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 AND PREDICTION OF SMART POWER SEMICONDUCTOR LIFETIME DATA USING A GAUSSIAN PROCESS PRIOR

Kathrin Plankensteiner
Olivia Bluder

KAI - Kompetenzzentrum für Automobil- und Industrieelektronik GmbH
Europastraße 8
Villach, 9524, AUSTRIA

Jürgen Pilz

Alpen-Adria-Universität Klagenfurt
Universitätsstraße 65-67
Klagenfurt, 9020, AUSTRIA

ABSTRACT

In automotive industry end-of-life tests are necessary to verify that semiconductor products operate reliably. Due to limited test resources it is not possible to test all devices and thus, accelerated stress tests in combination with statistical models are commonly applied to achieve reliable forecasts. Challenging thereby is the highly complex data that shows mixture distributions and censoring. For the main purpose, the extrapolation to other test conditions or designs, neither frequently used acceleration models like Arrhenius, nor complex models like Bayesian Mixtures-of-Experts or Bayesian networks give accurate lifetime predictions, although, the latter two are precise in case of interpolation. To compensate the limitations of ordinary linear based regression models, we propose the application of a Gaussian process prior. The proposed model shows a high degree of flexibility by exploiting sums or products of appropriate covariance functions, e.g. linear or exponential, and serves as a reliable alternative to currently applied methods.

1 INTRODUCTION

The development in automotive industry increases fast due to new advances in material science and extensive use of smart power semiconductor devices in daily life. At the same time reliability has to be guaranteed since in automotive applications the developed devices are used for passengers' safety. Therefore, extensive lifetime testing and modeling based on known physical relationships, e.g. Arrhenius or Coffin-Manson (NIST 2010), was performed in the past for each device type to cover these requirements. Due to steadily increasing demands on device lifetime, the testing time continues to grow immeasurably, but based on limited resources, it is neither possible to test, nor to inspect all devices anymore. Thus, reliable forecasts for the lifetime are needed.

Nowadays, accelerated stress tests in combination with statistical models are commonly applied to model the lifetime of semiconductor devices. Frequently used acceleration models like Arrhenius are insufficient since the lifetime behavior of today's semiconductor products is highly complex, showing different failure mechanisms resulting in a mixture distributed lifetime (Bluder et al. 2012). The application of a Bayesian Mixtures-of-Experts (MoE) Norris-Landzberg (Bluder 2011) or a Bayesian network (BN) model (Plankensteiner, Bluder, and Pilz 2013) leads to a satisfying interpolation quality, but their performance is insufficient in case of extrapolation.

It is assumed that ordinary linear based regression models cannot capture the complex behavior of the observed data. Therefore, a Gaussian process (GP) prior is proposed to model the mean lifetime of each mixture component. Therewith, a new approach to model semiconductor lifetime is gained.

2 TEST PARAMETERS, ELECTRO-THERMAL & THERMO-MECHANICAL EFFECTS

The analyzed data was obtained under different electrical and thermal stress conditions produced by a cycle stress test system (Glavanovics et al. 2007). To stress the semiconductor devices, a combination of five test parameters is defined by the operator:

1. the ambient temperature (T_{amb}),
2. the stress pulse shape (triangle or rectangle),
3. the induced peak current (I),
4. the pulse length (t_p), and
5. the repetition time (t_{rep}).

Another significant factor is the clamping voltage (V), which is a property of the device itself and cannot be adjusted by the operator. An example of an oscilloscope measurement during a repetitive stress test is shown in Figure 1, with measured V in magenta and induced I in green. As illustrated, the main stress pulse is defined by the shape, the length (t_p), and the height (I) of the pulse as well as the time between two stress pulses (t_{rep}). The combination of these four factors with T_{amb} defines one test setting (test condition).

