APPLICATION-AWARE ESTIMATION OF THE JUNCTION TEMPERATURE SWING UNDER ACTIVE CYCLING

Ciprian V. Pop[1], Andi Buzo[2], Cristian Diaconu[3], Georg Pelz[4], Horia Cucu[5], Corneliu Burileanu[6]

This paper presents a methodology for application-aware estimation of the junction temperature swing of smart power devices under active cycling, based on electro-thermal simulation. According to the most known lifetime model, i.e. the Coffin-Manson law, the maximum junction temperature swing within a fast thermal cycle (ΔT) represents the main stress-factor of the lifetime. Consequently, for application-aware lifetime estimation of power devices an important step is to estimate ΔT in the space of application operating conditions. The proposed application-aware ΔT model is fitted from data on a grid of electro-thermal simulations based on estimated power profiles over a wide space of operating conditions. The model is validated on electro-thermal simulation data based on the power profiles corresponding to the active cycling experimental scenarios, showing a maximum unsigned relative prediction error of 6.56%. With the proposed model, the dependencies of ΔT with the operating parameters can be observed and predictions on different application operating conditions can be done.

Keywords: active cycling, junction temperature swing, lifetime estimation, reliability, smart power devices.

1. Introduction

The reliability assessment of power devices is an essential concern. In automotive, for instance, very strong safety regulations are imposed and "zero defects" on the entire vehicle's life are required. On the other hand, as the

[1] Ph.D. candidate, Doctoral School of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: cprv.pop@gmail.com

[2] Ph.D., Infineon Technologies AG, Neubiberg, Germany, e-mail: andi.buzo@infineon.com

[3] Ph.D., Infineon Technologies Romania & CO SCS, Bucharest, Romania, e-mail: cristian-vasile.diaconu@infineon.com

[4] Professor, Infineon Technologies AG, Neubiberg, Germany, e-mail: georg.pelz@infineon.com

[5] Lecturer, Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: horia.cucu@upb.ro

[6] Professor, Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: corneliu.burileanu@upb.ro

power products are used in many different applications, the lifetime estimation of the power devices is required in the space of application operating conditions.

According to the classical Coffin-Manson lifetime model [1], presented in (1), the maximum junction temperature swing within a fast thermal cycle (ΔT) is the main stress-factor of the lifetime. So, for estimation of the lifetime at different operating conditions, an important step is to analyse and model the ΔT in the space of application operating conditions (application-aware).

$$N_f \sim [\Delta T\,]^{-q} \tag{1}$$

where: N_f is the lifetime (expressed in Cycles-To-Failure), ΔT represents the maximum junction temperature swing and q is a constant. In logarithm, this relation shows a linear dependence of the lifetime with ΔT .

In general, a limited amount of reliability data is available. Consequently, it is not possible to develop a complex and accurate ΔT model directly from the available active cycling scenarios.

The goal of this research is to develop a methodology for application-aware estimation of ΔT (estimation that depends on the application operating conditions), based on data from a grid of electro-thermal simulations performed in the space of the operating conditions. For that, the power profiles parameters required for the electro-thermal simulations are modeled in the space of operating conditions, based on the power pulses parameters corresponding to the active cycling scenarios. The electro-thermal simulations corresponding to different power profiles are performed based on the methodology described in [2]. The presented methodology is applied on smart power devices comprising DMOS (double-diffused MOS) structures.

The paper is organized as follows: Section 2 presents an overview of the related work, the description of the proposed methodology is provided in Section 3, then Section 4 describes how the methodology is applied and provides the results. In Section 5 the conclusions are drawn.

2. **Related Work**

In the literature, there are several methods for measurement or estimation of ΔT. We present the features and limitations of the most widespread and practical solutions. An *analytical approach* is proposed in [3], where ΔT in DMOS corresponding to a triangular-shape power pulse is estimated with 2.

$$\Delta T_{peak} = \frac{P_{peak} \cdot k_{therm}}{A_{DMOS}} \cdot \frac{\sqrt{2}}{3} \cdot \sqrt{t_{pulse}} \tag{2}$$

where: P_{peak} is the power pulse peak amplitude, k_{therm} represents a material coefficient, A_{DMOS} is DMOS active area and t_{pulse} is the power pulse duration. The model was used for power switches in mixed bipolar-CMOS-DMOS (BCD) technology, with a prediction accuracy of 10%. The main limitation of this method is that it is valid only for short pulses and moderate temperature rise.

