

# Remaining Useful Lifetime Prediction of Discrete Power Devices by Means of Artificial Neural Networks

ALESSANDRO VACCARO  (Member, IEEE), DAVIDE BIADENE  (Member, IEEE),  
AND PAOLO MAGNONE  (Senior Member, IEEE)

Department of Management and Engineering, Università degli Studi di Padova, 36100 Vicenza, Italy

CORRESPONDING AUTHOR: PAOLO MAGNONE (e-mail: paolo.magnone@unipd.it)

This work was supported by the research project “Interdisciplinary Strategy for the Development of Advanced Mechatronics Technologies (SISTEMA),” DTG, University of Padova-Project code CUP-C36C18000400001.

**ABSTRACT** This work proposes a deep learning-based model for predicting the lifetime of power devices subjected to power cycling. To this purpose, a neural network based on bidirectional long short-term memory is adopted. The neural network is trained with experimental on-voltage degradation profiles. The application of the proposed method is based on the monitoring of a precursor, that is the on-voltage degradation. According to considered precursor, the model allows predicting the remaining useful lifetime (RUL) of power components. In order to prove the accuracy of the model, TO-247 power devices are stressed under power cycling and their wear-out is experimentally investigated. RUL predicted by the neural network is then compared with the experimental lifetime of power devices. Thanks to the proposed deep learning model, the accuracy in the lifetime estimation improves as long as more information about the state of health of the device under test is acquired.

**INDEX TERMS** Power cycling, IGBT, semiconductor power device reliability, remaining useful lifetime, artificial neural network.

## I. INTRODUCTION

Many industrial, healthcare, automotive, energy, transportation, and aerospace applications rely on power electronic circuits [1]. The requirement for reliability in this field has increased considerably [2], [3]. For instance, in some applications such as avionics, the demand for failure tolerance is even zero [1]. Moreover, the sustainability of a power electronic circuit/system is closely related to its durability. Consequently, it has a significant impact from both economic and safety perspectives [1], [4], [5], [6].

Among the failure mechanisms occurring with greater probability in power electronic circuits, those affecting semiconductor power devices are of high relevance. The power dissipated in electronic devices leads to self-heating effects, which in turn bring to thermo-mechanical stress at the interface of materials with different coefficients of thermal expansion [7]. This phenomenon is particularly severe in the case of varying power dissipation, and it is then referred to as power cycling. Two main failure mechanisms can occur, both

in discrete devices and modules: solder joint fatigue and wire bonds degradation [8].

In general, the lifetime of power components can be estimated by considering model-driven and data-driven approaches. Model-driven approach can be either empirical [9], [10], [11], [12], [13], [14], [15] (i.e., calibrated according to accelerated lifetime tests), or physics-based [16], [17]. Models, in combinations with Miner’s rule [18], allow estimating the lifetime consumption by considering a given mission profile in terms of temperature swing, average temperature, heating time and current density [19]. Data-driven approach is based on the monitoring of the State of Health (SoH) of the component. In the case of wire bonds degradation, on-voltage is usually adopted as a precursor, while in the case of solder joint fatigue the thermal impedance gives a better indication of the SoH [20]. The knowledge of the SoH allows implementing prognostic techniques and hence estimating the Remaining Useful Lifetime (RUL). The implementation of prognostics techniques is the key to achieve predictive maintenance and

hence to avoid catastrophic failure events [1], [21], [22], [23], [24], [25]. In fact, failure phenomena are intrinsically random events, being modeled with the Weibull statistic distribution in the case of power cycling effects [26]. As a result, the assessment of lifetime by means of a model-driven approach, allows estimating the number of cycles to failure for a given Probability of Failure (PoF). The selection of a low value of PoF ( $\leq 10\%$ ) is a conservative approach in the estimation of component/circuit lifetime [27]. Prognostics, based on SoH monitoring, allows overcoming this limitation, since RUL is estimated online for each specific device under test.

Some data-driven methods use the on-voltage to predict the fault event, by implementing particle filter algorithms [28], [29], [30]. In [29] the Mahalanobis distance algorithm is also used for anomaly detection, which, however can be affected by signal noise [31]. According to [32], [33], imprecise knowledge of parameters of the function describing the SoH, as well as the inaccurate initialization of the filter, can lead to inconsistent results in the prognosis.

Neural networks (NNs) represent a viable solution for data-driven prognostic methods, allowing to avoid the definition of models, to learn online and adopt itself to the degradation profile [33]. In [33], a time delay neural network (TDNN) was developed to monitor the SoH of insulated gate bipolar transistors (IGBTs) through the on-voltage and it was combined with a stochastic approach for the prediction of RUL. In [34], a feedforward neural network (FFNN) was considered in order to estimate the RUL based on the evaluation of the on-resistance. However, the above-mentioned NN approaches considered a limited dataset for training the model. In particular, in [33], four profiles were considered, with three utilized for training and one for testing, along with their respective combinations. On the other hand, in [34], the focus was on only two profiles, one for training and the other for testing. However, neither TDNN nor FFNN take into account memory effects, which become especially relevant when the SoH at a given moment is influenced by preceding events.

In [35], lifetime prediction was addressed using a memory-effect-incorporating network called LSTM (Long Short-Term Memory). In this study, a total of six samples were used to train the network. Specifically, leave-one-out cross-validation methodology was employed for the training phase. This form of training entails partitioning individual profile data into training and validation samples. Nonetheless, this strategy results in the omission of some profile data points during the training dataset. Potentially, this approach could curtail the statistical significance and robustness of the outcomes. A similar approach was proposed in [36], where, despite improving the network performance through a physics-informed approach, the training method further fragmented individual profiles into data to be used for the training and for testing phases.

In this work, a data-driven method, based on NNs, has been implemented, allowing to estimate the RUL of semiconductor power devices under power cycling stress. More specifically, the prognostics technique is based on a bLSTM (bidirectional LSTM) network and the on-voltage degradation is adopted

as a precursor parameter. Discrete IGBT devices are stressed under constant power cycling conditions and experimental on-voltage ( $V_{ce,on}$ ) degradation profiles are used to train the NN. The trained bLSTM network allows estimating the End of Life (EoL) of components also based on the real-time acquisition of the on-voltage degradation. By exploiting the memory capability of the bLSTM network, the accuracy of the RUL prediction is improved, allowing to account for the intrinsic statistical distribution of the failure phenomenon. In contrast with the methodology proposed in [35], [36], in this work a sliding window approach is employed to consider the entire on-voltage profile independently of the chosen inputs number. Furthermore, the approach pursued in this work is broader and more comprehensive compared to previous studies. While prior studies considered a single stress condition with a limited number of samples and combinations, this work explores multiple datasets and stress conditions. Specifically, it incorporates outcomes from a comprehensive set of 28 networks, corresponding to all possible combinations for each stress condition.

