Remaining Useful Life (RUL) Regression Using Long-Short Term Memory (LSTM) Networks

Sofia Yousuf and Salman A. Khan
Karachi Institute of Economics and Technology
Pakistan
Email: sofia.yousuf@kiet.edu.pk, sakhan@kiet.edu.pk

Saqib Khursheed
Dept. of Electrical Engineering and Electronics
University of Liverpool
UK
Email: S.Khursheed@liverpool.ac.uk

***Abstract*—The accurate prediction of the remaining useful life (RUL) of components is a major concern in electronic circuits. The RUL-based health diagnostics plays an important role in the determination of time-of-failure of a device as an early warning in industrial applications. In this paper, a Long Short Term Memory (LSTM) based regression model is proposed for the prediction of RUL of a Ring Oscillator (RO) circuit utilizing the most essential extracted electrical features of the device. LSTM networks are capable of capturing the temporal dependencies in the time-series data and eliminating the vanishing gradient problem encountered in the conventional recurrent neural networks (RNNs). From Cadence simulations, utilizing the 22 nm CMOS technology library, it has been demonstrated that the RO frequency degradation essentially depends on three major factors including the working temperature, voltage, and most importantly, the device aging parameter. The results show that more than 90% of the cases of the RUL prediction for the 13 and 21 stage constrained under the supply voltage variation from 0.7 V to 0.9 V with the least prediction deviation of 2 days to 6 days.**

KEYWORDS: Aging, Remaining Useful Life, Machine Learning, Online Prediction, Reliability

I. Introduction

With the rapidly decreasing size of the electronic circuits, the remaining useful life (RUL) of the electronic devices is becoming a major concern. The detection of RUL as an early warning indicator is essential for preventive maintenance and reliability of electronic systems [1]. The RUL indicator is therefore, a useful parameter determining the durability of the electronic components and systems [2].

With today's miniaturization systems incorporating nanometer technology, the RUL prediction problem has become more critical. An important factor that is directly related to the RUL degradation of the device is the 'path delay' [3]. Therefore, determination of accurate RUL parameter of these devices is critical. Due to the lack of efficient RUL prediction methods, it cannot be determined as for how much time the system would operate satisfactorily without failure. This solicits the need to explore more efficient methods for RUL prediction.

In the current literature, the most common indicator for determining the health of electronic components and its reliability is the Mean Time to Failure (MTTF). However, MTTF cannot be utilized for estimating the reliability of lifetimes of discrete electronic components. The RUL estimation based on parametric estimation method such as Particle Filter (PF) was also reported in the literature [4]. In this method, PF was utilized for RUL monitoring of an Insulated Gate Bipolar Transistor (IGBT). The RUL prediction variance with this method was closer to 20%. The limitation of the method was that the RMS error was found to be decreasing with the increasing number of particles in Auxiliary Particle Filter (APF). For the prediction of failure in electrolytic capacitors, the models that were based on equivalent series resistance and capacitance were developed by Long et. al. [5]. This method utilized a hybrid scheme that was based on the ARMA (Autoregressive Moving Average) model incorporating the Genetic Algorithm (GA) for parameter optimization. The advantage is that the GA essentially requires lesser information related to the problem at hand. However, for efficient implementation of GA, the optimization function needs to be effectively formulated. The RUL prediction method for solder joints was developed based on an ARMA prediction model using 'resistivity' as an input to the model [6]. In a study by Lall et. al. [7], the RUL prediction based on visual inspections was performed. A Radio Frequency based RUL prediction strategy was designed utilizing the Gaussian Regression (GR) technique [8]. The limitation of the scheme was that the impact of GP model parameters such as the covariance functions as well as the maximum look-ahead time was not investigated. Liu et. al. [9] proposed a generalized Cauchy degradation model with Long-Range dependence and maximum Lyapunov exponent for RUL. Duan et. al. [10] proposed product technical life prediction based on multi-modes and fractional Levy stable motion.

Besides parametric approaches, methods based on machine learning have also been reported for RUL estimation. A method utilizing neural networks considering the deviations in electronic resistor-capacitor (RC) signals was reported [11], [12]. The essential features were extracted from the RC signals information. The techniques based on neural networks are good at non-linear function approximation problems. However, these techniques encounter two major problems which are (1) the network over-fitting issues and (2) the neural network parameter selection for which there are no standard methods for tuning the network parameters. More recently, a device aging detection technique [13] was proposed that utilizes the RO along-with the electronic chip

Fig. 1. Flow Diagram of the Proposed LSTM-Based RUL Method.

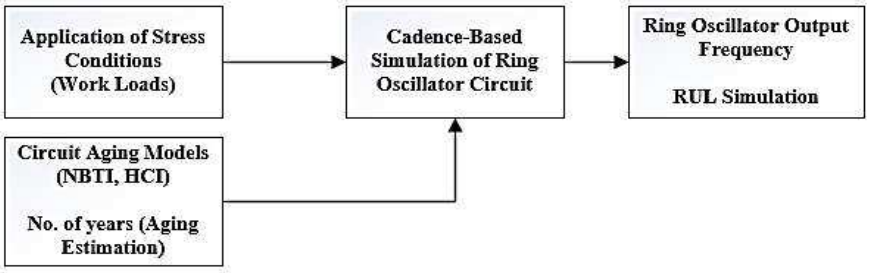

Fig. 2. Flow of RUL Simulation Process

ID (ECID). The RO frequency measurement was stored in the RFID tag. Furthermore, the high standard applications require electronic systems to operate under different temperature conditions. Therefore, another factor that should be considered is the device in-use time, more commonly known as the 'stress' parameter which depends on the workload of the device. Other notable factors for device aging include the Hot Carrier Injection (HCI) and the Bias Temperature Instability (BTI) introducing the delays in the circuit. Studies have highlighted the path delays, temperature, and the drifts in frequencies inherent in the operation of the Ring Oscillators. The use of support vector regression (SVR) was reported in Martinez et. al. [14], which evaluated the impact of SVR on the 13 and 21 RO circuits.

