

Optimized Intelligent Auto-Regressive Neural Network Model (ARNN) for Prediction of Non-Linear Exogenous Signals

Lubna Farhi (✉ lubnafarhi@yahoo.com)

Sir Syed University of Engineering and Technology

Agha Yasir

Sir Syed University of Engineering and Technology

Research Article

Keywords: Autoregressive Moving Average with Exogenous Excitation (ARMX), Autoregressive with Exogenous Variables (ARX), Non-linear Autoregressive with Exogenous Variables (NARX), Neural Network

Posted Date: March 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-263891/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Wireless Personal Communications on January 26th, 2022. See the published version at <https://doi.org/10.1007/s11277-021-09399-8>.

Optimized Intelligent Auto-Regressive Neural Network Model (ARNN) for Prediction of Non-Linear Exogenous Signals

Lubna Farhi · Agha Yasir

Received: date / Accepted: date

Abstract The paper presents a prediction of non-linear exogenous signal by optimized intelligent auto-regressive neural network model (ARNN). A signal comprises of two sets of data called deterministic and error. The former type of data represents the degradation index of a signal, while the error is the uncertainties associated with the signal. To understand and predict signals, a intelligent approach is taken through the use of ARNN model. In this approach, the first step is to diagnose whether a time series signal is normally distributed or not by utilizing the Jarque-Bera test. The high and low volatility data elements can be separated via kurtosis hypothesis. The deterministic component of the signal is also predicted by developing a neural network based non-linear autoregressive model (NN-NARX) and the error component by using a linear model. The final forecast is formed by combining the results determined from each of the models and evaluated using the mean square error results. Validation of the prediction is obtained through a comparison of the results with other models such as ARNN, traditional ARMX, and NARX models. The results show that the proposed model provides improved predictions, minimize high dependence on design parameters with low computational cost.

Keywords Autoregressive Moving Average with Exogenous Excitation (ARMX) Autoregressive with Exogenous Variables (ARX) Non-linear Autoregressive with Exogenous Variables (NARX) Neural Network

Sir Syed University of Engineering and Technology
University road, Karachi
Tel.: +92-21-34988000
Fax: +92-21-34988005
E-mail: lubnafarhi@yahoo.com, lfarhi@ssuet.edu.pk

Sir Syed University of Engineering and Technology
University road, Karachi
Tel.: +92-21-34988000
Fax: +92-21-34988005
E-mail: agyasir@ssuet.edu.pk

1 Introduction

The field of acoustics studies the propagation and interaction of mechanical waves such as sound waves through different media like air or water. Various methods can be used to model and understand the behavior of those waves including stochastic, deterministic, and statistical methods based on underlying physical mechanics of sound waves. Many prediction models have been presented in literature that are most either model based or data-driven or both. However, model based techniques rely on mathematical models that can realistically predict behaviour based on fundamental physical parameters while also being able to predict the linear component of time series. On the other hand, data-driven techniques are statistically intensive where they rely on artificial intelligence to sift through a huge amount of data and create a model that forecast the desired phenomenon. This allows data-driven techniques to be highly flexible despite the complexity of the system being modeled. Some of these data-driven techniques have been proposed by researchers [3],[4],[5]. Since these techniques are reliant on historical data to build the predicting model, they require a vast amount of failure data to accurately interpolate and infer a model. Although this might be hard to achieve in some cases, these models are more flexible and can adapt accurately with complex non-linear components of any system. However, as compared to model-based techniques, data-driven techniques are more prone to error as any lapses in coherent information can lead to an inaccurate prediction model. In contrast with the former technique where physics and mathematics dictate the behavior of model reality. Despite being more inaccurate, data-driven techniques are more flexible and can model a complex system with greater ease. This allows them to be more favorable when used with artificial intelligence to extrapolate models using historical data regardless of the complexity of the system.

Predicting time series has become a focal point of research where several unique techniques have been proposed including a method of local modeling by [2]. Similarly, models for time series prediction were proposed by [10] that utilized an ensemble of methods. Building upon this, [9] described a methodology based on Lyapunov exponents to provide multidimensional predictions, however, considering the complexity and dynamic time series, the modeling process has been difficult. To tackle this issue, researchers have turned towards non-linear prediction techniques that have shown to be more accurate regarding prediction accuracy [12],[13]. One of the most prominent methodologies is neural networks which has potential to improve the prediction accuracy of non-linear time series. Recurrent neural networks can be used to develop multi-step-ahead prediction due to flexible internal memory structure that accounts the historical data when predicting future values [14],[15],[11]. In addition, researchers have begun looking at hybrid models that further improve prediction accuracies for investors such as the Combined autoregressive integrated moving average model proposed by [5] combined with neural network model. These models were used in conjunction to accurately forecast popular time-series data. Another interesting hybrid methodology incorporated both

parametric and nonparametric techniques to predict the exchange market and it was proposed by [16],[22]. These methodologies follow a similar procedure in that, they first predict a linear relation using a model-based technique and then they compare the actual values and the predicted results using a data-driven technique that assesses the residuals between the two values. The result of this comparison is compounded within the model and the procedure is repeated thus reducing error between actual and forecast values. However, some hybrid approaches are only suitable for a short-term future prediction such as [6],[7],[8].

In this paper, a optimized intelligent prediction model is proposed to predict time series signals. This recursive model involves a neural network based nonlinear autoregressive and an autoregressive moving average. In the proposed approach, deterministic and error components are determined from actual time series data separately. Neural network based nonlinear autoregressive (NN-NARX) is developed to find the dimension of embedded network while autoregressive moving average(ARMX) is simply utilized to calculate the error that can be defined as the error between residual and the actual data. The NN-NARX and ARMX models are deployed in unison to predict the past and the future respectively while the final prediction results are the summative of results attained from a single proposed dynamic model. However, one issue that is often experienced with time series prediction techniques is embedding the dimension of net. This problem mitigate by utilizing either Cao method [24] to estimate the embedding dimension. It separates the deterministic and stochastic signals elements and determine the embedded network dimension and generate the phase space that should be synchronize with hidden units of neural network.

2 Methodology

2.1 Error Estimation Methods

Error estimation methodology is used to identify the system model and minimize the cost function that calculates the difference between the actual and estimated output. The fundamental model used for error estimation that the system output is measured between two input signal μ stand the effect of disturbances e_t . Mathematically, model is used:

$$y_t = b_0\mu_t + b_1\mu_{t-1} + \dots + b_{m_b}\mu_{t-m_b} + e_t = B(q^{-1})\mu_t + e_t \quad (1)$$

where input vector is μ_t and $B(q^{-1})$ is filter kernel,

$$B(q^{-1}) = b_0 + b_1q^{-1} + b_{m_b}q^{-m_b} \quad (2)$$

Characteristics of models are depending on the structure of $B(q^{-1})$ and adaptive noise e_t [9].

2.2 Autoregressive with Exogenous Inputs model(ARX)

ARX model is used to describe the global properties of the system and determine the FIR model coefficients. An Auto-Regressive model with exogenous inputs is specified as [21]

$$A(q^{-1})y_t = B(q^{-1})\mu_t + e_t \quad (3)$$

where

$$A(q^{-1}) = 1 + a_1q^{-1} + a_{m_a}q^{-m_a} \quad (4)$$

and $B(q^{-1})$ is given as in (3). This implies that the model equals

$$y_t = \frac{B(q^{-1})}{A(q^{-1})}\mu_t + \frac{1}{A(q^{-1})}e_t \quad (5)$$

The noise is modelled by a factor of $1/A$ and multiplied with the dynamics model. It is important to note that ARX does not model noise and dynamics independently [18].