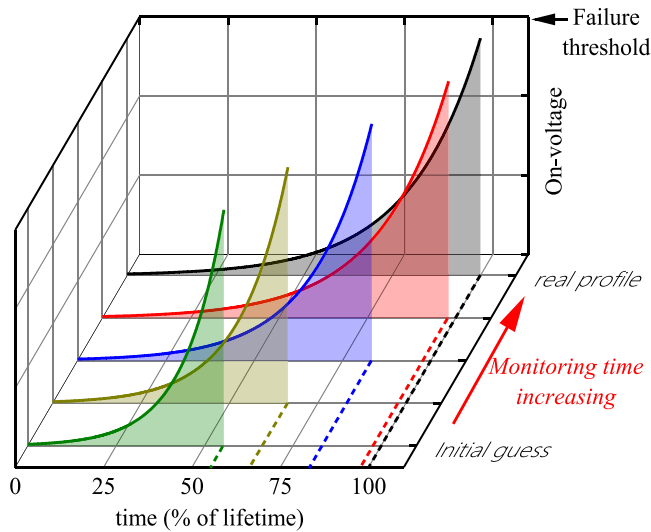
The remainder of this work is organized as follows. In Section II, the data-driven model is described, by focusing on the methodology considered for the training process of the bLSTM network and for its adoption in the RUL estimation. Section III presents the experimental setup along with power cycling experimental tests. In Section IV, several test cases are considered for the evaluation of RUL based on the developed model. Finally, in the conclusive section the main achievements are summarized.

## II. METHODOLOGY FOR RUL ESTIMATION BASED ON BLTSM NETWORK

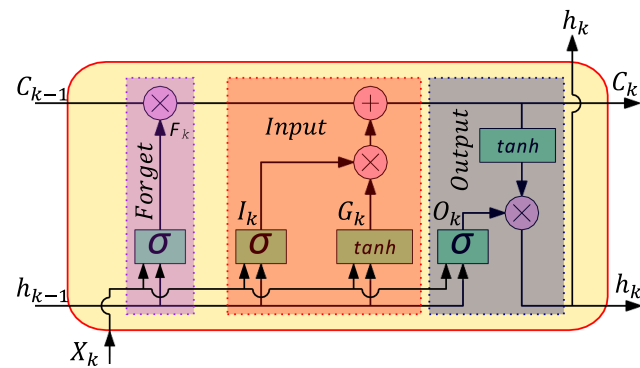
The proposed approach aims at developing a deep learning-based model for predicting the degradation profile of the on-voltage of switching devices under fixed stress conditions. Being the failure event a stochastic phenomenon, NN models are the most suitable to account for the variability in the degradation process. Fig. 1 illustrates the expected outcome of the data-driven model, with the predicted on-voltage profile over time as the model output. In power cycling stress scenarios, the on-voltage is expected to increase due to wire bonds degradation, and a 5% increment is considered as the failure threshold [20]. The estimated on-voltage profile, and consequently the lifetime prediction, relies on the real-time on-voltage acquisition. Initially, the prediction is mainly based on the off-line training of the model, resulting in an approximation close to the average value of the voltage profiles used in the training phase. However, as the monitoring time increases and the on-voltage of the tested device is experimentally measured, the accuracy of the lifetime prediction improves. Consequently, the RUL estimation approaches the ideal value.

### A. ARTIFICIAL NEURAL NETWORK MODEL

To tackle time-sequence forecasting, recurrent neural networks (RNNs) are designed to effectively process sequential data. Compared to traditional feedforward NNs, where inputs are propagated and processed through the hidden layer stack,



**FIGURE 1.** Graphic representation of the expected outcome of the data-driven model.



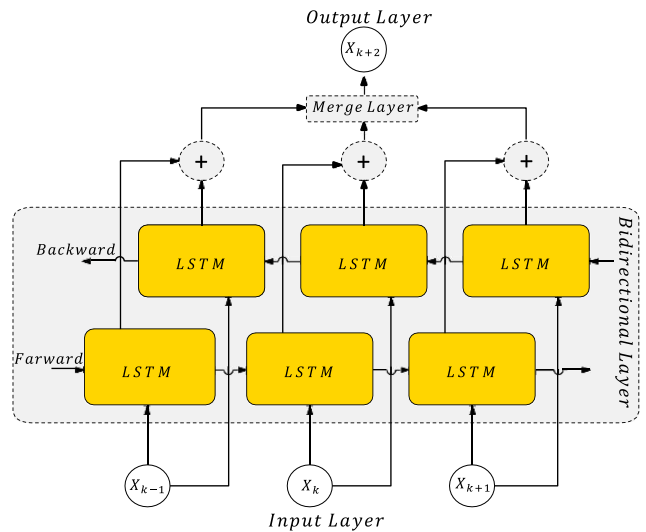
**FIGURE 2.** Schematic description of a gated cell (LSTM network).  $\sigma$  and  $\tanh$  are the sigmoid and hyperbolic functions, respectively.

RNNs allow previous outputs to be used as inputs. The key feature of RNNs is their ability to maintain an internal memory or hidden state that can capture temporal dependencies in the input data. This memory enables RNNs to process sequences of variable length and make predictions based on previous elements in the sequence.

RNNs are affected by the vanishing gradient issue, making it challenging for RNNs to learn and capture long-term dependencies effectively. LSTM can be considered to overcome this problem, thanks to its ability to ignore or retain information to remember [37]. The atomic element of an LSTM network is the gated cell shown in Fig. 2.

The cell is supplied with three gates, namely forget, input and output, regulating the flow of information into and out of the cell. Each gate processes the linear combination of its inputs through a non-linear function (i.e., the activation function) and returns a value between 0 and 1 used to weigh the desired information. The forget gate combines the input  $x_k$  and the output of previous  $h_{k-1}$ :

$$F_k = \sigma(W_{F,h}[h_{k-1}], W_{F,x}[x_k], b_F) \quad (1)$$



**FIGURE 3.** Bidirectional long-short term memory (bLSTM) network.

Where  $W_{F,h}$  and  $W_{F,x}$  are weight matrices,  $b_F$  is a bias constant and  $\sigma$  is the sigmoid activation function.

The input gate  $I_k$  regulates the amount of new information, i.e.,  $G_k$ , that has to be added to the LSTM cell's memory.  $I_k$  and  $G_k$  components are non-linear functions of  $x_k$  and  $h_{k-1}$ , each one with its respective activation function: sigmoid ( $\sigma$ ) and  $\tanh$  [36]:

$$I_k = \sigma(W_{I,h}[h_{k-1}], W_{I,x}[x_k], b_I) \quad (2)$$

$$G_k = \tanh(W_{G,h}[h_{k-1}], W_{G,x}[x_k], b_G) \quad (3)$$

$W_{I,h}$ ,  $W_{I,x}$ ,  $W_{G,h}$ ,  $W_{G,x}$  refer to the weight matrix expressions associated with these two layers, and  $b_I$  and  $b_G$  are bias constants. The outcomes of (2) and (3) are combined with the contribution of the previous state  $c_{k-1}$  and with (1) to define the state  $c_k$  as follows:

$$c_k = F_k \cdot c_{k-1} + I_k \cdot G_k \quad (4)$$

The output gate  $O_k$  is related to  $x_k$  and  $h_{k-1}$  as:

$$O_k = \sigma(W_{O,h}[h_{k-1}], W_{O,x}[x_k], b_O) \quad (5)$$

Where  $W_{O,h}$ ,  $W_{O,x}$  and  $b_O$  represent the weight matrix and bias constant associated with the output gate. Ultimately, the cell output  $h_k$  is governed through the following equation:

$$h_k = O_k \cdot \tanh(c_k) \quad (6)$$

The input gate  $I_k$  regulates the amount of new information, i.e.,  $G_k$ , that has to be added to the LSTM cell's memory.  $I_k$  and  $G_k$  components are non-linear functions of  $x_k$  and  $h_{k-1}$ , each one with its respective activation function: sigmoid ( $\sigma$ ) and  $\tanh$  [36]: Remarkably, inputs and states are both processed using the  $\tanh$  function to mitigate the vanishing or exploding gradient issues.