The current study is part of an ongoing work on runtime estimation of aging and predicting failure of CMOS designs with the aid of light weight semi-autonomous sensor and machine learning algorithm. The sensor can be activated at any time during the lifetime of the device and may be used for estimating its remaining useful lifetime (RUL) and improve upon the MTTF (Mean Time to Failure) parameter that is currently used. This aspect is discussed in a wider context in several previous studies [14], [15], [16], [17] by the third author.

Citing the limitations of several previous studies mentioned above, the present study utilizes a special class of RNNs called the Long-Short-Term-Memory (LSTM) networks to build a regression model for the RUL prediction, which was previously solved using SVR [14]. The method utilizes essential features from the 13 and 21 RO circuit. The trained LSTM network aims to provide early warnings for the device failure conditions based on RUL predictions. The system would also fundamentally provide an easy online inspection of electronic devices with reduced cost of testing.

The rest of the paper is structured as follows. In Section II, the proposed methodology for RUL is presented. Section III provides details of the data organization consisting of preprocessing stage and the restructuring of data before creating the learned LSTM RUL regression model. The background on LSTM networks is given in Section IV. In Section V, the results of the RUL prediction under different temperature and voltage conditions are presented for the 13 and 21 stage ROs. Finally, concluding remarks are presented in Section VI.

II. Proposed Methodology for RUL Prediction

The proposed scheme is depicted in Fig. 1. The raw information from the 22 nm technology based on CMOS device (based on Cadence simulation using aging models) was introduced to the data acquisition block. This data was further structured in the form of a matrix. For training the LSTM network, the data was further arranged in cell arrays. In the next block, the LSTM network is trained and the learned regression model for RUL is developed.

A. CMOS Parameterization

In this work, a 22 nm CMOS-based technology is considered. The frequency of the Ring Oscillator circuits is extracted for different conditions of temperature as well as voltages. The objective is to investigate whether the RO frequency parameters have a strong relationship with the device aging, operating temperature, and voltage. The CMOS data utilized in this research was taken from a previous study [14]. This dataset was generated through the Cadence ADE-L software utilizing the CMOS models based on the temperature and voltage variations and device aging phenomena due to BTI and HCI. This data was generated using 22 nm CMOS library provided by GlobalFoundries (GF).

B. Characteristics of Dataset

The dataset utilized in this work comprises of data of 20 years of device degradation considering the various conditions of temperature and voltage. The characteristics of the data utilized in this study are described in Fig. 2.

The data was collected considering the various models for device aging including NBTI, HCI and aging phenomena. The data consisted of 630 samples (considering 20 years of degradation) for various operating temperatures (0°C, 5°C, 10°C,...., 100°C) and voltages between V = 0.7 V and V = 0.9 V, with a step size of V = 0.05 V.

The dataset used in this study reflects the dependency of RO frequency degradation with respect to the temperature effects of BTI and HCI [18], [19], [20], [21]. In various studies, it has been shown that the changes in the threshold voltages in the CMOS transistors due to above factors of BTI and HCI are responsible for path delays in the CMOS gates. The degradation trend of the RO frequency variation with respect to the circuit lifetime has a strong relationship. This relationship is depicted in Fig. 3 and Fig. 4. The non-linear decrease of RO frequency with respect to time can be easily observed from the figures and is therefore utilized as a pseudo-signal to train the LSTM regression network. The ROs were designed to output same frequency using internal capacitance and output load capacitances. This allowed direct degradation comparison between the ROs over the aging period. The numerical values for the frequency degradation of 13-RO are additionally reported in Table I with respect to the operating time (i.e., described for a period of 20 years). Similar observations were made for the 21-RO circuit as seen in Table II.

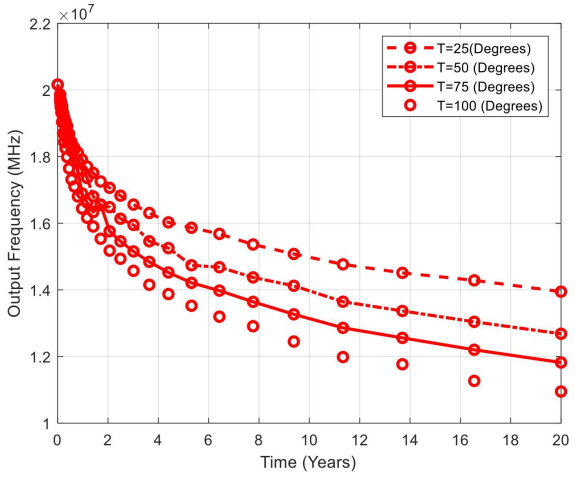

Fig. 3. Frequency Degradation Trend for 20 Years for the RO-13 Circuit (for 0.9 V). Temperature is in centigrade

TABLE I
FREQUENCY DEGRADATION (13-RO).

Working Temperature (°C)	5 Years	10 Years	15 Years	20 Years
25	18.62%	23.12%	25.88%	28.03%
50	22.73 %	27.60 %	31.54 %	34.57 %
75	26.77 %	32.46 %	36.41 %	39.71 %
100	31.51 %	37.88 %	41.95 %	45.38 %

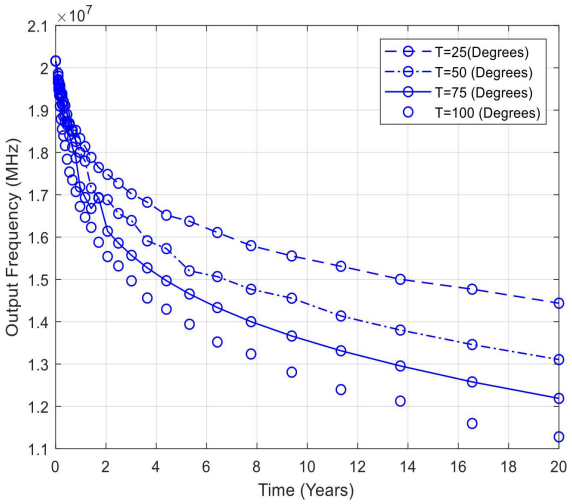

Fig. 4. Frequency Degradation Trend for 20 Years for the RO-21 Circuit (for 0.9 V). Temperature is in centigrade

TABLE II
FREQUENCY DEGRADATION (21-RO).

