Analysis of Recurrent Neural Networks for Henon Simulated Time-Series Forecasting

Said Jadid Abdulkadir[1], Hitham Alhussian[1], Ahmed Ibrahim Alzahrani[2]
[1]*Department of Computer and Information Sciences, Universiti Teknologi PETRONAS.*
[2]*Department of Computer Science, King Saud University.*
saidjadid.a@utp.edu.my

***Abstract*—Forecasting of chaotic time-series has increasingly become a challenging subject. Non-linear models such as recurrent neural networks have been successfully applied in generating short term forecasts, but perform poorly in long term forecasts due to the vanishing gradient problem when the forecasting period increases. This study proposes a robust model that can be applied in long term forecasting of henon chaotic time-series whilst reducing the vanishing gradient problem through enhancing the models ability in learning of long-term dependencies. The proposed hybrid model is tested using henon simulated chaotic time-series data. Empirical analysis is performed using quantitative forecasting metrics and comparative model performance on the generated forecasts. Performance evaluation results confirm that the proposed recurrent model performs long term forecasts on henon chaotic time-series effectively in terms of error metrics compared to existing forecasting models.**

***Index Terms*—Chaotic Time-Series; Recurrent Networks; Henon Time-Series.**

I. INTRODUCTION

Chaotic forecasting is a huge problem in many real-world applications. Performance of chaotic time-series models is built on historic data which is used for model training and accuracy is solely applied to validate the performance of any time-series forecasting model.

For more than a decade, statistical models have been utilized in time-series forecasting. However their ability to forecast is limited to the nature of time-series applied which is of linear behavior unlike real world problems which are non-linear. Box-Jenkins [1], Multi Regressions [2-3] and Exponential Smoothing [4-5] are examples of statistical methods applied in time-series forecasting.

Non-statistical models which are non-linear models have also been applied in time-series forecasting. Examples include fuzzy logic [6], genetic algorithm [7], support vector machines [8] and artificial neural networks [9-11].

Forecasting chaotic time-series is classified as long term or short term. In the latter, neural network models have no loop that provides a feedback between the network output and the input regressor [12]. Only actual data samples are used in this model unlike the long-term forecasting model that utilizes the actual and target data as inputs to the model. The long-term forecast has a feedback loop that feeds the input for a number of time steps depending on the horizon set by the forecaster [13]. The input regressor is constituted of replaced actual data from previously forecasted values of the time-series, nonetheless when the time steps tend to infinity the values in the input regressor consist of only estimated time-series values. The infinity process makes long term forecasting a much more complex dynamic modelling process than short term forecasting [14].

Forecaster's aim is to apply various approaches in successfully forecasting the applicable data using legal forecasting policies. "The central idea to successful chaotic time-series forecasting is achieving the best results using minimum required input data and the least complex model" [15]. With a focus on this idea it is evident that as a result of the complex nature of chaotic time-series forecasting, there is need for the application of dynamic forecasting models which would help in estimating future trends and reducing risks of decision making.

The outline of this paper is as follows: the problem statement is highlighted in Section 2. In Section 3, the proposed hybrid recurrent network model is proposed, and Section 4 presents experimental results. Finally, the findings are summarized in the last section.

II. PROBLEM STATEMENT

Recurrent neural networks are very powerful sequence models proposed for modelling time-series, however they do not enjoy widespread use because it is extremely difficult to train them properly due to the vanishing gradient problem [16-19]

In an RNN trained over long sequences (e.g. minimum of 100 time steps) the gradients can easily vanish due to the magnitude of gradients being back propagated through the recurrent layers [16-19] This problem affects the performance of recurrent networks in generating long term forecasts.

In previous models [16-20] this problem still persists which hinders the ability of the models in generating long term forecasts when applied on chaotic cases. Therefore, the proposed hybrid model provides procedures in enhancing learning of long-term dependencies with the aim of reducing the vanishing gradient problem as a result of the network being trained over long sequences (more than 100 time steps); hence generating robust long term forecasts.