An extension and improvement of LSTM performance is achieved with the bidirectional LSTM (bLSTM) [38]. As illustrated in Fig. 3, bLSTM consists of two chains of LSTM

cells that consider both time directions. According to the temporal input order  $x_k$ , gated cells connected in ascendent order define the forward state. On the contrary, the ones associated with the descending order give the backward state. The output layer (i.e., the output sequence  $y_k$ ) is then given by a combination of both forward and backward states.

### B. TRAINING OF BLSTM MODEL

The architecture of the artificial neural network (ANN) model is based on a time-series forecasting structure. Multiple bLSTM layers are connected in cascade to catch the trend of the target on-voltage profile through the selected activation functions in each layer [39].

The target output is the on-voltage profile of a power device under the effect of power cycling stress. The device voltage is measured after each temperature cycle, meaning the profile is a function of the number of applied cycles. To hold down the complexity of the ANN, the samples are filtered and down-sampled (i.e., 100:1).

The following approach is based on a single-step time-series forecasting model. A fixed window, containing  $m$  samples, from the input sequence  $x$  is selected as the model's input (i.e.,  $x_k, \dots, x_{k-m+1}$ ). The neural network predicts the subsequent value  $\tilde{x}_{k+1}$ , where  $k$  is the index of the last input value.

The learning process is aimed at tuning the parameters of the non-linear function  $f_{NN}$  associated with the ANN architecture minimizing the loss function (e.g., RMSE) of the predicted value:

$$\tilde{x}_{k+1} = f_{NN}(x_k, x_{k-1}, \dots, x_{k-m+1}) \quad (7)$$

with respect to the real one  $x_{k+1}$ . To this purpose, the input dataset used for the training is composed of portions of the on-voltage profiles arising from different samples. The corresponding next value of the sequence window is the target output.

### C. RUL ESTIMATION

The proposed approach is aimed at estimating the RUL of a device under constant power cycling stress. The forecast is based on recursive iterations of the bLSTM model to obtain the on-voltage profile along the thermal cycles, as schematically reported in Fig. 4. At the first iteration (initial guess),  $m$  samples of the experimental profile are provided to the NN model to guess the subsequent value  $\tilde{x}_{k+1}$ . At the next iteration, the predicted value  $\tilde{x}_{k+1}$  is used as the model's input discarding the oldest sample  $x_{k-m+1}$  and sliding one step forward the  $m$ -length window. At the  $i$ -th iteration, with  $i \geq 1$ , the on-voltage is predicted through both experimental and predicted samples if  $i < m$ , or only predicted values if  $i \geq m$

$$\begin{aligned} \tilde{x}_{k+i} &= f_{NN}(\tilde{x}_{k+i-1}, \dots, \tilde{x}_{k+1}, x_k, \dots, x_{k-m+i+1}), \quad i < m \\ \tilde{x}_{k+i} &= f_{NN}(\tilde{x}_{k+i-1}, \dots, \tilde{x}_{k-m+i+1}), \quad i \geq m \end{aligned} \quad (8)$$

This process is iterated until  $\tilde{x}_{k+i}$  reaches the EoL condition (i.e., an increase of 5% of the initial on-voltage value). From

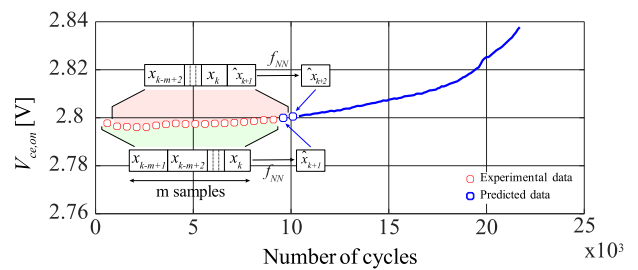


FIGURE 4. On-voltage prediction according to the proposed methodology.  $m$  samples are considered ( $x_k, \dots, x_{k-m+1}$ ) as the input of the NN and allow calculating  $\tilde{x}_{k+1}$ . Subsequently, the vector ( $\tilde{x}_{k+1}, \dots, x_{k-m+2}$ ) is considered as a new input of the NN and another value ( $\tilde{x}_{k+2}$ ) is estimated. This process is repeated until the EoL condition is reached.

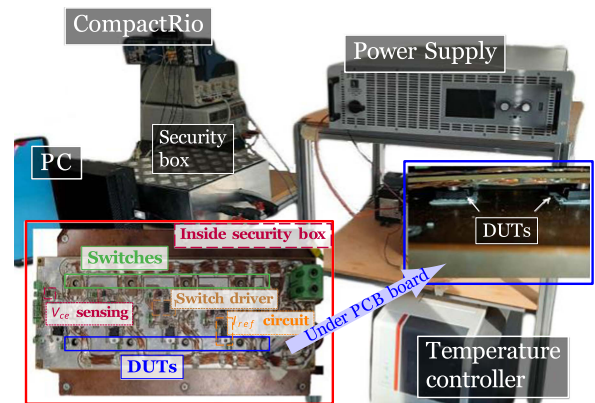


FIGURE 5. Picture of the experimental setup for power cycling tests.

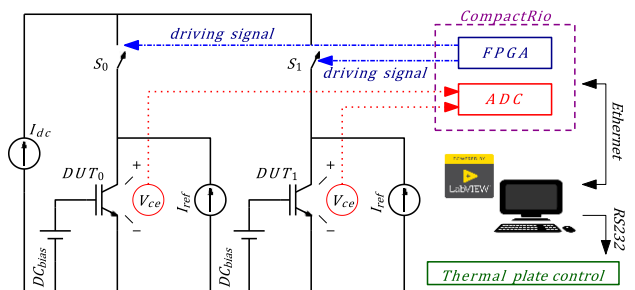


FIGURE 6. Schematic description of power cycling tests.

this definition, the RUL can be expressed as

$$RUL(k) = i \mid \tilde{x}_{k+i} \geq x_{EoL} \text{ AND } \tilde{x}_{k+i-1} < x_{EoL} \quad (9)$$

where  $k$  and  $i$  represent the number of monitored cycles and the remaining number of cycles to failure, respectively.  $x_{EoL}$  is the failure threshold.