Other methods are based on *thermal imaging*. The works [4, 5] present the experimental results of dynamic thermal mapping based on a radiometric 2D measurement system. The measurement setup consists of a radiometric microscope with a cooled InSb sensor, working in the first Infrared window. A x-y step motor controlled by a PC is used for scanning of the mapped surface.

The main shortcoming of the thermal mapping system based on radiometric sensor is that the fast thermal transients cannot be acquired. Also, if a high spatial resolution is required, the measurement time could be very long, because of the point by point acquisition of the targeted grid area.

Another thermal mapping method, presented in [6, 7], is based on thermoreflectance measurement system, which measures the reflection coefficient variation due to a temperature change in metals and semiconductors. This method permits the analysis of fast thermal transients. However, the difficulties in preparation of the samples and the complexity of the experimental setup do not allow this thermal mapping system to be an easy-to-use solution.

The work [8] introduces a novel temperature mapping system using an Infrared thermo-camera, capable to map ultrafast temperature transients distribution on power devices. In [9] the temperature measurements based on the same system are compared with the results of Finite Elements Method (FEM) thermal simulations, resulting less than 10% inaccuracies.

As in the case of previously presented systems based on Infrared sensors, this approach has the problem of the emissivity contrast, due to the presence of different metal and passivation layers. So, an initial stage for preparation of the samples is required to overcome this problem. Moreover, when DMOS structures consist of thick power metal on top the junction temperature is obscured so, the Infrared measurements can not be used.

Other methods are based on *temperature sensors*. The approach presented in [10] estimates ΔT by means of the intrinsic body diode of MOSFETs. For that, a specific test structure is required, in order to forward bias the body diode, after applying the ramp power pulse. For estimation of ΔT of smart power devices under repetitive clamping this approach is not proper because the back body diode method does not actually provide the maximum ΔT, but a weighted average temperature of the MOS. Moreover, the temperature measurement within the energy pulse is not possible as the required forward current for the back body diode cannot be provided within the clamping period.

A similar approach is presented in [11], where the parasitic n-p-n transistor of DMOS power devices fabricated in BCD technology is used as temperature sensor. Access to the sensor is gained by cutting the connection to the n^+ source region of one DMOS cell and connecting a dedicated terminal to the emitter. The temperature is measured based on the voltage of the base-emitter junction, while it is forward biased with a constant current by means of a dedicated setup. Also, dedicated test structures were built and measured on wafer. So, using this concept on a real assembled power device is not appropriate.

3. The Proposed Approach

For estimation of the lifetime main stress-factor, i.e. the junction temperature swing, in the space of application operating conditions (load current - I_L, ambient temperature - T_{amb}, repetitive energy - E_R), the following ΔT modeling methodology is proposed. Because the available data corresponding to the active cycling scenarios are not enough for fitting an accurate model, the idea is to fit the ΔT model from data on a grid of electro-thermal simulations, performed in the space of the operating conditions.

First, the electro-thermal simulations corresponding to active cycling scenarios are performed, based on the methodology described in [2]. As inputs, the electro-thermal simulator requires the power pulse corresponding to the experimental power profile and the ambient temperature.

The parameters that characterize the power pulse within a fast thermal cycle are extracted from the measurements corresponding to the active cycling experiment. An example of the power pulse shape is presented in Fig. 1. The power pulse is provided to the electro-thermal simulator in Piece-Wise Linear (PWL) format. It consists of consecutive (time, power) pairs. For an optimal compromise between the simulation time and the results accuracy, six (time, power) pairs are considered in this case. They are the red points of the power pulse, marked by letters (A - F) in Fig. 1. The points A and C correspond to the moments when the power switch is turned on and off, respectively. The point D corresponds to the peak power (P_{peak}). E is the point when I_L reaches the zero value and the point F corresponds to the end of the cycle. The point B is added only for improvement of the simulation accuracy.

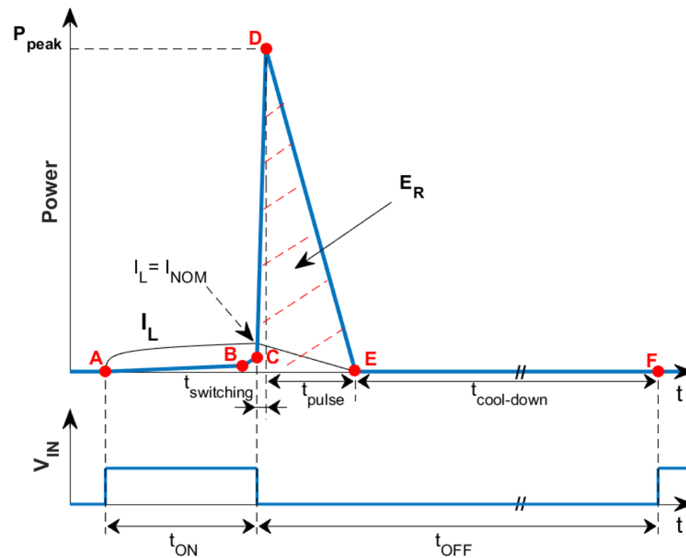

FIGURE 1. The power pulse within a fast thermal cycle.

From the electro-thermal simulations the maximum junction temperature swings, defined as in (3), are determined.

$$\Delta T = T_{peak} - T_{min} \tag{3}$$

where: T_{peak} is the maximum junction temperature within one thermal cycle; T_{min} is the steady-state case temperature, from the end of the thermal cycle.

3.1. ΔT Modeling Methodology Flow

The methodology flow for ΔT modeling is presented in Fig. 2.

The first step consists in the extraction of the power profiles for all scenarios of active cycling experiments and of the corresponding PWL parameters, required for the electro-thermal simulations.

In the second step the electro-thermal simulations for the experimental scenarios are performed as presented in the introductory part of this section, resulting the corresponding ΔT values from DMOS.

Then, in order to run electro-thermal simulations in points of operating conditions where no power pulse parameters are available from experimental measurements, the PWL parameters required for the electro-thermal simulations need to be modeled in the space of operating conditions. For that, based on the PWL parameters of the experimental operating conditions scenarios, data-driven models or analytical approximations for PWL parameters in the space of operating conditions are built, in the third step.

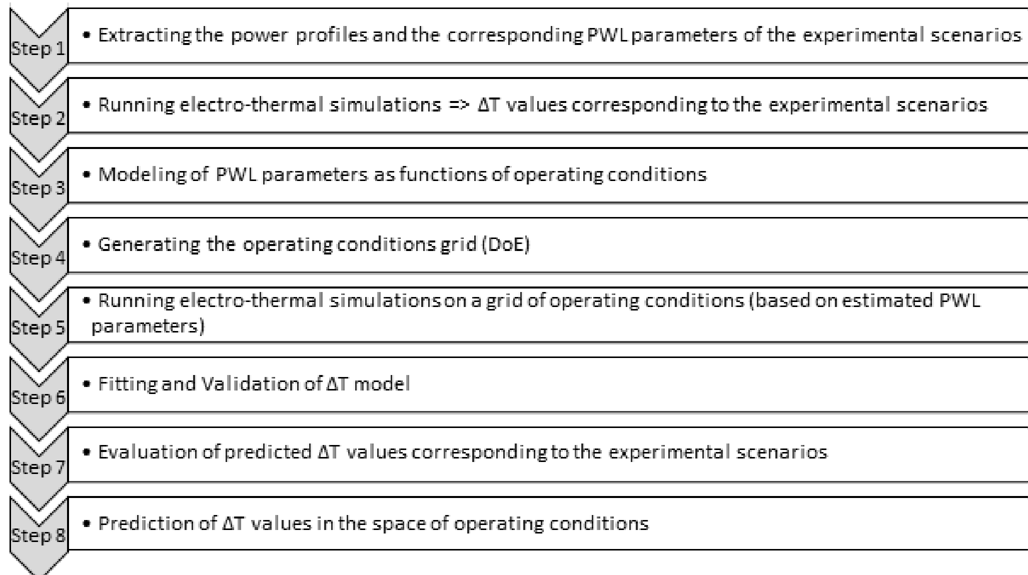

FIGURE 2. The methodology flow for ΔT modeling.

The forth step consists of generating the operating conditions grid (Design of Experiments - DoE). On the resulted grid the PWL parameters (required for the electro-thermal simulations) are predicted with the models or analytical approximations built in the previous step.

In step five, electro-thermal simulations on the grid in the space of operating conditions are performed based on the estimated PWL parameters, resulting the corresponding ΔT values.

The sixth step consists of fitting the ΔT model as a function of the operating conditions, based on the grid data. Also, the model validation on the training data set is done by means of leave-k-out and bootstrapping methods.

In step seven the ΔT model is evaluated on the experimental data. Also, the maximum relative error is determined.

In the last step, the influences of the operating conditions on ΔT can be observed and predictions of the junction temperature swing in the space of application operating conditions can be made, based on the developed model.

The most important methodology stages will be described in details, in the following subsections.

3.2. Estimation of PWL Parameters for Grid Data

For running electro-thermal simulations on a grid of operating conditions, including in points where no power pulse parameters are available from experimental measurements, the PWL parameters are estimated in the space of operating conditions by means of direct data-driven models or analytically.

Switching-ON Time (t_{ON}) and Cool-Down Time ($t_{cool-down}$). These two parameters depend on all operating conditions. Consequently, it has been chosen to model them with metamodels of type interaction, as presented in 4, having the operating conditions as factors. Apart from the main effects, this type of metamodel takes into account also the interactions of the factors.

$$t_x = c_0 + c_1 \cdot I_L + c_2 \cdot T_{amb} + c_3 \cdot E_R + c_{12} \cdot I_L \cdot T_{amb} + c_{13} \cdot I_L \cdot E_R + c_{23} \cdot T_{amb} \cdot E_R \quad (4)$$

The coefficients of the models are fitted from experimental data (the PWL parameters of the experimental scenarios), with the least squares method.

Pulse Time (t_{pulse}). The pulse time represents the clamping period. It can be estimated from the approximation formula of the energy discharged (E_R) when the power switch is turned off. As E_R is the integral of the triangle-shape power pulse in clamping (see Fig. 1), t_{pulse} is analytically approximated by 5:

$$t_{pulse} \approx \frac{2 \cdot E_R}{P_{peak}} \approx \frac{2 \cdot E_R}{I_L \cdot V_{clamp}} \quad (5)$$

Switching Time ($t_{switching}$). It is very small in comparison with the others time parameters. For this reason, it is considered constant and is computed as the average of the switching times from all active cycling experiments scenarios.

Intermediate Time ($t_{intermediate}$). The intermediate point B on the power pulse (see Fig. 1) is added only for improvement of the accuracy of the input power pulse and, consequently, of the electro-thermal simulation. The intermediate time is chosen at an arbitrary percentage from t_{ON} period.

Power Values. The power values in points A - F of the power pulse (see Fig. 1) are given by the following relations: $P_A = P_E = P_F = 0$; $P_D = P_{peak} = I_L \cdot V_{clamp}$; P_C and P_B are estimated based of the average percentage of the ON-period energy from the total energy (within the total cycle period).

3.3. Generating the Grid of Operating Conditions (DoE)

The DoE for the grid of operating conditions, where electro-thermal simulations will be performed, has to be done taking into account that E_R can not take any value at specified (I_L, T_{amb}) operating conditions. For (I_L, T_{amb}) factors, a p^2 full factorial design [12] can be used. For the 3rd factor (the repetitive energy), a fixed set of values at different (I_L, T_{amb}) pairs can not be considered, because E_R is limited by the value of the Single Pulse Energy (E_{AS}) corresponding to each (I_L, T_{amb}) scenario. The solution is to build the interaction model (6) of the maximum allowed E_R in active cycling experiments, as function of I_L and T_{amb}, in order to predict the maximum E_R value for any (I_L, T_{amb}) data-point. The rest E_R values for each (I_L, T_{amb}) data-point are chosen relative to the predicted maximum E_R.

$$E_{R-max} = c_0 + c_1 \cdot I_L + c_2 \cdot T_{amb} + c_{12} \cdot I_L \cdot T_{amb} \tag{6}$$

where: the model coefficients (c_0, c_1, c_2, c_{12}) are extracted from the experimental data with the least squares method.

3.4. ΔT Model Fitting based on Grid Data

Based on the metamodels and approximations described in the subsection 3.**2**, a grid of power pulses in the space of operating conditions is generated. Then, electro-thermal simulations are performed on this grid, resulting the corresponding ΔT values.

The model of ΔT in DMOS, having the operating conditions as factors, is build by fitting from grid data the metamodel 7 (3rd order polynomial with interactions).

$$\begin{aligned}\Delta T = c_0 + c_1 I_L + c_2 T_{amb} + c_3 E_R + c_4 I_L T_{amb} + c_5 I_L E_R + c_6 T_{amb} E_R + c_7 I_L^2 + \\ + c_8 T_{amb}^2 + c_9 E_R^2 + c_{10} I_L T_{amb} E_R + c_{11} I_L^2 T_{amb} + c_{12} I_L^2 E_R + c_{13} T_{amb}^2 I_L + \\ + c_{14} T_{amb}^2 E_R + c_{15} E_R^2 I_L + c_{16} E_R^2 T_{amb} + c_{17} I_L^3 + c_{18} T_{amb}^3 + c_{19} E_R^3\end{aligned} \tag{7}$$

where: the coefficients c_0 - c_{19} are extracted with the least squares method.

3.5. ΔT Model Validation

With Leave-K-Out Method. The leave-k-out method is used for validation of ΔT model with the training data set (grid data). It consists in getting-out a number (k) of data points from the available observations and fitting the model based on the rest data points. Then, the model is evaluated on the left-out data points and the corresponding prediction relative errors are computed. These steps are repeated for all combinations of k data points, out of all available ones. At the end, the maximum unsigned relative error is considered.

With Bootstrapping Method. The confidence interval of ΔT model prediction is estimated with the bootstrapping technique. Instead of the original set of observations, a bootstrapp sample is generated from the original data set by randomly picking observations, with replacement. The size of the bootstrapp sample is 1÷3 times the size of the original data set. Typically, a number of 10 000 iterations are performed. The number of left-out observations for each iteration is not fixed, but randomly chosen. The confidence limits are computed from the histogram of the prediction relative errors, corresponding to the desired confidence level (typically 95%).

On the Experimental Scenarios - based Electro-Thermal Simulations Data. The model is also evaluated on the operating conditions of the active cycling scenarios and the prediction relative errors of the corresponding ΔT values are computed. Eventually, the maximum unsigned relative error is considered.

4. Experimental Results

4.1. Electro-Thermal Simulations on Active Cycling Scenarios

The active cycling experiments and the corresponding electro-thermal simulations are performed as described in [2]. The active cycling experiments are made on 12 different operating conditions scenarios, consisting of 5 (I_L, T_{amb}) pairs, each on 2-3 values of E_R. The load current takes values from the device nominal current (I_{NOM}) to double of I_{NOM}. The ambient temperatures are $5°C$, $45°C$ and $85°C$.

The electro-thermal simulations are performed based on the power profiles extracted from the active cycling experiments.

4.2. The Application of the Proposed Methodology

Generating the Operating Conditions for Grid Data. The DoE for the electro-thermal simulations on the grid of operating conditions consists in 125 data-points (5 values for each factor). In terms of (I_L, T_{amb}) factors, a 5^2 (2 factors, each at 5 levels) full factorial design is used. I_L takes equally-spaced values from the nominal current (I_{NOM}) to the double of I_{NOM}. For each value of the current, the ambient temperature takes values from $5°C$ to $85°C$, with an increment of $20°C$. For E_R (the 3rd factor), the interaction model (6) of

the maximum E_R is fitted from the experimental data. Based on the E_{R-max} model, the maximum E_R is estimated for all (I_L, T_{amb}) data-points from the grid. The rest 4 values (of E_R) for each (I_L, T_{amb}) data-point are considered the 60%, 70%, 80% and 90% of the estimated E_{R-max}.

Electro-Thermal Simulations on the Grid Scenarios. Electro-thermal simulations are performed based on the estimated PWL parameters on the grid in the space of the operating conditions, resulting the corresponding ΔT values. Fig. 3 shows the simulated values of ΔT corresponding to the grid of power profiles scenarios, as well as the ΔT values resulted from the simulations on the active cycling measurements power profiles. The grid simulations values are drawn with circles, while the values corresponding to active cycling scenarios are represented with dots. The ΔT values are displayed in the space of operating conditions and are coded by the color of the markers.

ΔT Model Fitting. The ΔT values resulted from the grid simulations are used to fit the ΔT model (7), which has the operating conditions as factors. Before fitting, because the operating conditions span on different magnitude orders, the factors are normed in [-1, 1] interval. The model coefficients $(C_0 - C_{19})$, which are extracted from the simulation data with the least squares method, are presented in Table 1. From the resulted values of the coefficients, one can observe that E_R (by coefficient c_3) has the biggest impact on ΔT and the following significant terms are I_L (by coefficient c_1) and the interaction between I_L and E_R (by coefficient c_5). Note that c_0 is the coefficient of the constant term.

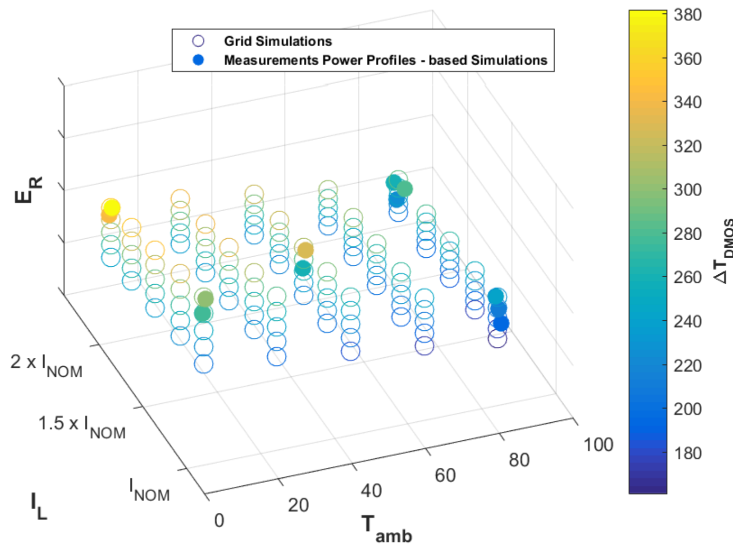

FIGURE 3. The resulted ΔT values for the grid of power profiles scenarios.

TABLE 1. The coefficients of the ΔT model

Coeff.	Value	Coeff.	Value	Coeff.	Value	Coeff.	Value
c_0	287.56	c_5	52.11	c_{10}	20.79	c_{15}	29.91
c_1	58.90	c_6	37.67	c_{11}	2.23	c_{16}	29.03
c_2	17.95	c_7	-5.08	c_{12}	2.38	c_{17}	1.94
c_3	133.69	c_8	5.46	c_{13}	2.61	c_{18}	1.08
c_4	11.94	c_9	38.63	c_{14}	9.86	c_{19}	25.74

Validation of the ΔT Model.

With Leave-K-Out Method. In this case, the value chosen for k is 3. In total, a number of 317 750 (the number of combinations of 125 chosen by 3) metamodels are built, each of them being evaluated on 3 different observations. Fig. 4 presents the histogram of the resulted prediction relative errors of the ΔT model, with leave-3-out validation method. From the distribution of the relative errors resulted the maximum unsigned relative error of 1.34%.

