Letters

Artificial Intelligence-Aided Thermal Model Considering Cross-Coupling Effects

Yi Zhang, *Student Member, IEEE*, Zhongxu Wang, *Student Member, IEEE*, Huai Wang, *Senior Member, IEEE*, and Frede Blaabjerg, *Fellow, IEEE*

***Abstract*—This letter proposes an artificial intelligence-aided thermal model for power electronic devices/systems considering thermal cross-coupling effects. Since multiple heat sources can be applied simultaneously in the thermal system, the proposed method is able to characterize model parameters more conveniently compared to existing methods where only single heat source is allowed at a time. By employing simultaneous cooling curves, linear-to-logarithmic data re-sampling, and differentiated power losses, the proposed artificial neural network-based thermal model can be trained with better data richness and diversity while using fewer measurements. Finally, experimental verifications are conducted to validate the model capabilities.**

***Index Terms*—Artificial intelligence, power electronic devices and systems, thermal cross-coupling effects, thermal modeling.**

I. Introduction

THERMAL modeling is essential to reliability analysis and thermal management of power electronics [1]. With an increasing request for higher power density and lower parasitics in power electronic systems, a more compact power device packaging with multichips is a trend. One of the challenges that come with it is the thermal cross-coupling (TCC) effects among different semiconductor chips. Junction temperature of a chip is not only affected by its own power losses but also that of neighboring devices. Conventional one-dimensional (1-D) RC lumped models (e.g., Cauer and Foster models [2]) neglect the TCC effects, and might result in misleading results for multichip modules. Finite-element method can provide detailed thermal results, but it is restrained to a limited time span. Beyond multichip components, a power electronic system consisting of multiple components also has the TCC effects [3], [4].

Thermal impedance matrix is a method to model the TCC effects in thermal systems. However, the dimension of the thermal matrix increases significantly with the number of heat sources (e.g., the number of chips). For instance, a typical 1.7-kV power module (e.g., Infineon FF1000R17IE4 with 24 chips) requires a 24×24 thermal matrix, which imposes heavy computation burden for long-term thermal stress estimation. Moreover, the parameter characterization of the thermal matrix is also time-consuming. The thermal matrix in [3] with 24 heat sources requires 24 FEM simulations or experiments, and curve fittings to obtain the parameters. To simplify the calculation, only thermal resistances are considered in [3], without the thermal capacitances which represent thermal dynamics.

Artificial neural networks (ANNs) are used to predict lifetime of a photovoltaic system in [5]. The established work has validated the potential of artificial intelligence-aided methods supporting the analysis and design of power electronic systems. However, the application of these methods to establish and characterize a thermal model is not discussed. Thermal model is a time-series problem essentially. Many artificial intelligence methods have been applied to time-series applications, such as grid phasor detection [6] and health-state estimation [7]. One of the assumptions of these applications is that the training data and the prediction have an identical sampling frequency. However, a thermal model is necessary to be adaptive to different sampling frequencies of thermal estimation (e.g., different mission-profile resolutions). Moreover, artificial intelligence-aided methods usually need a large number of training data to improve the model accuracy. Thermal model characterization, however, tends to reduce the measurement time and the number of tests. This conflict imposes a challenge in how to obtain training data with rich diversity while using as few measurements as possible.

This letter proposes a thermal model with the aid of an ANN algorithm. The novel aspects are three folds: 1) the proposed simultaneous cooling curve simplifies the thermal characterization with less measurements, 2) the utilization of time interval makes the trained model to be adaptive to different sampling frequencies, and 3) the logarithmic transformation and the differential power losses improve the data diversity.