## III. EXPERIMENTAL POWER CYCLING TESTS

### A. EXPERIMENTAL SETUP

The experimental investigation of power cycling phenomenon requires the application of controlled temperature cycles in the device under tests (DUTs), along with the capability of real-time monitoring the on-voltage. The experimental setup adopted for this goal is reported in Fig. 5 [40]. It consists

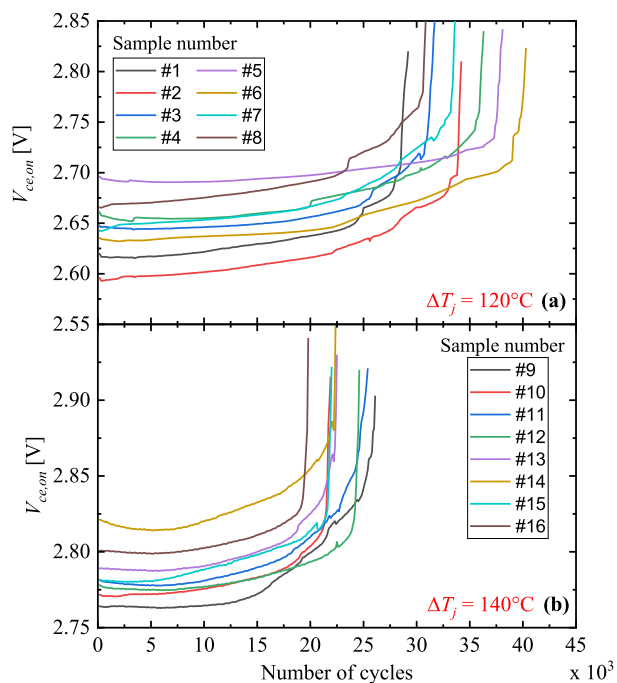
of a power supply (EA-PSB 9080-120), a custom board with DUTs placed on liquid-cooled thermal plate, a temperature controller (Julabo Presto A40) and a compactRIO system. As reported in Fig. 6, two IGBT devices are stressed within the same experiment. The power supply provides a high current ( $I_{dc}$ ), flowing alternately in the two DUTs. The compactRIO generates control signals for switches  $S_0$  and  $S_1$ , and allows for  $V_{ce}$  measurements on the DUTs. In order to measure  $V_{ce}$ , an amplifier with voltage gain of 3 is adopted. The conditioned signal is acquired by the compactRIO's analog-to-digital converter (voltage range of  $\pm 10$ V, sampling frequency of 1 MS/s and resolution of 16 bits). The thermal cycling across the DUT arises from a heating-up phase and a cooling-down phase, lasting a time  $t_{on}$  and  $t_{off}$ , respectively. The desired temperature swing ( $\Delta T_j$ ) is achieved by properly selecting  $I_{dc}$ ,  $t_{on}/t_{off}$  times and the temperature of the thermal plate. Although the current in both DUTs is the same, the temperature swings can be slightly different, because of mismatches in the thermal pads and intrinsic devices characteristics, or because of mutual heating effects. In order to achieve the same  $\Delta T_j$  on both devices, slightly different  $t_{on}$  times are selected. According to the guidelines for the qualification of power devices, such as [41], heating current and  $t_{on}/t_{off}$  times are kept constant during the entire experiment. Since the component degrades during the power cycling test, changes of  $\Delta T_j$  are possible.

The gate of DUTs is biased with a DC voltage of 15V, hence devices are in conduction state for the entire experiment. During the on-phase, the on-voltage across the IGBT ( $V_{ce,on}$ ) is acquired and used to monitor the degradation state of the component. Typically, increases in  $V_{ce,on}$  ranging from 5% to 20% are regarded as EoL thresholds for determining device failure due to wire bond degradation (the sole failure effect considered in this work) [20]. In this study, an increase of 5% in  $V_{ce,on}$  is considered as EoL condition. During the off-phase, a small current  $I_{ref} = 50$  mA is injected in the device. The measured  $V_{ce,off}$  voltage is used as a Temperature Sensitive Electrical Parameter, allowing to estimate the junction temperature of the component.

DUTs used in the experiments are commercial IGBTs in TO-247 packages, with a rated pulsed current of 120A, rated voltage of 600V, typical on-resistance of 10m $\Omega$ , and maximum junction temperature of 175  $^{\circ}$ C.

## B. POWER CYCLING EXPERIMENTS

Power cycling tests are carried out in this work by considering two different types of stress:  $\Delta T_j = 120$   $^{\circ}$ C ( $I_{dc} = 70.5$ A) and  $\Delta T_j = 140$   $^{\circ}$ C ( $I_{dc} = 68.5$ A). In both cases, the minimum junction temperature is 25  $^{\circ}$ C. For each stress condition, eight different DUTs were considered. Experimental  $V_{ce,on}$  profiles as a function of the number of cycles are extrapolated from [42] and reported in Fig. 7. Initially,  $V_{ce,on}$  is almost constant, while for a large number of cycles an increase of  $V_{ce,on}$  is observed, which can be ascribed to wire bonds degradation. The increase of  $V_{ce,on}$  by 5%, with respect to the initial value, is commonly considered as a failure criterion for the device.



**FIGURE 7.** Experimental on-voltage profiles as a function of the number of cycles.  $V_{ce,on}$  profiles are obtained for (a)  $\Delta T_j = 120$   $^{\circ}$ C and (b)  $\Delta T_j = 140$   $^{\circ}$ C.

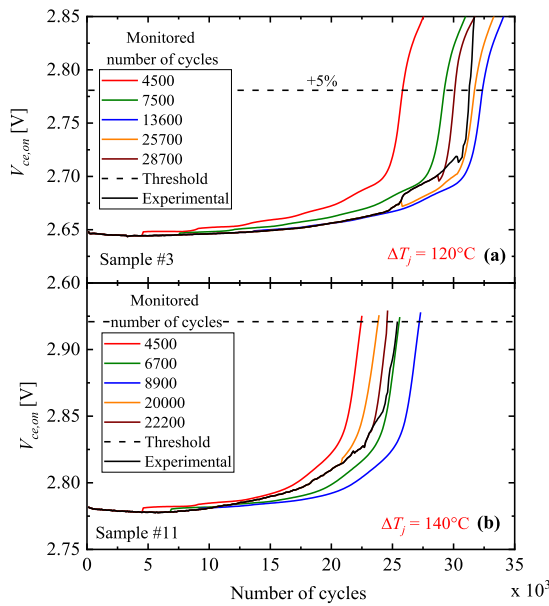
It is worth noting that the increase of  $V_{ce,on}$  slightly changes the temperature in the device. In fact, at the end of each experiment,  $\Delta T_j$  exceeds the nominal value of about 10  $^{\circ}$ C (not shown here). This temperature increase is expected to modify the number of cycles to failure. More specifically, according to [7], [43], [44], [45], [46], a lower lifetime is foreseen with respect to the case of a constant  $\Delta T_j$  for the entire experiment.

The application of a given thermal cycling stress (either 120  $^{\circ}$ C or 140  $^{\circ}$ C), leads to a significant randomness in the device lifetime (in terms of the number of cycles to failure), which is well described by a Weibull distribution [40]. It is therefore fundamental that the proposed neural network model is trained by considering an adequate number of samples, having different lifetimes. This allows the neural network to be robust against the intrinsic variability of failure events.

## IV. RESULTS AND DISCUSSION

### A. TEST RESULTS OF THE NEURAL NETWORK

The proposed neural network has been trained according to the procedure reported in Section II-B, by using the experimental  $V_{ce,on}$  profiles reported in Fig. 7. These profiles are decimated by a factor 100 in order to reduce the complexity of the neural network while maintaining good performances. A window size ( $m$ ) of 45 elements (which also corresponds to the batch size of the bLSTM) is considered for both training and testing phases, corresponding to 4500 cycles for the chosen decimation factor. The network structure consists of a sequence of bLSTM layers, with the initial layer comprising 16 units, followed by a subsequent layer with 36 units.