2.3 Autoregressive Moving Average with Exogenous inputs model (ARMX)

In ARMX the dynamics of the noise that are parameterized are more flexible than the ARX model. The ARMX extends the ARX structure by providing more flexibility for modeling noise using the C parameters (a moving average of white noise). This allows ARMX to be the preferable option when the input is dominated by disturbances known as load disturbances. ARMX model an auto-regressive model with exogenous inputs and Moving Average model for the disturbances is given as [1]

$$A(q^{-1})y_t = B(q^{-1})\mu_t + C(q^{-1})e_t \quad (6)$$

and $A(q^{-1})$ is as in (4) and $B(q^{-1})$ is given as in (2). This model output is:

$$y_t = \frac{B(q^{-1})}{A(q^{-1})}\mu_t + \frac{C(q^{-1})}{A(q^{-1})}e_t \quad (7)$$

ARMX methodology follows the equation error models. In these models, the observed output is a sum of three regression expressions: previous inputs, outputs and white noise. It is conceivable to infer ARMX methods as describe in Figure 1, i.e. by considering the parallel connection of a deterministic part determined by the observed input u_t and of a stochastic part determined by a remote white process e_t . The deterministic part is characterized by the transfer function $\frac{B(q^{-1})}{A(q^{-1})}$ and its output $y_0(t)$ is not accessible. The stochastic part is characterized by the transfer function $\frac{C(q^{-1})}{A(q^{-1})}$ and its output is a noise e_t that can signify the effect of white noise on the state of the deterministic part. The observed output is:

$$y_t = y_0(t) + n_{(d)} \quad (8)$$

process noise and measurement noise are assumed to be independent, with zero mean. The white and Gaussian noises are represented as $e_t \sim N(0, \sigma^2)$, $n_{(d)} \sim N(0, \lambda^2)$. The autocorrelations is used to error estimation, denoted as $\rho(n)$, can be derived only depend on the lag:

$$\rho(n) = \frac{Cov(y_t, y_{t+n})}{\sqrt{Var(y_t)}\sqrt{Var(y_t + n)}} \quad (9)$$

2.4 Neural Network-Nonlinear Autoregressive Model(NN-NARX)

One of the more flexible architectures of nonlinear modelling technique is the neural network. This network is built with two or more layers . Generally a network comprises of three layers: input, hidden and an output layer. The number of inputs along with the number of neurons in every hidden layer also can be varied. The neurons are processing units which are cyclically linked . In proposed approach, the neural network is utilizing time series data to develop a Neural Network-Nonlinear Autoregressive Model. The sequence y_t is reflected as a nonlinear function of $y_{(t-1)}, \dots, y_{(t-N)}$. The corresponding equation is shown in (10) where the function g is a nonlinear function, and n_d is a noise or error term.

$$y_t = g(y_{(t-1)}, y_{(t-2)} \dots y_{(t-n)}) + e_t \quad (10)$$

NN model is also suitable for modeling nonlinear systems and time series is the Nonlinear Auto-regressive with Exogenous Input (NARX) neural network. It is considered as a dynamic neural network because it connects several hidden layers of the network to the input layer through recurrent feedback [20]. The NARX model is used to predict the deterministic component as it is the preferable option for nonlinear issues. A single-input single-output (SISO) with true output $y_0(t) \in r$ and measurement noise corrupted output $y_t \in r$, mathematically is:

$$y(t+1) = g(y_t, \mu_t) + e_t \quad (11)$$

and true output is $y_0(t)$, therefore eq(12) can be write:

$$y_t = y_0(t) + n_{(d)} \quad (12)$$

where input represent by $\mu_{(t)}$ and output by $y_{(t)}$ respectively at time t . The state vector is composed of the true system outputs: $y_t = [y_t, \dots, t-n_y+1]^T$; similarly state vector of inputs is: $\mu_{(t)} = [\mu_t, \dots, \mu_{t-n}, \mu_{t+1}]$. $g()$ is some nonlinear function, and $\mu_{(t)} \in R$ is the known system input. The model orders n_μ and n_y are order dynamics of input and output, respectively.

Therefore, NARX model dynamics model in eq(11) as the state equation is:

$$y_{(t+1)} = Ay_t + Bg(y_t, \mu_t) + e_t \quad (13)$$

The measurement equation, which describes the mapping from the unobserved states to the observed output signal, is

$$y(t) = Cy_t + n_{(d)} \quad (14)$$

Discrete-time nonlinear mathematical equation of NARX can be represented as

$$y(n+1) = g[(y(n) \dots y(n-d_y+1); \mu(n-k), \mu(n-k+1), \dots, \mu(n-d_\mu-k+1))] \quad (15)$$

where input $\mu(n)$ and output $y(n)$ of the model at discrete time step n , while $d_y \geq 1$ and $d_\mu \geq 1$ and $d_\mu \leq d_y$ are the memory orders of input and output respectively. The parameter k is a delay or process dead-time.

NARX architecture with single input and output time series signals and Series-Parallel (SP) network are used in training mode. In this case, actual values of the systems output are used to formed the outputs regressors [10].

$$y_{\hat{sp}}(n+1) = \hat{g}[(y(n) \dots y(n-d_y+1); \mu(n), \mu(n+1) \dots \mu(n-d_\mu+1))] \quad (16)$$

where estimate values are represented by symbol ($\hat{}$). A Levenberg-Marquardt (LM) training algorithm was used to train the NN model in this paper.

3 Proposed Optimized Intelligent Auto-Regressive Neural Network Model(ARNN)

For the prediction of time series signals different methodologies are utilized, however these techniques are not appropriate for time series signals due to its complexity. Hence, for prediction of time series data, none is suitable model, either ARMX or NARX. By utilizing neural network algorithm for top noise data, challenged to convergence problem if the amount of neurons are less or over fitting thanks to vast quantity of neurons. On the other hand, ARMX model isn't suitable for those data which have nonlinearity and is insufficient the stationary condition.

In this paper, an optimized intelligent auto-Regressive neural network model(ARNN) is proposed, in which the data information is partitioned into two components: deterministic and error. The deterministic component is generated by filtering method and the error component is the residual between time series data and deterministic components. The deterministic component is calculated by Neural Network-NARX algorithm and ARMX model is utilized to find the error component which depicts the presence of uncertainties.

The improved dynamic ARNN model equation is the sum of error estimation $\rho(n)$ from ARMX model and deterministic component $E(n)$ from NN-NARX model. The deterministic component $E_{(n)} = [E_1, E_2, \dots, E_n, \dots, E_{(t-1)}]$ is obtained from a time series data $y = [y_1, y_2, \dots, y_n, \dots, y_{(t)}]$, therefore E_n is:

$$E_{(n)} = \frac{y_{n-1} + y_n + y_{n+1}}{3}, \quad n = 1, 2, 3 \dots, t_1 \quad (17)$$

and $\rho(n)$ is error component, find from ARMX model:

$$\rho(\hat{n}) = y_{(n)} - E_{(n)} \quad (18)$$

where $f'(\hat{n} + 1)$ is an estimated embedding dimension of NN-ARMX network is:

$$f'(\hat{n} + 1) = f'_0 \left[b_0 + \sum_{h=1}^{N_h} w_{h_0} g_h \left(b_h + \sum_{i=1}^{d_\mu} w_{ih} \mu(n-i) + \sum_{j=1}^{d_y} w_{jh} y(n-j) \right) \right] \quad (19)$$

where $i = 1, 2, \dots, d_\mu, j = 1, 2, \dots, d_y$ and $h = 1, 2, \dots, N_h$

w_{ih}, w_{jh} and w_{h_0} , are the weights, b_h and b_0 are biases and $f'_0(\cdot)$ and $f'_h(\cdot)$ are hidden and output functions. This network is trained by back propagation algorithm and stochastic gradient decent is used to update the weights. This technique reduces the variance in the estimate of the gradient and adaptive learning rate used to minimizes the computation time.

Additionally, the number of observations used as the contribution for prediction model, called embedding dimension. Embedding dimension is the dispute often experienced in time series estimating techniques. In proposed approach, embedding dimension is pre-determined by Ca'o method to separate the deterministic and stochastic signals [24].

4 Performance Evaluation

The ARNN intelligent model is trained by creating a correlation between the predicted value of time series and its residuals, and the original time series. This is done by providing the predicted values and the residuals as the input and the original time series as the output. Then the prediction performance is evaluated with respect to previous algorithms such as the mean squared error (MSE), root mean squared error (RMSE) and normalized mean squared error (NMSE) according to eqs. (20), (21) and (22) respectively.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (20)$$

$$RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (21)$$

$$NMSE = \left(\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \right) \quad (22)$$

where N is the length of data and y_i, \hat{y}_i and \bar{y}_i are observation of data, predicted data and the average-observed data respectively.