II. Conventional Thermal Model With TCC Effects

In a multichip power module, any device dissipating power leads to a temperature rise not only for itself but also for neighboring devices. The conventional method uses the following

Manuscript received January 18, 2020; revised February 19, 2020; accepted March 4, 2020. Date of publication March 11, 2020; date of current version June 23, 2020. This work was supported by the Innovation Fund Denmark through the Advanced Power Electronic Technology and Tools (APETT) under Project 877302. *(Corresponding author: Zhongxu Wang.)*

The authors are with the Department of Energy Technology, Aalborg University, 9220 Aalborg, Denmark (e-mail: yiz@et.aau.dk; zhongxu2020@gmail.com; hwa@et.aau.dk; fbl@et.aau.dk).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPEL.2020.2980240

0885-8993 © 2020 EU

thermal matrix to consider the TCC effect:

$$\begin{bmatrix} \Delta T_1 \\ \Delta T_2 \\ \vdots \\ \Delta T_n \end{bmatrix} = \begin{bmatrix} Z_{\text{th}1,1} & Z_{\text{th}1,2} & \cdots & Z_{\text{th}1,n} \\ Z_{\text{th}2,1} & Z_{\text{th}2,2} & \cdots & Z_{\text{th}2,n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{\text{th}n,1} & Z_{\text{th}n,2} & \cdots & Z_{\text{th}n,n} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} \tag{1}$$

where ΔT_i and P_i are the temperature rise of a predefined point (e.g., junction temperature of a chip) and power losses of the ith device, respectively, and $Z_{\text{th}i,j}$ is the thermal impedance.

The characterization of the thermal impedance typically relies on cooling curves with a single power loss input at a time. For instance, when a step power loss is applied to the first device, i.e., $[P_1 \ P_2 \ \cdots \ P_n]' = [P_{\text{loss}} \ 0 \ \cdots \ 0]'$, the thermal impedance in the first column of (1) is obtained as

$$Z_{\text{th}i,1}(t) = \frac{\Delta T_i(t)}{P_{\text{loss}}}. \tag{2}$$

Applying the same procedure to other devices, the remaining thermal impedance can be derived. Afterwards, each thermal impedance is curve-fitted by the following expression as

$$Z_{\text{th}i,j}(t) = \sum_v R_{\text{th}i,j\_v} \left(1 - e^{-t/\tau_{\text{th}i,j\_v}}\right) \tag{3}$$

where $R_{\text{th}i,j\_v}$ and $\tau_{\text{th}i,j\_v}$ are the thermal resistance and time constant of the vth-order Foster network, respectively. The number of v is dependent on the accuracy of the fitted curve to $Z_{\text{th}}(t)$. Third- or fourth-order network is typically employed.

Briefly, the characterization of a thermal matrix requires two steps: 1) applying separated power loss steps to each chip to obtain the thermal responses, and 2) curve-fitting for parameter extraction. Two major challenges are thus involved in this process. First, the testing time and the number of measurements increase proportionally with the number of heat sources. For instance, a power module with 24 chips needs 24 separated experiments. Second, the number of R_{th}–τ_{th} parameters increases with the square of the number of devices. If a fourth-order network is employed (i.e., eight parameters for each thermal impedance), the 24 $\times$ 24 thermal matrix has 4608 parameters, leading to considerably heavy characterization process and challenge in long-term thermal analysis.

III. Development of ANN-Based Thermal Model

By applying separated power loss steps, the conventional thermal matrix model in essence simplifies a multi-input–multi-output (MIMO) system into several single-input–multi-output (SIMO) systems. With this regard, multiple tests have to be conducted. Unlike the thermal matrix, artificial intelligence-aided methods are able to deal with MIMO systems directly. Thus, this section proposes an ANN-based thermal model considering TCC effects with less characterization efforts.

A. ANN Principle

Various types of ANNs have been summarized in [8]. Particular network selection depends mostly on the relationship between inputs and outputs. According to [9], the temperature estimation of a first-order network is expressed as

$$\begin{aligned} \Delta T(t_n) = {} & \Delta T(t_{n-1})\, e^{-\Delta t/\tau_{\text{th}}} \\ & + P(t_n)\, R_{\text{th}} \left(1 - e^{-\Delta t/\tau_{\text{th}}}\right) \end{aligned} \tag{4}$$

where the present temperature $\Delta T(t_n)$ is determined by the previous temperature $\Delta T(t_{n-1})$ and the present dissipation $P(t_n)$. Thus, a dynamic ANN method for time-series problem is selected, which is

$$\begin{aligned} y(k) = f[& y(k-1), y(k-2), \ldots, y(k-n_y), \\ & x(k), x(k-1), \ldots, x(k-n_x)] \end{aligned} \tag{5}$$

where $y(k)$ is the predicted value, $y(k-1)$ to $y(k-n_y)$ are the previous values of the output, $x(k)$ to $x(k-n_x)$ are the input data, and $f[\cdot]$ represents the network of the ANN.

In many applications of the ANN algorithm of (5), data sampling frequencies used in the training and prediction are assumed to be identical [6], [7]. However, a thermal model with a fixed sampling frequency could limit its scope for different mission profiles. To take into account the impact of different sampling frequencies, the proposed method not only uses the power loss information as the input but also the time interval Δt as inspired by (4). As shown in Fig. 1, the proposed ANN-based thermal model is composed of an input layer, one or more hidden layers, and an output layer. The inputs of each layer are multiplied by weights w and added with the bias b. The results are processed through an activation function f, which is usually a sigmoid function.

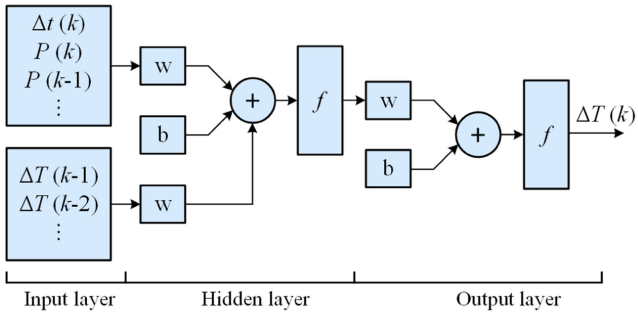

Fig. 1. Structure of the ANN to convert power losses into temperature variations (w: weights, b: bias, and f: activation function).

B. ANN-Based Thermal Model Trained by Simultaneous Cooling Curves

As shown in Fig. 2, a common half-bridge module with two insulated-gate bipolar transistor (IGBT) chips and two diode chips is chosen as a case study. Their thermal responses are measured by thermo-optical fibers [10].

Unlike the conventional thermal matrix with separated power loss steps [2], in the proposed ANN-based thermal model, the four chips in the power module are stressed by an ac current, i.e., $i_{\text{ac}}(t) = 5 + 35\sin(2\pi \times 100t)$ A, as shown in Fig. 3. The power losses are injected into four power devices simultaneously. The obtained results are thus defined as simultaneous cooling curves. The frequency of 100 Hz is selected to ensure

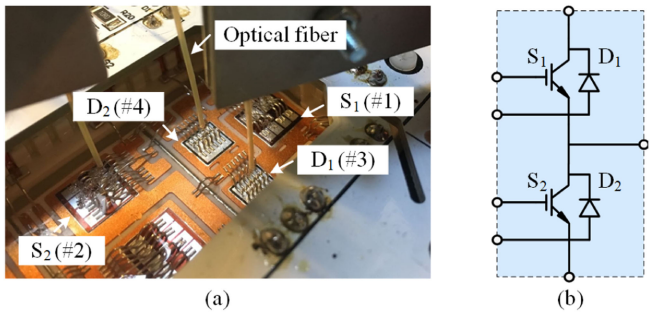

Fig. 2. Power module used in the case study. (a) Temperature measurement via four thermo-optical fibers. (b) Circuit diagram.

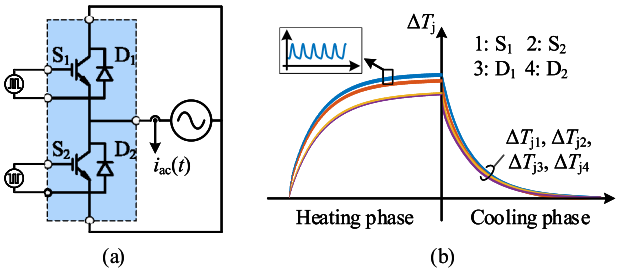

Fig. 3. The proposed simultaneous cooling method. (a) The circuit and method to apply the power losses to the four power devices simultaneously [$i_{ac}(t) = 5 + 35\sin(2\pi \times 100t)$ A in this case]. (b) Diagram of temperature measuring process.

