Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

MODELING FATIGUE LIFE OF POWER SEMICONDUCTOR DEVICES WITH ε-N FIELDS

Olivia Bluder
Kathrin Plankensteiner
Michael Nelhiebel
Walther Heinz
Christian Leitner

KAI – Kompetenzzentrum für Automobil- und Industrieelektronik
Europastrasse 8
Villach, 9524, AUSTRIA

ABSTRACT

In this study, fatigue life of power semiconductor devices measured in cycles to failure during an accelerated stress test in a climate chamber is analyzed. The tested devices fail mainly in a short circuit event and their physical inspection reveals cracks in the power metallization. Commonly, the time till fracture of macroscopic metal layers is modeled with S-N or ε-N fields, this means that the lifetime (N) depends on the mechanical stress (S) or the strain (ε), respectively. Metal layers of semiconductor devices are microscopic (≤ 20µm) and, in general, their ageing mechanisms are different than for macroscopic layers, nevertheless the application of the macroscopic based ε-N model to semiconductor lifetime data shows good results. Hence, fatigue life due to micro-mechanisms can be described by parameters representing the mechanical load (strain) in the device.

1 INTRODUCTION

Modeling and predicting the fatigue life of power semiconductor devices is a challenging part of reliability analysis. It requires not only to understand the dominant failure mechanism and to be able to measure lifetime, but also to identify the relevant parameter representing the stress that leads to the device failure. If the device fails due to cracks in the metal layers, in literature two approaches are used: either modeling the crack growth directly or the measured lifetime dependent on parameters representing the mechanical stress in the device. Both approaches were developed and verified for macroscopic structures. The objective of this paper is to apply the second modeling approach to semiconductor lifetime data and evaluate whether the model is also applicable to microscopic crack growth in thin metal layers (≤ 20µm). Since the time-effort for investigations of the fatigue life under real load conditions is too high (up to one year), accelerated stress tests are widely used.

2 ACCELERATED STRESS TESTING OF SEMICONDUCTOR DEVICES

The lifetime of semiconductor devices is measured in Cycles to Failure (CTF). Commonly, the average lifetime of such devices lies between 10^6 and 10^9 cycles, dependent on the product and the application. For this investigation short-circuit cycle testing is performed according to the AEC-Q100-012 (Automotive Electronics Council 2006) standard with the ACUTE test system (Steinwender 2010). With this real time test system it is possible to measure and record the lifetime of up to256 Devices under Test (DUTs) simultaneously, tested in groups with maximum 16 different operating conditions. An operating condition is thereby defined by the inductance, resistance and the pulse width of the load pulse. Test acceleration can be achieved by increasing the electrical load (e.g. the energy).

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

The real time test system counts the applied load cycles and records the state of all DUTs after each cycle. Therefore, exact CTF values are available for the statistical evaluation. However, testing until all devices failed is not always possible. If a test is stopped before all DUTs failed, for the remaining DUTs only a time until surviving is recorded. These DUTs are so-called right censored data, they need special attention at the modeling process. Although for one operating condition the electrical load is nominally the same, in reality slight variations caused by the test system hardware and the DUT itself are present. Due to these variations and different applied operating conditions, it is possible to observe the lifetime of DUTs for a variety of energies within a pre-defined interval.

Figure 1 shows the measured CTF under varying energies for one device type of interest, where data points in black rectangles mark censored observations. The non-linear dependence between energy and CTF is clearly visible. In fracture mechanics such relationships are commonly modeled by a Wöhler line (Wöhler, 1870), also called S-N or ε-N line (for a detailed description see Section 3.1). At this representation, on the y-axis the applied "stress" is plotted, where "stress" can be any quantity describing the load, e.g. energy, temperature rise, mechanical stress or strain.