Algorithm of Dynamic ARNN Model

Result: Evaluate the "Final Predicted Values"

Data: Time Series Signal y_t , where $y = [y_1, y_2, y_3, \dots, y_t]$

Initialization ARMX Model ;

For $n = 1 : n - 1$ **do**

1. Calculate "error term" by using by using Eq.(18)
2. Check "the stationary conditions"
3. Build "the ARMX model"
4. Predict "the error $\rho(\hat{n}) = \rho(1), \rho(2) \dots \rho(t - 1)$ "

Initialization NN-NARX Model ;

For $n = 1 : n - 1$ **do**

5. Calculate "Deterministic term" by using "Eq.(17)"
6. Calculate "Embedded dimension of Net" by using Eq.(19)
7. Build " the NN-NARX model"
8. Evaluate the "back-propagation error term $\delta_j^{(i+1)}$ "
9. $\delta_j^{(l)} = f'(n_j^l) \sum_l W_{ij}^{(i+1)} \delta_j^{(i+1)}$
10. Calculate the gradient function with respect to the weight vector w_{ij}
11. $\frac{\partial E(w,b)}{\partial w_{ij}^l} = \delta_j^{(l+1)} f(n)_i^l$,

where $f(n)$ is logistic sigmoid activation function

12. Obtain "the Prediction values of $E(\hat{n})$ " by utilizing Eq. (18)"

End

13. **Final Predicted Values**= $E(\hat{n}) + \rho(\hat{n})$

End

5 Result and Analysis

Dynamic nature of time series complex signal was utilized to calculate the proposed optimized ARNN model. Time series filtered data were a set of 901 points as shown in Figure 2 In general, to diagnose whether a given sequence is normally distributed or not, the JarqueBera normality test can be performed. This test analysis determined the kurtosis value of the data, it is 3 or not:

$$kurtosis(k) = \frac{E[y - E(y)]^4}{\left(E[y - (E(y))]^2\right)^2} \quad (23)$$

In Eq.(23), y is the variable data for which the kurtosis is being computed, and E stands for the expected value.

Therefore, if the sequence has $k=3$ of low-volatility data elements, ARMX model is suitable for prediction or $k \geq 3$ for high-volatility data's elements, the predictions are obtained by using an ANN model. However, if the sequence has $k = 3$ or $k \geq 3$, ARNN proposed approach can be used for prediction. This is done to not only reduce computation time but to also reduce error as

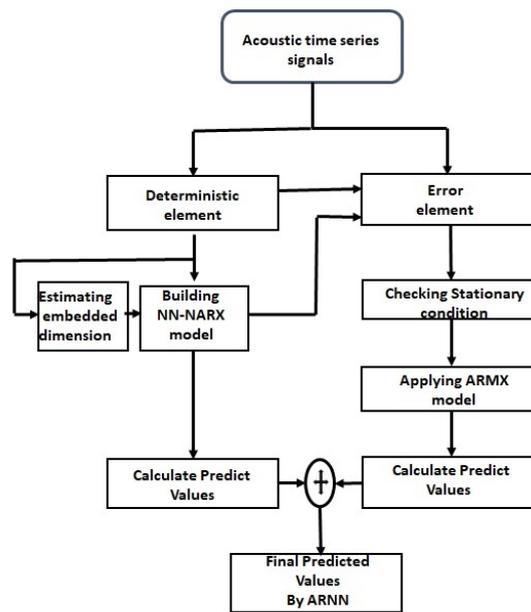


Fig. 1: Prediction Process of ARNN Model.

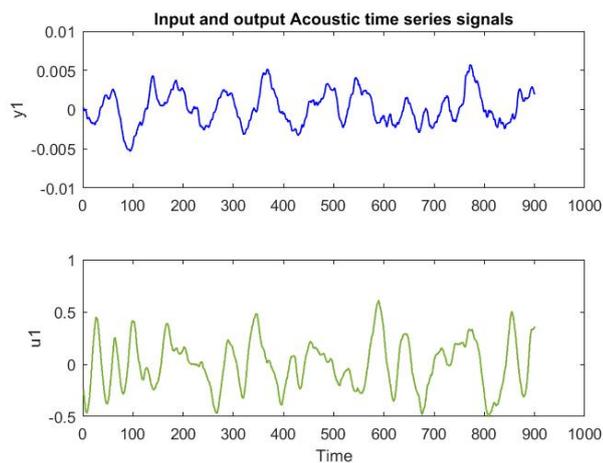


Fig. 2: Input Output time series data with noise.

the linear component can be modeled by a nonlinear model if the sequence is not separated. Simulation results show that when Jarque Bera normality test is applied on 900 samples of time series then the kurtosis value of $k=5.2$

is obtained. However, the proposed ARNN optimized technique provided a $k=2.5$ of the predicted data for the same time series.

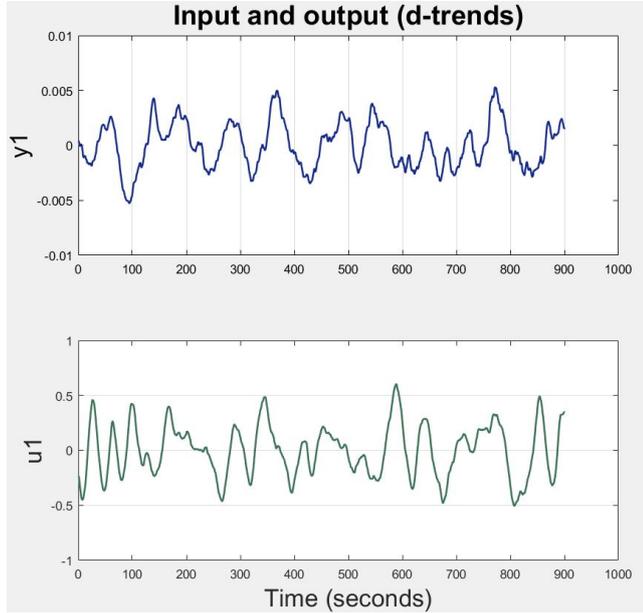


Fig. 3: Filtered data.

The NARX estimation model is then build by utilizing data. To generate NARX model, first determined the embedding dimension by Chaos method. It used to estimate embedding dimension of time series Figure 4 and extricate deterministic signals and stochastic signals clearly. There are two significant qualities that are first and second deterministic components $E1(d)$ and $E2(d)$ should have been determined respectively as shown in Figure 5 $E1(d)$ is utilized to find the minimum embedding dimension when it extends the saturation point.

$E2(d)$ is employed for the difficulty in real computations wherever $E1(d)$ is gradually increasing or has quit ever-changing if embedding dimensions adequately huge.

In this paper were applied 900 sampled of time series sequence, the prediction performance results are shown in Figure 6. From the results, the proposed model ARNN showed significant improvement in prediction response as compared to the ARMX and NARMX models separately.

The predicted errors are then calculated using an error histogram that represents the distribution of predicted error in every instance as shown in Figure 7 which represents the error histogram for the proposed scaled Armx, NARX and ARNN models.

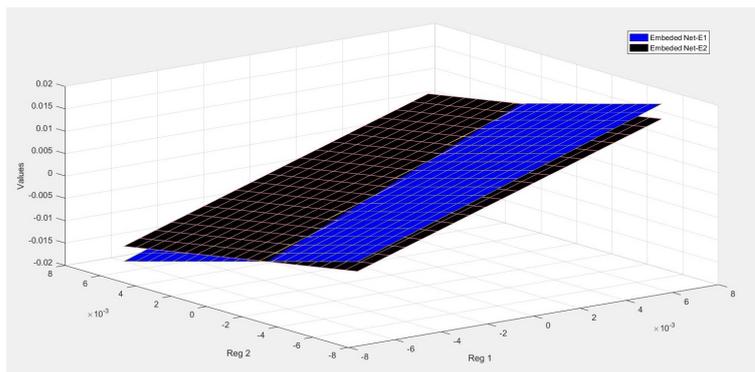


Fig. 4: Represent the embedding dimension useful for deterministic component.