With Bootstrapping Method. For this case, the bootstrapp sample size is chosen to be the same as the original data set. The mean number of left-out data points is 46 (out of 125), meaning that, on average, 63% of the data-points are used for metamodels fitting, while 37% of them are used for validation. Fig. 5 presents the histogram of the resulted prediction relative errors of the ΔT model. According to this, the 95% confidence level interval spreads over relative errors between -0.65% (lower limit) and 0.64% (upper limit).

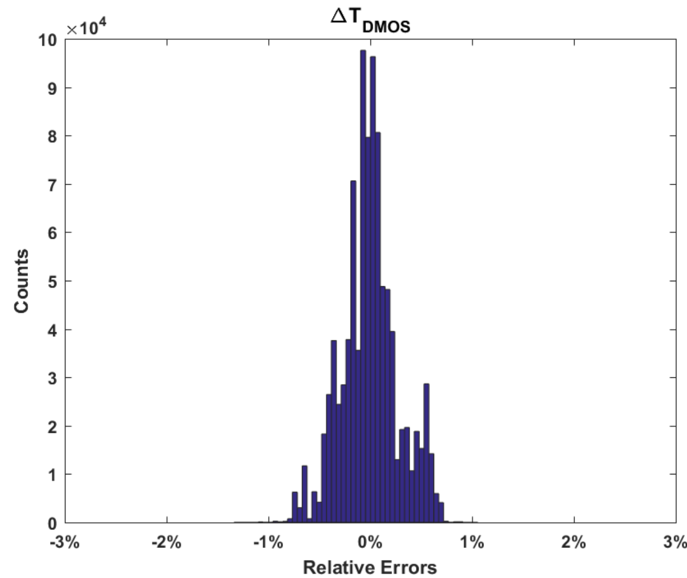

FIGURE 4. The histogram of relative errors with leave-3-out method.

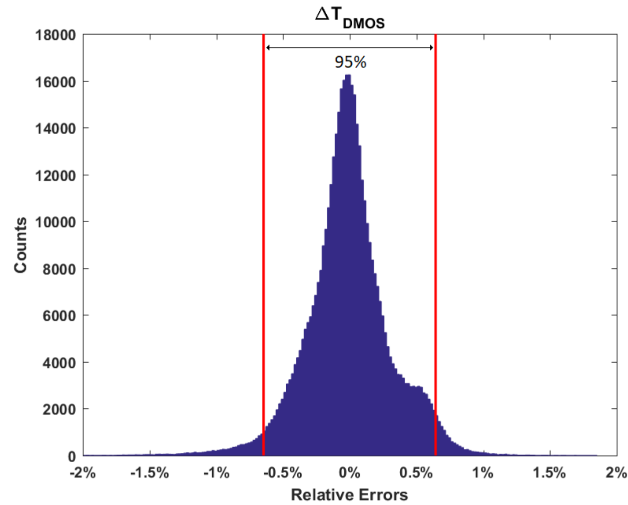

FIGURE 5. The histogram of relative errors with bootstrapping method.

On the Experimental Scenarios - based Electro-Thermal Simulation Data. Apart from the training data set, the ΔT model is evaluated on the operating conditions of the active cycling scenarios (validation data set) and the prediction relative errors of the corresponding ΔT values are computed. The results are shown in Fig. 6. The maximum unsigned relative error obtained is 6.56%.

Discussion. The ΔT model is fitted on a training data set obtained by modeling and approximating the power profiles parameters with the experimental operating conditions. The model validation on this training data set (with leave-k-out and bootstrapping methods) shows small prediction errors (less than 2%). The evaluation of the model on the experimental operating conditions (power profiles) reveals higher prediction errors (up to 6.56%) because of additional errors introduced by the models of the power profiles parameters.