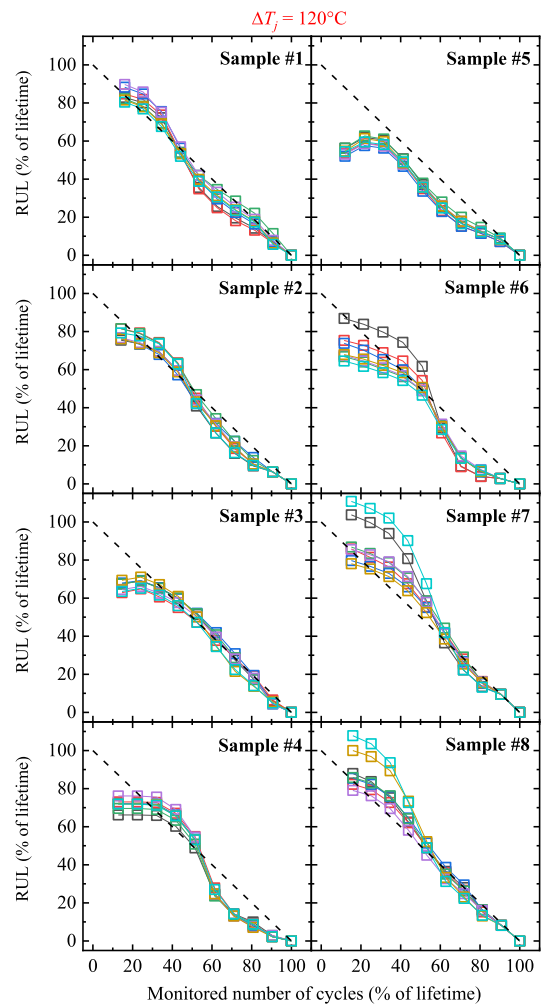


**FIGURE 8.**  $V_{ce,on}$  profiles estimated by the neural network in the case of (a)  $\Delta T_j = 120^\circ\text{C}$  and (b)  $\Delta T_j = 140^\circ\text{C}$ . Each curve arises from the experimental observation of a given number of power cycles (as reported in the legend) and from the application of the proposed recursive algorithm. As a result, the accuracy in the lifetime estimation improves as long as the monitored number of cycles increases.

Additionally, two individual units of bLSTM are present, employing *tanh* and exponential activation functions to enhance the understanding of the  $V_{ce,on}$  profile behaviour from the mentioned 16 and 36-units bLSTM layers. The outputs of these supplementary units are ultimately combined in the last layer of the network, which performs summation.

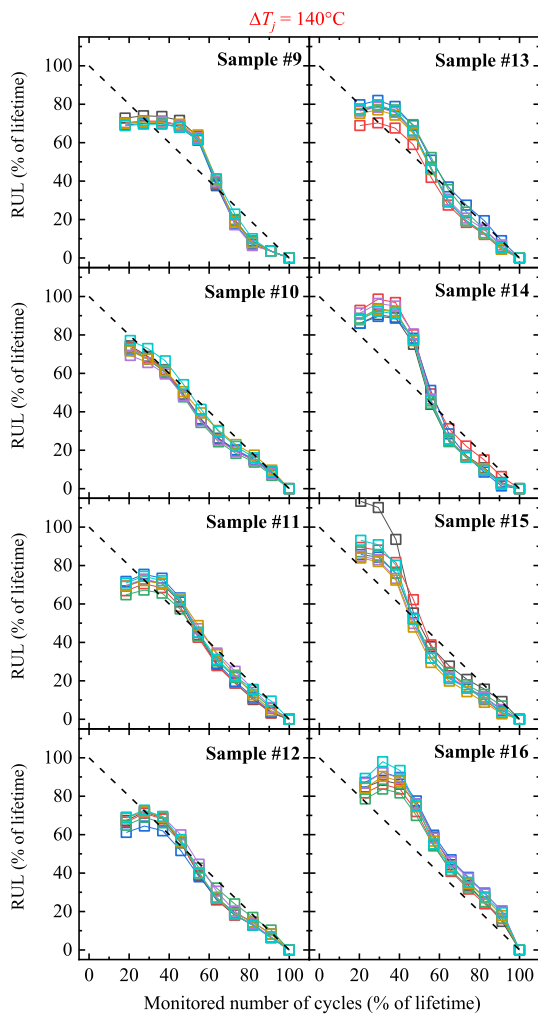
Regularization techniques have been implemented to improve the network's learning ability, and the Adam algorithm with a learning rate of 0.1 has been used to train the neural network [47]. The dataset is split into a training subset (6 profiles) and a test subset (2 profiles). To verify the robustness of the model concerning the partition of the available dataset, the model is trained using every possible unique combination of the 8 available samples, resulting in a total of 28 distinct neural networks. This number (28) is determined by the binomial coefficient (8, 6), being 8 the number of available experimental samples and 6 the number of samples included in the training subset. It is worth mentioning that all 28 NNs share the same architecture but are individually trained with a different selection of 6 samples and are tested with the remaining 2 samples, ensuring a unique combination of training/test subset.

Two different conditions are considered for the training phase:  $\Delta T_j = 120^\circ\text{C}$  and  $\Delta T_j = 140^\circ\text{C}$ . An example of  $V_{ce,on}$  profiles estimated by means of the neural network is reported in Fig. 8. In particular, Fig. 8(a) (or 8b) considers a neural network trained at  $\Delta T_j = 120^\circ\text{C}$  (or  $\Delta T_j = 140^\circ\text{C}$ ) with samples #1, #2, #4, #5, #6, and #7 (or samples #9, #10, #12, #13, #14 and #15) and tested on sample #3 (or #11). Experimental  $V_{ce,on}$  profiles as a function of the number of cycles are reported in black (solid lines), along with the



**FIGURE 9.** Predicted RULs in comparison with ideal RULs (dashed curves) for all 8 samples stressed at  $\Delta T_j = 120^\circ\text{C}$ . The 7 curves reported in each subplot arise from different neural networks, each one trained with a different selection of the 6 (out of 8) training samples.

thresholds assumed for the failure criterion (dashed lines). The other curves are those predicted by the neural network according to the selected observation windows, i.e., the monitored number of cycles indicated as  $k$  in (8) and (9). After an observation of 4500 cycles, predicted lifetimes are relatively different from those experimentally evaluated. However, the predicted values are within the range of values adopted for the neural network training. It is worth noting that, the training phase is based on power cycling experiments carried out with a constant current stress, where  $\Delta T_j$  slightly increases over the wear out phase. As a consequence, the proposed model is affected by inaccuracy in the initial stage of the monitoring phase, being the predicted lifetime mainly based on the average profiles adopted for the training phase. Moreover, when a limited number of cycles is monitored, the degradation of  $V_{ce,on}$  can be negligible and the SoH cannot be quantified by the model. As a result, the accuracy does not necessarily improve in this case. As the monitored number of cycles increases, and more knowledge is available about the SoH of the component, the predicted  $V_{ce,on}$  profiles get



**FIGURE 10.** Predicted RULs in comparison with ideal RULs (dashed curves) for all 8 samples stressed at  $\Delta T_j = 140^\circ\text{C}$ . The 7 curves reported in each subplot arise from different neural networks, each one trained with a different selection of the 6 (out of 8) training samples.

closer to the expected ones, hence improving the accuracy in the lifetime estimation.