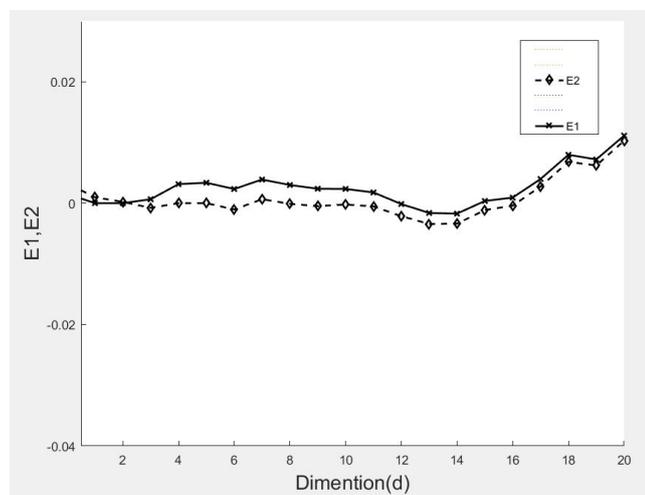


Fig. 5: The value E1 and E2 of deterministic component.

The estimating and prediction methods are used to forecast the future states of data. The prediction results are calculated by the error between forecasted values and actual values in the testing set. Table 1 shows the comparison results of MSE, RMSE and NMSE of three models applied for time series data. In this table, all error parameters of the ARNN model is minimum then other traditional algorithms.

The proposed model has a regression value that approaches to 1 as shown in Figure 8, with a regression value of 1 is considered as a best fit model.

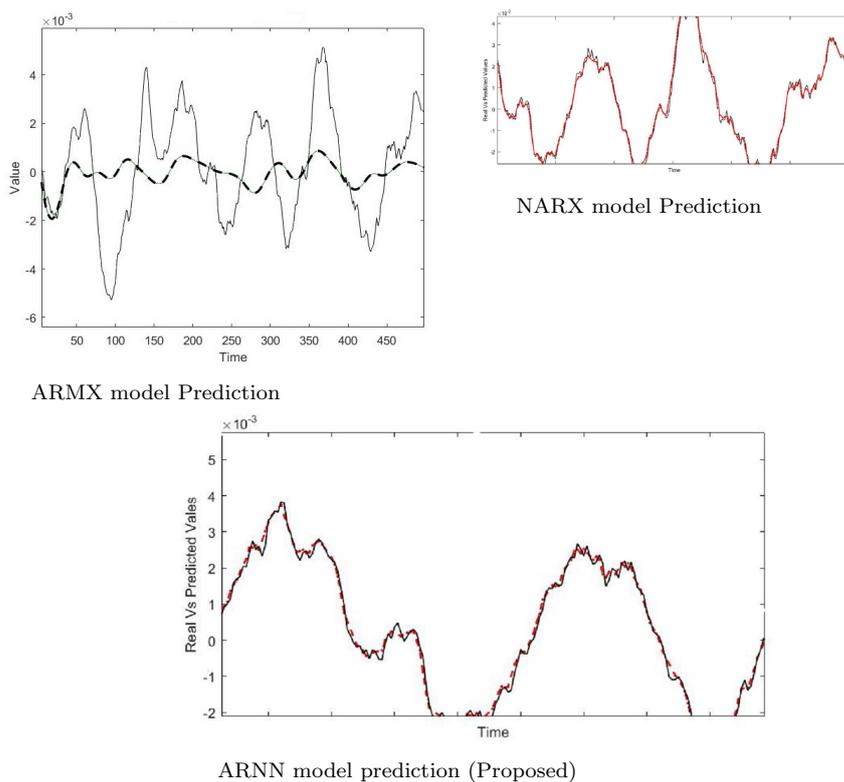


Fig. 6: Compression of Prediction responses

Table 1: Comparison between the prediction errors reported in the literature and the proposed method

Prediction Models	Prediction Errors		
	MSE	RMSE (%)	NMSE(%)
ARMA-Neural Network	-	2.5e-03	-
Hybrid neuralnetwork(HNN)	-	-	5.30e-02
AutoRegressive model	-	0.19	-
ARNN	1.3e-8	2.79e-8	2.77e-8

6 Conclusion

The proposed technique is an amalgamation of an embedding theorem and Neural Network-NARX to predict the deterministic component and an ARMX model to verify the stationary condition along with the error components of a time series. First, it is determined if the sequence is normally distributed or not by conducting the Jarque-Bera and Kurtosis Hypothesis tests. These tests provide information regarding the normality of the sequence and separate the high and low volatility data elements. Then, embedding Chaos theorem is utilized to extract the embedded dimension and time delay from the original time

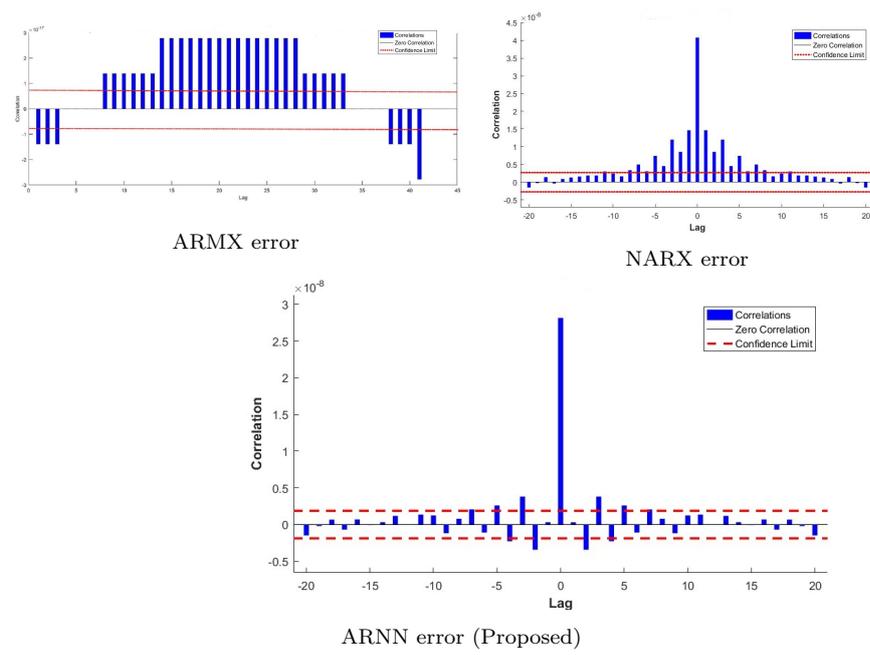


Fig. 7: Prediction Histogram error compression

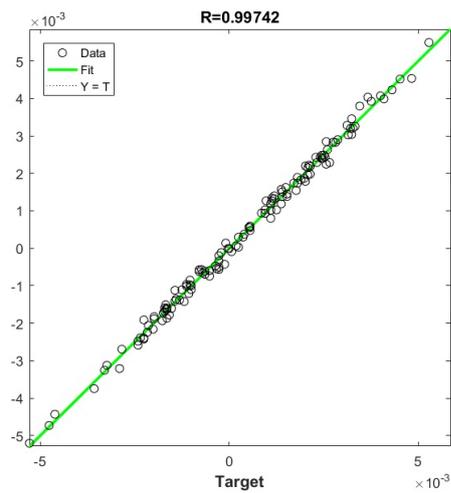


Fig. 8: Regression value for Prediction of time series data.

series. Consequently, the embedded dimension and time delay are developed into phase space and synchronized with the number of neurons of optimized neural network. It overcomes the dependency on design parameters and min-

imizes the computational cost. Finally, the NN-NARX algorithm is trained using learning rate, gradient descent and momentum to estimate the values of embedded phase space points. This analysis was repeated multiple times to average the result and reduce statistical errors. The Levenberg- Marquardt method is also used to train NN-NARX to accurately predict the original time series and residuals. This procedure allowed the proposed technique to provide more accurate results as compare to other experimental models. The results were validated and the models performance was evaluated by using an error index that compared the predictions with other models in the literature.