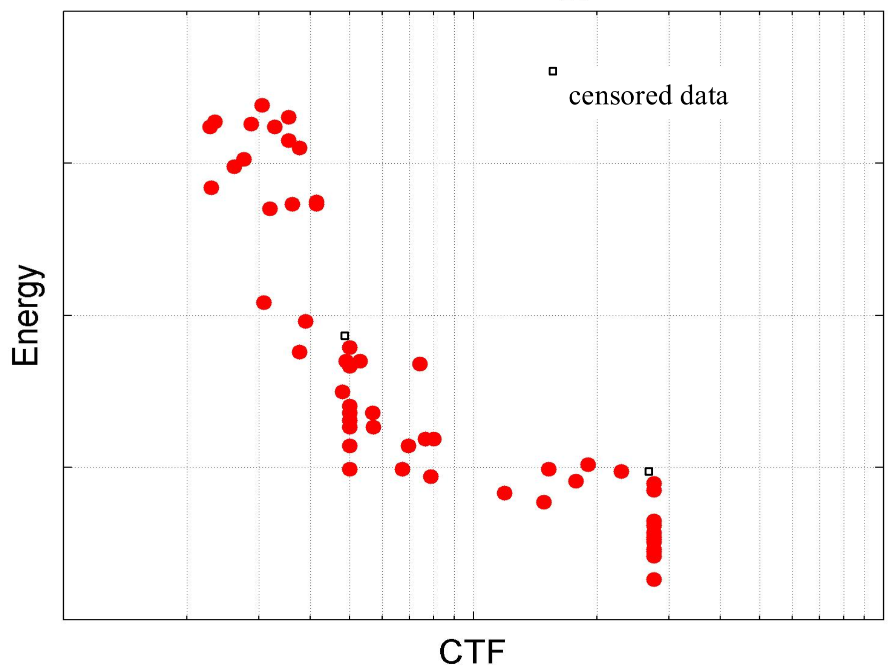

Figure 1: Measured CTF vs. measured energy for devices with equal power metallization. DUTs in black boxes did not fail within testing time (censored data).

The applied cyclic stress and the resulting temperature rise during the load pulse, activate a degradation process in the metal layers of the DUT. The Power Metal (PM) and interconnect degradation leads to the loss of cooling capacity and therefore to device failure (Nelhiebel et al. 2011).

3 FATIGUE MODELING OF SEMICONDUCTOR DEVICES

In Castillo et al. (2008) and Castillo and Fernández-Canteli (2009) a statistical S-N model for the crack growth in materials of structural and mechanical components is introduced, where the applied mechanical stress (S) was used as a covariate. To verify the model, fatigue sample data from specimens made of steel alloy bars were used. The same approach can also be used for an ε-N model, meaning that instead of the mechanical stress the strain (ε) is used as covariate. If the investigated test conditions lead to plastic deformation in the specimen, the ε-N model should be used, since it takes the elastic ΔT range into account which is assumed to be damage-free.

Based on the data characteristic shown in Figure 1 and the failure mechanism of the semiconductor devices explained in Section 2, a Wöhler line approach to model the measured CTF dependent on parameters representing the mechanical load is justified. Whether the model proposed by Castillo et al. 2008 is also applicable to microscopic crack growth is part of this investigation. Electro-thermal FEM simulations (Bernardoni 2012) of the applied load pulses to the given semiconductor device indicate that the applied energies lead to temperature rises that cause plastic deformation, therefore a strain based (ε-N) instead of a stress based (S-N) model is used.

3.1 The ε-N Model

The idea of the ε-N modeling approach (Castillo et al. 2008) is to transform the data non-linearly based on a parameter representing the mechanical strain. Therefore, the measured lifetime (CTF) of the semiconductor device is transformed dependent on the applied strain range (Δε) in the following way

$$X = (log_{10}CTF - B)(log_{10}(\Delta\varepsilon/\Delta\varepsilon_0) - C)\,, \tag{1}$$

where *B* is the threshold for the lifetime ($log_{10}CTF$), *C* is the endurance limit for Δε and $\Delta\varepsilon_0$ is the fatigue limit (the Δε for the maximum observed median lifetime). For interpretation purposes, especially for the CTF, the decadic logarithm is here used. *B* and *C* are parameters that need to be estimated and $\Delta\varepsilon_0$ is defined by the given data. With the weakest link principle it can then be shown that the transformed values follow an extreme value distribution (Castillo, Fernández-Cantelí 2009), e.g. Gumbel $X \sim G(\lambda, \delta)$.