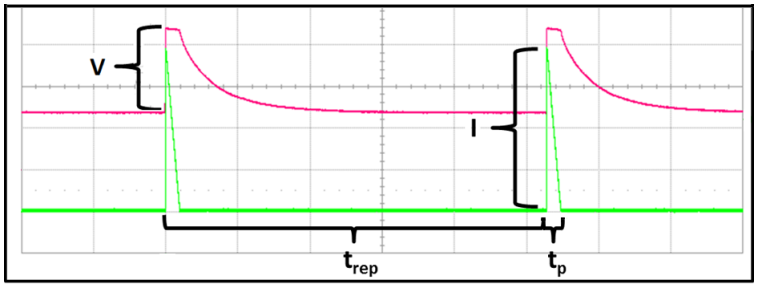


Figure 1: Oscilloscope measurement of I and V on different y-axes during a test.

Based on the definition of these test parameters, the Device Under Test (DUT) heats up during the first minutes to a stable case temperature (T_{case}). Further, caused by electric power dissipation the DUT heats up and cools down during each stress pulse. The thereby induced temperature rise (ΔT) can either be simulated with the Finite-Element-Method (FEM) (Kosel et al. 2007, Riccio et al. 2009) or it can be approximated analytically (Glavanovics and Zitta 2001). Given this temperature rise, the corresponding peak temperature (T_{peak}) can be calculated by $T_{peak} = T_{case} + \Delta T$.

Since the DUT is built of various metal layers with different thermal expansion coefficients, the temperature cycling causes, additionally to electro-thermal, also thermo-mechanical effects. Similarly, the thereby induced mechanical stress can either be simulated using FEM (Kravchenko et al. 2014) or it can be analytically determined by approximation formulas (Chen and Neson 1979, Suresh 1998).

Summarized, the observed lifetime is the result of interacting electro-thermal and thermo-mechanical effects.

3 OBSERVED DATA & CHALLENGES

During the repetitive stress test, the system counts the number of applied stress cycles. If the DUT fails during the test, the exact lifetime, measured in cycles to failure (CTF), is logged by the test system. Whenever a DUT does not fail by the end of the test, it gets the status survivor (SURV) and the total number of applied stress cycles is logged for this DUT. In this case, the lifetime for the DUT is censored.

The logarithmic transformed CTF (logCTF) follow a mixture of two normal distributions (Bluder 2011, Plankensteiner 2011). Figure 2 shows the outcome of three tests in a normal probability plot with a logarithmic x-axis. Each stress test consists of 16 DUTs, illustrated by blue triangles (a high-stress test), green circles (a moderate-stress test) and red squares (a low-stress test). As expected, increasing the stress level leads to a significantly shorter lifetime on average, cf. 50% quantiles. Taking a closer look at the lifetime of DUTs tested with the moderate-stress level, a double distribution with CTF $< 10^6$ for the first component and CTF $> 10^6$ for the second component is evident. The DUTs of the other two tests follow a single component distribution. For the high-stress test, all DUTs fail within the first 10^4 stress cycles, whereas for the low-stress test the first fail does not occur before 10^6 cycles and moreover, three DUTs of the low-stress test even survive, highlighted by the black rectangle.