References

1. Menezes Jr, J. M. P., Barreto, G. A. Long-term time series prediction with the NARX network: An empirical evaluation. *Neurocomputing*, 71(16-18),(2008).
2. McNames, James. A nearest trajectory strategy for time series prediction, Proceedings of the International Workshop on Advanced Black-Box Techniques for Nonlinear Modeling. KU Leuven Belgium, (1998).
3. Liu J, Wang W, Golnaraghi F.,A multi-step predictor with a variable input pattern for system state forecasting, *Mechanical Systems and Signal Processing* ,Jul 1;23(5):1586-99,(2009)
4. Yang BS, Oh MS, Tan AC,Machine condition prognosis based on regression trees and one-step-ahead prediction, *Mechanical Systems and Signal Processing*,22(5):1179-93, (2008)
5. Wang W, Vrbanek Jr J., A multi-step predictor for dynamic system property forecasting, *Measurement Science and Technology*,18(12):3673,(2007)
6. Valenzuela O, Rojas I, Rojas F, Pomares H, Herrera LJ, Guilln A, Marquez L, Pasadas M. ,Hybridization of intelligent techniques and ARIMA models for time series prediction, *Fuzzy sets and systems*, 159(7),821-45,(2008)
7. Rojas I, Valenzuela O, Rojas F, Guilln A, Herrera LJ, Pomares H, Marquez L, Pasadas M. ,Soft-computing techniques and ARMA model for time series prediction, *Neurocomputing*. ,71(4-6),519-37.(2008)
8. Pai PF, Lin CS., A hybrid ARIMA and support vector machines model in stock price forecasting, *Omega*, 33(6),497-505,(2005)
9. Zhang J, Chung HS, Lo WL.,Chaotic time series prediction using a neuro-fuzzy system with time-delay coordinates, *IEEE Transactions on Knowledge and Data Engineering*, 20(7),956-64,(2008)
10. Wichard JD, Ogorzalek M., Time series prediction with ensemble models, *IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*, Vol. 2, pp. 1625-1630,(2004)
11. Zhang J, Lam KC, Yan WJ, Gao H, Li Y.,Time series prediction using Lyapunov exponents in embedding phase space, *Computers and Electrical Engineering*,30(1):1-5,(2004)
12. Karray F, Karray FO, De Silva CW.,Soft computing and intelligent systems design: theory, tools, and applications, Pearson Education, (2004).
13. Pratama M, Er MJ, Li X, Oentaryo RJ, Lughofer E, Arifin I., Data driven modeling based on dynamic parsimonious fuzzy neural network, *Neurocomputing*, 110:18-28,(2013)
14. Andalib A, Atry F., Multi-step ahead forecasts for electricity prices using NARX: a new approach, a critical analysis of one-step ahead forecasts, *Energy Conversion and Management*, 50(3):739-47,(2009)
15. Renart A., Recurrent networks learn to tell time, *Nature neuro-science*,(7):772,(2013)
16. Ince H, Trafalis TB., A hybrid model for exchange rate prediction. *Decision Support Systems*, 42(2),1054-62,(2006)
17. Leontaritis IJ, Billings SA., Input-output parametric models for non-linear systems part I: deterministic non-linear systems, *International journal of control*, 41(2),303-28,(2005)
18. Huang CM, Huang CJ, Wang ML.,A particle swarm optimization to identifying the ARMAX model for short-term load forecasting., *IEEE Transactions on Power Systems*, 20(2),1126-33,(2005)

19. Cao SG, Rees NW, Feng G. *Analysis and design for a class of complex control systems part I: fuzzy modelling and identification*. Automatica, 33(6):1017-28,(2001)
20. Kennel MB, Brown R, Abarbanel HD.,Determining embedding dimension for phase-space reconstruction using a geometrical construction., Physical review A., 45(6),3403,(2008)
21. Rojas I, Valenzuela O, Rojas F, Guilln A, Herrera LJ, Pomares H, Marquez L, Pasadas M. ,Soft-computing techniques and ARMA model for time series prediction.,Neurocomputing, 71(4-6),519-37,(2008)
22. Inoue H, Fukunaga Y, Narihisa H., Efficient hybrid neural network for chaotic time series prediction., InInternational Conference on Artificial Neural Networks, 712-718,(2001)
23. Rojas I, Pomares H, Bernier JL, Ortega J, Pino B, Pelayo FJ, Prieto A., Time series analysis using normalized PG-RBF network with regression weights.,Neurocomputing., 42(1-4),267-85,(2002)
24. Xu, Xiaoke, Xiaoming Liu, and Xiaonan Chen., The Cao method for determining the minimum embedding dimension of sea clutter., CIE International Conference on Radar,1-4, (2006)

Figures

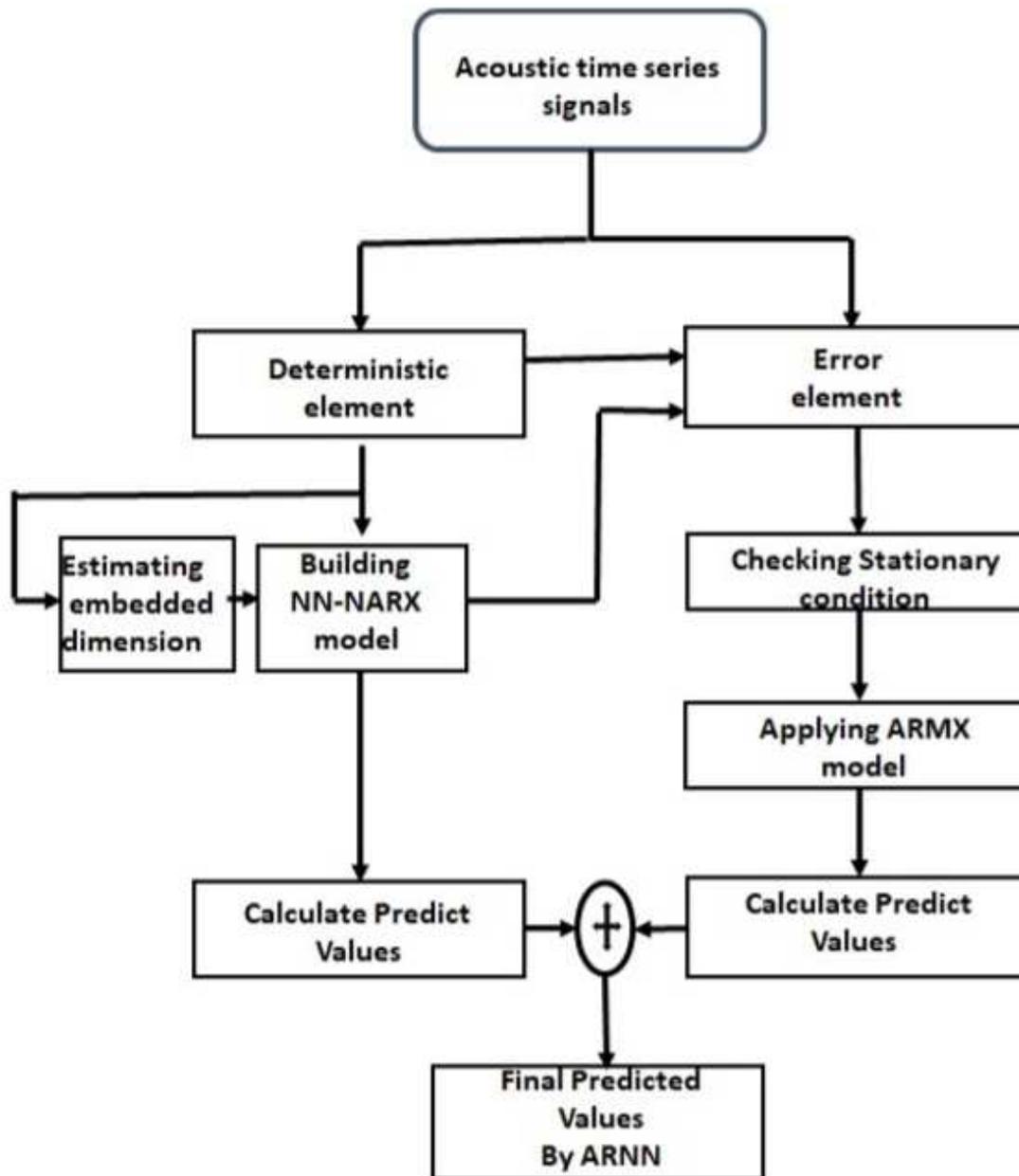


Figure 1

Prediction Process of ARNN Model.