Working Temperature (°C)	5 Years	10 Years	15 Years	20 Years
25	21.37 %	25.57 %	28.55 %	30.96 %
50	25.31 %	29.98 %	34.00 %	37.15 %
75	28.95 %	34.35 %	38.37 %	41.79 %
100	33.12 %	39.03 %	43.55 %	47.42 %

III. DATA ORGANIZATION

In order to utilize data for empirical analysis, the data need to be prepared in a proper structure. This is done via the data pre-processing and data restructuring. Details of these two stages are given below.

A. Data Pre-processing

Fundamentally, the data extracted from the dataset consisted of a matrix consisting of the output frequencies of the Ring Oscillator. In order to train the LSTM Regression Network, the data was restructured from matrix into the cell arrays of separate training and testing sets. The dataset consisted of Ring Oscillator data collected for different frequencies under various operating conditions of temperatures and voltages. Furthermore, the data essentially highlights the dependence of a ring oscillator circuit frequency on four major factors. These factors include the working temperature, voltage, circuit load capacitance (C_k), and aging time (considering duration of 20 years). The format of data matrix that was structured in the original dataset is shown in Eq. 1

$$\mathrm{F}_m(T_m, V_t) = \{\mathrm{freq}_{m1}, \mathrm{freq}_{m2}, \mathrm{freq}_{m3}, \ldots, \mathrm{freq}_{mn}\} \quad (1)$$

In the above equation, T_m represents the working temperature, where 'm = 21'. Furthermore, V_t represents the number

of different operating voltages with 't = 5', ranging from V = 0.7 V to V = 0.9 V. The symbol 'n' represents the number of samples where 'n = 30', and the load capacitance is assigned a constant value of C = 200 fF. Thus, the total number of data samples was 630. Note that these were simulation samples taken from aging degradation simulated using EDA tools. The corresponding data structure is illustrated in Figure 5.

In Fig. 6, the data processing stage shows how the data was obtained considering variations in the various electrical variables such as the operating voltage and the working temperature. The continuous workload conditions degrade the operating frequencies of the Ring Oscillators. This effect was then captured using the Cadence based simulation in the form of 'Remaining Useful Life (RUL)' predictions. This data was finally trained using the Long-Short-Term-Memory (LSTM) model.

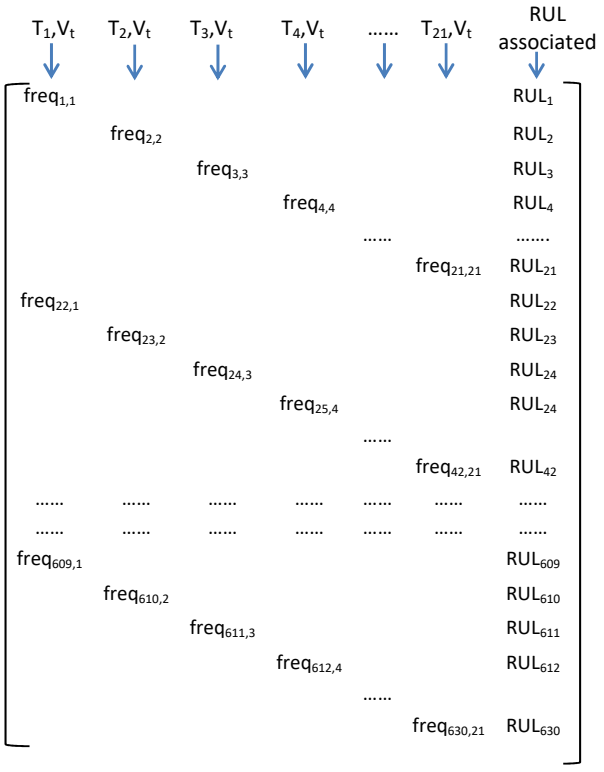

Fig. 5. Data matrix consisting of ring oscillator output frequencies.

B. Dataset Restructuring

To generate an LSTM-based regression model, first the dataset is prepared in the pre-processing stage that is in suitable format for training the LSTM regression model. The sparsity in the dataset was removed by adding small Gaussian noise to the data. The Gaussian noise model employed for this purpose is given in Eq. 2.

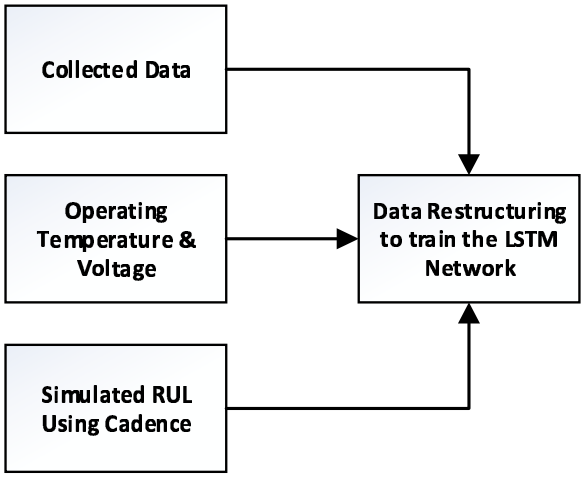

Fig. 6. Data Preprocessing Stage

$$G = a * rand(n) + b \tag{2}$$

where 'a' is the standard deviation and 'b' is the mean value. This random noise is added to each sample to remove sparsity in the data.

IV. Learning with Long-Short Term Memory (LSTM) Networks

This section discusses the motivation to use the LSTM networks for the underlying problem. The section further discusses the details of LSTM structure as well as training of LSTM networks.

A. Motivation

The Recurrent Neural Networks (RNNs) are a class of deep-learning methods that have been shown to have a great potential in capturing both the long-term and the short-term temporal dependencies within the time-series data. The LSTMs are a particular class of RNNs that utilize their inherent memory capabilities, and use the past information to make efficient predictions.