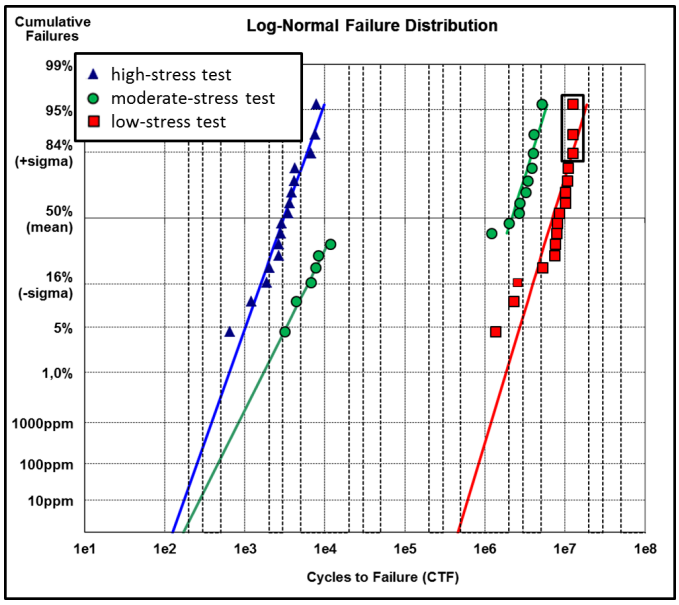
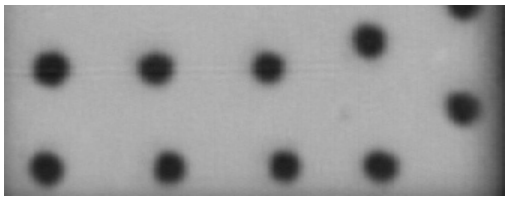


Figure 2: Semiconductor lifetime probability plots from three stress tests.

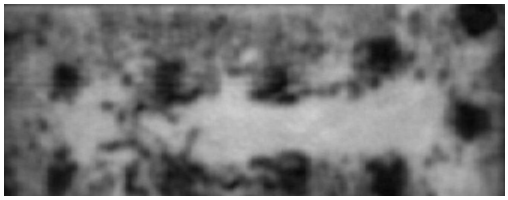
The physical failure inspections after testing indicate that the mathematically identified mixture components represent different physical failure mechanisms. Pictures of a Scanning Acoustic Microscopy (SAM) and a Focused Ion Beam (FIB) cut of stressed DUTs can be taken from Figure 3. DUTs failing before 10^6 stress cycles neither show any degradation (e.g. cracks or voids) nor delamination effects, whereas the physical failure description of DUTs failing after 10^6 stress cycles is characterized by degradation and delamination. A detailed discussion of the FIB cuts is given in (Nelhiebel et al. 2011).

For the investigation in this paper lifetime data from five different device types, A, B, C, D, and E, are available. Device types A to D are similar devices, only differing in the size and shape of the active DMOS area, whereas device type E belongs to another device technology. With this data set, not only the model and interpolation quality is investigated, but also the prediction power to another device type.

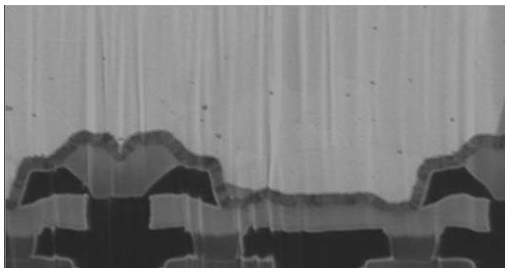
Since logCTF are a mixture of two normally distributed components representing two different physical failure mechanisms, the data set is divided into two subsets. Based on the fact that it is neither possible to inspect all DUTs nor to identify the actual physical failure cause in each of them, this separation is based on the characteristics of the measured lifetime distribution and the physical inspection of some representative DUTs from each mixture component. Thereby it is assumed, that within one component always two different (competing) failure mechanisms are present. One of them is dominant and leads to the DUT's failure. With this, the statistical problem of group identification and allocation is derived, which is modelled by so called mixture weights related to switching regression models (Frühwirth-Schnatter 2006).



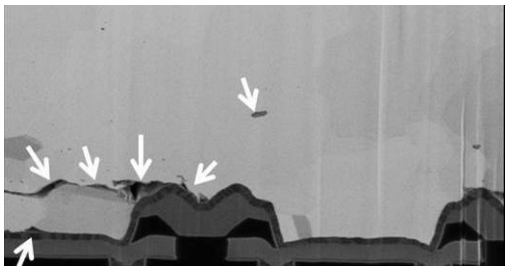
(a) SAM of DUT failed before 10^6 CTF.



(b) SAM of DUT failed after 10^6 CTF.