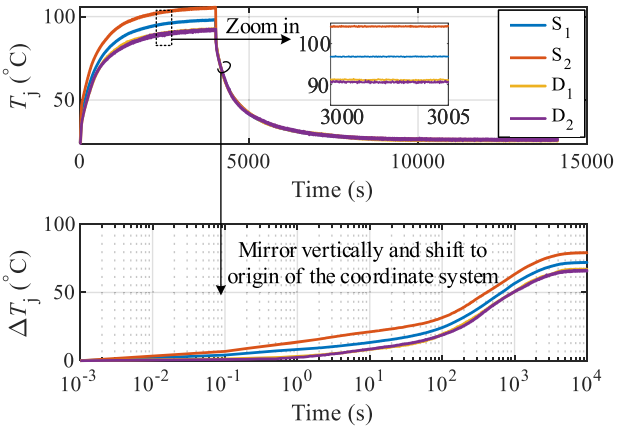

Fig. 4. The measured temperature responses based on the proposed simultaneous cooling method.

a relatively small steady-state temperature variation. The measured simultaneous cooling curves are shown in Fig. 4. The four power devices are heated up simultaneously with smooth curves (see the zoom-in area of Fig. 4). Moreover, the additional 5-A dc bias current is utilized to differentiate the power losses among different devices, producing more diversified training data for the ANN.

Afterwards, instead of being curve-fitted to characterize thermal matrix, the measured temperature responses are used to train the proposed ANN-based thermal model directly. Notably, the performance of the ANN method is largely dependent on the

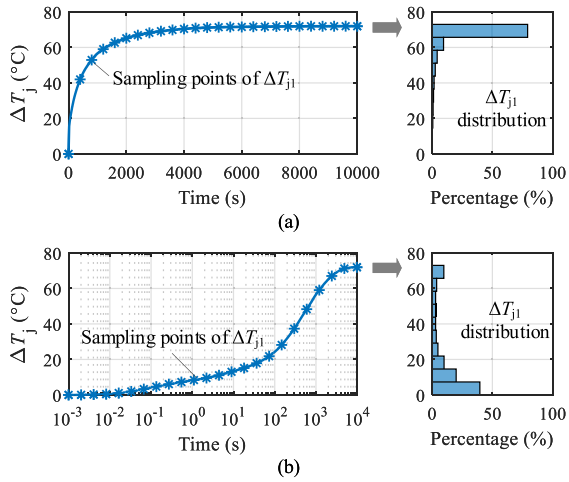

Fig. 5. Comparison of two sampling methods. (a) Temperature distribution in the original linear measurement. (d) Distribution after logarithmic transformation.

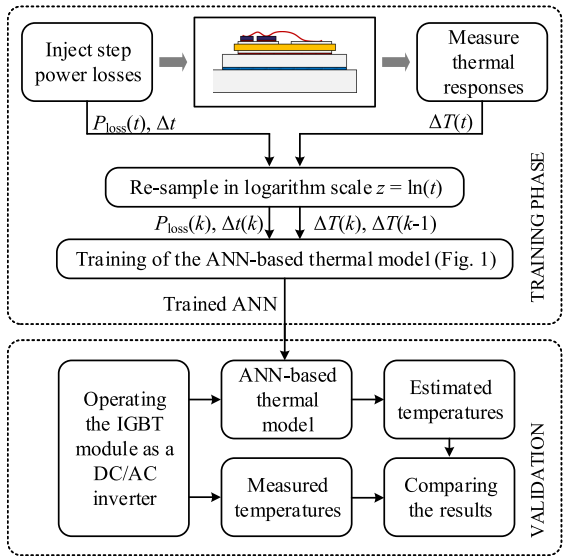

Fig. 6. Flowchart of the proposed ANN-based thermal modeling (training phase) and its validation process (validation).