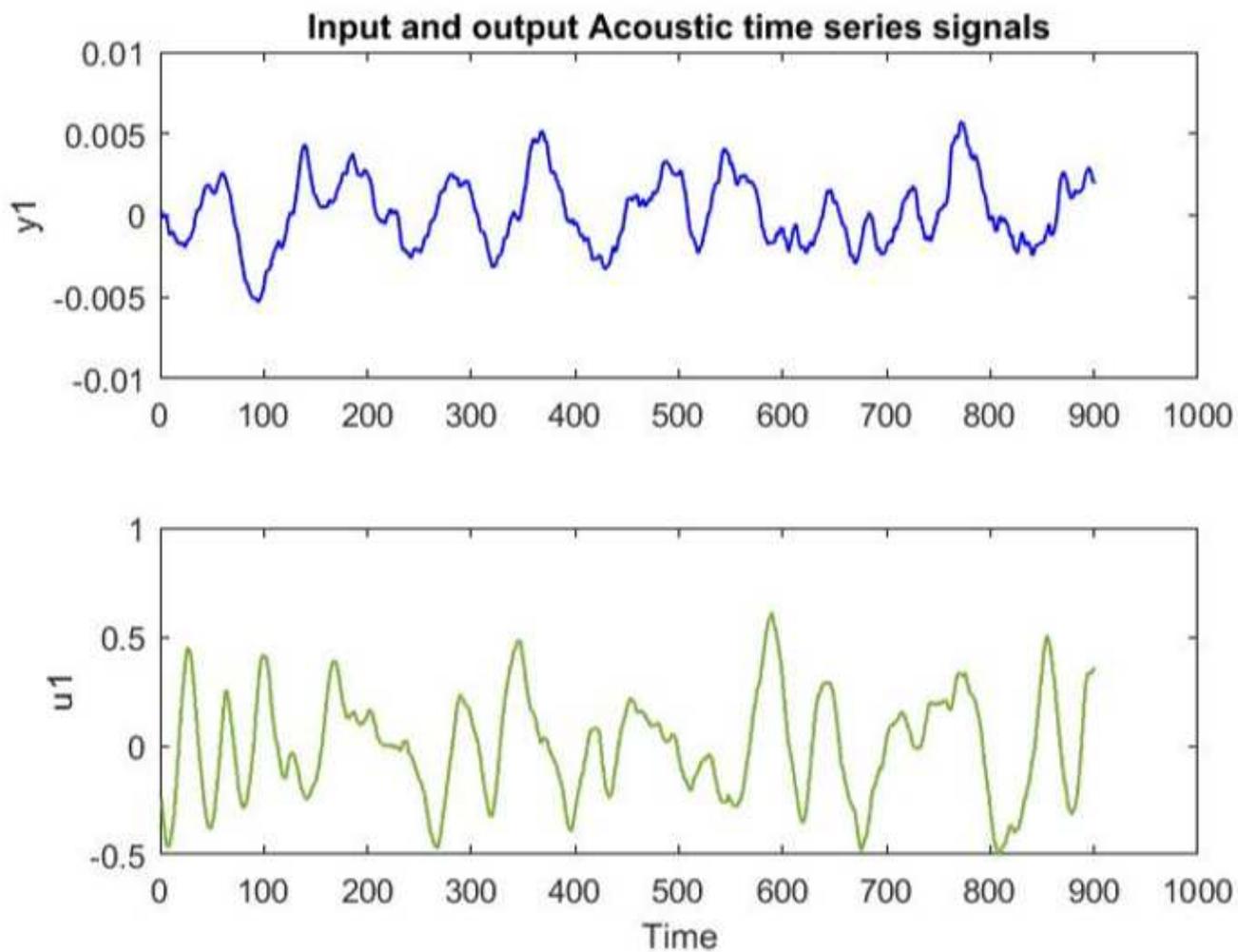


Figure 2

Input Output time series data with noise.

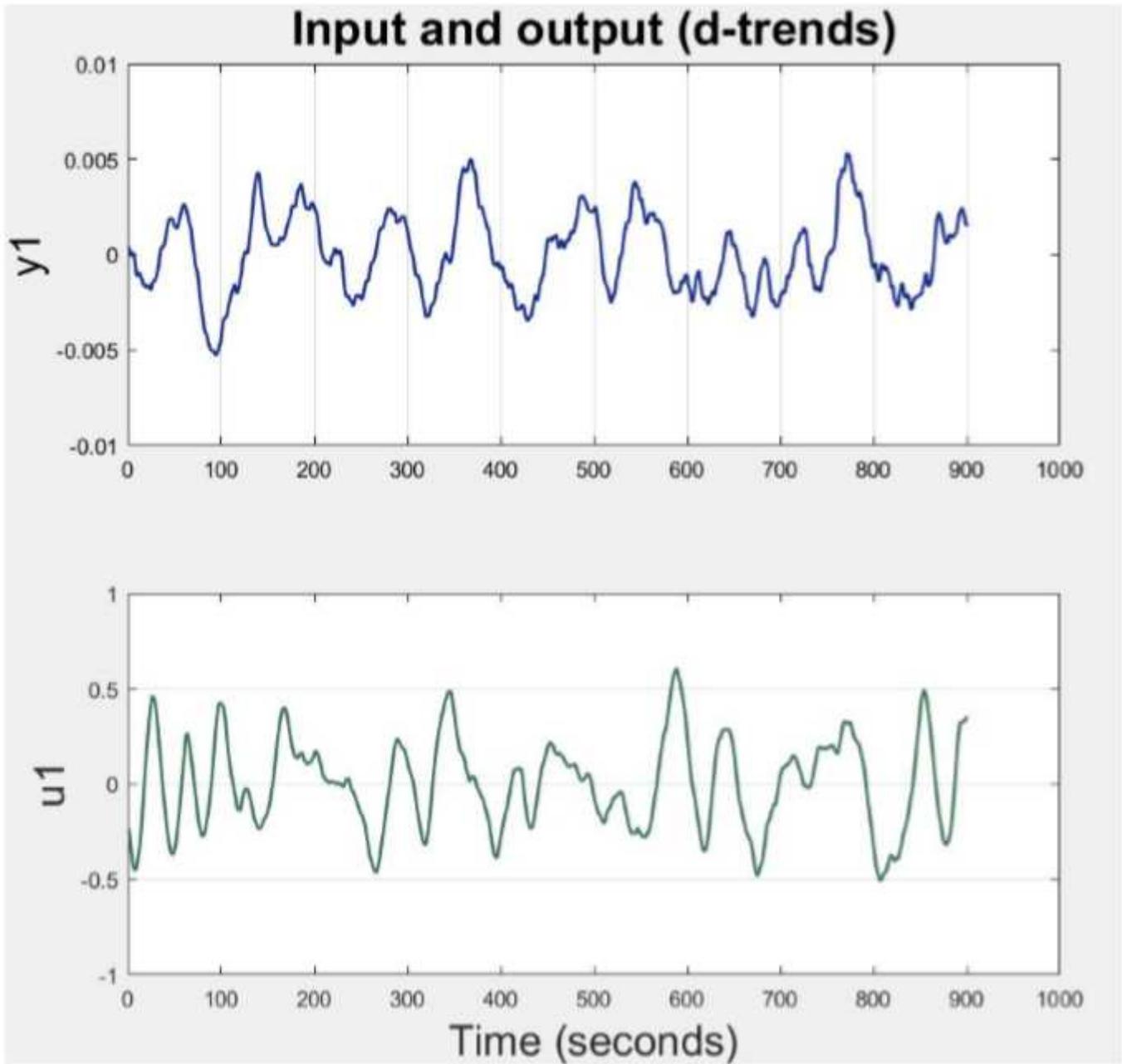


Figure 3

Filtered data.