**B. REMAINING USEFUL LIFETIME**

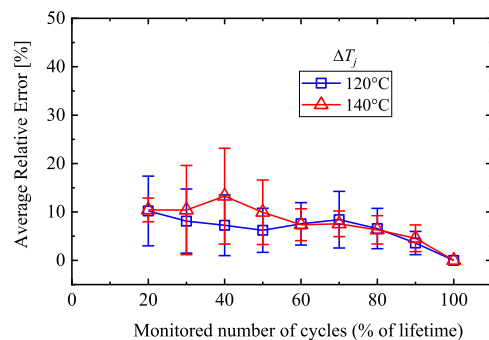
The remaining useful lifetime represents the difference between the predicted lifetime and the monitoring time, both expressed as number of cycles. In Fig. 8, the predicted lifetime is calculated as the number of cycles required to reach an increase of  $V_{ce,on}$  by 5%. Hence, the RUL can be easily calculated as a function of the monitored number of cycles. For the given dataset, by considering the selection of 6 out of 8 samples for the training phase, the testing on each sample foresees 7 differently trained neural networks.

The results of the RUL analysis are reported in Figs. 9 and 10 for  $\Delta T_j = 120^\circ\text{C}$  and  $\Delta T_j = 140^\circ\text{C}$ , respectively. Both RUL and monitored number of cycles are expressed as a percentage value of the effective lifetime. The RUL is estimated for all the 16 samples (8 for each stress condition) considered in this work as summarized in Table I. As mentioned above,

**TABLE 1** Average Value of RULs Predicted in Figs. 9 and 10 as a Function of the Monitored Number of Cycles: 25%, 50%, and 75% of the Expected Lifetime

DUTs	$\Delta T_j = 120^\circ\text{C}$		
	25%	50%	75%
Ideal RUL	75%	50%	25%
Sample #1	80.88%	44.53%	21.88%
Sample #2	75.10%	47.45%	20.80%
Sample #3	67.41%	52.19%	23.28%
Sample #4	71.44%	54.53%	14.90%
Sample #5	59.61%	37.42%	17.62%
Sample #6	67.48%	52.46%	13.22%
Sample #7	86.47%	62.69%	22.32%
Sample #8	86.49%	54.78%	24.44%

DUTs	$\Delta T_j = 140^\circ\text{C}$		
	25%	50%	75%
Ideal RUL	75%	50%	25%
Sample #9	70.87%	65.65%	25.38%
Sample #10	70.60%	45.50%	22.84%
Sample #11	71.28%	53.05%	24.41%
Sample #12	69.35%	49.03%	22.19%
Sample #13	76.98%	59.24%	25.00%
Sample #14	91.13%	67.25%	20.97%
Sample #15	90.09%	46.48%	19.51%
Sample #16	89.51%	54.40%	26.19%



**FIGURE 11.** Relative error between predicted and experimental lifetime as a function of the monitored number of cycles. Errors are averaged over the 56 tests for both  $\Delta T_j = 120^\circ\text{C}$  and  $\Delta T_j = 140^\circ\text{C}$ . Error bars represent the standard deviations around the average values.

the 7 different curves reported in each sub-plot refer to different neural networks, trained with a different combination of samples. For each  $\Delta T_j$  stress condition, 28 neural networks are trained in total, which are used to test the 2 samples not adopted in the training phase of the specific neural network. As a result, 56 RUL curves are visible in each figure. Although the estimated RULs can be initially different with respect to the ideal ones (black dashed lines), in general the accuracy of the RUL prediction improves with the monitored number of cycles. For each sample, reported in Figs. 9 and 10, the RULs predicted with 7 different NNs are averaged and the results are summarized in Table I.

In order to assess the performance of the proposed neural network model, the relative error, defined as the relative difference between the predicted and the experimental lifetime, is averaged for all the 56 tests performed at a given  $\Delta T_j$ . The results are reported in Fig. 11. Regarding the relative error, in the range of the monitored number of cycles, comprised between 20% and 100% of the device lifetime, its average value

is always lower than 15%, although there are individual cases in which a larger error can be found (according to the error bar of Fig. 11). As long as the number of cycles increases, the relative error, along with the standard deviation associated to the averaging process, tends to decrease. For example, by exceeding 80% of the device lifetime, the average relative error is below 7%, with a standard deviation lower than 5%. This is a remarkable result for predictive maintenance, since the EoL can be accurately predicted well before the failure event.

## V. CONCLUSION

In this article, the development of a deep learning-based model for the lifetime prediction of semiconductor power devices is discussed. The proposed NN model is composed of bidirectional LSTM blocks. The model is trained with experimental on-voltage degradation profiles arising from power cycling stresses and featuring a temperature swing  $\Delta T_j$  of 120 °C and 140 °C. Eight samples are considered for each stress condition, representing the dataset adopted to train and to test the proposed neural network.

A fundamental peculiarity of the model is that the training phase is carried out by considering a significant number of experimental on-voltage profiles arising from different samples stressed under the same conditions. More specifically, 6 out of 8 samples are adopted for the training phase.

The application of the model consists in the prediction of the lifetime based on the monitoring of the on-voltage profile. When a limited amount of data is available, the lifetime prediction is within experimental range of samples adopted in the training phase. As long as more data are acquired, concerning the SoH of the device under test, the accuracy of the model improves.

In order to understand the impact of dataset partitioning on the NN performance, the model is trained with all the possible combinations of subsets. Therefore, 28 neural networks are trained for each  $\Delta T_j$  stress condition. Those networks are hence used in this work to evaluate the RUL of test samples as a function of the monitored number of cycles. The relative error between the lifetime predicted by the NN and the actual experimental lifetime tends to decrease by increasing the monitored number of cycles. Its average value (among all the trained neural networks) is always lower than 13% and it becomes as low as 5% when the monitoring time is above 80% of the device lifetime. The accuracy of the model is influenced by the size of the training dataset. Therefore, a larger number of experiments is expected to improve the capability of the model to recognize any on-voltage degradation profile.