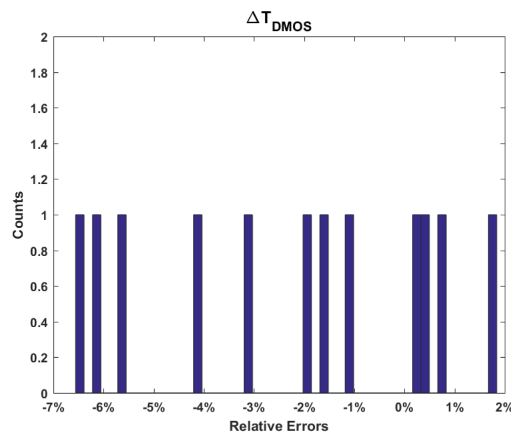

FIGURE 6. The prediction relative errors on validation data set.

5. Conclusions

To meet the current requirements of estimating the lifetime of the power devices within the application operating conditions, the paper presents a methodology for application-aware estimation of the lifetime main stress-factor, i.e. the junction temperature swing within a fast thermal cycle (ΔT), based on electro-thermal simulation. The proposed application-aware ΔT model is fitted from data on a grid of electro-thermal simulations based on estimated power profiles over a wide space of operating conditions. The methodology requires the electro-thermal model of the device and involves the development of several data-driven models. The overall prediction error, computed on data corresponding to the experimental scenarios, is 6.56%. The prediction of ΔT on different operating conditions with the proposed methodology can help estimating the lifetime of the power devices used in a wide range of applications.

REFERENCES

[1] *S. S. Manson*, Thermal stress and low cycle fatigue, McGraw-Hill Book Co., 1966.

[2] *Ciprian V. Pop, Andi Buzo, Cristian V. Diaconu, Georg Pelz, Horia Cucu and Corneliu Burileanu*, Application-Aware Lifetime Model for Power Devices based on Electro-Thermal Simulation, 42th International Semiconductor Conference, 2019, 177-180.

[3] *M. Glavanovics and H. Zitta*, Thermal destruction testing: An indirect approach to a simple dynamic thermal model of smart power switches, 27th European Solid-State Circuits Conference, 2001, 221-224.

[4] *G. Breglio, F. Frisina, A. Magri and P. Spirito*, Electro-thermal instability in low voltage power MOS: Experimental characterization, 11th International Symposium on Power Semiconductor Devices and ICs (ISPSD), 1999, 233-236.

[5] *P. Spirito, G. Breglio, V. d'Alessandro and N. Rinaldi*, Analytical model for thermal instability of low voltage power MOS and SOA in pulse operation, 14th International Symposium on Power Semiconductor Devices and ICs, 2002, 269-272.

[6] *L. Rossi, G. Breglio, A. Irace and P. Spirito*, Thermal Mapping of Power Devices with a Completely Automated Thermoreflectance Measurement System, Ph.D. Research in Microelectronics and Electronics, 2006, 41-44.

[7] *G. Breglio, A. Irace, P. Spirito, R. Letor and S. Russo*, Fast transient infrared thermal analysis of smart Power MOSFETS in permanent short circuit operation, IEEE International Symposium on Power Semiconductor Devices and ICs, 2006, 1-4.

[8] *M. Riccio, G. Breglio, A. Irace and P. Spirito*, A dynamic temperature mapping system with a 320x256 pixels frame size and 100khz sampling rate, 26th International Conference on Microelectronics, 2008, 371-374.

[9] *M. Riccio et al.*, Thermal simulation and ultrafast IR temperature mapping of a Smart Power Switch for automotive applications, 21st International Symposium on Power Semiconductor Devices ICs, 2009, 200-203.

[10] *J. M. Bosc et al.*, Reliability characterization of LDMOS transistors submitted to multiple energy discharges, 12th International Symposium on Power Semiconductor Devices ICs, 2000, 165-168.

[11] *M. Pfost, C. Boianceanu, H. Lohmeyer and M. Stecher*, Electrothermal Simulation of Self-Heating in DMOS Transistors up to Thermal Runaway, IEEE Transactions on Electron Devices, **60**(2013), No. 2, 699-707.

[12] *D. Montgomery*, Design and Analysis of Experiments, John Wiley and Sons, 2001.