The estimation of the parameters (*B, C, λ, δ*) is divided into two steps. First, the threshold parameters *B* and *C* are estimated and second the parameters of the extreme value distribution λ and δ. To define the minimization problem for *B* and *C*, Equation (1) is treated as the expectation of *X*. Rearranging (1) with respect to *logCTF* leads to

$$E[log_{10}CTF | log_{10}(\Delta\varepsilon/\Delta\varepsilon_0) - C)] = B + \frac{\mu}{log_{10}(\Delta\varepsilon/\Delta\varepsilon_0) - C},$$

where $\mu = E(X)$ is the mean of the Gumbel distribution. Then, *B* and *C* can be estimated by solving the least squares (LS) minimization problem

$$\sum_{i=1}^{n} \left(log_{10}CTF_i - B - \frac{\mu}{log_{10}(\Delta\varepsilon_i/\Delta\varepsilon_0) - C} \right)^2 \underset{B,C,\mu}{\longrightarrow} min\,, \tag{2}$$

where *n* is the sample size. The mean μ is only estimated as intermediate result, it is not needed for further calculations. To avoid convergence problems at solving, the results of several initial values should be obtained and analyzed. With estimated *B* and *C* from (2), the data transformation given in (1) can be performed and next, the parameters of the Gumbel distribution (λ and δ) can be estimated.

As mentioned before, the given set of semiconductor lifetime data contains censored observations. To include this information properly the Maximum Likelihood (ML) is used to estimate the distribution parameters, where censored observations are included via the survivor function *S(x) = 1-F(x)*. After that confidence intervals for all percentiles (*p*) of interest can be calculated by

$$x_p = \lambda + \delta\left[log\left(-log(1-p)\right)\right], \quad 0 \le p \le 1.$$

3.2 Application to Semiconductor Fatigue Lifetime

To apply the proposed model in Section 3.1 to the given semiconductor lifetime data, strain ranges corresponding to the applied load pulses are needed. Since the plastic deformation is responsible for the degradation process, for the lifetime model only the plastic strain is used, not the total strain. Thermo-mechanical simulations of the plastic strain are not available, but an approximation based on electro-thermal FEM simulations of the temperature rise (ΔT) in a DUT during a load pulse can be calculated. In general, $\Delta\varepsilon_{plas}$ can be approximated by (Ohring 2001).

$$\Delta\varepsilon_{plas} = \Delta\alpha \cdot (\Delta T - \Delta T_{elas}), \tag{3}$$

where $\Delta\alpha$ is the coefficient of thermal expansion for the used PM on Silicon (e.g. for Cu on Si $\Delta\alpha$=13.6*10^{-6}). ΔT_{elas} is the maximum temperature rise to remain in the elastic regime, meaning the maximum ΔT such that the applied stress is only elastic and not plastic (see Figure 2). This value can be gained by wafer bow experiments, which measure mechanical stress vs. temperature hysteresis on wafers level with the PM of interest (Heinz 2014). Figure 2 shows the results of such a wafer bow experiment where the red line indicates the linear-elastic behavior during heating. The starting point of the plastic regime during heating is not clearly defined, but most common is the point where the hysteresis deviates first from the linear-elastic line. Knowing ΔT_{elas} from experiments and FEM simulated ΔT for all DUTs, $\Delta\varepsilon_{plas}$ for all DUTs can be approximated by (3) and the data can be transformed according to (1). Thereby, for $\Delta\varepsilon_0$ the lowest strain rate in the dataset was used.

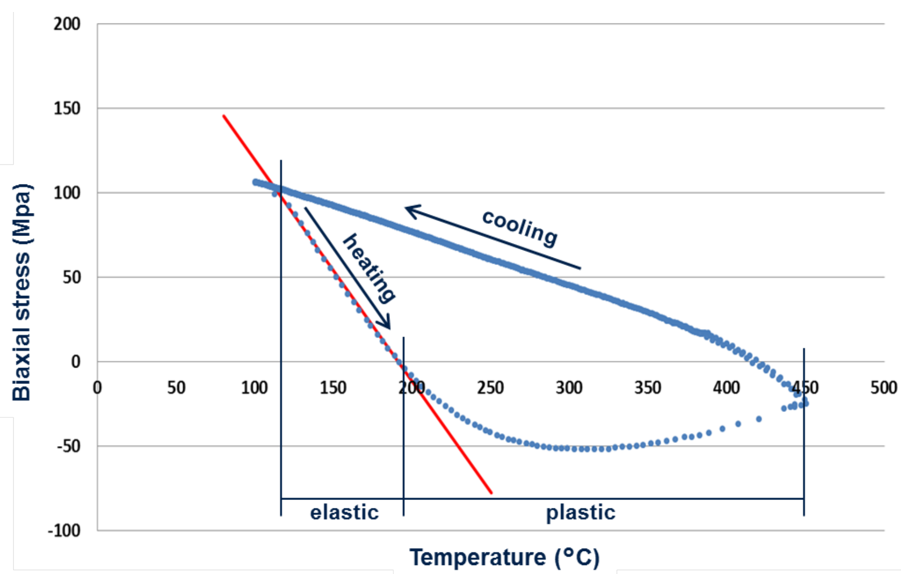

Figure 2: Stress vs. temperature hysteresis from wafer bow measurements.

