sciendo

Int. J. Appl. Math. Comput. Sci., 2021, Vol. 31, No. 3, 387–398
DOI: 10.34768/amcs-2021-0026

amcs

FORECASTING MODELS FOR CHAOTIC FRACTIONAL–ORDER OSCILLATORS USING NEURAL NETWORKS

KISHORE BINGI [a,*], B RAJANARAYAN PRUSTY [a]

[a]School of Electrical Engineering
Vellore Institute of Technology
Tiruvalam Rd, Katpadi, Vellore, 632014, Tamil Nadu, India
e-mail: {bingi.kishore,b.r.prusty}@ieee.org

This paper proposes novel forecasting models for fractional-order chaotic oscillators, such as Duffing's, Van der Pol's, Tamaševičius's and Chua's, using feedforward neural networks. The models predict a change in the state values which bears a weighted relationship with the oscillator states. Such an arrangement is a suitable candidate model for out-of-sample forecasting of system states. The proposed neural network-assisted weighted model is applied to the above oscillators. The improved out-of-sample forecasting results of the proposed modeling strategy compared with the literature are comprehensively analyzed. The proposed models corresponding to the optimal weights result in the least mean square error (MSE) for all the system states. Further, the MSE for the proposed model is less in most of the oscillators compared with the one reported in the literature. The proposed prediction model's out-of-sample forecasting plots show the best tracking ability to approximate future state values.

Keywords: chaotic oscillators, data-driven forecasting, fractional-order systems, model-free analysis, neural networks, time-series prediction.

1. Introduction

1.1. Context of research. The study of the chaotic behavior of nonlinear dynamic systems such as oscillators has been a popular research interest (Liang *et al*., 2020; Wang *et al*., 2020; Corinto *et al*., 2021). Extensive investigation of some popular oscillators proposed by Chua (Petras, 2010), Duffing (Luo and Cui, 2020; Kabziński, 2018), Lorenz (Kanchana *et al*., 2020), Rayleigh (Pan and Das, 2018), Van der Pol (Giresse and Crépin, 2017), or Tamaševičius (Ueta and Tamura, 2012) help mimic dynamic characteristics embedded in physical systems. Van der Pol's and Duffing's oscillators are chaotic oscillators having two states. The former can be modeled with a single parameter and is considered the simplest chaotic oscillator ever. On the other hand, the latter completely characterizes the aperiodic nonlinear behavior of physical systems with five model parameters. Three-state Tamaševičius and non-linear resistor-based Chua oscillators are also widely used in chaotic behavior studies.

As this research aims to start with lower order systems, Duffing's, Van der Pol's, Tamaševičius's and Chua's oscillators are considered. In general, the above oscillators act as the benchmark for models developed for mimicking various physical systems. For instance, an elastic beam's dynamic behavior (Cao *et al*., 2010), a particle in plasma (Miwadinou *et al*., 2015) and forced double-well (Sun *et al*., 2006) exhibit similar behavior to Duffing's oscillator. Although the above oscillators' chaotic behavior is bounded, it is challenging to model their behavior as they are aperiodic and susceptible to initial conditions (Vaidyanathan and Azar, 2020; Azar and Vaidyanathan, 2015).

For a chaotic behavior study, a practical model design for the oscillator is imperative. The realistic modeling features supported by the fractional-order concept (Petrávš, 2011; Bingi *et al.*, 2020; 2019; Mainardi, 2018; Kaczorek and Sajewski 2020) provide more degrees of freedom in deciding the orders and hence are adopted in this paper. Owing to the oscillators' aperiodic nature, a developed prediction model for its chaotic behavior assesses the system's

*Corresponding author

stability beforehand. Therefore, this paper's research is aimed at creating a model for predicting a fractional-order oscillator's chaotic behavior. Henceforth, the word "oscillator" used in the subsequent sections refers to a fractional-order oscillator.

1.2. Literature review. Over the years, researchers have developed several fractional-order models for chaotic oscillators (Petráš, 2011; Bingi *et al*., 2020; Cattani *et al*., 2015). For instance, the Van der Pol oscillator's behavior has been used to study the real phenomena of heartbeat and neurons (Shen *et al*., 2014; Kuiate *et al*., 2018). Similarly, in the study of vibration analysis of heavy machinery (Cao *et al*., 2010), Duffing's oscillator is used. The behaviour of Chua's and Tamaševičius's oscillators has been used respectively to create intelligent autonomous mobile chaotic robots (Zang *et al*., 2016) and obtain secure communications in wireless networks (Ueta and Tamura, 2012). Their chaotic behavior is obtained through numerical solutions of the governing equations with assumed initial conditions (Bingi *et al.*, 2019; 2020; Petráš, 2011; Salas and El-Tantawy, 2021).

To list a few, the numerical solution methods are Cauchy, Grünwald–Letnikov, Riemann–Liouville (De Oliveira and Tenreiro Machado, 2014). However, applying such an approach to obtain system states' values for a desired future time frame is computationally burdensome, further affected by the step size (Petráš, 2011; Bingi *et al*., 2020). In this note, a data-driven approach (Lu *et al*., 2017) is a choice which has been explored recently. The data-driven approach is limited in application to Lorenz's chaotic system. In the work of Lu *et al*. (2018), a short-term prediction model is successfully developed for attractor reconstruction. A recurrent neural network was applied in Vlachas *et al*. (2018) and claimed to have better performance than Gaussian process regression. Besides, a hybrid strategy, named the self-constructing recurrent fuzzy neural network proposed by Li and Lin (2016) achieved superior prediction accuracy. Another strategy that hybridizes deep neural networks and differential evolution effectively characterized the spatio-temporal correlations and improved the convergence speed (Huang *et al*., 2020). Though generative adversarial networks used in Wu *et al*. (2020) have better performance, such a network's training is challenging compared with the neural network-based approach. In the above studies, the data-driven approaches are found to be time-efficient. Further, a less informative process with complete data suffices to establish an overall better prediction performance using data-driven models (Lu *et al.*, 2017; 2018; Vlachas *et al.*, 2018; Yang *et al.*, 2020).

1.3. Research gaps and contributions. From the above literature review, the following research gaps are worth highlighting for neural network-based data-driven prediction models:

- The neural network-based models used in the literature predicted fewer system states requiring multiple analysis techniques to determine all states.
- Neural network-based models cannot forecast outputs at a future time if the information about future values of inputs is absent. This necessitates an extra modeling strategy for generating suitable future inputs from the present output for out-of-sample predictions.
- A higher-order system that takes huge computational time to ascertain all the states' chaotic behavior using existing numerical models fails to assess system stability without prolonged simulation.

In this paper, feedforward neural network (FNN)-based prediction models are developed for fractional-order oscillators proposed by Duffing, Van der Pol, Tamaševičius, and Chua, which address the above limitations. The proposed models predict all the states of the oscillators through a single trained FNN algorithm. The models are designed to predict a change in the state values, thereby not letting the system lead the unstable region because the states' movement decides on the system stability. In the absence of training data, the system's governing equations can generate required initial training samples. For the neural network training, the Levenberg–Marquardt algorithm is used because of its capability to yield a robust optimal solution (Smith *et al*., 2018). In this paper, the selected activation functions for the hidden and output layer are respectively *tansig* and *purelin*; such a selection is usual in the literature. The nonlinear *tansig* function is more efficient because of its more comprehensive range for fast learning (Abdullah *et al*., 2019). Further, it is proven in the literature that effective training necessitates a linear and nonlinear combination of activation functions (Abdullah *et al*., 2019). And it is customary to select the output layer activation function as a linear one, so it is adopted in this research.