## REFERENCES

- [1] A. Hanif, Y. Yu, D. Devoto, and F. Khan, "A comprehensive review toward the state-of-the-art in failure and lifetime predictions of power electronic devices," *IEEE Trans. Power Electron.*, vol. 34, no. 5, pp. 4729–4746, May 2019, doi: [10.1109/TPEL.2018.2860587](https://doi.org/10.1109/TPEL.2018.2860587).
- [2] W. Huai et al., "Transitioning to physics-of-failure as a reliability driver in power electronics," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 2, no. 1, pp. 97–114, Mar. 2014, doi: [10.1109/JESTPE.2013.2290282](https://doi.org/10.1109/JESTPE.2013.2290282).
- [3] M. Novak, A. Sangwongwanich, and F. Blaabjerg, "Monte Carlo-based reliability estimation methods for power devices in power electronics systems," *IEEE Open J. Power Electron.*, vol. 2, pp. 523–534, 2021, doi: [10.1109/OJPEL.2021.3116070](https://doi.org/10.1109/OJPEL.2021.3116070).
- [4] A. Sangwongwanich and F. Blaabjerg, "Reliability assessment of fault-tolerant power converters including wear-out failure," in *Proc. IEEE Conf. Appl. Power Electron. Conf. Expo.*, 2022, pp. 300–306, doi: [10.1109/APEC43599.2022.9773367](https://doi.org/10.1109/APEC43599.2022.9773367).
- [5] M. Demir, G. Kahramanoglu, and A. B. Yildiz, "Importance of reliability for power electronic circuits, case study: Inrush current test and calculating of fuse melting point," in *Proc. IEEE Int. Power Electron. Motion Control Conf.*, 2016, pp. 830–834, doi: [10.1109/EPEPEMC.2016.7752101](https://doi.org/10.1109/EPEPEMC.2016.7752101).
- [6] S. Peyghami, P. Palensky, and F. Blaabjerg, "An overview on the reliability of modern power electronic based power systems," *IEEE Open J. Power Electron.*, vol. 1, pp. 34–50, 2020, doi: [10.1109/OJPEL.2020.2973926](https://doi.org/10.1109/OJPEL.2020.2973926).
- [7] S.-H. Tran et al., "Constant  $\Delta T_j$  power cycling strategy in DC mode for top-metal and bond-wire contacts degradation investigations," *IEEE Trans. Power Electron.*, vol. 34, no. 3, pp. 2171–2180, Mar. 2019, doi: [10.1109/TPEL.2018.2847234](https://doi.org/10.1109/TPEL.2018.2847234).
- [8] H. S. H. Chung, H. Wang, F. Blaabjerg, and M. Pecht, *Reliability of Power Electronic Converter Systems*. Stevenage, U.K.: Inst. Eng. Technol., 2016.
- [9] B. Ji, V. Pickert, B. Zahawi, and M. Zhang, "In-situ bond wire health monitoring circuit for IGBT power modules," in *Proc. IEEE Sixth IET Conf. Power Electron., Mach. Drives.*, 2012, vol. 2012, pp. 1–6, doi: [10.1049/cp.2012.0239](https://doi.org/10.1049/cp.2012.0239).
- [10] W. Lai et al., "Study on the lifetime characteristics of power modules under power cycling conditions," *Inst. Eng. Technol. Power Electron.*, vol. 9, no. 5, pp. 1045–1052, 2016, doi: [10.1049/iet-pel.2015.0225](https://doi.org/10.1049/iet-pel.2015.0225).
- [11] M. Held, P. Jacob, G. Nicoletti, P. Scacco, and M. H. Pösch, "Fast power cycling test for IGBT modules in traction application," in *Proc. Second Int. Conf. Power Electron. Drive Syst.*, 1997, vol. 1, pp. 425–430, doi: [10.1109/peds.1997.618742](https://doi.org/10.1109/peds.1997.618742).
- [12] R. Bayerer, T. Herrmann, T. Licht, J. Lutz, and M. Feller, "Model for power cycling lifetime of IGBT modules-various factors influencing lifetime," in *Proc. CIPS -5th Int. Conf. Integr. Power Electron. Syst.*, 2008, pp. 37–42.
- [13] U. Scheuermann and R. Schmidt, "A new lifetime model for advanced power modules with sintered chips and optimized Al wire bonds," in *Proc. PCIM Europe Conf.*, 2013, pp. 810–817.
- [14] G. Zeng, L. Borucki, O. Wenzel, O. Schilling, and J. Lutz, "First results of development of a lifetime model for transfer molded discrete power devices," in *Proc. PCIM Europe Conf.*, 2018, pp. 706–713.
- [15] A. Vaccaro, P. Magnone, A. Zilio, and P. Mattavelli, "Predicting lifetime of semiconductor power devices under power cycling stress using artificial neural network," *IEEE J. Emerg. Sel. Topics Power Electron.*, early access, Jul. 27, 2022, doi: [10.1109/JESTPE.2022.3194189](https://doi.org/10.1109/JESTPE.2022.3194189).
- [16] N. Dornic et al., "Stress-based model for lifetime estimation of bond wire contacts using power cycling tests and finite-element modeling," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 7, no. 3, pp. 1659–1667, Sep. 2019, doi: [10.1109/JESTPE.2019.2918941](https://doi.org/10.1109/JESTPE.2019.2918941).
- [17] K. Sasaki et al., "Thermal and structural simulation techniques for estimating fatigue life of an IGBT module," in *Proc. 20th Int. Symp. Power Semicond. Devices ICs*, 2008, pp. 181–184, doi: [10.1109/ISPSD.2008.4538928](https://doi.org/10.1109/ISPSD.2008.4538928).
- [18] M. A. Miner, "Cumulative damage in fatigue," *J. Appl. Mech.*, vol. 12, no. 3, pp. A159–A164, 1945, doi: [10.1115/1.4009458](https://doi.org/10.1115/1.4009458).
- [19] P. D. Reigosa, H. Wang, Y. Yang, and F. Blaabjerg, "Prediction of bond wire fatigue of IGBTs in a PV inverter under a long-term operation," *IEEE Trans. Power Electron.*, vol. 31, no. 10, pp. 7171–7182, Oct. 2016, doi: [10.1109/TPEL.2015.2509643](https://doi.org/10.1109/TPEL.2015.2509643).
- [20] J. Lutz, H. Schlangenotto, U. Scheuermann, and R. De Doncker, *Semiconductor Power Devices*, vol. 4. Berlin, Germany: Springer, 2011.
- [21] J. Poon, P. Jain, C. Spanos, S. K. Panda, and S. R. Sanders, "Fault prognosis for power electronics systems using adaptive parameter identification," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2862–2870, May/Jun. 2017, doi: [10.1109/TIA.2017.2664052](https://doi.org/10.1109/TIA.2017.2664052).
- [22] D. Astigarraga et al., "Analysis of the results of accelerated aging tests in insulated gate bipolar transistors," *IEEE Trans. Power Electron.*, vol. 31, no. 11, pp. 7953–7962, Nov. 2016, doi: [10.1109/TPEL.2015.2512923](https://doi.org/10.1109/TPEL.2015.2512923).