Due to the weakest link principle and the compatibility condition introduced by Castillo and Fernández-Canteli (2009), the Gumbel, the Weibull or the Fréchet distribution are the only physically suitable distributions. Fréchet can be rejected, since its support is not restricted to the positive values, which is needed for lifetime data. To define whether the Gumbel or the Weibull distribution fits the given data best, the Anderson-Darling (AD) test is used. Table 1 shows that for the transformed data X the AD test accepts both proposed distributions, hence, the physically right extreme value distributions fit the data well. Although the p-value obtained with the AD test statistics for the Gumbel distribution is higher than for the Weibull, the Gumbel will be used for the fatigue model because it contains less parameter.

Table 1: Results from Anderson-Darling test (α = 5%) on transformed data *X*.

Distribution	# of param.	AD stat	critical value	p-value
Gumbel	2	0.92	2.45	0.39
Weibull	3	0.74	2.45	0.48

3.3 Analysis of Modeling Results

Figure 3 visualizes the median ε-N curve for the given semiconductor fatigue measurements (blue), as well as 90% (dashed blue) and the lower 99% (dashed red) confidence bounds. The plot shows that the Gumbel model fits the data well and that the characteristic fatigue life of the semiconductor devices under investigation can be explained by the approximated strain range. Narrow 90% confidence bounds underline the good model quality. Lower confidence bounds are especially interesting since reliability in semiconductor industries is defined by these thresholds.

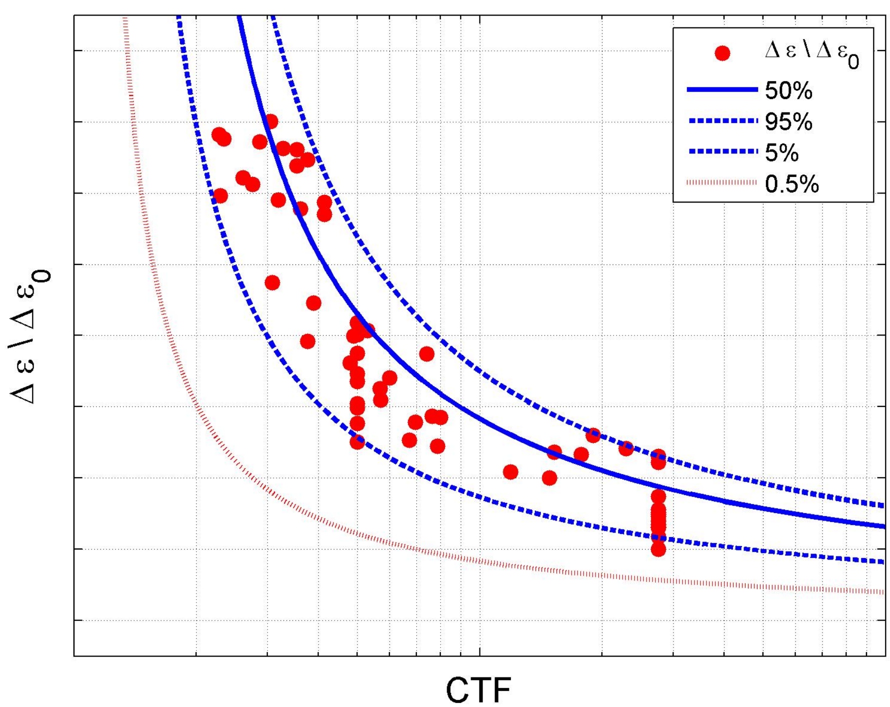

Figure 3: ε-N fatigue model for semiconductor fatigue life data with 90% and lower 99% confidence bounds.

In general, the plots of ε-N curves can be divided into a low-cycle (also called Coffin-Manson part) and high-cycle (also called Basquin part) fatigue. For example, steel has a horizontal line at high-cycle fatigue, this means a real endurance limit is present. All other materials show flat, but not horizontal, lines. However, commonly also for them an endurance limit is specified, e.g. region where the change in the ε-N curve is small. This behavior is also visible for the ε-N curve of the given semiconductor fatigue data (Figure 3), hence an endurance limit can be specified and extrapolations to lower strain rates are possible with only a small expected prediction error.