This hypothesizes that the new non-linear hybrid model is able to cope with multifactor in chaotic time-series for robust multi-step-ahead forecasts as a result of enhanced learning in long-term dependencies.

III. HYBRID RECURRENT NETWORK MODEL

In the proposed recurrent model, as shown in Figure 1, a time-series filter is added onto the input layer for noise reduction [21]. An unscented kalman filter is used whereby a set of identified points is selected to represent the applied data distribution. To control the distribution, a scaling parameter is used within the filter.

With the dataset, points referred to as sigma points are chosen using the function $y^{(i)} = g(x^{(i)})$. The distribution of the states are obtained using the following:

For $k = 1,2, \dots \dots.., \infty$:

1. Using the state covariance, the ideal number of sigma points are obtained by (γ is a scaling parameter given by):

$$\gamma = \sqrt{N + \lambda}, \tag{1}$$

$$\lambda = \alpha^2 (N + \kappa) - N, \tag{2}$$

where α and κ are tuning parameters. The parameter λ, controls the size of the sigma point distribution.

2. Time-update equations: Using the state-update function, apriori state estimate and apriori covariance transform sigma points using:

$$X^x_{i,k/k-1} = f(X^x_{i,k-1}, X^v_{i,k-1}, u_{k-1}), \tag{3}$$

3. Measurement-update equations: The generated sigma points are used in transformation through the measurement-update function:

$$Y_{i,k/k-1} = h(X^x_{i,k/k-1}, X^n_{k-1}, u_k), \tag{4}$$

The Kalman gain is given by:

$$K_k = P_{x_k y_k} P^{-1}_{\bar{y}_k}, \tag{5}$$

and the Scaled kalman filter estimate and its covariance are generated by:

$$\hat{x}_k = \hat{x}^-_k + K_k (y_k - \hat{y}^-_k), \tag{6}$$

$$P_{x_k} = P^-_{x_k} - K_k P_{\bar{y}_k} K^T_k \tag{7}$$

Filtered chaotic henon data is then fed into the autoregressive recurrent network. In the training process, historic data is split into training and testing i.e. 70-30%. To perform long term forecasting, the 30% of henon data is used in closed loop and there would be no input updates within the network:

$$y(n+1) = f[y(n), \dots, y(n - d_y + 1); \\ u(n), \dots, u(n - d_u + 1)] \tag{8}$$

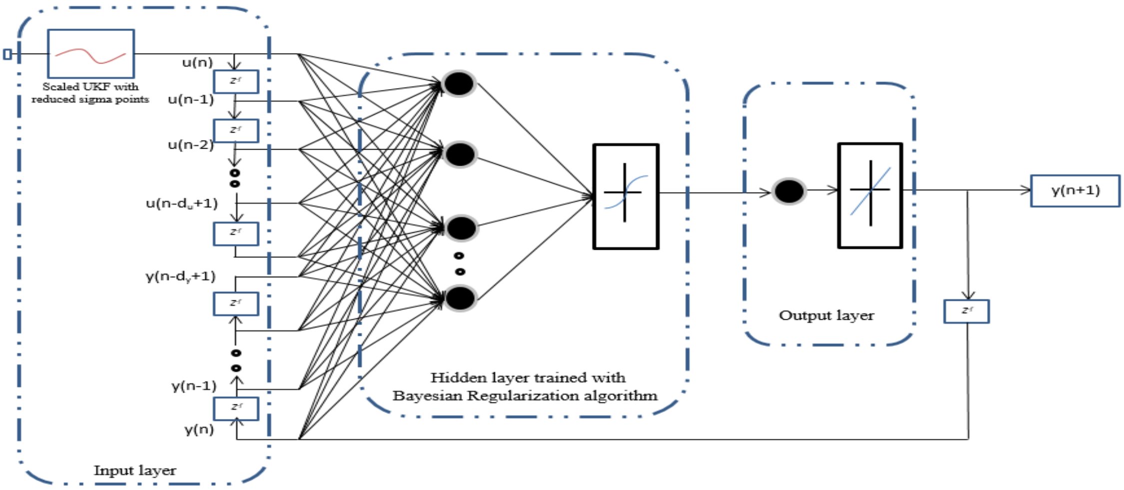

Figure 1: Proposed recurrent hybrid neural network model with time-series filter

A. *Bayes' Training Algorithm*

The training data (chaotic henon) is assigned as $R = (x_i, t_i)$ and $i = 1,2, \dots, n$. Where, n is the number of training samples, w is the parameters within the network, m is the number of parameters and (α, β) are regularization parameters. Given the network framework Q, supervised model is expressed as $y_i = f(x_i, w; Q)$. Without training data, the prior distribution is $P(w \mid \alpha, Q)$, once given sample data R, according to Bayes theorem the posterior distribution $P(w \mid R, \alpha, \beta, Q)$, is written as:

$$P(w \mid R, \alpha, \beta, Q) = \frac{P(R \mid w, \beta, Q)\, P(w \mid \alpha, Q)}{P(R \mid \alpha, \beta, Q)} \tag{9}$$

where, $P(R \mid w, \beta, Q)$ is a likelihood function, $P(R \mid \alpha, \beta, Q)$ is a normalization factor and is expressed as:

$$P(R \mid \alpha, \beta, Q) = \int_{-\infty}^{+\infty} P(R \mid \alpha, \beta, Q)\, P(w \mid \alpha, Q) \tag{10}$$

Prior distribution $P(w \mid \alpha, Q)$ takes index distribution and is expressed as:

$$P(w \mid \alpha, Q) = \frac{1}{Z_W(\alpha)} \exp(-\alpha E_W) \tag{11}$$

Z_W represents the normalization factor if an assumption is made that the prior distribution is Gaussian with zero mean and variance $1/\alpha$.

$$Z_W = \int \exp(-\alpha E_W)\, dw = (\tfrac{2\pi}{\alpha})^{m/2} \tag{12}$$

With respect to the distribution, the probability of occurrence of data is achieved by:

$$P(R|\, w, \beta, Q) = \prod_{i=1}^{n} P(t_i|\, x_i, w, Q) \quad (13)$$

Hence, the error outcome of the network is:

$$e = y(x_i\,|w, Q) - t_i \quad (14)$$

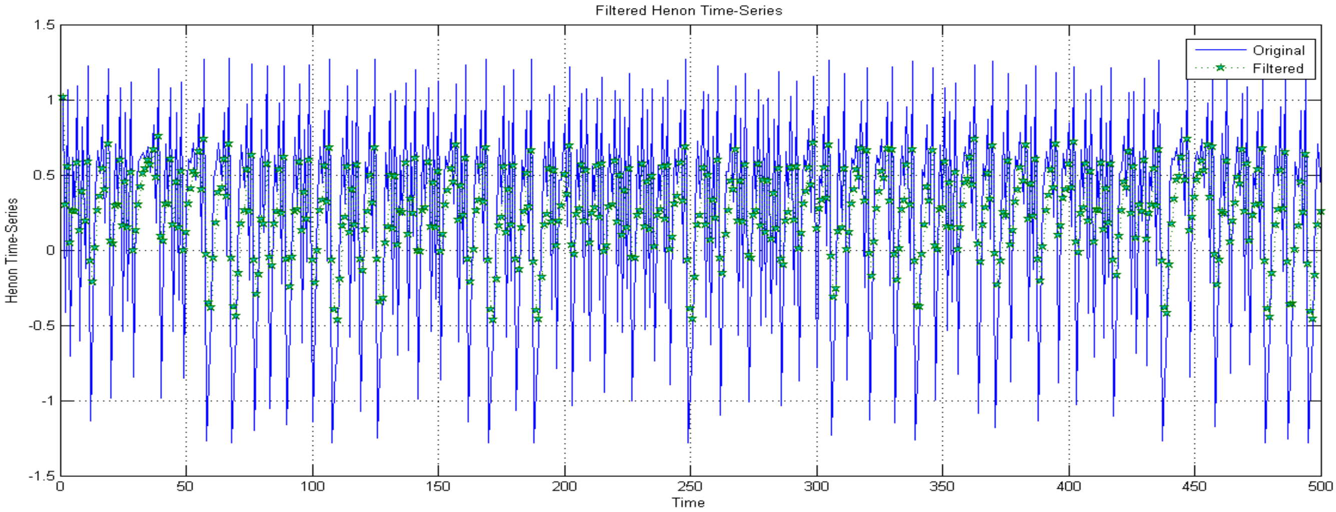

Figure 2: Filtered henon simulated time-series (500 samples)

IV. Experimental Results

A. Simulated Dataset

In the experimental setup, a dynamic discrete time-series system is simulated using mathematical formulae. The simulated test case is a common example of systems that displays chaotic behavior. In previous studies the henon map originates from a single point to a new projected position which is dependent on two variables a and b. Specific values of the parameters can result in a chaotic scenario which are a = 1.4 and b = 0.3.

In this case of simulation, new values of parameters are used and tested to provide a new case study for henon time-series. Using Equations 11 and 12, 500 samples of simulated henon time-series are generated as shown in Figure 2.

$$x_{t+1} = 1 - ax_t^2 + by_t \quad (15)$$

$$y_{t+1} = x_t \quad (16)$$

B. Recurrent network analysis

Largest lyapunov exponent (LEE) is used to check whether the generated Henon time-series (500 samples) are chaotic. When tested, Lorenz time series have an LEE value of 22.21 (positive) which demonstrates the existence of chaotic nature in the selected time-series.

To optimize the time-series input selection by reducing the level of noise associated with the selected simulated henon time-series, a time-series filter is used to lower and reduce noise levels in the proposed recurrent hybrid model. The selected value parameters for measurement and process noises are 0.001 and 1 respectively [18]. For optimal scaling values κ, numerical analysis using NMSE error metric is used to adjust and obtain the scaling value within the rage of 0-12.

Table 1
Analysis of optimal scaling factor in time-series filter

Scaling Factor	Normalized MSE	Scaling Factor	Normalized MSE
0	1.02e-11	5	1.14e-12
1	3.57e-11	8	1.05e-12
2	2.24e-11	9	1.92e-12
4	1.99e-12	12	2.12e-12

Using a forecasting metric that normalizes the mean squared error (NMSE), optimal scaling parameter is selected. Table 1 shows the normalized mean square error (NMSE) for applied scaling values with the optimal number being 8 sigma points as shown in Table 1.

To avoid model overfitting, the total number of henon time-series samples of 500 is divided into 300 and 200 samples for training and testing respectively. Division of data samples is done using block validation whereby the 200 samples of data are used in closed loop performance to verify the models ability in performing long-term forecasting. In the structural setup of the recurrent model, the number of delays is increased with the aim of further overcoming the problem of vanishing gradients during the training process.

Using the network structure, the model produces forecasted outputs for simulated henon time-series using the proposed recurrent hybrid model as shown in Figure 3. For comparative purposes the same structure is applied to the normal recurrent model as shown in Figure 4. The difference between the two models is the error outcome which is higher in normal recurrent model due to the level of noise associated with an unfiltered henon time-series data.

For model evaluation, both short and long-term forecasts obtained from the proposed recurrent hybrid model and normal recurrent model are evaluated using an error histogram as shown in Figures 5 and 6. Based on the inclination and level of error per histogram bin, the proposed recurrent hybrid model produces lower error rates as compared to the normal recurrent model.