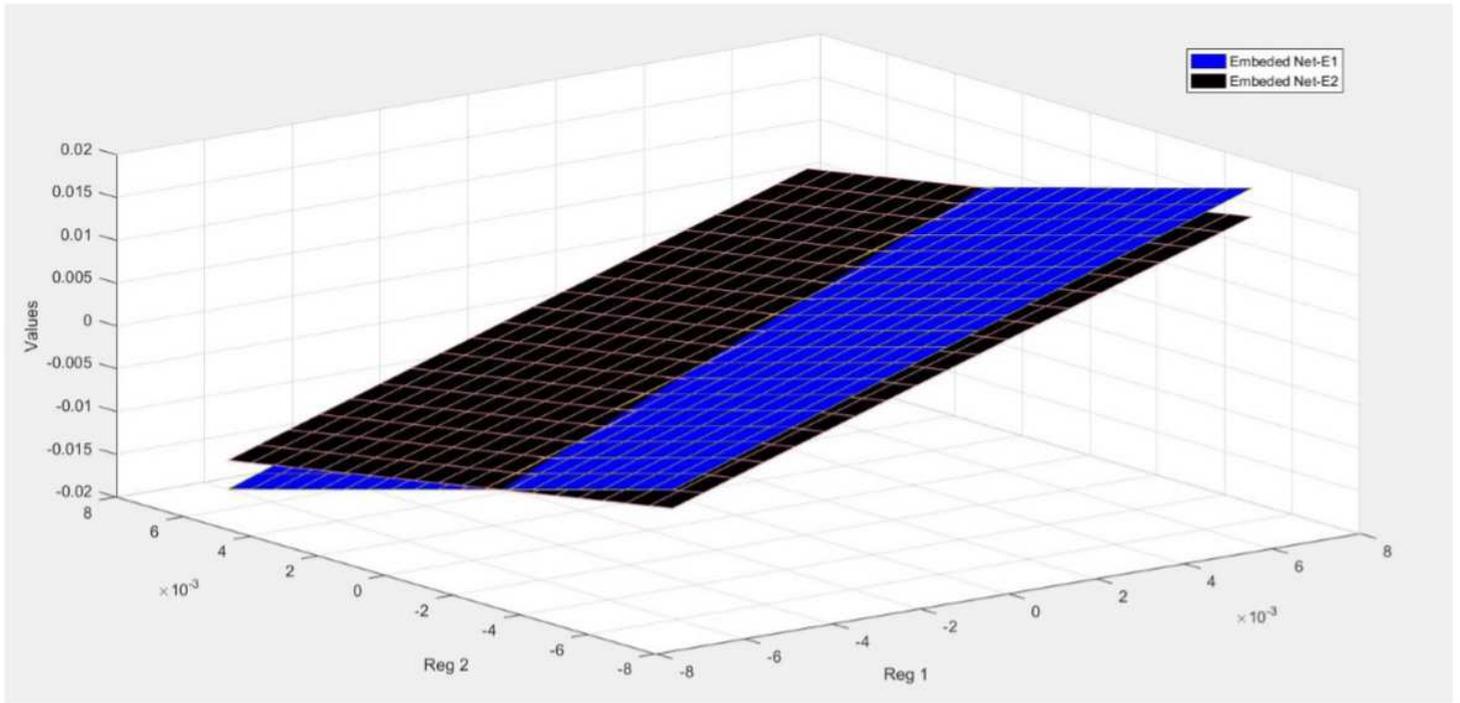


Figure 4

Represent the embedding dimension useful for deterministic component.

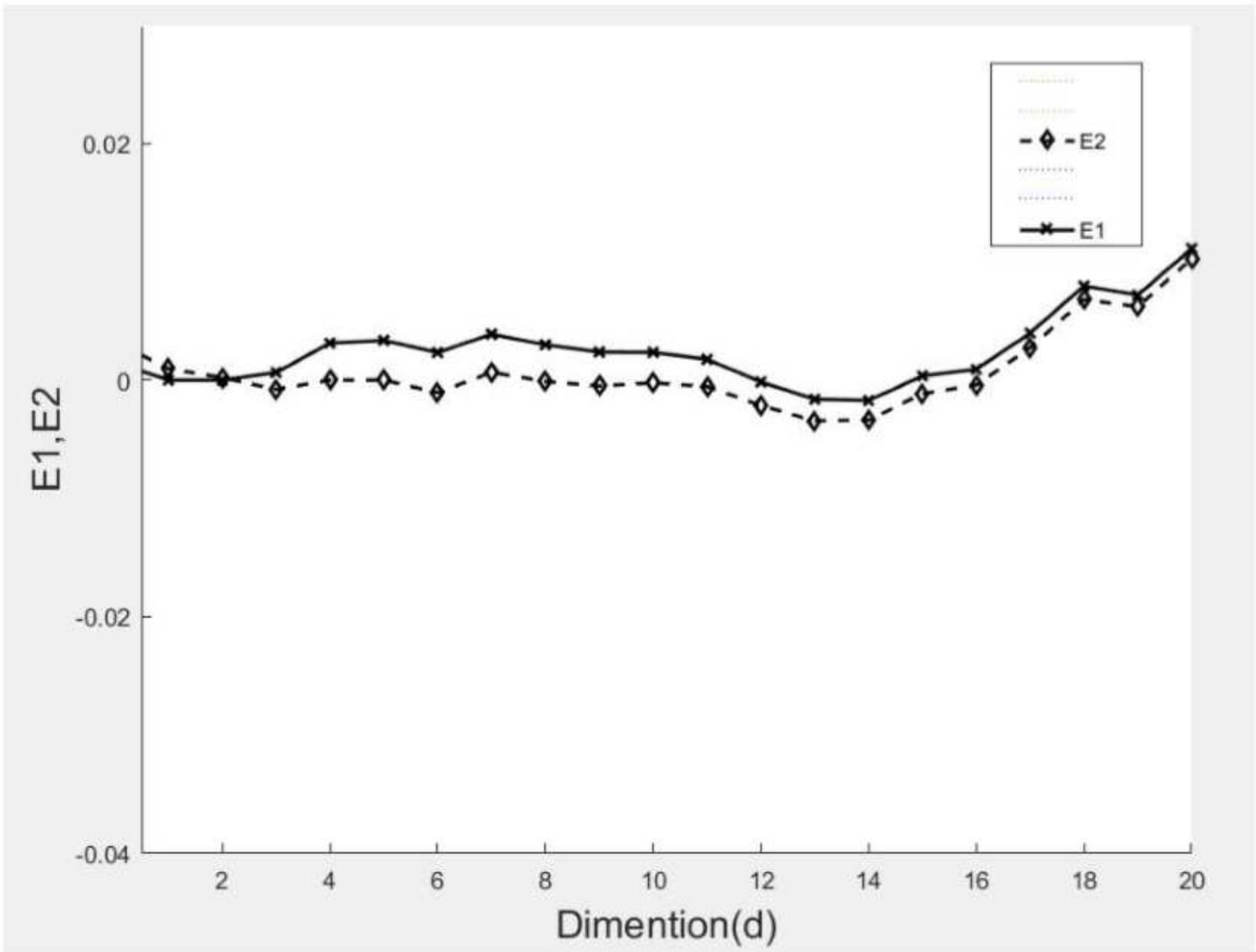
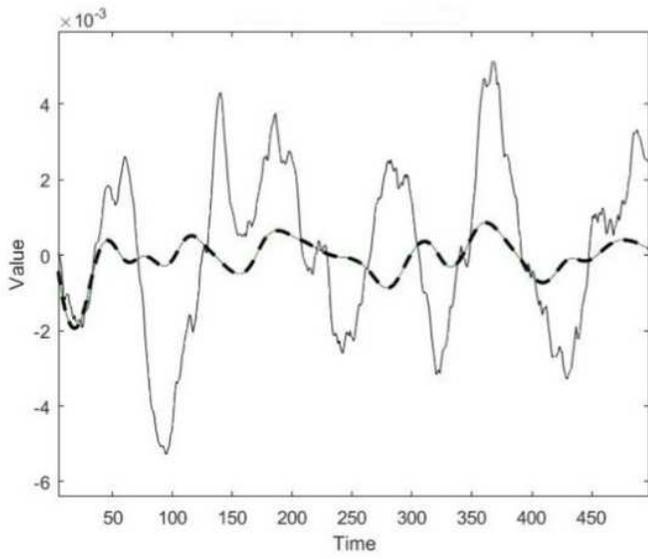
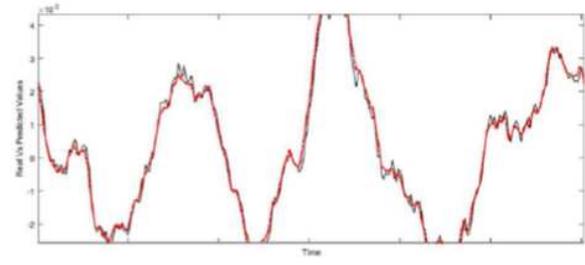