diversity and richness of the training data. As shown in Fig. 5(a), although the sampling frequency is 1 kHz (i.e., 10^7 sampled data in total), approximately 90% measured ΔT_{j1} information is from the steady state (around 60–80 °C). The sampled data from the fast dynamic range (i.e., 0–1 s) accounts for less than 0.1% of the whole data set. The over-concentrated training data might be challenging for ANN algorithm to establish an effective thermal model. To tackle this problem, a logarithmic variable is introduced, inspired by the exponential thermal transient expression in (4), that is, $z = \ln(t)$. The measured temperatures are resampled by a piece-wise linear interpolation in the logarithmic scale, i.e., $\Delta T_j(z) = \Delta T_j(t = \exp(z))$. The resampling period is selected as $\Delta z = 0.01$. The corresponding temperature distribution is more even now, as shown in Fig. 5(b).

The flowchart of the proposed method is summarized in Fig. 6. The power loss information, time intervals, and corresponding

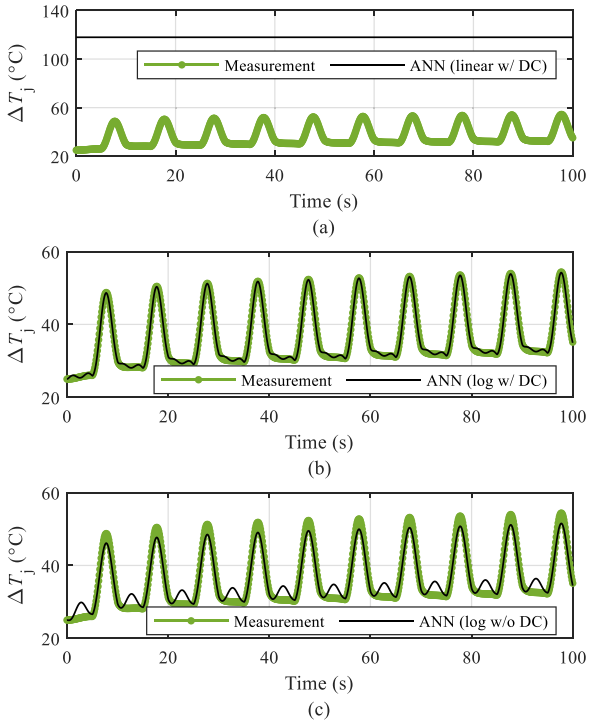

Fig. 7. Comparison between the measured junction temperature of S_1 and the ANN-based thermal model with different training data. (a) Trained by the originally linear-sampled data with the dc-bias current. (b) Trained by logarithmic transformation and the dc-bias current. (c) Logarithmic transformation but without the dc-bias current.

temperature responses are obtained by the proposed simultaneous cooling curve. The originally measured data is converted by the logarithmic transformation and then used to train the ANN model, as shown in Fig. 1. In this case, the hidden layer is a single layer with ten neurons. Since the thermal estimation does not require the previous information of the power loss or the time interval, $\Delta t(k)$ and $P(k)$ are selected as inputs, while $\Delta T(k-1)$ and $\Delta T(k-2)$ are also used. The higher order of the previous temperature than (4) could benefit the accuracy of the ANN-based thermal model.

IV. Experimental Verifications

As illustrated in Fig. 6, after training the ANN-based thermal model, its validation process is conducted based on the same IGBT module but operating as a dc/ac inverter. The operational conditions are as follows: $i_{ac}(t) = 20\sin(2\pi \times 0.1t)$ A, 1-kHz switching frequency, and 100-V dc voltage. The junction temperature of the four devices is measured and compared with the results from the proposed method. Fig. 7 presents the measured temperatures and the estimated results from ANN-based thermal model with different training methods. In Fig. 7(a), the ANN-based thermal model is trained by the originally linear-sampled data. The dc-bias current is applied to differentiate the power losses of devices in the IGBT module. It can be seen that the estimated temperature is almost a constant without fluctuations. The estimation error is as large as 80 °C compared with the

TABLE I
Comparison of the Numerical Results of Fig. 7 (90–100 s)

Unit: (°C)	T_{jmax}	Error	T_{jmin}	Error
Measurement	54.4		32.2	
ANN linear w/ DC	117.9	-63.5	117.9	-85.7
ANN log w/ DC	54.1	0.3	32.1	0.1
ANN log w/o DC	51.5	2.9	32.5	-0.3

*Note: Error = Measurement − Estimation.