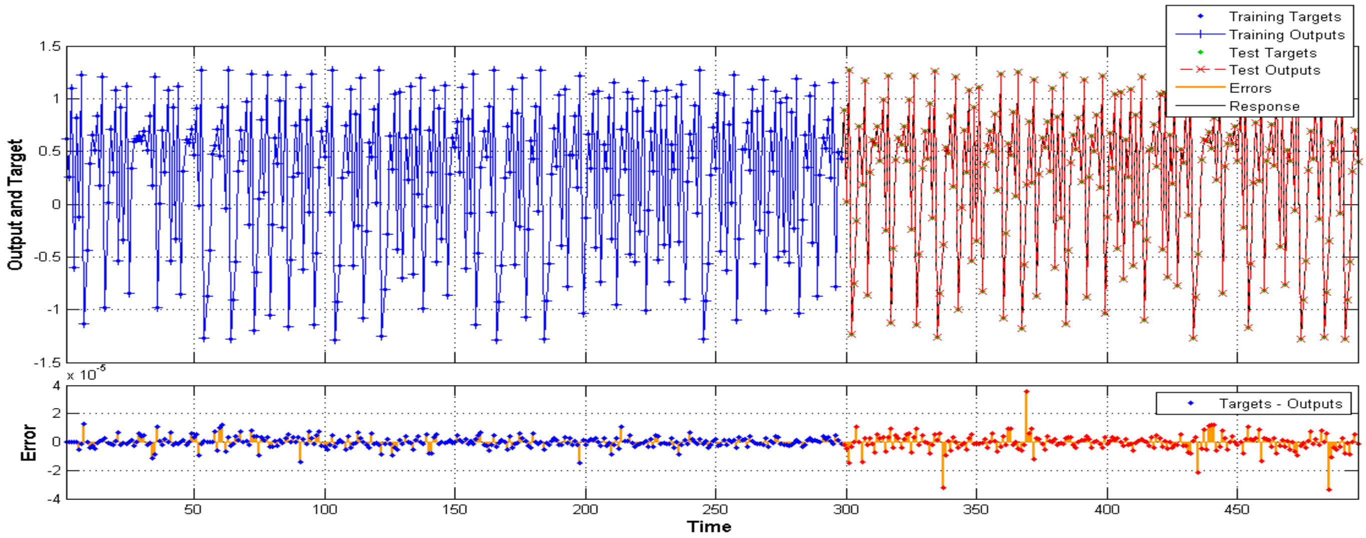

Figure 3: Henon training and testing outcomes using proposed recurrent hybrid model