Figure 5

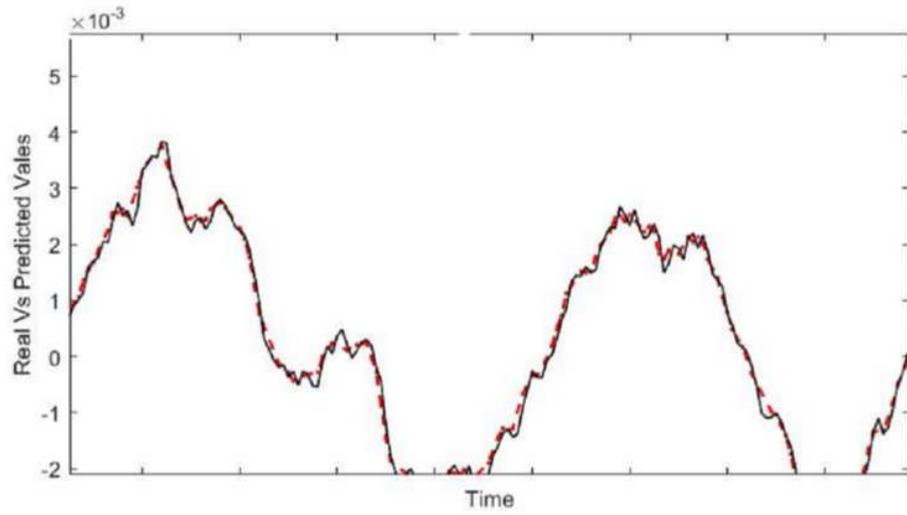
The value E1 and E2 of deterministic component.



ARMX model Prediction



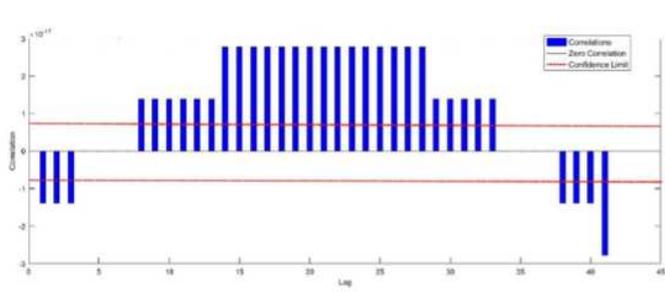
NARX model Prediction



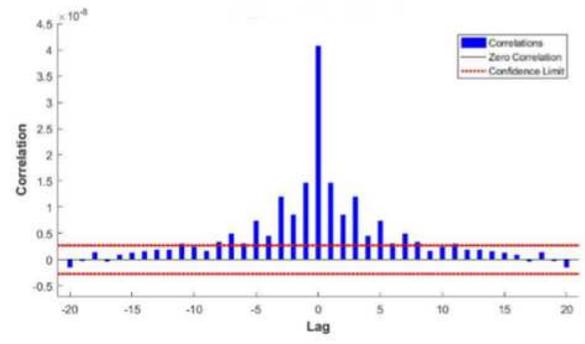
ARNN model prediction (Proposed)

Figure 6

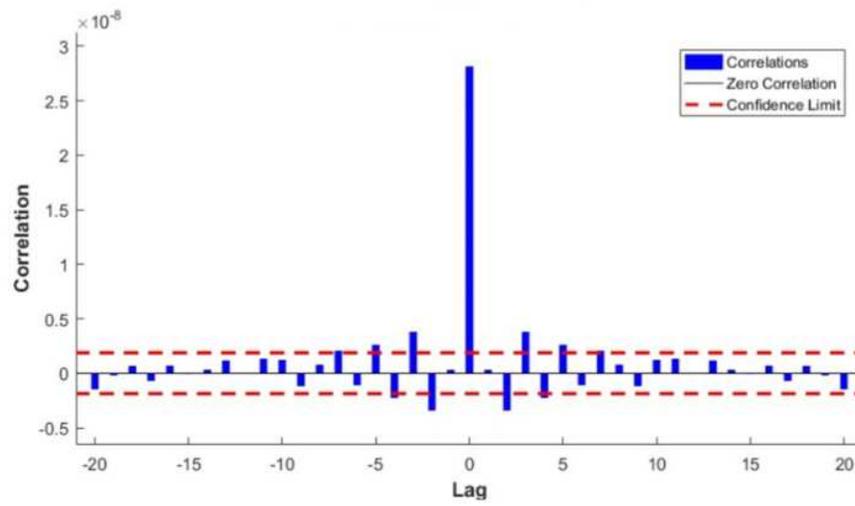
Compression of Prediction responses



ARMX error



NARX error



ARNN error (Proposed)

Figure 7

Prediction Histogram error compression

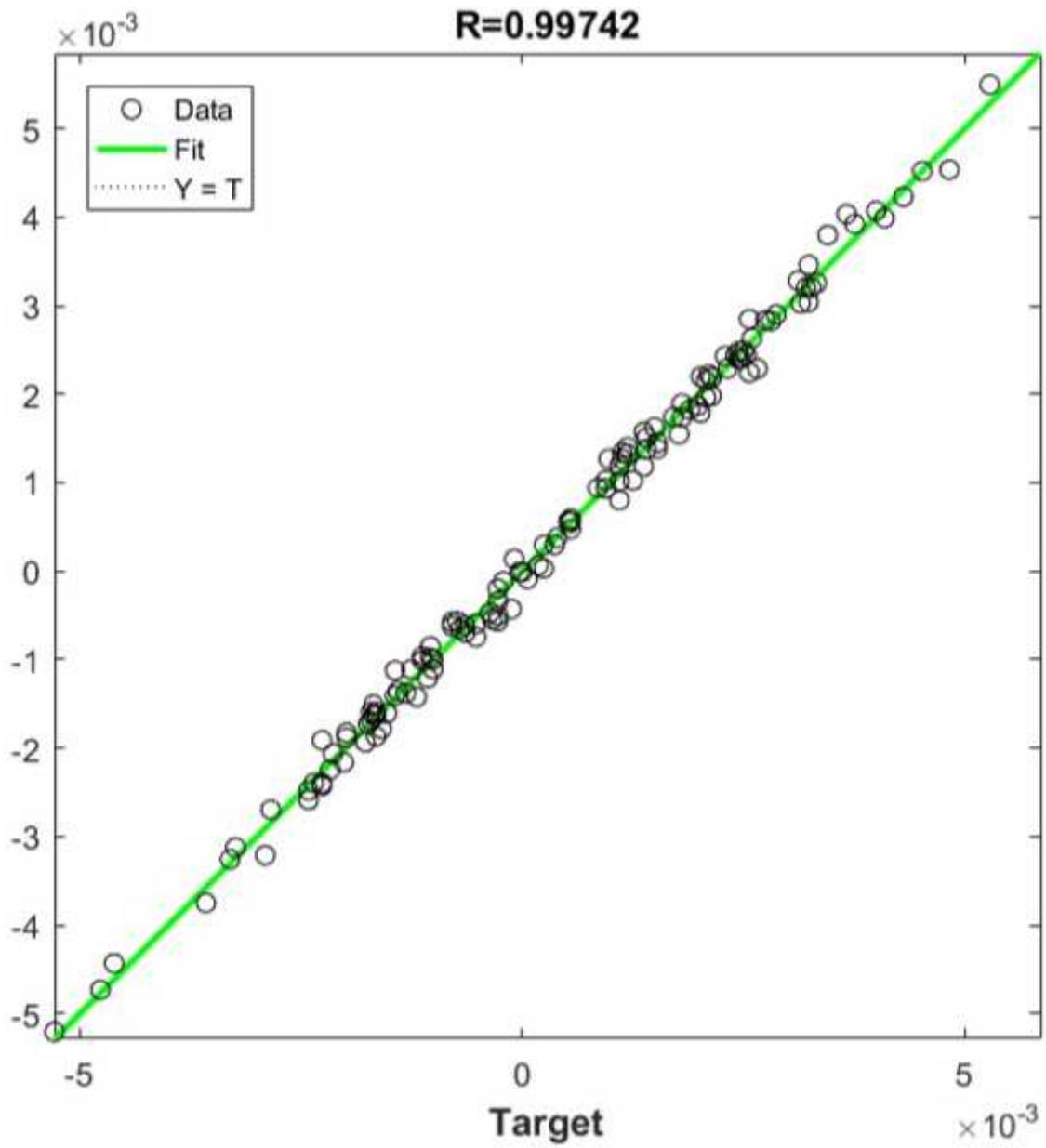


Figure 8

Regression value for Prediction of time series data.