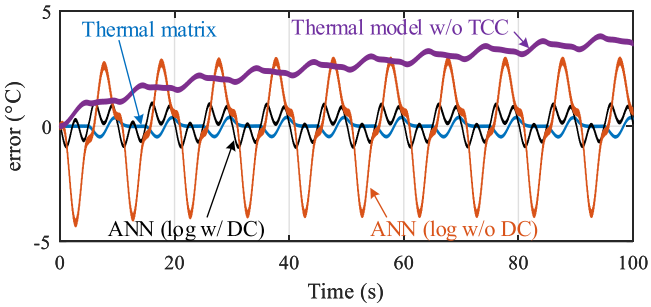

Fig. 8. Estimated errors of different methods [error = measurement − estimation, ANN (log with dc): ANN trained by the logarithmic data and the dc-bias current, ANN (linear without dc): trained by the logarithmic data but without the dc-bias current, thermal model without TCC: conventional thermal model without considering TCC effects, and thermal matrix: refers to (1)].

measurement. By contrast, when the logarithmic transformation is applied, as shown in Fig. 7(b), the junction temperature can be more accurately estimated by the proposed method and it is in a good agreement with the measured result. It verifies that the proposed ANN-based method is able to characterize the thermal model of an IGBT module by a single measurement, and the logarithmic transformation is necessary to avoid over-concentrated data with better data richness and diversity. In addition, various power losses for different devices (i.e., achieved by a dc-bias current in this case study) can contribute to a better model training performance compared with Fig. 7(c) where the dc-bias current is not employed. Furthermore, the numerical results during the period of 90–100 s in Fig. 7 are summarized, as listed in Table I. From the maximum and minimum temperatures (i.e., T_{jmax} and T_{jmin}), the differences between the proposed method and the measurement are no more than 1 °C. The above results verify the effectiveness of the proposed method again.

Temperature estimation errors of different methods are compared in Fig. 8. Both the proposed ANN-based thermal model with dc bias current and the thermal matrix method can achieve a good thermal estimation compared to the measurement with the error being less than 1 °C. Although the error of the proposed ANN-based method is slightly larger, the minor temperature fluctuations have a negligible impact on the reliability evaluation according to [11]. For the ANN model trained without a dc-bias current, a larger but stable estimation error (i.e., around 4 °C) can be observed. Furthermore, when it comes to the conventional thermal model without considering TCC effects, the estimation error increases over time (not stable) and becomes the largest among the four methods at $t = 100$ s. The conclusion can

TABLE II
COMPARISON OF THE THERMAL SYSTEM CHARACTERIZATION TIME AND THE COMPUTATIONAL EFFICIENCIES

	ANN method	Thermal matrix (4 chips)	Thermal matrix (n chips)
Thermal measurement time (hours)	4×1	4×4	$4 \times n$
Computational time (s)	109	1,005	–

*Conditions: Intel Xeon CPU@2.40 GHz, 128-GB RAM, 64-bit system.

be reached that the proposed ANN-based thermal model can substitute the thermal matrix to consider the TCC effect with less testing efforts.

Furthermore, the thermal system characterization time and the computational efficiency are compared as listed in Table II. The characterization of the thermal system (i.e., the IGBT module plus its heat sink) requires around 4 h to complete the heating and cooling processes. Due to the simultaneous cooling method, the proposed ANN-based thermal model needs one test. By contrast, the conventional thermal matrix has to inject power losses in each semiconductor chip separately. In this case study with four chips, 4×4 h are needed. In a power module with n chips, much more characterization time (e.g., $4 \times n$ h) is required. Moreover, computational efficiency comparison is conducted by converting the power loss profiles (data size: $800\,019 \times 4$, double type) into thermal profiles. The proposed method costs around 109 s while the CPU time of the conventional thermal matrix is 1005 s in total. Both of these comparisons reveal the potentiality of the proposed method to do efficient thermal characterizations and estimations.

V. DISCUSSION

The core idea of the proposed method is to establish a thermal model consuming less time and efforts, while considering the TCC effects. However, several limitations still need further studies. First, the discussion of this letter is based on the power module that each switch is composed of a single chip. The feasibility for a power module in which each switch consists of multiple chips in parallel has not been discussed. Moreover, the thermo-optical fibers are used in this letter with a sampling frequency of 1 kHz. The feasibility of the different thermal measuring methods or sampling frequencies could be studied in the future. Finally, it should be pointed out that the discussion in this letter does not take into account the impact of temperature-dependent material property on the thermal model parameters.

VI. CONCLUSION

This letter proposes an artificial intelligence-aided thermal model considering TCC effects. The target is to simplify the characterization process of thermal modeling compared with the conventional thermal matrix method. The ANN-based thermal model can be established with the aid of the proposed simultaneous cooling curves. The testing time of a multichip power module can be reduced to only once. Moreover, by using the sampling time interval Δt as one of the inputs, the proposed method is applicable to different data sampling frequencies. The logarithmic transformation of the originally linear-sampled data and the additional dc-bias current further improve the training performance. The experimental measurements show that the proposed method can provide thermal estimation with an error less than 1 °C in the case study.

REFERENCES

[1] M. Andresen, K. Ma, G. Buticchi, J. Falck, F. Blaabjerg, and M. Liserre, "Junction temperature control for more reliable power electronics," *IEEE Trans. Power Electron.*, vol. 33, no. 1, pp. 765–776, Jan. 2018.
[2] Infineon Technologies, "Transient thermal measurements and thermal equivalent circuit models," *Application Note AN 2015-10*, 2018, pp. 1–13.
[3] Y. Shen, A. Chub, H. Wang, D. Vinnikov, E. Liivik, and F. Blaabjerg, "Wear-out failure analysis of an impedance-source PV microinverter based on system-level electrothermal modeling," *IEEE Trans. Ind. Electron.*, vol. 66, no. 5, pp. 3914–3927, May 2019.
[4] Y. Zhang, H. Wang, Z. Wang, F. Blaabjerg, and M. Saeedifard, "Mission profile-based system-level reliability prediction method for modular multilevel converters," *IEEE Trans. Power Electron.*, vol. 35, no. 7, pp. 6916–6930, Jul. 2020.
[5] T. Dragicevic, P. Wheeler, and F. Blaabjerg, "Artificial intelligence aided automated design for reliability of power electronic systems," *IEEE Trans. Power Electron.*, vol. 34, no. 8, pp. 7161–7171, Aug. 2019.
[6] I. Kamwa, R. Grondin, V. K. Sood, C. Gagnon, V. T. Nguyen, and J. Mereb, "Recurrent neural networks for phasor detection and adaptive identification in power system control and protection," *IEEE Trans. Instrum. Meas.*, vol. 45, no. 2, pp. 657–664, Apr. 1996.
[7] M. Landi and G. Gross, "Measurement techniques for online battery state of health estimation in vehicle-to-grid applications," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 5, pp. 1224–1234, May 2014.
[8] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.
[9] Y. Zhang, H. Wang, Z. Wang, Y. Yang, and F. Blaabjerg, "Simplified thermal modeling for IGBT modules with periodic power loss profiles in modular multilevel converters," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2323–2332, Apr. 2018.
[10] U. M. Choi, F. Blaabjerg, and S. Jørgensen, "Power cycling test methods for reliability assessment of power device modules in respect to temperature stress," *IEEE Trans. Power Electron.*, vol. 33, no. 3, pp. 2531–2551, May 2018.
[11] H. Wang, K. Ma, and F. Blaabjerg, "Design for reliability of power electronic systems," in *Proc. 38th Annu. Conf. IEEE Ind. Electron. Soc.*, 2012, pp. 33–44.