The LSTM structure was first introduced in 1997 by Reiter and Schmidhuber [22]. The structure fundamentally is comprised of the memory units. The flow of information through the LSTM network is controlled by the network's internal gating mechanism. One major advantage of the LSTM network is that it overcomes the vanishing as well as the exploding gradient problem inherent in the conventional RNNs. An LSTM network accepts a sequential time-series data as input.

B. LSTM Structure Details

The basic LSTM cell structure is shown in Fig. 7. The LSTM architecture resembles RNNs. However, the LSTM networks include the memory block instead of the summation blocks in RNNs. The architecture of an LSTM network comprises of the input layer, forget gate, output gate, memory cell, and memory state. The memory cells hold the past information which is further utilized in the training of the LSTM network when the present samples are available for training. The forget gate handles the amount of the past information for training the LSTM network. For a specific time *t*, the LSTM equations are represented as:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{4}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{5}$$

$$z_t = \tanh(W_z x_t + U_z h_{t-1} + b_z) \tag{6}$$

$$c_t = i_t \otimes z_t + f_t \otimes c_{t-1} \tag{7}$$

$$h_t = o_t \otimes \tanh c_t \tag{8}$$

The performance of the LSTM network is governed by the three gates, i.e., the input gate *i*, the forget gate *f*, and the output gate *o*. Equations (3) to (7) are the update equations for the LSTM. In the above equations, the operator $\otimes$ represents element-wise multiplication, and the variable h_t represents the hidden state. At any time instance *t*, the forget gate is equipped with the input *x* as well as the hidden state h_{t-1}. Furthermore, z_t represents the candidate for cell state (memory) at time t, while c_t defines the memory at time t.

C. Training the LSTM Network and Performance Evaluation

In this section, the LSTM parameter optimization process is briefly described with respect to the 13-RO Ring oscillator. The similar process applies for the 21-RO circuit. For training the regression model, the LSTMs with different number of hidden units are considered. After proper tuning, the LSTM with 30 hidden units were selected based on the goal to achieve minimum errors in the RUL prediction. The training pattern of the LSTM network is depicted in Fig. 8. For training the LSTM network, various experiments were performed. Fig. 8 shows the decreasing trend of the 'loss function' and the decreasing 'RMSE' values while training of the LSTM model.

To determine the effectiveness of the regression model, the prediction results are evaluated using three different metrics. These metrics are the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Squared Error (MSE). These metrics are the most widely used metrics for the evaluation of the regression models according to several studies [14]. The mathematical representation of these metrics is given below.

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\Big(\frac{d_i - f_i}{\sigma_i}\Big)^2} \tag{9}$$

$$MAE = (\frac{1}{n})\sum_{i=1}^{n}|y_i - x_i| \tag{10}$$

$$MSE = \sum_{i=1}^{n}(x_i - y_i)^2 \tag{11}$$

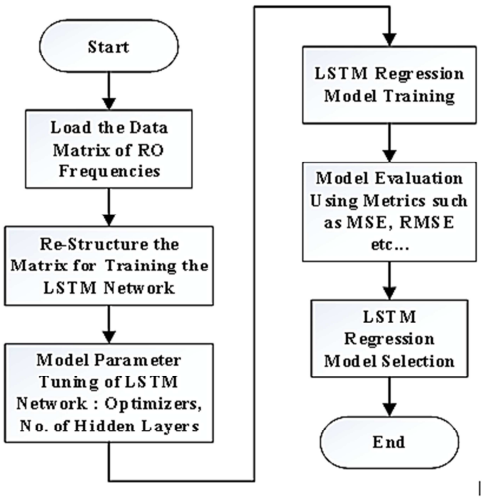

Fig. 9. LSTM Training Algorithm Flow.

In Fig. 9, the major steps of the LSTM algorithm flow are highlighted. Starting from the data acquisition and the pre-processing stage towards the model evaluation, the six main steps are presented. First, the data is restructured in the form in which the LSTM model can be learned. The LSTM network is then structured by defining different number of layers and optimizers. Finally, the model is evaluated based on regression metrics MSE, RMSE and MAE. These metrics are known as the 'primary metrics'. Also, for the model evaluation, other metrics known as secondary metrics can also be used.

In Fig. 10, the optimization process of the LSTM network is presented. The LSTM parameters include the number of hidden layers and parameters such as the other hyper-parameters for the optimization. In this study, the 'Adam Optimizer' is used for training the LSTM model. It should be noted that the LSTM training is an iterative process based on achieving a certain criterion for the prediction accuracies.

V. Results and Discussion

In this section, the simulation results of 13-RO and the 21-RO circuits for RUL prediction are presented. The training was performed utilizing the LSTM network. While setting the parameters, it should be noted that the LSTM

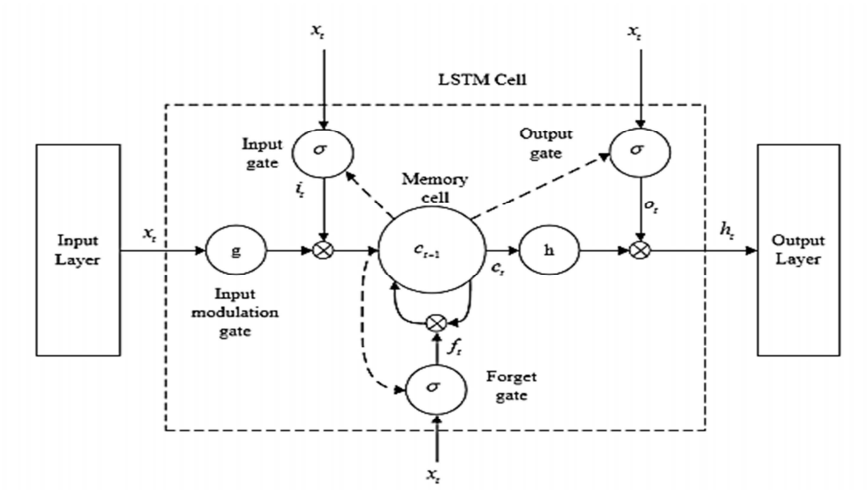

Fig. 7. Basic LSTM cell structure [2]