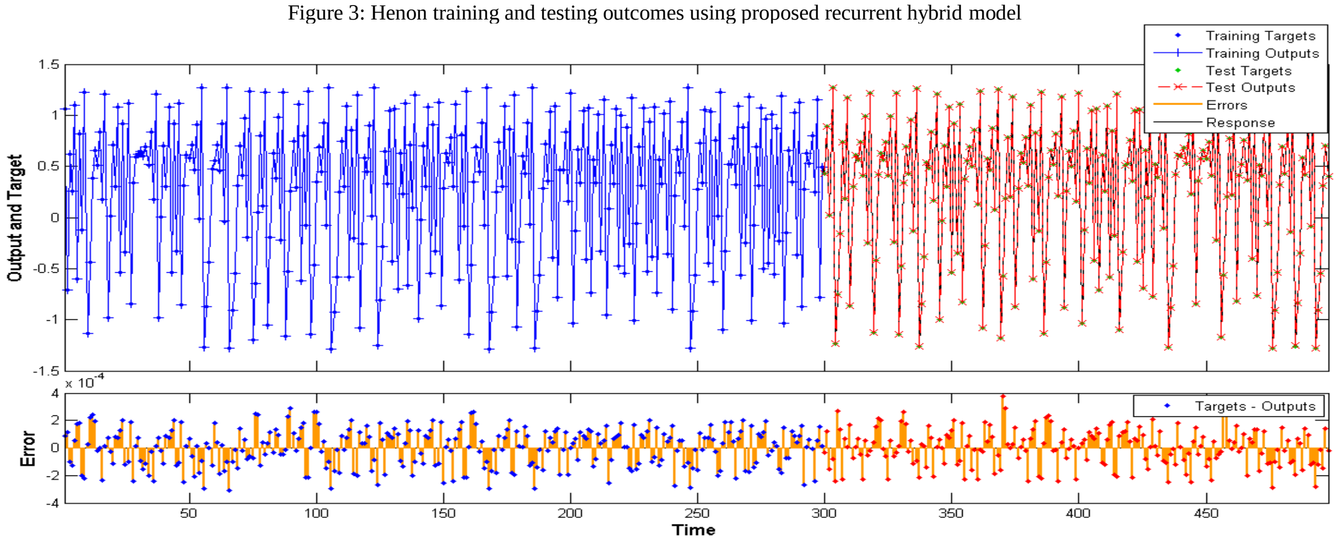

Figure 4: Henon training and testing outcomes using normal recurrent model

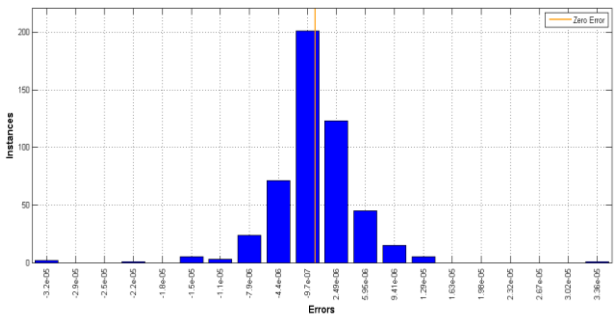

Figure 5: Error histogram for proposed recurrent hybrid model

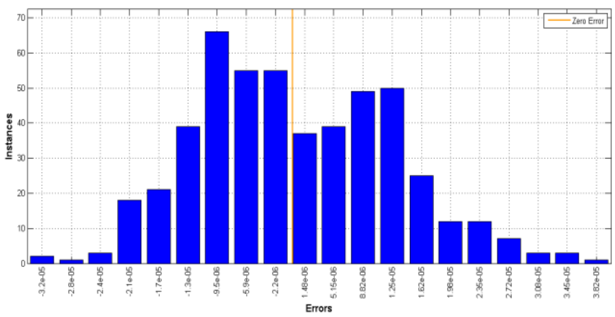

Figure 6: Error histogram for normal recurrent model

Experimental analysis based on regression value showed that the proposed hybrid recurrent model with a regression value of 1 which translates to a fit model. However for the normal singular recurrent model, it had a regression value of 0.95963. In the case for normal recurrent model, lower levels of accuracy in both the training and testing phases resulted in a low regression value as compared to the proposed hybrid recurrent model.

V. CONCLUSION

The central idea to successful chaotic time-series forecasting is attaining robust results from non-complex model in terms of computational complexity and robust results in terms of accuracy. With a focus on this idea it is evident that there is the need for the application of dynamic forecasting models which would help in estimating future trends and reducing risks of decision making.

For more than a decade, statistical models have been utilized in time-series forecasting; however, their ability to forecast is limited to the nature of time-series applied which is of linear behavior, unlike real world problems which are non-linear. In this study, simulated henon time-series is applied in forecasting with the aim of trying to improve the performance of recurrent neural networks.

Based on the proposed modification to the structural network of recurrent networks, the addition of a time-series filter in the input layer provided increases forecasting performance by reducing the noise levels that affects network training process. Empirical analysis is performed using quantitative forecasting metrics and comparative model performance on the generated forecasts. Performance evaluation results confirm that the proposed recurrent model performs long term forecasts on henon simulated time-series efficiently as per the experimental outcomes of error metrics compared to the current recurrent model.

ACKNOWLEDGMENT

This research was supported by STIRF grant (0153AA-F73), funded by Universiti Teknologi PETRONAS (UTP).

REFERENCES

[1] Guarin, Diego L., and Robert E. Kearney. "Identification of a Time-Varying, Box-Jenkins Model of Intrinsic Joint Compliance." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016.

[2] Chen, Shangyuan, Jinfeng Mao, and Xu Han. "Heat transfer analysis of a vertical ground heat exchanger using numerical simulation and multiple regression model." *Energy and Buildings* 129:81-91, 2016.

[3] Levine, D. M. Basic business statistics. Pearson Australia Group, 2013.

[4] Taylor, J.W., "Short-term load forecasting with exponentially weighted methods". *IEEE Transactions on Power Systems*, 27(1):458–464, 2012.

[5] Y., M., "A fuzzy logic fog forecasting model for perth airport". *Pure and Applied Geophysics*, 169(5-6):1107–1119, 2012.

[6] Egrioglu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks". *Expert Systems with Applications*, 40(3):854–857, 2013

[7] Shen, Xin, et al. "Support Vector Machine Classifier with Truncated Pinball Loss." *Pattern Recognition*, 2017.

[8] Oscar and Torra, S., "Forecasting tourism demand to catalonia: Neural networks vs. time series models". *Economic Modelling*, 36:220–228, 2014

[9] Xue, "A novel hybrid approach for wind power forecasting". In *Unifying Electrical Engineering and Electronics Engineering*, pages 1019–1027. Springer, 2014

[10] Liu, Y. and Jiang, "An optimal atm cash replenishment solution using ann-based bagging algorithm". *In Proceedings of the 9th International Symposium on Linear Drives for Industry Applications*, Volume 1, pages 217–224. Springer, 2014

[11] Menezes Jr, J. M. P. and Barreto, G. A., "Long-term time series prediction with the narx network: An empirical evaluation". *Neurocomputing*, 71(16):3335–3343, 2008

[12] Méndez, Eduardo, Omar Lugo, and Patricia Melin. "A Competitive Modular Neural Network for Long-Term Time Series Forecasting." *Nature-Inspired Design of Hybrid Intelligent Systems*. Springer International Publishing, 243-254, 2017.

[13] Fink, O., Zio, E., and Weidmann, U., "Predicting component reliability and level of degradation with complex-valued neural networks". *Reliability Engineering & System Safety*, 121:198–206, 2014

[14] Haykin, S. S., "Neural networks and learning machines", volume 3. *Pearson Education Upper Saddle River*, 2009

[15] Ardalani-Farsa, M., "Chaotic time series prediction with residual analysis method using hybrid elman–narx neural networks". *Neurocomputing*, 73(13):2540–2553, 2011

[16] Pascanu, R., Mikolov, T., and Bengio, Y., "On the difficulty of training recurrent neural networks*". arXiv preprint*., 2012

[17] Bengio, Y., Boulanger-Lewandowski, N., and Pascanu, R., "Advances in optimizing recurrent networks". *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pages 8624–8628. IEEE, 2013

[18] Abdulkadir, S. J., & Yong, S. P., "Scaled UKF–NARX hybrid model for multi-step-ahead forecasting of chaotic time series data". *Soft Computing*, 19(12), 3479-3496, 2015

[19] Abdulkadir, S. J., Yong, S.-P., Marimuthu, M., and Lai, F.-W., "Hybridization of unscented kalman filter and non-linear auto-regressive neural network for financial forecasting". *In Mining Intelligence and Knowledge Exploration*, pages 72–81, 2014

[20] Ma, Q.-L., Zheng, Q.-L., Peng, H., Zhong, T.-W., and Xu, L.-Q., "Chaotic time series prediction based on evolving recurrent neural networks". *In International Conference on Machine Learning and Cybernetics*, volume 6, pages 3496–3500. IEEE, 2007

[21] Abdulkadir, S. J. and Yong, S.-P., "Unscented kalman filter for noisy multivariate financial time-series data". *In Multi-disciplinary Trends in Artificial Intelligence*, pages 87–96. Springer, 2013

[22] Sivakumar, Bellie. "Modern Nonlinear Time Series Methods." *Chaos in Hydrology*. Springer Netherlands, 111-145, 2017