1.4. Paper organization. The rest of the paper is organized as follows. Section 2 presents the chaotic behavior of fractional-order oscillators by neatly describing the governing equations and their numerical solution using Grünwald–Letnikov's definition. Section 3 demonstrates the development of the proposed modeling steps using the FNN. Section 4 discusses the analysis of various results in a comprehensive manner. Finally, Section 5 summarizes the paper.

2. Fractional-order oscillators and their characteristics

This section elaborates modeling strategies adopted for the oscillators considered, followed by a detailed analysis of their chaotic behaviors.

2.1. Modeling and numerical solutions of fractional-order oscillators. The governing equations for all the four oscillators considered are discussed in integer and fractional-order forms in Fig. 1. The oscillators' states are represented as x, y, and z. For each oscillator described in Fig. 1, the model parameters are highlighted. Detailed modeling and understanding of these parameters' significance can be found in the works of Petráš (2011; 2010), Ueta and Tamura (2012), Shen *et al*. (2014), and Bingi *et al*. (2020). Further, the equations for their numerical solution are also included. The numerical simulation helps obtain the initial training data for the proposed prediction models in the absence of any training data. The equations for numerical solutions are obtained from their corresponding fraction-order model based on Grünwald–Letnikov's definition.

According to Grünwald and Letnikov (De Oliveira and Tenreiro Machado, 2014), the fractional-order derivative of a function $f(t)$ for the order q $(0 \leq q \leq 1)$ is defined as

$$\frac{\mathrm{d}^q}{\mathrm{d}t^q}(f(t)) \approx \frac{1}{\mathrm{h}^q}\sum_{i=0}^{\mathrm{T/h}} c_i^q f(t - j\mathrm{h}), \tag{1}$$

where "h" is the step size, "T" is the simulation time, the value of coefficient c_0^q is one and the other binomial coefficients are calculated as

$$c_i^q = \left(1 - \frac{q+1}{i}\right)c_{i-1}^q, \quad i = 1, 2, \ldots, \frac{\mathrm{T}}{\mathrm{h}}. \tag{2}$$

For all the oscillators, the values of "T" and "h" are kept constant and are respectively set to 200 and 0.005. The above selection of a larger "T" value ensures the visualization of all the oscillators' chaotic behavior for a longer time span. The chosen smaller value of "h" yields smoother chaotic attractors for all cases considered. Hence, the total number of samples "T/h" for all the numerical solutions is 40,000. The counting variable k in Fig. 1 ranges from 1 to 40,000.

2.2. Study of oscillators' chaotic behavior. The motivation here is to observe the studied oscillators' chaotic behavior. In this note, a detailed analysis is devoted to Duffing's oscillator, along with a brief note on similar observations in the case of the remaining oscillators.

Firstly, for Duffing's oscillator model parameters $\delta = 0.15$ Ns/m, $\rho = -1$ N/m, $\mu = 1$ N/m^3, $\lambda = 0.3$ N, and $\omega = 1$ rad/s (Petráš, 2011), four different cases for different fractional orders and initial conditions are examined in Fig. 2. It is evident from the attractors that, though the behavior is bounded, it is aperiodic. It is to be noted from Figs. 2(a) and (b) that a slight change in fractional order values causes a dramatic change in the attractor behavior. This indicates higher flexibility in characterizing the chaotic behavior of different physical systems. Further, it can be noticed from Figs. 2(c) and (d) that, even with a small change in the initial conditions, the attractor is significantly affected, attesting that the oscillator's behavior is sensitive to initial conditions.

Secondly, the attractors of the Van der Pol oscillator with the model parameter $\epsilon = 1$, initial conditions $x(0) = 0.2$, and $y(0) = -0.2$ (Petráš, 2011) with different randomly set values of fractional orders are examined in Fig. 3. The attractor behavior is bounded and aperiodic for the orders less than or greater than one.

Finally, the bounded and aperiodic attractor chaotic behavior is also observed in the case of the remaining two oscillators. A few specific cases with differently set fractional orders are elucidated underneath. The attractor of the Tamaševičius oscillator with the model parameters $a = 0.705$, $b = 20$, $c = 4 \times 10^{-9}$ and $d = 0.13$, initial conditions $x(0) = y(0) = z(0) = 0.1$ (Ueta and Tamura, 2012) and fractional-orders $\alpha_3 = 0.985$, $\beta_3 = 0.995$ and $\gamma_3 = 0.975$ is shown in Fig. 4. Further, the attractor of Chua's oscillator with the model parameters $p = 10.725$, $q = 10.593$, $r = 0.268$, $m_0 = -0.7872$ and $m_1 = -1.1726$, initial conditions $x(0) = 0.6$, $y(0) = 0.1$ and $z(0) = -0.6$ (Petráš, 2011), fractional-orders $\alpha_4 = 0.92$, $\beta_4 = 0.99$ and $\gamma_4 = 0.93$ is shown in Fig. 5, highlighting an interesting double-scroll attractor behavior.

3. Proposed modeling strategy

In this section, the modeling steps of the proposed prediction models for two-state and three-state oscillators are comprehensively detailed, followed by a brief note on the performance assessment metrics.

3.1. Development of the prediction model. The proposed prediction models' inputs are the states of the oscillators, and the outputs correspond to a change in the values of the states. The data set is divided into training, validation, and testing subsets with a data division of 60%:20%:20%. This paper adopts an FNN model for future predictions. Time series data of oscillators' system states is the primary requirement for FNN training. The required training data is obtained by a numerical solution of the system's governing equations using Grünwald–Letnikov's definition. These numerical data

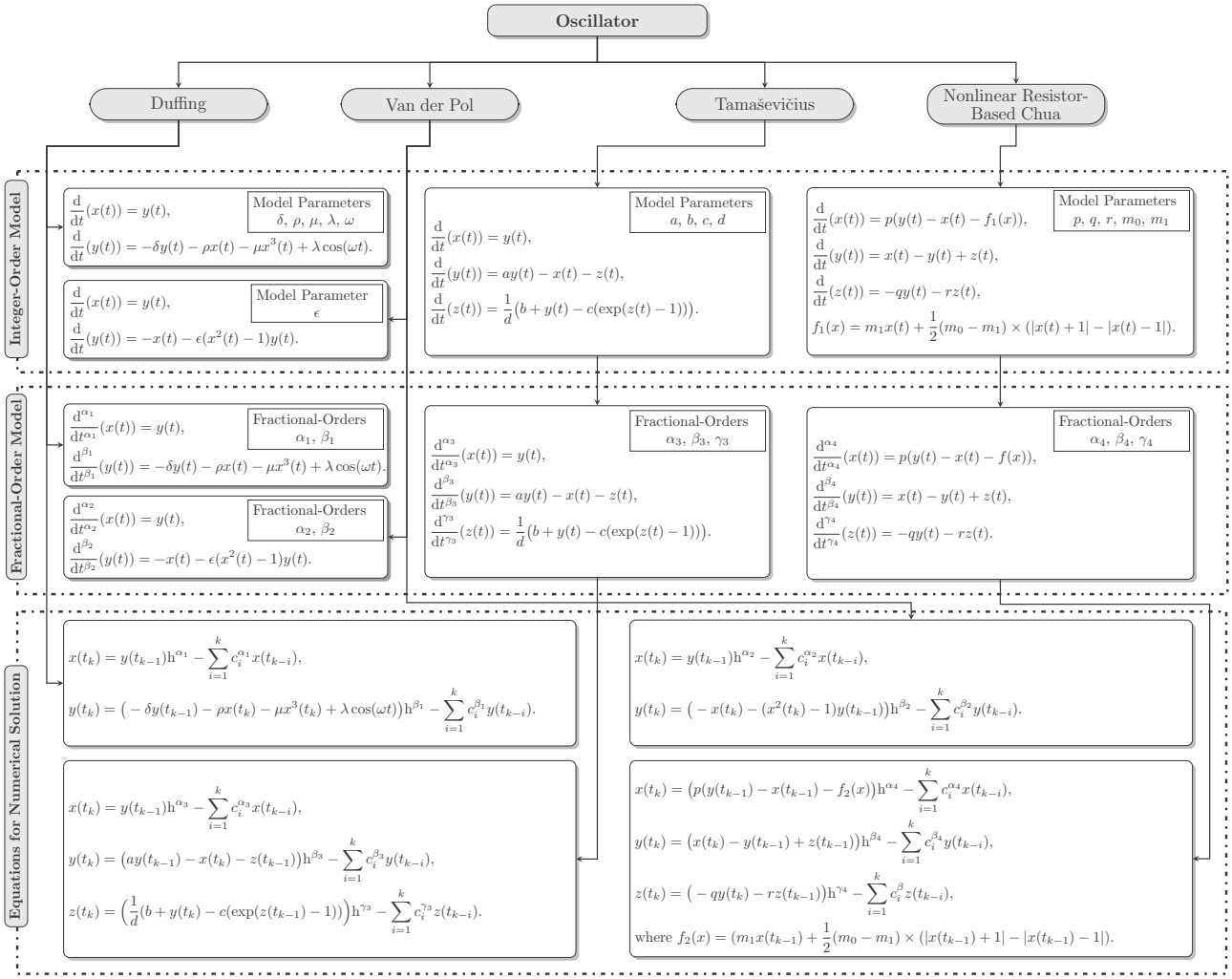


Fig. 1. Pictorial display of modeling steps and numerical simulations for the fractional-order oscillators under study.

are considered training data for developing the proposed prediction model. Thus, the input data for training are obtained using the steps suggested in Section 2.1. The output data for training can be obtained from the inputs by using the following proposed formulations:

- for two-state oscillators,

$$\begin{aligned}\Delta x(t) &= w\big(x(t+\mathrm{h})-x(t)\big)\\ &\quad+(1-w)\big(y(t+\mathrm{h})-y(t)\big),\\ \Delta y(t) &= w\big(y(t+\mathrm{h})-y(t)\big)\\ &\quad+(1-w)\big(x(t+\mathrm{h})-x(t)\big);\end{aligned}\tag{3}$$

- for three-state oscillators,

$$\begin{aligned}\Delta x(t) &= w\big(x(t+\mathrm{h})-x(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(y(t+\mathrm{h})-y(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(z(t+\mathrm{h})-z(t)\big),\\ \Delta y(t) &= w\big(y(t+\mathrm{h})-y(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(x(t+\mathrm{h})-x(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(z(t+\mathrm{h})-z(t)\big),\\ \Delta z(t) &= w\big(z(t+\mathrm{h})-z(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(x(t+\mathrm{h})-x(t)\big)\\ &\quad+\Big(\frac{1-w}{2}\Big)\big(y(t+\mathrm{h})-y(t)\big).\end{aligned}\tag{4}$$

A generalized FNN model of the 3:6:3 structure is shown in Fig. 6. The number of neurons in the hidden layer, $\mathrm{N_h}$, is obtained using the thumb rule as stated by Sheela and Deepa (2013) which is given as

$$\mathrm{N_h}=\frac{4.5\times \mathrm{N_i}}{\mathrm{N_i}-1},\tag{5}$$

where $\mathrm{N_i}$ is the number of input nodes.

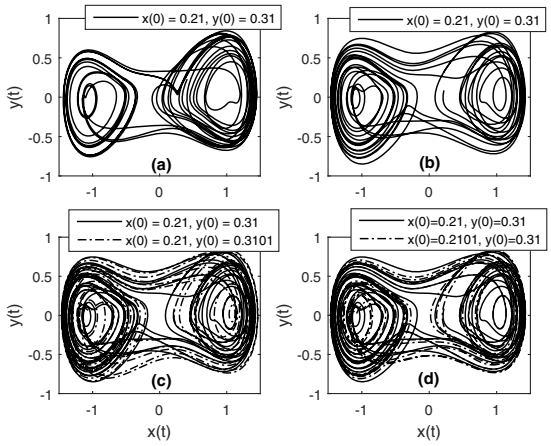


Fig. 2. Chaotic behaviors of the fractional-order Duffing oscillator for $\alpha_1 = \beta_1 = 0.95$ in (a) and $\alpha_1 = 0.97$, $\beta_1 = 0.96$ in (b), (c) and (d).

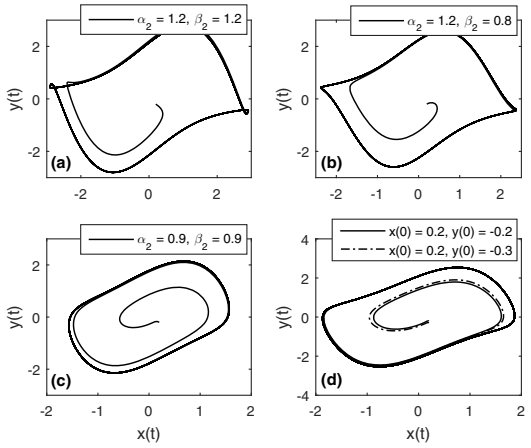


Fig. 3. Chaotic behaviors of the fractional-order Van der Pol oscillator for $x(0) = 0.2$ and $y(0) = -0.2$.

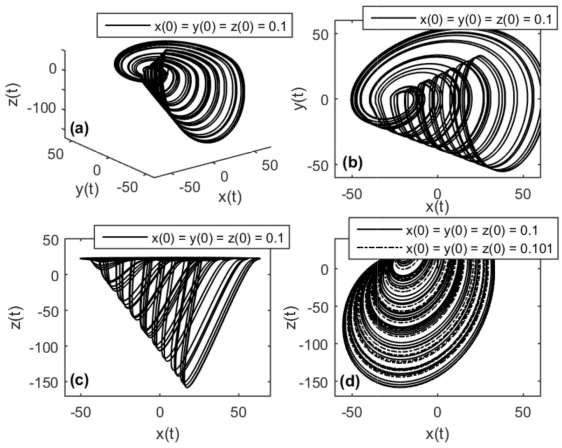


Fig. 4. Chaotic behavior of the fractional-order Tamaševičius oscillator for $\alpha_3 = 0.985$, $\beta_3 = 0.995$ and $\gamma_3 = 0.975$ in (a), (b), (c) and (d).

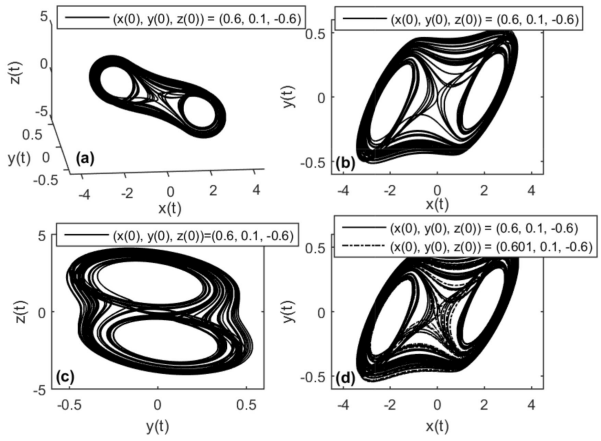


Fig. 5. Chaotic behavior of the fractional-order Chua oscillator for $\alpha_4 = 0.92$, $\beta_4 = 0.99$ and $\gamma_4 = 0.93$ in (a), (b), (c) and (d).

The activation function for the hidden layer is chosen to be the hyperbolic tangent sigmoid (*tansig*) function. This function is more efficient because of its wider range for fast learning. To have a nonlinear and linear combination of activation functions for weight updating during the training process, the output layer's activation function is chosen as a pure linear (*purelin*) function.

The Levenberg–Marquardt algorithm is used for training the proposed prediction model. The algorithm uses a gradient descent approach to find the initial guess, which is relatively close to the optimal solution. Then, the Gauss–Newton method with this initial guess can find the potential solution area and then the final robust optimum solution (Shaik *et al.*, 2021; Smith *et al.*, 2018). The weight update rule using the Levenberg–Marquardt algorithm is given as follows:

$$W_{n+1} = W_n - \left[H - l_p I\right]^{-1} g, \tag{6}$$

where W_n is the current weight calculated using the Gauss–Newton method, W_{n+1} is a new weight calculated using the gradient descent approach, I is the identity matrix, and l_p is the learning parameter. Further, the gradient vector, g, and the approximated Hessian matrix, H, of (6) are expressed as

$$g = J^T e, \tag{7}$$

$$H \approx J^T J, \tag{8}$$

where J is the Jacobian matrix and e is the cumulative error vector.

The model is trained using training and validation subsets for the update of weight w of (3) and (4). The

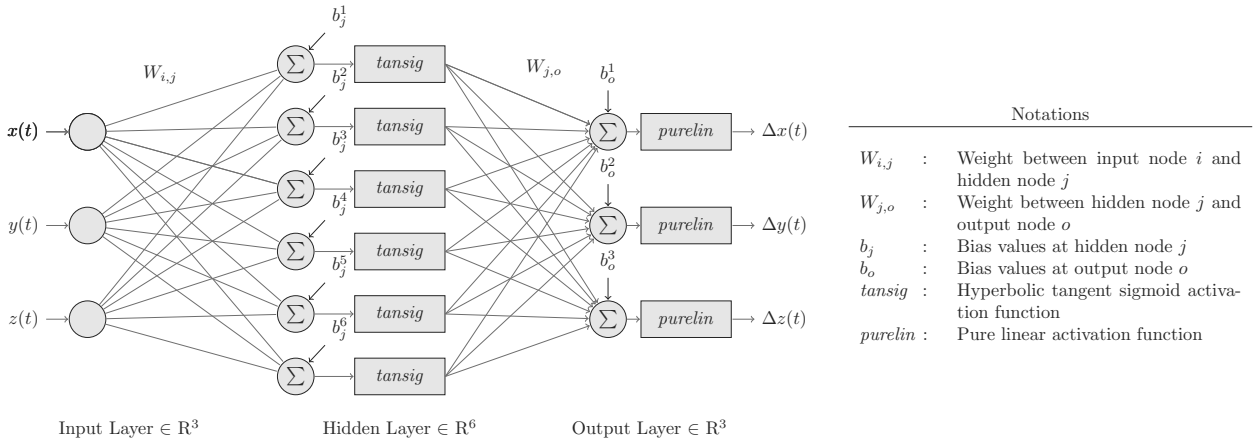


Fig. 6. Structure of the neural network model with six hidden layer nodes.

value of hyperparameter w is selected on a trial and error basis corresponding to the minimum mean square error (MSE) during the validation. The MSE is calculated as (Bingi *et al.*, 2021)

$$\text{MSE} = \frac{1}{\text{m}} \sum_{i=1}^{\text{m}} (Y_{\text{Act},i} - Y_{\text{Pred},i})^2. \tag{9}$$

In (9), "m" is number of samples, Y_{Act} stands for the actual outputs and Y_{Pred} denotes the predicted outputs by the prediction model.

During out-of-sample testing, at time "t + h" the values of the states are obtained using (i) suitable FNN inputs at time t and (ii) a change in the values of states at time t, referred to as the state reconstruction approach (SRA) as discussed underneath:

- for two-state oscillators,

$$\begin{bmatrix} x(t+h) \\ y(t+h) \end{bmatrix} = \begin{bmatrix} w & 1-w \\ 1-w & w \end{bmatrix}^{-1} \times \begin{bmatrix} \Delta x(t) + wx(t) + (1-w)y(t) \\ \Delta y(t) + wy(t) + (1-w)x(t) \end{bmatrix}, \tag{10}$$

- for three-state oscillators,

$$\begin{bmatrix} x(t+\text{h}) \\ y(t+\text{h}) \\ z(t+\text{h}) \end{bmatrix} = \begin{bmatrix} w & \frac{1-w}{2} & \frac{1-w}{2} \\ \frac{1-w}{2} & w & \frac{1-w}{2} \\ \frac{1-w}{2} & \frac{1-w}{2} & w \end{bmatrix}^{-1} \times \begin{bmatrix} \Delta x(t) + wx(t) + \left(\frac{1-w}{2}\right)y(t) + \left(\frac{1-w}{2}\right)z(t) \\ \Delta y(t) + wy(t) + \left(\frac{1-w}{2}\right)x(t) + \left(\frac{1-w}{2}\right)z(t) \\ \Delta z(t) + wz(t) + \left(\frac{1-w}{2}\right)x(t) + \left(\frac{1-w}{2}\right)y(t) \end{bmatrix}, \tag{11}$$

where w is the weight value, $x(t)$, $y(t)$ and $z(t)$ are the state values at time t and $\Delta x(t)$, $\Delta y(t)$ and $\Delta z(t)$ are the values of the change in states at time t.

3.2. Numerical assessment of the prediction model. The performance measures used in this paper are MSE and R^2. The latter is calculated as (Bingi *et al.*, 2021)

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^{\text{m}} (Y_{\text{Act},i} - Y_{\text{Pred},i})^2}{\sum_{i=1}^{\text{m}} (Y_{\text{Act},i} - Y_{\text{Avg},i})^2}, \tag{12}$$

where "m" is number of samples, Y_{Act} stands for the actual outputs, Y_{Avg} denotes the average values of Y_{Act}, and Y_{Pred} marks the predicted outputs by the prediction model.

In (12), R^2 measures the model's predictive ability in fitting the actual data. Thus, the goodness of its fit ranges from zero to one. A value of fit equal to 1.0 indicates a perfect fit. Similarly, a smaller MSE indicates a better estimation of the predicted model.

4. Results and discussions

This section first explains the data preparation steps followed by calculating the proposed models' hyperparameter ω and performance comparison with the literature.

4.1. Data preparation. The oscillators, as discussed in Figs. 2(b), 3(b), 4(a), and 5(a), are taken into consideration to study the proposed prediction models' forecasting accuracy. For all the above cases, the values of "T" and "h" are respectively set to 200 s and 0.005 s. This results in the generation of 40,000 samples, of which the first 24,000 samples (corresponding to 0–120 s) are used for model training, followed by 8,000 samples

Table 1. Comparison of model performance during training and validation.

Oscillator	Performance measure	Training	Validation
Duffing	R^2	0.9674	0.9669
	MSE	0.0037	0.0037
Van der Pol	R^2	0.9920	0.9919
	MSE	0.0009	0.0009
Tamaševičius	R^2	0.9444	0.9428
	MSE	0.0009	0.0009
Chua	R^2	0.9986	0.9987
	MSE	0.0008	0.0008

for validation (corresponding to 120–160 s) and 8,000 samples (160–200 s) for testing. The forecasting models' performance evaluation for the oscillators is carried out comprehensively in the subsequent subsection.

4.2. Selection of the proposed models' hyperparameters. The forecasting models are developed for all four oscillators by selecting a single hidden layer comprising node numbers estimated using (5). The hidden layer nodes are calculated to be nine and seven, respectively, for the two-state and three-state oscillators. The attractiveness in the proposed models is the relationship between the inputs and outputs, which are based on a proper selection of weight (the hyperparameter of the model) as defined in (3) and (4).

The appropriate weight w selected for the proposed models is based on the obtained minimum MSE during validation through a trial-and-error mechanism. The obtained weight values are 0.985, 0.995, 0.986, and 0.805, respectively, for the Duffing, Van der Pol, Tamaševičius and Chua oscillators forecasting models. With this hyperparameter selection, the training and validation performances are summarized in Table 1. In all cases, the R^2 values are close to one and the MSE values are close to zero, indicating excellent training and validation. Further, the comparison of model outputs with the ones produced by numerical simulation during validation for all the four cases is shown in Fig. 7. The results shown in the table and various plots in the figure indicate the proposed forecasting models' superior training and validation performance.

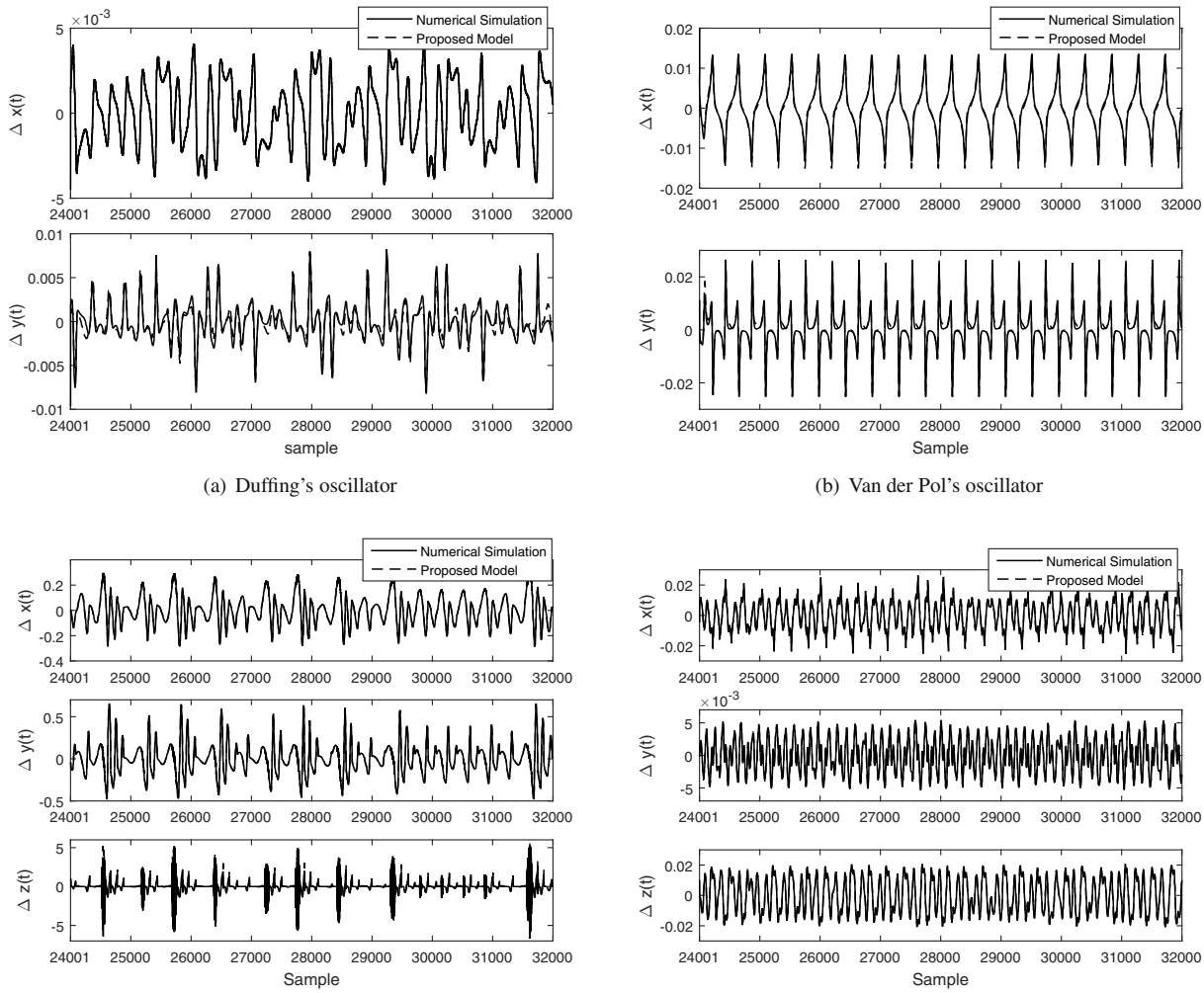
To observe the forecasting results of the proposed models, out-of-sample predictions are performed for the testing period. To obtain a future value of a state at time "t + h", the knowledge of ω and the previous value of the states, i.e., values at time t, is essential. Hence, (10) or (11) can be applied to determine future state values at any given instant of time. Figure 8 indicates the proposed forecasting models' tracking ability in approximating the future values of the states. It is to be noted from the various plots in the figure that any randomly chosen weight leads to erratic prediction behavior, which highlights importance of appropriate weight selection. For all the oscillator cases, plots corresponding to optimal weight provide the improved forecasting performance.

Table 2. Comparison of model performance during out-of-sample prediction with different values of ω.

Oscillator	w	$MSE_{\Delta x}$	$MSE_{\Delta y}$	$MSE_{\Delta z}$
Duffing	1	0.0031	0.2800	–
	0.985	**0.0025**	**0.1329**	–
	0.925	0.0027	0.1441	–
Van der Pol	1	0.7351	0.9336	–
	0.995	**0.0310**	**0.0364**	–
	0.985	0.2053	0.2423	–
Tamaševičius	1	16.6704	7.9626	40.4298
	0.986	**1.9455**	**6.3924**	**16.9050**
	0.970	3.6611	6.6962	16.9551
Chua	1	1.9138	0.0829	2.8564
	0.805	**0.2917**	**0.0200**	**0.4158**
	0.780	3.9071	0.0485	4.0490

Table 2 summarizes the MSE values for the oscillator states. It can be noted from the table that the MSE value corresponding to the optimally set weight is the lowest, i.e., 0.0025, 0.031, 1.9455 and 0.02 are the obtained lowest MSE respectively for Duffing, Van der Pol, Tamaševičius and Chua. It is highlighted here that any small error in the obtained change in state values propagates through the entire out-of-sample predictions, finally resulting in erroneous forecasting of oscillators' states, demanding the estimation of weight with high precision.

4.3. Performance evaluation of the proposed models. The forecasting performance of the proposed SRA-based FNN models is compared with that of the model suggested by Li and Lin (2016) (refer to Fig. 9). Further, the importance of the suggested SRA is highlighted by comparing the proposed approaches' performance for all the oscillators in the plots of Fig. 9 with an arbitrarily chosen SRA. Unlike the proposed SRA, the arbitrarily chosen SRA adds the predicted change in state values to their corresponding immediate past input state values for forecasting future state values. It can be inferred from the comparison of various plots for all the four oscillators that the proposed SRA with the FNN structure yields accurate forecasting results compared with the arbitrarily chosen SRA because of the benefit gained due to the weighted model discussed in Section 3.1. For this comparison, the hyperparameter ω is set to the optimal values as obtained in Section 4.2 for the oscillators. Table 3 summarizes forecasting errors in all the states using the proposed SRA and arbitrary SRA for all the oscillators. As expected, the



(a) Duffing's oscillator

(b) Van der Pol's oscillator

(c) Tamaševičius's oscillator

(d) Chua's oscillator

Fig. 7. Comparison of prediction performance of the proposed models during validation.

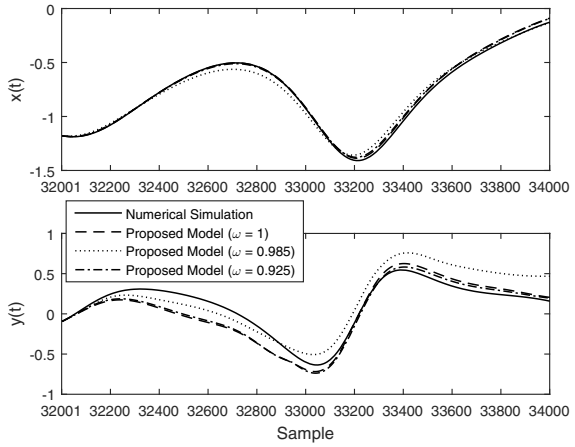
MSE value is the least for the proposed SRA for all the cases considered.

Further, it is worth noting that, in all the oscillators except for the Tamaševičius, the proposed SRA-based forecasting model yields accurate forecasting for all the states compared with the results of Li and Lin (2016). This might be because the proposed model is sensitive to the hyperparameter ω. A slight deviation in a weight from the optimal value causes a drastic change in the forecasting performance, as can be seen from Section 4.2.

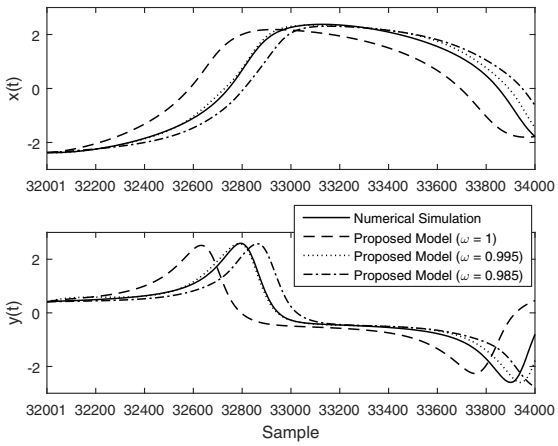
Although the proposed SRA-based FNN is a potential candidate for chaotic time series prediction, it suffers from certain limitations in its present form as elucidated underneath. In the process of the SRA for the three-state oscillators, the weight equal to $(1-\omega)/2$ is imposed on the secondary states (cf. (4)). In reality, the secondary states might have a different level of sensitivity. Therefore, deciding on an appropriate weight value for the secondary states is a challenge. Furthermore, the value of ω is selected on a trail-and-error basis corresponding to the minimum MSE during the validation. The trial and error process takes numerous simulations to determine an optimal value of ω.

5. Conclusion

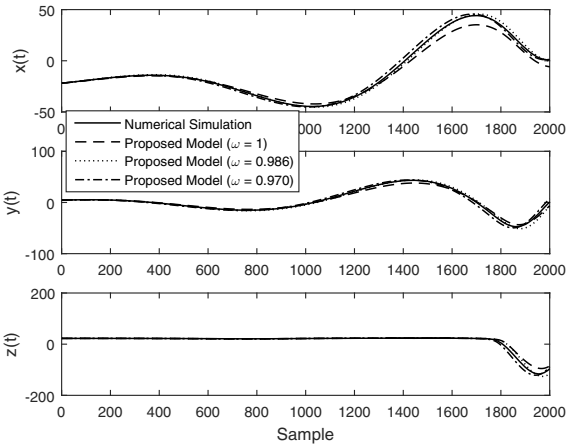
This paper proposed FNN-based forecasting models for out-of-sample prediction of chaotic fractional-order oscillator states using the corresponding predicted values of the change in states. The higher R^2 and lower MSE values during the validation and testing phase assert the proposed predictions models' suitability for forecasting future states of chaotic oscillators. The forecasting models' weight-dependent performance is comprehensively studied and results are compared with a self-constructing recurrent neural network model reported



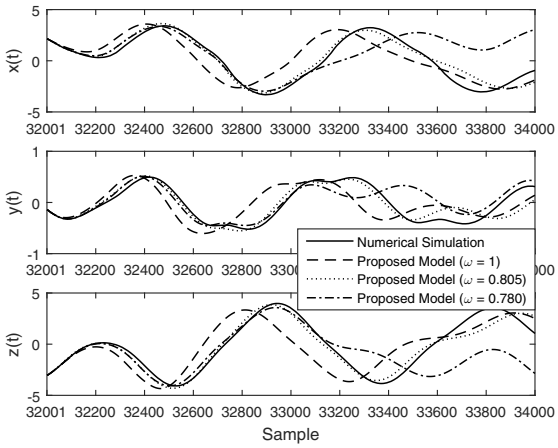
(a) Duffing's oscillator



(b) Van der Pol's oscillator



(c) Tamaševičius's oscillator



(d) Chua's oscillator

Fig. 8. Comparison of out-of-sample forecasting results of the proposed models with different values of ω.

in the literature. The obtained results established that the forecasting performance of the proposed model is outmoded for a specific, appropriately chosen weight value for all four cases.

The proposed modeling framework can be effortlessly extended to any higher-order oscillator with a proper selection of weight. The forecasting performance is susceptible to the weight value, and the choice of an accurate weight value (up to four decimal points) on a trial-and-error basis is computationally cumbersome. The future research scope would aim to develop a variable weight-based model or a fixed weight model with a suitable weight calculation strategy.

References

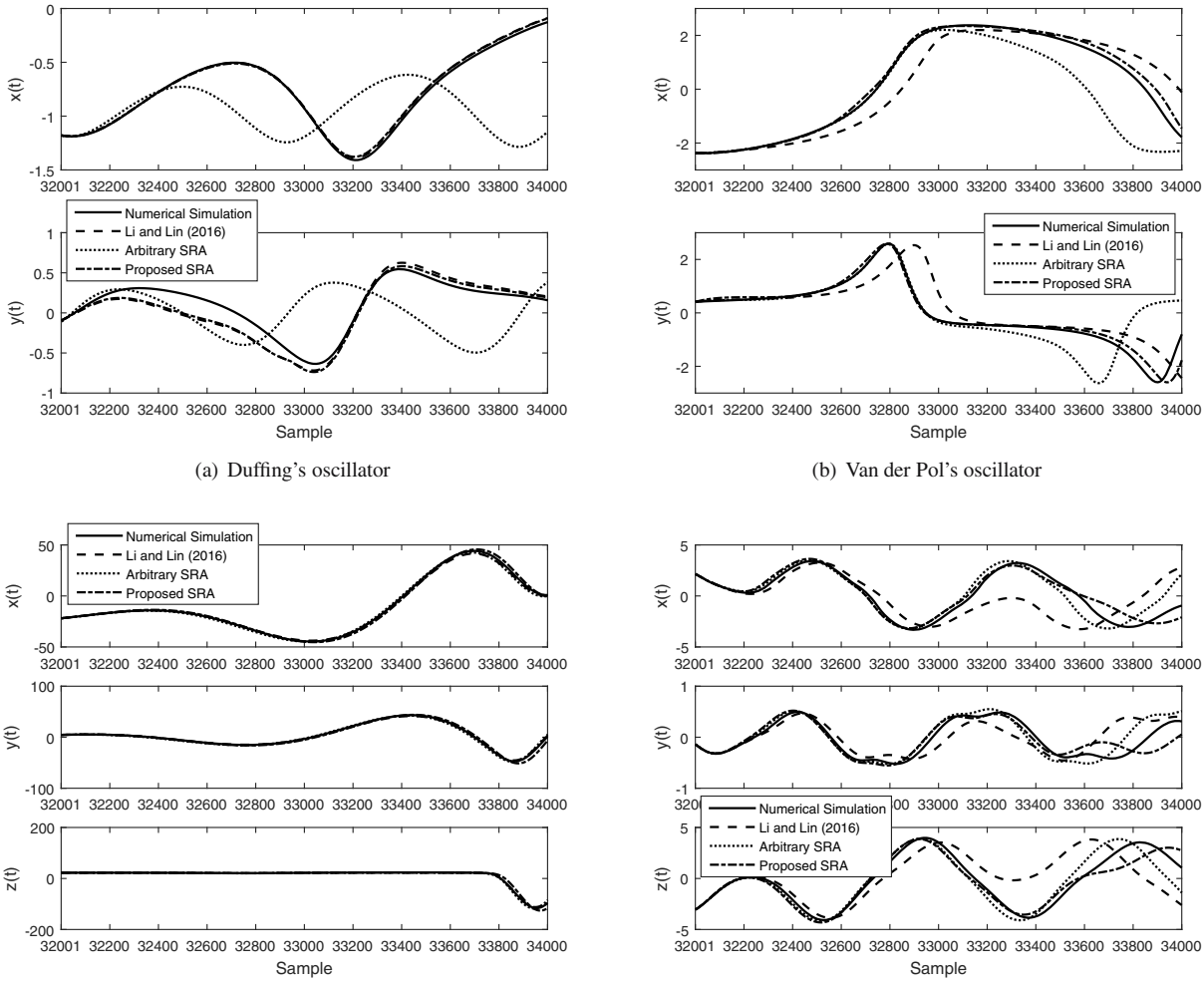
Abdullah, S., Ismail, M., Ahmed, A.N. and Abdullah, A.M. (2019). Forecasting particulate matter concentration using linear and non-linear approaches for air quality decision support, *Atmosphere* **10**(11): 667.

Azar, A.T. and Vaidyanathan, S. (2015). *Chaos Modeling and Control Systems Design*, Springer, Cham.

Bingi, K., Ibrahim, R., Karsiti, M.N., Hassam, S.M. and Harindran, V.R. (2019a). Frequency response based curve fitting approximation of fractional-order PID controllers, *International Journal of Applied Mathematics and Computer Science* **29**(2): 311–326, DOI: 10.2478/amcs-2019-0023.

Bingi, K., Ibrahim, R., Karsiti, M.N., Hassan, S.M., Elamvazuthi, I. and Devan, A.M. (2019b). Design and analysis of fractional-order oscillators using SCILAB, *2019 IEEE Student Conference on Research and Development (SCOReD), Bandar Seri Iskandar, Malaysia*, pp. 311–316.

Bingi, K., Ibrahim, R., Karsiti, M.N., Hassan, S.M. and Harindran, V.R. (2020). *Fractional-order Systems and PID Controllers*, Springer, Cham.



(a) Duffing's oscillator

(b) Van der Pol's oscillator

(c) Tamaševičius's oscillator

(d) Chua's oscillator

Fig. 9. Accuracy comparison of out-of-sample forecasting results of various models.

Bingi, K., Prusty, B.R., Kumra, A. and Chawla, A. (2021). Torque and temperature prediction for permanent magnet synchronous motor using neural networks, *3rd International Conference on Energy, Power and Environment: Towards Clean Energy Technologies, Shillong, Meghalaya, India*, pp. 1–6.

Cao, J., Ma, C., Xie, H. and Jiang, Z. (2010). Nonlinear dynamics of Duffing system with fractional order damping, *Journal of Computational and Nonlinear Dynamics* **5**(4), Article ID: 041012, DOI: 10.1115/1.4002092.

Cattani, C., Srivastava, H.M. and Yang, X.-J. (2015). *Fractional Dynamics*, De Gruyter, Warsaw.

Corinto, F., Forti, M. and Chua, L.O. (2021). *Nonlinear Circuits and Systems with Memristors: Nonlinear Dynamics and Analogue Computing via the Flux-Charge Analysis Method*, Springer, Cham.

De Oliveira, E.C. and Tenreiro Machado, J.A. (2014). A review of definitions for fractional derivatives and integral, *Mathematical Problems in Engineering* **2014**, Article ID: 238459, DOI: 10.1155/2014/238459.

Giresse, T.A. and Crépin, K.T. (2017). Chaos generalized synchronization of coupled Mathieu–Van der Pol and coupled Duffing–Van der Pol systems using fractional order-derivative, *Chaos, Solitons & Fractals* **98**: 88–100, DOI: 10.1016/j.chaos.2017.03.012.

Huang, W., Li, Y. and Huang, Y. (2020). Deep hybrid neural network and improved differential neuroevolution for chaotic time series prediction, *IEEE Access* **8**: 159552–159565, DOI: 10.1109/ACCESS.2020.3020801.

Kabziński, J. (2018). Synchronization of an uncertain Duffing oscillator with higher order chaotic systems, *International Journal of Applied Mathematics and Computer Science* **28**(4): 625–634, DOI: 10.2478/amcs-2018-0048.

Kaczorek, T. and Sajewski, Ł. (2020). Pointwise completeness and pointwise degeneracy of fractional standard and descriptor linear continuous-time systems with different

Table 3. Accuracy comparison of out-of-sample forecasting results of the proposed models with the literature.

Oscillator	Model	$MSE_{\Delta x}$	$MSE_{\Delta y}$	$MSE_{\Delta z}$
Duffing	Li and Lin, 2016	0.0042	0.1970	–
	Arbitrary SRA	0.2371	0.2197	–
	Proposed SRA	**0.0025**	**0.1329**	–
Van der Pol	Li and Lin, 2016	0.4089	0.4027	–
	Arbitrary SRA	0.9398	0.8183	–
	Proposed SRA	**0.0310**	**0.0364**	–
Tamaševičius	Li and Lin, 2016	**1.3835**	**1.0238**	**1.4125**
	Arbitrary SRA	2.7438	6.5161	19.0773
	Proposed SRA	1.9455	6.3924	16.9050
Chua	Li and Lin, 2016	4.2065	0.0624	4.6165
	Arbitrary SRA	0.6467	0.0211	0.7150
	Proposed SRA	**0.2917**	**0.0200**	**0.4158**

fractional orders, *International Journal of Applied Mathematics and Computer Science* **30**(4): 641–647, DOI: 10.34768/amcs-2020-0047.

Kanchana, C., Siddheshwar, P. and Yi, Z. (2020). The effect of boundary conditions on the onset of chaos in Rayleigh–Bénard convection using energy-conserving Lorenz models, *Applied Mathematical Modelling* **88**: 349–366, DOI: 10.1016/j.apm.2020.06.062.

Kuiate, G.F., Kingni, S.T., Tamba, V.K. and Talla, P.K. (2018). Three-dimensional chaotic autonomous Van der Pol–Duffing type oscillator and its fractional-order form, *Chinese Journal of Physics* **56**(5): 2560–2573.

Li, Q. and Lin, R.-C. (2016). A new approach for chaotic time series prediction using recurrent neural network, *Mathematical Problems in Engineering* **2016**, Article ID: 3542898, DOI: 10.1155/2016/3542898.

Liang, Y., Wang, G., Chen, G., Dong, Y., Yu, D. and Iu, H.H.-C. (2020). S-type locally active memristor-based periodic and chaotic oscillators, *IEEE Transactions on Circuits and Systems I: Regular Papers* **67**(12): 5139–5152.

Lu, Z., Hunt, B.R. and Ott, E. (2018). Attractor reconstruction by machine learning, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28**(6): 061104.

Lu, Z., Pathak, J., Hunt, B., Girvan, M., Brockett, R. and Ott, E. (2017). Reservoir observers: Model-free inference of unmeasured variables in chaotic systems, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**(4): 041102.

Luo, W. and Cui, Y. (2020). Signal denoising based on Duffing oscillators system, *IEEE Access* **8**: 86554–86563, DOI: 10.1109/ACCESS.2020.2992503.

Mainardi, F. (2018). *Fractional Calculus: Theory and Applications*, Multidisciplinary Digital Publishing Institute, Basel.

Miwadinou, C., Monwanou, A. and Chabi Orou, J. (2015). Effect of nonlinear dissipation on the basin boundaries of a driven two-well modified Rayleigh–Duffing oscillator, *International Journal of Bifurcation and Chaos* **25**(02): 1550024.

Pan, I. and Das, S. (2018). Evolving chaos: Identifying new attractors of the generalised Lorenz family, *Applied Mathematical Modelling* **57**: 391–405, DOI: 10.1016/j.apm.2018.01.015.

Petras, I. (2010). Fractional-order memristor-based Chua's circuit, *IEEE Transactions on Circuits and Systems II: Express Briefs* **57**(12): 975–979.

Petráš, I. (2011). *Fractional-Order Nonlinear Systems: Modeling, Analysis and Simulation*, Springer Science & Business Media, Berlin.

Salas, A.H. and El-Tantawy, S.A.E.-H. (2021). *Analytical Solutions of Some Strong Nonlinear Oscillators*, IntechOpen, London, DOI: 10.5772/intechopen.97677.

Shaik, N.B., Pedapati, S.R., Othman, A., Bingi, K. and Abd Dzubir, F.A. (2021). An intelligent model to predict the life condition of crude oil pipelines using artificial neural networks, *Neural Computing and Applications*, DOI: 10.1007/s00521-021-06116-1.

Sheela, K.G. and Deepa, S.N. (2013). Review on methods to fix number of hidden neurons in neural networks, *Mathematical Problems in Engineering* **2013**, Article ID: 425740, DOI: 10.1155/2013/425740.

Shen, Y.-J., Wei, P. and Yang, S.-P. (2014). Primary resonance of fractional-order Van der Pol oscillator, *Nonlinear Dynamics* **77**(4): 1629–1642.

Smith, J.S., Wu, B. and Wilamowski, B.M. (2018). Neural network training with Levenberg–Marquardt and adaptable weight compression, *IEEE Transactions on Neural Networks and Learning Systems* **30**(2): 580–587.

Sun, Z., Xu, W., Yang, X. and Fang, T. (2006). Inducing or suppressing chaos in a double-well Duffing oscillator by time delay feedback, *Chaos, Solitons & Fractals* **27**(3): 705–714.

Ueta, T. and Tamura, A. (2012). Bifurcation analysis of a simple 3D oscillator and chaos synchronization of its coupled systems, *Chaos, Solitons & Fractals* **45**(12): 1460–1468.

Vaidyanathan, S. and Azar, A.T. (2020). *Backstepping Control of Nonlinear Dynamical Systems*, Academic Press, Cambridge.

Vlachas, P.R., Byeon, W., Wan, Z.Y., Sapsis, T.P. and Koumoutsakos, P. (2018). Data-driven forecasting of

high-dimensional chaotic systems with long short-term memory networks, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **474**(2213): 20170844.

Wang, X., Jin, C., Min, X., Yu, D. and Iu, H.H.C. (2020). An exponential chaotic oscillator design and its dynamic analysis, *IEEE/CAA Journal of Automatica Sinica* **7**(4): 1081–1086.

Wu, J.-L., Kashinath, K., Albert, A., Chirila, D., Prabhat and Xiao, H. (2020). Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems, *Journal of Computational Physics* **406**, Article ID: 109209, DOI: 10.1016/j.jcp.2019.109209.

Yang, Q., Sing-Long, C. and Reed, E. (2020). Rapid data-driven model reduction of nonlinear dynamical systems including chemical reaction networks using l1-regularization, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30**(5): 053122.

Zang, X., Iqbal, S., Zhu, Y., Liu, X. and Zhao, J. (2016). Applications of chaotic dynamics in robotics, *International Journal of Advanced Robotic Systems* **13**(2): 60.

Kishore Bingi currently works as an assistant professor the Department of Control and Automation, School of Electrical Engineering (SELECT), VIT University, Vellore, India. He obtained his PhD degree in the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Perak, Malaysia, in 2019, his MTech degree in instrumentation and control systems from the National Institute of Technology (NIT), Calicut, Kerala, India, in 2014, and his BTech degree in electrical and electronics engineering from Bapatla Engineering College, Andhra Pradesh, India, in 2012. He worked as a research scientist and postdoctoral researcher in the Institute of Autonomous Systems, Universiti Teknologi PETRONAS, from 2019 to 2020. He also worked with TATA Consultancy Service (TCS) as an assistant systems engineer from 2014 to 2015. His current research interests include chaos prediction, non-linear process modeling, fractional-order control, and optimization.

B Rajanarayayn Prusty presently works as an assistant professor in the School of Electrical Engineering, Vellore Institute of Technology. He obtained his PhD degree from the National Institute of Technology Karnataka, Surathkal, India. He is the recipient of the prestigious POSOCO power system award (PPSA) for 2019 under the doctoral category. He co-authored the textbook entitled *Power System Analysis: Operation and Control* (I.K. International Publishing House). His research interests include time series preprocessing and forecasting, high-dimensional dependence modeling and probabilistic power system analysis.

Received: 17 March 2021
Revised: 5 June 2021
Accepted: 6 July 2021