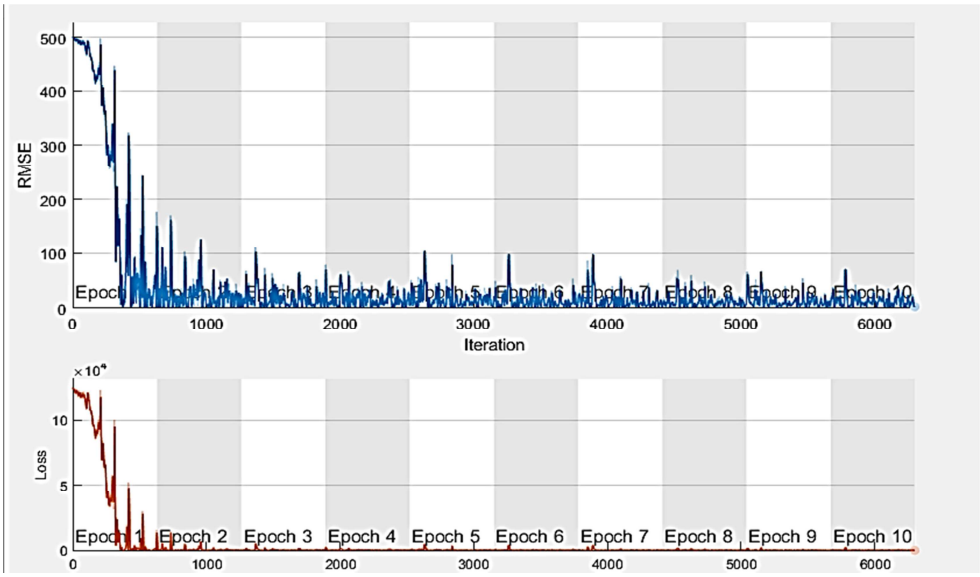

Fig. 8. LSTM Training Pattern.

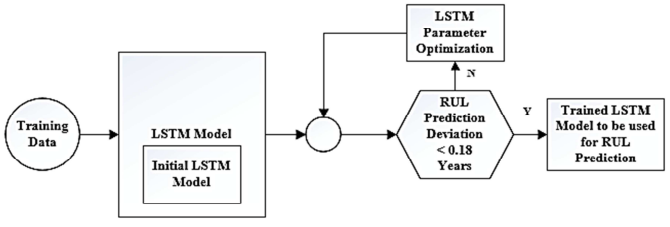

Fig. 10. LSTM Parameter Optimization.

performance depends upon multiple factors including (1) the method of data sampling (2) the time duration (intervals) between the individual samples and (3) the missing samples in the existing dataset. Based on the above factors, the data utilized in this work was restructured as detailed in Section III-B.

A. RUL Prediction for 13 and 21 Ring Oscillator Circuits

For simulations, the circuits of 13-RO and 21-RO utilizing the 22 nm CMOS technology were built incorporating the Global Foundries (GF)-based aging model. Furthermore, the extreme condition of 100% stress was considered where each circuit was kept operational for its complete duration of the life-time. More specifically, we considered a normal waveform with 50% duty cycle and fully stressed the ROs by exercising them over the aging duration of 20 years. Data was generated using Cadence EDA tools and 22-nm GF design library. Duty cycle can be altered, however 50% duty cycle

was considered a good representation for this study. From the various electrical parameters of the Ring Oscillator, the RO frequency degradation has been selected as the main proxy for the determination of the RUL for the two circuits.

The experiments using the LSTM network are performed in an off-line mode by training the network. However, it has been shown that on average, the testing-time of the network is less than 1 second. Therefore, once trained, the model can be effectively utilized for real-time on-line testing by deploying on the actual hardware. The major parameter considered here is the RO working temperature. It was found in a previous study that the similar measurements of the output frequencies are produced even with different variations in the temperature [23]. Furthermore, in the online testing, the prediction of the RUL-based prediction can be made using the information of the temperature, the device lifetime, and the voltage irrespective of the operating condition of the RO as described in [24]. In this work, the generated model on the LSTM network is based not on the failure condition but on the information regarding the failure-time, i.e., in this case, is the cut-off frequency parameter of the Ring-Oscillator. Furthermore, the time-series generated from the operating conditions of the RO (considering 20 years of degradation) for various operating temperatures ($0°$C, $5°$C, $10°$C,...., $100°$C) are considered. It has been shown in previous studies [25][26][15] that the RUL prediction of a Ring-Oscillator is dependent on parameters such as voltage, temperature and the device aging phenomenon. Therefore, the variations in these variables have been exploited to create the LSTM-based regression model and can be proved to be effective set of features for training any regression model.

In order to train the LSTM model, the data was structured in the form of Fig. 5 similar to the one by Martinez et. al. [14]. However, based on the requirements of the LSTM model, variations to the existing data have been made to make it compatible for the LSTM training. The RUL data to be predicted for 20 years of degradation is shown in Figs. 11 and 12 for 13-RO and 21-RO, respectively. These RUL values were generated considering the operating temperature ranges from $0°$C to $100°$C. The data used for simulation was captured with a temperature difference of $5°$C. For each case, the RUL curve is plotted in order to observe the RUL degradation trend for each case. From the plots, it can be clearly inferred that the RUL of the RO depends upon the amount of stress applied. That is, the curve for $100°$C shows less value of RUL for the period of 20 years as compared to the curve at $0°$C. It can also be noticed further that as the working temperature increases, the slope of the RUL curve also increases. Note that Fig. 11 shows one of 630 data points using voltage and temperature variations.