- [23] S. H. Ali, S. Dusmez, and B. Akin, "Investigation of collector emitter voltage characteristics in thermally stressed discrete IGBT devices," in *Proc. IEEE Energy Convers. Congr. Expo.*, 2016, pp. 1–6, doi: [10.1109/ECCE.2016.7855216](https://doi.org/10.1109/ECCE.2016.7855216).
- [24] N. Patil, J. Celaya, D. Das, K. Goebel, and M. Pecht, "Precursor parameter identification for insulated gate bipolar transistor (IGBT) prognostics," *IEEE Trans. Rel.*, vol. 58, no. 2, pp. 271–276, Jun. 2009, doi: [10.1109/TR.2009.2020134](https://doi.org/10.1109/TR.2009.2020134).
- [25] K. Hu et al., "Cost-effective prognostics of IGBT bond wires with consideration of temperature swing," *IEEE Trans. Power Electron.*, vol. 35, no. 7, pp. 6773–6784, Jul. 2020, doi: [10.1109/TPEL.2019.2959953](https://doi.org/10.1109/TPEL.2019.2959953).
- [26] R. Amro, J. Lutz, and A. Lindemann, "Power cycling with high temperature swing of discrete components based on different technologies," in *Proc. IEEE Rec. Annu. Power Electron. Specialists Conf.*, 2004, vol. 4, pp. 2593–2598, doi: [10.1109/PESC.2004.1355239](https://doi.org/10.1109/PESC.2004.1355239).
- [27] Y. Zhang, H. Wang, Z. Wang, Y. Yang, and F. Blaabjerg, "Impact of lifetime model selections on the reliability prediction of IGBT modules in modular multilevel converters," in *Proc. IEEE Energy Convers. Congr. Expo.*, 2017, pp. 4202–4207, doi: [10.1109/ECCE.2017.8096728](https://doi.org/10.1109/ECCE.2017.8096728).
- [28] Z. Rao, M. Huang, and X. Zha, "IGBT remaining useful life prediction based on particle filter with fusing precursor," *IEEE Access*, vol. 8, pp. 154281–154289, 2020, doi: [10.1109/ACCESS.2020.3017949](https://doi.org/10.1109/ACCESS.2020.3017949).
- [29] N. Patil, D. Das, and M. Pecht, "A prognostic approach for non-punch through and field stop IGBTs," *Microelectronics Rel.*, vol. 52, no. 3, pp. 482–488, 2012, doi: [10.1016/j.microrel.2011.10.017](https://doi.org/10.1016/j.microrel.2011.10.017).
- [30] M. S. Haque, S. Choi, and J. Baek, "Auxiliary particle filtering-based estimation of remaining useful life of IGBT," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2693–2703, Mar. 2018, doi: [10.1109/TIE.2017.2740856](https://doi.org/10.1109/TIE.2017.2740856).
- [31] W. Chen, L. Zhang, K. Pattipati, A. M. Bazzi, S. Joshi, and E. M. Dede, "Data-driven approach for fault prognosis of SiC MOSFETs," *IEEE Trans. Power Electron.*, vol. 35, no. 4, pp. 4048–4062, Apr. 2020, doi: [10.1109/TPEL.2019.2936850](https://doi.org/10.1109/TPEL.2019.2936850).
- [32] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, 2nd ed. Oxford, U.K.: OUP Oxford, 2012.
- [33] A. Alghassi, S. Perinpanayagam, and M. Samie, "Stochastic RUL calculation enhanced with TDNN-based IGBT failure modeling," *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 558–573, Jun. 2016, doi: [10.1109/TR.2015.2499960](https://doi.org/10.1109/TR.2015.2499960).
- [34] K. Pugalenth, H. Park, and N. Raghavan, "Prognosis of power MOSFET resistance degradation trend using artificial neural network approach," *Microelectronics Rel.*, vol. 100, 2019, Art. no. 113467, doi: [10.1016/j.microrel.2019.113467](https://doi.org/10.1016/j.microrel.2019.113467).
- [35] W. Li, B. Wang, J. Liu, G. Zhang, and J. Wang, "IGBT aging monitoring and remaining lifetime prediction based on long short-term memory (LSTM) networks," *Microelectronics Rel.*, vol. 114, 2020, Art. no. 113902, doi: [10.1016/j.microrel.2020.113902](https://doi.org/10.1016/j.microrel.2020.113902).
- [36] Z. Lu, C. Guo, M. Liu, and R. Shi, "Remaining useful lifetime estimation for discrete power electronic devices using physics-informed neural network," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 10167, doi: [10.1038/s41598-023-37154-5](https://doi.org/10.1038/s41598-023-37154-5).
- [37] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of LSTM and BiLSTM in forecasting time series," in *Proc. IEEE Int. Conf. Big Data*, 2019, pp. 3285–3292, doi: [10.1109/Big-Data47090.2019.9005997](https://doi.org/10.1109/Big-Data47090.2019.9005997).
- [38] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997, doi: [10.1109/78.650093](https://doi.org/10.1109/78.650093).
- [39] T. Xia, Y. Song, Y. Zheng, E. Pan, and L. Xi, "An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation," *Comput. Ind.*, vol. 115, 2020, Art. no. 103182, doi: [10.1016/j.compind.2019.103182](https://doi.org/10.1016/j.compind.2019.103182).
- [40] A. Vaccaro and P. Magnone, "Influence of power cycling test methodology on the applicability of the linear damage accumulation rule for the lifetime estimation in power devices," *IEEE Trans. Power Electron.*, vol. 38, no. 5, pp. 6545–6554, May 2023, doi: [10.1109/TPEL.2023.3242314](https://doi.org/10.1109/TPEL.2023.3242314).
- [41] ECPE, "Qualification of power modules for use in power electronics converter units in motor vehicles," ECPE guideline AQG 324, Rev. 03.1/2021, 2019.
- [42] A. Vaccaro and P. Magnone, "Analysis of thermal cycling effects in power devices under non-constant cumulative stress," in *Proc. IEEE Appl. Power Electron. Conf. Expo.*, 2022, pp. 330–335, doi: [10.1109/APEC43599.2022.9773598](https://doi.org/10.1109/APEC43599.2022.9773598).
- [43] G. Zeng, F. Wenisch-Kober, and J. Lutz, "Study on power cycling test with different control strategies," *Microelectronics Rel.*, vol. 88, pp. 756–761, 2018, doi: [10.1016/j.microrel.2018.07.088](https://doi.org/10.1016/j.microrel.2018.07.088).
- [44] U. Scheuermann and S. Schuler, "Power cycling results for different control strategies," *Microelectronics Rel.*, vol. 50, no. 9–11, pp. 1203–1209, 2010, doi: [10.1016/j.microrel.2010.07.135](https://doi.org/10.1016/j.microrel.2010.07.135).
- [45] Z. Sarkany, A. Vass-Varnai, and M. Rencz, "Comparison of different power cycling strategies for accelerated lifetime testing of power devices," in *Proc. IEEE 5th Electron. Syst.-Integration Technol. Conf.*, 2014, pp. 1–5, doi: [10.1109/ESTC.2014.6962833](https://doi.org/10.1109/ESTC.2014.6962833).
- [46] S. Schuler and U. Scheuermann, "Impact of test control strategy on power cycling lifetime," in *Proc. PCIM*, 2010, pp. 355–360.
- [47] P. Murugan and S. Durairaj, "Regularization and optimization strategies in deep convolutional neural network," 2017, *arXiv:1712.04711*.



**ALESSANDRO VACCARO** (Member, IEEE) was born in Cariati, Italy, in 1996. He received the B.S. (*cum laude*) and M.S. (*cum laude*) degrees in electronics engineering from the University of Calabria, Rende, Italy, in 2018 and 2020, respectively. He is currently working toward the Ph.D. degree in mechatronics with the University of Padova, Vicenza, Italy. His research interests include oriented to the experimental analysis and reliability of semiconductor power devices.



**DAVIDE BIADENE** (Graduate Student Member, IEEE) received the M.S. degree in electronic engineering and the Ph.D. degree in information engineering from the University of Padova, Padova, Italy, in 2014 and 2017, respectively. He is currently a Research Fellow with the Department of Management and Engineering, University of Padova, Vicenza, Italy. From 2017 to 2021, he was with Infineon Technologies Italia, employed as an R&D test Engineer in the automotive business line team. His research interests include dc-dc converters for renewables and energy storage devices, and artificial intelligence techniques applied to control and reliability of power electronics converters.



**PAOLO MAGNONE** (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from the University of Calabria, Rende, Italy, in 2003 and 2005, respectively, and the Ph.D. degree in electronic engineering from the University of Reggio Calabria, Reggio Calabria, Italy, in 2009.

In the period 2006–2008, he joined for one year the Interuniversity MicroElectronics Center (IMEC), Leuven, Belgium, within the "Advanced PROcess Technologies for Horizontal Integration" project (Marie Curie Actions), where he worked on parameters extraction and matching analysis of FinFET devices. From 2009 to 2010, he was a Postdoctoral Researcher with the University of Calabria. From 2010 to 2014, he was with the Advanced Research Center on Electronic Systems for Information and Communication Technologies "E. De Castro" (ARCES), University of Bologna, Bologna, Italy. In 2014, he was appointed Associate Professor of electronics with the University of Padova, Padua, Italy. His current research interests include the electrical characterization and modeling of semiconductor devices and circuits for power applications.

Dr. Magnone is the Editor of IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS.