting, arXiv Prepr arXiv13026613. 2013.
4. Meryem O, Ismail J, Mohammed E-M. A Comparative Study of Predictive Algorithms for Time Series Forecasting. In: Information Science and Technology (CIST), 2014 Third IEEE International Colloquium; 2014. p. 68−73.
5. Khashei M, Bijari M. An Artificial Neural Network (p,d,q) Model for Timeseries Forecasting, Expert Syst. Appl. 2010; 37(1):479−89. Doi: 10.1016/j.eswa.2009.05.044.
6. Khashei M, Bijari M. A Novel Hybridization of Artificial Neural Networks and ARIMA Models for Time Series Forecasting, Appl. Soft. Comput. J. 2011; 11(2):2664−75. DOI: 10.1016/j.asoc.2010.10.015.
7. Khashei M, Bijari M. A New Class of Hybrid Models for Time Series Forecasting, Expert Syst. Appl. 2012; 39(4):4344−57.
8. De Gooijer JG, Hyndman RJ. 25 Years of Time Series Forecasting, Int. J. Forecast. 2006; 22(3):443−73. DOI: 10.1016/j.ijforecast.2006.01.001.
9. Babu CN, Reddy BE. A Moving-Average Filter based Hybrid ARIMA–ANN Model for Forecasting Time Series Data, Appl. Soft. Comput. 2014; 23:27−38. DOI: 10.1016/j.asoc.2014.05.028.
A Review of Soft Computing Techniques for Time Series Forecasting

10. Hyndman RJ, Athanasopoulos G. Forecasting: Principles and Practice. OTexts; 2014.
11. Montgomery DC, Jennings CL, Kulahci M. Introduction to Time Series Analysis and Forecasting. John Wiley & Sons; 2015.
12. Wang L, Zou H, Su J, Li L, Chaudhry S. An ARIMA-ANN Hybrid Model for Time Series Forecasting. Syst. Res. Behav. Sci. 2013; 30(3):244–59. DOI: 10.1002/sres.2179.
13. Mandal SN, Choudhury JP, Chaudhuri SRB, De D. Soft Computing Approach in Prediction of a Time Series Data, J Theor. Appl. Inf. Technol. 2008; 8(3):1131–41.
14. Rojas I, Palomares H. Soft-Computing Techniques for Time Series Forecasting. In: Proc. of the European Symposium on Artificial Neural Networks; 2004. p. 93–102.
15. Agami N, Atiya A, Saleh M, El-Shishiny H. A Neural Network based Dynamic Forecasting Model for Trend Impact Analysis, Technol. Forecast Soc. Change. 2009; 76(7):952–62.
16. Chen Y, Yang B, Dong J, Abraham A. Time-Series Forecasting using Flexible Neural Tree Model, Inf. Sci. (Ny). 2005; 174(3):219–35.
17. Piotrowski AP, Napiorkowski MJ, Napiorkowski JJ, Osuch M. Comparing Various Artificial Neural Network Types for Water Temperature Prediction in Rivers, J. Hydrol. 2015; 529:302–15.
18. Samsudin R. A Comparison of Time Series Forecasting using Support Vector Machine and Artificial Neural Network Model, J. Appl. Sci. 2010; 10(11):950–58.
19. Gunn SR. Support Vector Machines for Classification and Regression, ISIS Technical Report. 1998 May 10. p. 14.
20. Sapankevych N, Sankar R. Time Series Prediction using Support Vector Machines: A Survey, IEEE Comput. Intell. Mag. 2009; 4(May):24–38. DOI: 10.1109/MCI.2009.932254.
21. Xiang L, Tang G, Zhang C. Simulation of Time Series Prediction based on Hybrid Support Vector Regression. In: Natural Computation; 2008. ICNC’08. Fourth International Conference. 2008; 2:167–71.
22. Xiang L, Zhu Y, Tang G. A Hybrid Support Vector Regression for Time Series Forecasting. In: Software Engineering, 2009. WCSE’09. WRI World Congress. 2009; 4:161–65.
23. Chen K-Y. Forecasting Systems Reliability based on Support Vector Regression with Genetic Algorithms, Reliab. Eng. Syst. Saf. 2007; 92(4):423–32.
24. Zhang X, Zhang T, Young AA, Li X. Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data, PLoS One. 2014; 9(2):e88075. DOI: 10.1371/journal.pone.0088075.
25. Zhang GP. Time Series Forecasting using a Hybrid ARIMA and Neural Network Model, Neurocomputing. 2003; 50:159–75.
26. Bao DN, Vy NDK, Anh DT. A Hybrid Method for Forecasting Trend and Seasonal Time Series. In: Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2013 IEEE RIVF International Conference; 2013. p. 203–08.
27. Reyhani R, Others. A Heuristic Method for Forecasting Chaotic Time Series based on Economic Variables. In: Digital Information Management (ICDIM), 2011 Sixth International Conference; 2011. p. 300–04.
28. Yolcu U, Egrioglu E, Aladag CH. A New Linear and Nonlinear Artificial Neural Network Model for Time Series Forecasting, Decis. Support Syst. 2013; 54(3):1340–47. DOI: 10.1016/j.dss.2012.12.006.
29. Vijayalaksmi DP, Babu KSJ. Water Supply System Demand Forecasting Using Adaptive Neuro-fuzzy Inference System. Aquat. Procedia. 2015; 4:950–56.
30. Zounemat-Kermani M, Teshnehlab M. Using Adaptive Neuro-Fuzzy Inference System for Hydrological Time Series Prediction, Appl. Soft. Comput. 2008; 8(2):928–36.