In the graphs shown in Figures 11 and 12, the RUL values initiate from 20 years. The time-scale is defined by the following exponential equation:

$$Time\ Scale = 0.068e^{(0.189 \times sample)} \tag{12}$$

The sampling rate in the above equation is used as obtained

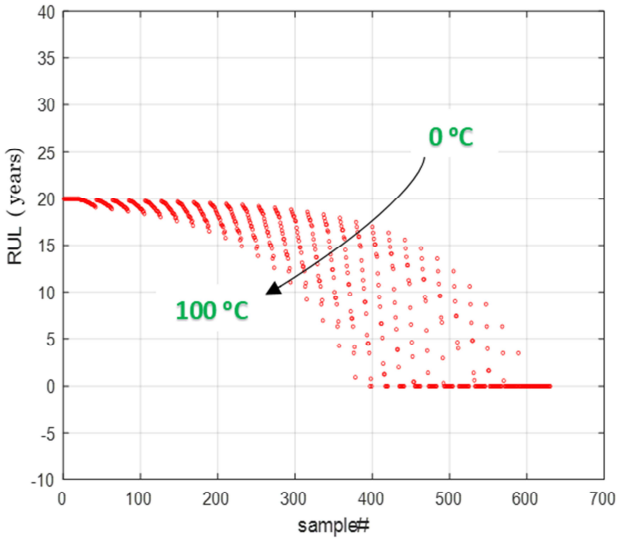

Fig. 11. RUL Data to be Predicted (Shown for 13-RO Oscillator Circuit).

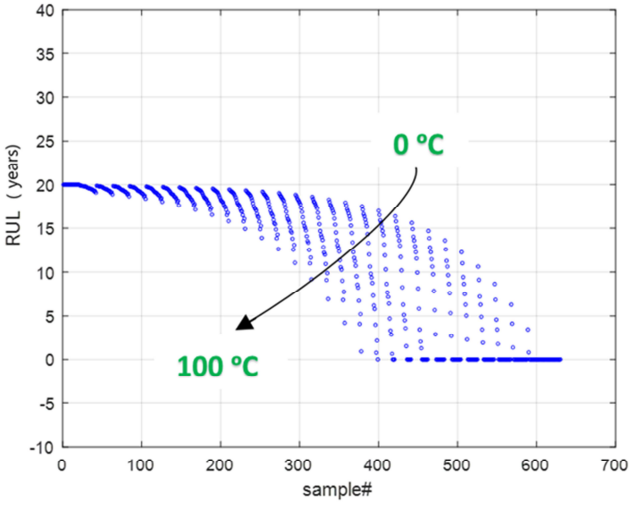

Fig. 12. RUL Data to be Predicted (Shown for 21-RO Oscillator Circuit).

by Martinez et. al. [14] using the Cadence based simulation. The values of the parameter samples are in the range 0 - 30 with a step-size of 0.0475. From the above equation, it can be inferred, for instance, that the RO frequency degradation for a working temperature of $5°$C is faster as compared to $0°$C and the same is true when other temperatures are compared. The efficacy of the proposed method is that the strategy based on ROs is effective according to other previous studies [27], [28], [29].

The results of RUL prediction for the 13-RO ring oscillator are presented in Table III below for the training data and for

testing data in IV respectively.

TABLE III
RESULTS ON TRAINING DATA FOR 13-RO

Error Metric	No. of hidden units	Epochs	Error values
RMSE	20	10	9.95940
MAE	20	10	50.1892
MSE	20	10	8.44930

TABLE IV
RESULTS ON TESTING DATA FOR 13-RO

Error Metric	No. of hidden units	Epochs	Error values
RMSE	20	10	11.0071
MAE	20	10	53.1841
MSE	20	10	9.24130

The performance of the LSTM regression algorithm was studied for different number of hidden layers in the architecture as well as the different number of training sessions (epochs). With large number of experiments, an optimal number of the hidden architectural layers for the regression model were selected to be 20. Similarly, the optimal number of epochs for which the model was trained is 20 epochs. Furthermore, a learning rate of 0.001 was selected for training the model. In general, it was observed that increasing the number of training epochs increases the values in the error.

Another important aspect affecting the performance of a deep learning strategy is the choice of an appropriate optimization algorithm. The Adam optimizer [30] is selected for the scheme developed in this paper using the LSTM. The Adam optimization approach is fundamentally an extension of the stochastic gradient descent algorithm and has been widely adopted for the optimization of deep-learning methodologies. The Adam optimizer is utilized in this study due to several advantages including (1) the optimization is computationally efficient (2) it is easy to implement (3) little requirements for hyper-parameter tuning and (4) good suitability for problems with sparse and noisy gradients. The settings for the training options for the Adam optimizer are listed in Table V which provided the best results in terms of RUL prediction errors.

In Fig. 13, the improvement can be observed in the RUL deviation of -0.005 years over the 0.02 years of deviation in the RUL. The worst case deviation observed is approximately 0.015 years.

Tables VI and VII give the results of 21 RO circuit with respect to training and testing data, respectively. The performance is reported in terms of RMSE, MSE, and MAE. The optimal number of hidden units was set at 30 and the number of epochs were set at 20. It can be observed from the results that the error values in the training phase are consistent with that of the testing phase. The graph for the remaining useful life prediction for the 20 years of degradation is depicted in Fig. 14.

TABLE V
ADAM OPTIMIZER TUNING PARAMETERS

Adam Optimizer Option	Value
Gradient Decay Factor	0.900
Squared Gradient Decay Factor	0.999
Epsilon	1.0e-08
Initial Learn Rate	1.0e-03
Learn Rate Drop Factor	0.1000
L2 Regularization	1.0e-04
Learn Rate Drop Period	10.00
Gradient Threshold	1.000

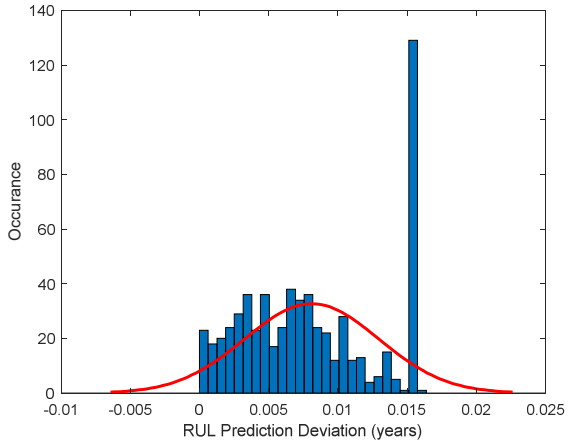

Fig. 13. Comparison to Normal Distribution Curve for LSTM Network for 13-RO (N=40; E = 20, Learning Rate = 0.001).