With this modeling approach it is possible to define load thresholds for a reliable operation of semiconductor devices given a confidence level α. Furthermore, the comparison of ε-N curves from e.g. devices tested under different ambient temperatures can be used to understand the influence on the lifetime. Investigations show that the influence need not be consistent for high and low-cycle fatigue region. The lifetime in the high-cycle fatigue region increases, but at the same time the lifetime in the low-cycle fatigue region decreases or vice versa. To account for this ε-N fields need to be investigated.

Extrapolations from the low-cycle fatigue region to even lower lifetime at higher load conditions are not possible with this model approach, because practically fatigue life follows an s-shaped curve and the Gumbel model covers only the second part of it. However, for semiconductor lifetime testing the value of interest in this fatigue region is the maximum energy which a device can sustain and this value can be found experimentally by energy ramp up tests. Therefore, the Gumbel ε-N model is sufficient for statements about the reliability of a semiconductor device.

For comparative purposes, an alternative to the ε-N Gumbel model is evaluated. The model consists of two separate regression models, one for the Coffin-Manson and one for the Basquin part of fatigue life (see Figure 4). The changing point between the models is held fix. Evaluation of the model quality with the Bayesian Information Criterion (BIC) shows that the ε-N Gumbel model performs significantly better (BIC = -130.1) compared to the alternative model (BIC = -26.1).

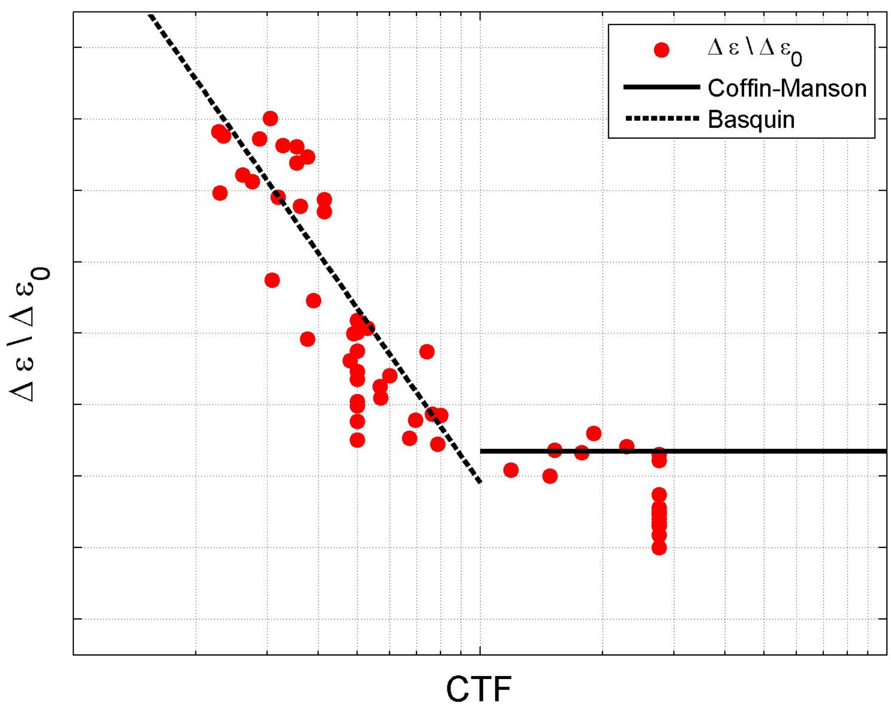

Figure 4: Alternative modeling approach with two separate regression models representing the Coffin-Manson and the Basquin part of fatigue life.

4 SUMMARY & CONCLUSION

In this paper we introduced the Gumbel ε-N model for semiconductor lifetime data, which relates lifetime to mechanical strain. For the investigation, the measured lifetime under varying load energies from cycle stress tests for devices with equal power metallization were available. The proposed fatigue model is based on a non-linear data transformation. To apply it, first the strain range ($\Delta\varepsilon$) had to be calculated based on wafer bow measurements. Next, the model and the Gumbel distribution parameters were estimated by a two-step procedure with LS and ML method. The analysis of the results shows that the assumed fatigue model fits the data well. Small 90% confidence bounds underline the good model quality.