B. Discussion

The performance of the proposed RUL prediction method based on the LSTM network is evaluated using MSE, RMSE and MAE metrics. Tables III to VII presented the training and testing results for the 13-RO and 21-RO circuits. From the tables, the consistencies of the predicted values for the training and testing data can be clearly observed. The execution time for training and testing the LSTM network was also calculated. The training time (averaged over the 13-Stage and 21-Stage) was found to be 140 s while the

TABLE VI
RESULTS ON TRAINING DATA FOR 21-RO

Error Metric	No. of hidden units	Epochs	Error values
RMSE	30	20	6.9744
MAE	30	20	48.6426
MSE	30	20	5.8537

TABLE VII
RESULTS ON TESTING DATA FOR 21-RO

Error Metric	No. of hidden units	Epochs	Error values
RMSE	30	20	7.0144
MAE	30	20	50.120
MSE	30	20	6.2001

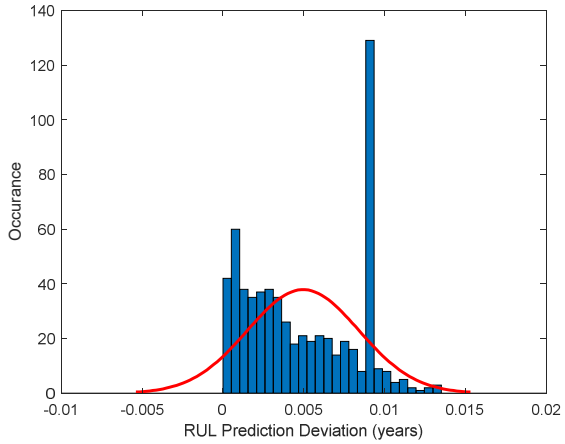

Fig. 14. Comparison to Normal Distribution Curve for LSTM Network for 21-RO (N=40; E = 20, Learning Rate = 0.001).

testing time was 0.79 s. It can be inferred from these values that since the LSTM model is a deep-learning technique, it requires more time for training compared to other methods. However, once trained, the testing time is low. For our case, the testing of 0.79 s made the LSTM model suitable for real-time applications.

VI. Concluding Remarks

The reliability of micro-electronic circuits is becoming a major issue with the rapid decrease in the size and the scaling of the circuits. This paper addressed this issue with respect to the 22 nm CMOS technology, and a Long-Short-Term-Memory (LSTM)-based RUL prediction solution is proposed utilizing the output frequency of the Ring Oscillator. Additionally, the RUL-based aging prediction of discrete components is made by the proposed method which is not effectively addressed by the MTTF indicator. The method utilizes a special class of RNNs called the Long-Short-Term-Memory (LSTM) networks to build a regression model for the Remaining Useful Life prediction. The LSTM is capable of overcoming the exploding as well as the vanishing gradient problem inherent in the conventional RNN-based deep learning approaches. A notable advantage is that despite the limited number of training samples, the LSTM model shows significant improvement in the prediction of RUL estimates for the Ring Oscillator incorporating a simple technique for Data Augmentation. The training of the regression model is conducted under the different scenarios of voltage and temperature. A Ring Oscillator aging dataset was used comprising of CMOS model which incorporates major factors for the device including voltage, temperature, and the BTI and HCI models. It was proven that the approach utilizing the output frequency parameter of the Ring Oscillator is effective in predicting the RUL estimate. The method introduced in this research can be used in systems that are critical to reliability issues. Furthermore, the results of 13 and 21 stage RO have also been presented in this paper. The testing is performed from the 13 and 21 stage RO circuit, under temperature variations from $0°C$ to $100°C$ with a supply voltage of 0.9 V. Finally, the above method can prove to be useful for detecting the recycled ICs in highly critical industrial applications enabling the manufacturing of reliable electronic systems.

Acknowledgement

This research was supported through a grant by Prince Mohammad bin Fahd Center for Futuristic Studies (PMFCFS) at the Prince Mohammad bin Fahd University, Saudi Arabia. We also acknowledge the support by ODA Research Seed Funding, University of Liverpool, UK.

List of Abbreviations

RUL Remaining Useful Life
LSTM Long Short Term Memory
RO Ring Oscillator
RNN Recurrent Neural Network
MTTF Mean Time To Failure
PF Particle Filter
IGBT Insulated Gate Bipolar Transistor
APF Auxiliary Particle Filter
ARMA Autoregressive Moving Average
GA Genetic Algorithm
GR Gaussian Regression
ECID Electronic Chip Identity
HCI Hot Carrier Injection
BTI Bias Temperature Instability
SVR Support Vector Regression
C_k Circuit load capacitance
T_m Working Temperature
V_c Number of different voltages

References

[1] Yanting Zhou, Yinuo Huang, Jinbo Pang, and Kai Wang. Remaining useful life prediction for supercapacitor based on long short-term memory neural network. *Journal of Power Sources*, 440:227149, 2019.

[2] Yuanhang Chen, Gaoliang Peng, Zhiyu Zhu, and Sijue Li. A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. *Applied Soft Computing*, 86:105919, 2020.

[3] Lin Yang, Feng Wang, Jiaojiao Zhang, and Weihao Ren. Remaining useful life prediction of ultrasonic motor based on elman neural network with improved particle swarm optimization. *Measurement*, 143:27–38, 2019.

[4] Moinul Shahidul Haque, Seungdeog Choi, and Jeihoon Baek. Auxiliary particle filtering-based estimation of remaining useful life of igbt. *IEEE Transactions on Industrial Electronics*, 65(3):2693–2703, 2017.

[5] Yiwen Long, Hongwei Luo, Yue Zhi, and Xiaoqiang Wang. Remaining useful life estimation of solder joints using an arma model optimized by genetic algorithm. In *2018 19th International Conference on Electronic Packaging Technology (ICEPT)*, pages 1108–1111. IEEE, 2018.

[6] YS Rana, Shantanab Banerjee, Tej Singh, and PV Varde. Experimental program for physics-of-failure modeling of electrolytic capacitors towards prognostics and health management. *Life Cycle Reliability and Safety Engineering*, 6(3):179–185, 2017.

[7] Pradeep Lall, Mahendra Harsha, Jeff Suhling, and Kai Goebel. Sustained damage and remaining useful life assessment in leadfree electronics subjected to sequential multiple thermal environments. In *2012 IEEE 62nd Electronic Components and Technology Conference*, pages 1695–1708. IEEE, 2012.

[8] Daeil Kwon, Michael H Azarian, and Michael Pecht. Remaining-life prediction of solder joints using rf impedance analysis and gaussian process regression. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 5(11):1602–1609, 2015.

[9] He Liu, Wanqing Song, Yujin Zhang, and Aleksey Kudreyko. Generalized cauchy degradation model with long-range dependence and maximum lyapunov exponent for remaining useful life. *IEEE Transactions on Instrumentation and Measurement*, 70:1–12, 2021.

[10] Shouwu Duan, Wanqing Song, Enrico Zio, Carlo Cattani, and Ming Li. Product technical life prediction based on multi-modes and fractional lévy stable motion. *Mechanical Systems and Signal Processing*, 161:107974, 2021.

[11] Chetan S Kulkarni, José R Celaya, Kai Goebel, and Gautam Biswas. Bayesian framework approach for prognostic studies in electrolytic capacitor under thermal overstress conditions. Technical report, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION MOFFETT FIELD CA AMES RESEARCH , 2012.

[12] Chetan Kulkarni, Gautam Biswas, Sankalita Saha, Kai Goebel, et al. A model-based prognostics methodology for electrolytic capacitors based on electrical overstress accelerated aging. In *Annual Conference of the PHM Society*, volume 3, 2011.

[13] Yuqiao Zhang and Ujjwal Guin. End-to-end traceability of ics in component supply chain for fighting against recycling. *IEEE Transactions on Information Forensics and Security*, 15:767–775, 2019.

[14] Antonio Hernandez Martinez, Saqib Khursheed, Turki Alnuayri, and Daniele Rossi. Online remaining useful lifetime prediction using support vector regression. *IEEE Transactions on Emerging Topics in Computing*, 2021.

[15] Antonio Leonel Hernández Martínez, Saqib Khursheed, and Daniele Rossi. Leveraging cmos aging for efficient microelectronics design. In *2020 IEEE 26th International Symposium on On-Line Testing and Robust System Design (IOLTS)*, pages 1–4. IEEE, 2020.

[16] Turki Alnuayri, Saqib Khursheed, Antonio Leonel Hernández Martínez, and Daniele Rossi. Differential aging sensor using subthreshold leakage current to detect recycled ics. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 29(12):2064–2075, 2021.

[17] Turki Alnuayri, A Leonel Hernández Martínez, Saqib Khursheed, and Daniele Rossi. A support vector regression based machine learning method for on-chip aging estimation. In *2021 4th International Conference on Computing & Information Sciences (ICCIS)*, pages 1–6. IEEE, 2021.

[18] Liting Yu, Jianguo Ren, Xian Lu, and Xiaoxiao Wang. Nbti and hci aging prediction and reliability screening during production test. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 39(10):3000–3011, 2019.

[19] Seelam V Sai Viswanada Prabhu Deva, Shyam Akashe, et al. Reliability analysis of comparator: Nbti, pbti, hci, ageing. In *International Conference on Communication, Networks and Computing*, pages 606–619. Springer, 2018.

[20] Vartika Pandey, Manisha Pattanaik, and RK Tiwari. Impact of nbti/hci/pbti on 6t sram reliability. In *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, pages 559–562. IEEE, 2020.

[21] S Udaya Shankar and P Kalpana. Reliability and circuit timing analysis with hci and nbti. In *Advances in VLSI, Communication, and Signal Processing*, pages 371–392. Springer, 2021.

[22] Xu Li, Mingming Sun, and Ping Li. Multi-agent discussion mechanism for natural language generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 6096–6103, 2019.

[23] Xiaoxiao Wang and Mohammad Tehranipoor. Novel physical unclonable function with process and environmental variations. In *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, pages 1065–1070. IEEE, 2010.

[24] Chuang Yang and Feng Feng. Multi-step-ahead prediction for a cmos low noise amplifier aging due to nbti and hci using neural networks. *Journal of Electronic Testing*, 35(6):797–808, 2019.

[25] AR Shamshiri, MB Ghaznavi-Ghoushchi, and AR Kariman. Ml-based aging monitoring and lifetime prediction of iot devices with cost-effective embedded tags for edge and cloud operability. *IEEE Internet of Things Journal*, 2021.

[26] Giuseppe Macera. A low jitter low power and wide tuning range differential ring oscillator topology in 28nm cmos technology for clock synthesizer applications. In *2018 29th Irish Signals and Systems Conference (ISSC)*, pages 1–4. IEEE, 2018.

[27] Sebastian Bock, Josef Goppold, and Martin Weiß. An improvement of the convergence proof of the adam-optimizer. *arXiv preprint arXiv:1804.10587*, 2018.

[28] Andrew R Sanderford, Andrew P McCoy, and Matthew J Keefe. Adoption of energy star certifications: theory and evidence compared. *Building research & information*, 46(2):207–219, 2018.

[29] Xiaoxiao Wang, Yueying Han, and Mark Tehranipoor. System-level counterfeit detection using on-chip ring oscillator array. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 27(12):2884–2896, 2019.

[30] Alexei Botchkarev. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv preprint arXiv:1809.03006*, 2018.