To sum up, with this approach the description of semiconductor lifetime at high-cycle and low-cycle fatigue can be described and the prediction for future load conditions is possible, even extrapolations in direction of the endurance limit. An increase in model quality may be achieved by using other mechanical load parameters or simulated instead of approximated strain ranges. However, further investigations of other power metallization are needed to verify if the Gumbel ε-N model is in general suitable for fatigue life of microstructures.

ACKNOWLEGMENT

This work was jointly funded by the Austrian Research Promotion Agency (FFG, Project No. 831163) and the Carinthian Economic Promotion Fund (KWF, contract KWF-1521|22741|34186).

REFERENCES

Automotive Electronics Council. 2006. Short Circuit Reliability Characterization of Smart Power Devices for 12V Systems. AEC - Q100-012.

Bernardoni, M. 2012. "Thermal and Electro-thermal Modeling of Electronic Devices and Systems for High-power and High-frequency Applications", Ph.D. thesis. Università degli Studi di Parma.

Castillo E., A. Fernández-Canteli. 2009. *A Unified Statistical Methodology for Modeling Fatigue Damage*. Springer Science + Business Media.

Castillo E., A. Fernández-Canteli, H. Pinto, M.L. Ruiz-Ripoll. 2008. "A Statistical Model for Crack Growth based Tension and Compression Wöhler Fields". *Engineering Facture Mechanics* 75:4439–4449.

Heinz, W. 2014. "Influence of Initial Microstructure on Thermo-mechanical Fatigue Behaviour of Cu Films on Substrates". In *Proceedings of Materials for Advanced Metallization MAM 2014.*

Nelhiebel, M., R. Illing, C. Schreiber, S. Wöhlert, S. Lanzerstorfer, M. Ladurner, C. Kadow, S. Decker, D. Dibra, H. Unterwalcher, M. Rogalli, W. Robl, T. Herzig, M. Poschgan, M. Inselsbacher, M. Glavanovics, S. Fraissé. 2011. "A Reliable Technology Concept for Active Power Cycling to Extreme Temperatures", *Microelectronics Reliability* 52: 1927–1932.

Ohring, M. 2001. *Materials Science of Thin Films*. Academic Press.

Steinwender, B. 2010. "Design of a Real-time System for in-situ Characterization of Smart Power Switches during Cycle Stress Testing". In Proceedings of VIP 2010, 155–159.

Wöhler, A. 1870. *Über die Festigkeitsversuche mit Eisen und Stahl*. Ernst & Korn.

AUTHOR BIOGRAPHIES

OLIVIA BLUDER is a Junior Researcher at KAI- Kompetenzzentrum für Automobil- und Industrieelektronik GmbH, Austria. She received her M.S and Ph.D. degrees in statistical analysis and Bayesian modeling of Smart Power semiconductor lifetime at the Alpen-Adria-Universität of Klagenfurt (Department of Statistics), Austria. Her email address is olivia.bluder@k-ai.at.

KATHRIN PLANKENSTEINER is a Ph.D. student at the Alpen-Adria-Universität of Klagenfurt (Department of Statistics), Austria and works at KAI- Kompetenzzentrum für Automobil- und

Industrieelektronik GmbH, Austria. Her research field is the reliability of smart power semiconductor lifetime data with Bayesian models. Her email address is kathrin.plankensteiner@k-ai.at.

MICHAEL NELHIEBEL is a Senior Researcher at KAI- Kompetenzzentrum für Automobil- und Industrieelektronik GmbH, Austria. He received his M.S. and Ph.D. degrees in physics from Technische Universität Wien and Ecole Centrale Paris, respectively. In 1999 he joined Infineon Technologies Austria. As a principal engineer, he currently focuses on reliability research at the Infineon-held competence center KAI, from MOS interface defects to thermo-mechanical degradation of power devices. His email address is michael.nelhiebel@k-ai.at.

WALTHER HEINZ is a Junior Researcher at KAI- Kompetenzzentrum für Automobil- und Industrieelektronik GmbH, Austria. He received his M.S and Ph.D. degrees in thermo-mechanical fatigue of thin films at the Montan University of Leoben (Department of Material Science), Austria. His email address is walther.heinz@k-ai.at.

CHRISTIAN LEITNER is a Lab Engineer at KAI- Kompetenzzentrum für Automobil- und Industrieelektronik GmbH, Austria. He received his Master degree in implementing industry measurement solutions at the University of Applied Sciences of Klagenfurt (Department of Communication Engineering for IT). His email address is christian.leitner@k-ai.at.