(c) FIB cut of DUT failed before 10^6 CTF.



(d) FIB cut of DUT failed after 10^6 CTF.

Figure 3: Physical failure inspection reveals different failure mechanisms for DUTs that fail in the first and second mixture component.

The mean lifetime of the first and the second mixture component is 3.5 logCTF and 7.5 logCTF, respectively. On average, the variance of the first subset is higher than the variance of the second one. The first and second subset contain 169 and 867 data points, respectively, tested under 65 different stress conditions. Table 1 shows the partition for all five device types, e.g. for device type A data of 34 different stress conditions with 485 data points are available. In total, 110 censored observations are included.

Table 1: Partition of 65 test conditions for five device types.

Device type	A	B	C	D	E
# stress condition	34	11	9	5	6
# lifetime data	485	169	214	72	96
# censored data	110	142	142	18	0

4 GAUSSIAN PROCESS PRIOR

To compensate the limitations of ordinary linear regression models, we propose the application of a GP prior (Gelman et al. 2014, Shi and Choi 2011) to model the mean function of the smart power semiconductor lifetime data of each mixture component. In this investigation a standard GP with zero mean function is used. Therefore assume that n observation pairs $(\boldsymbol{x}_1, logCTF_1), \ldots, (\boldsymbol{x}_n, logCTF_n)$ are available, where $\boldsymbol{x}_i$ denote the vector of d covariates $(x_{i1}, \ldots, x_{id})$. Since data is a mixture of two normally distributed components, $\boldsymbol{\eta} = (\eta_1, \eta_2)$ with $\eta_2 = 1 - \eta_1$, $\boldsymbol{f} = (f_1, f_2)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)$ define the mixture weight, the mean function and the standard deviation of the first and second component, respectively. With this, the model for the data of interest is defined by (Vanhatalo et al. 2013)

$$\text{observation model:} \quad \boldsymbol{logCTF}|\boldsymbol{\eta}, \boldsymbol{f}, \boldsymbol{\sigma}, \sim \prod_{i=1}^{n} \left[\sum_{k=1}^{2} \eta_k p(logCTF_i | f_k(\mathbf{x}_i), \sigma_k^2) \right] \tag{1}$$

$$\text{GP prior:} \quad f_k|\boldsymbol{x}_i, \boldsymbol{\theta}_k \sim GP(0, k(\boldsymbol{x}_i, \boldsymbol{x}_i'|\boldsymbol{\theta}_k)) \tag{2}$$

$$\text{hyperprior:} \quad \boldsymbol{\theta}_k, \sigma_k^2 \sim p(\boldsymbol{\theta}_k) p(\sigma_k^2), \tag{3}$$

where *GP* defines a standard Gaussian process with parameter vector $\boldsymbol{\theta}_k$ and $k(\boldsymbol{x}_i, \boldsymbol{x}_i'|\boldsymbol{\theta}_k)$ characterizes the correlation between $\boldsymbol{x}_i$ and any other input location $\boldsymbol{x}_i'$. Evaluating the combinations of appropriate covariates, using a Norris-Landzberg relationship gives the best fit according to the Bayesian Information Criterion (BIC) (Held 2008), where models with smaller BIC values are preferred. Therewith, the covariates are defined by $\boldsymbol{x} = (\log t_{rep}, \log \Delta T, T_{peak}^{-1})$.

The application of flexible covariance functions like squared exponential or Matern (Shi and Choi 2011) leads to non linear interpolations of the lifetime logCTF as shown in Figure 4. Based on the used covariates, t_{rep}, ΔT and T_{peak}^{-1}, Figure 4 shows that increasing t_{rep} (e.g. from 36ms to 60ms, red dashed line) at $T_{case} = 150°C$ and $\Delta T = 250°C$ leads to a shorter lifetime on average. Since increasing t_{rep} at constant T_{case} and ΔT means a reduction in stress, this does not fulfill expectations gained from previous device technologies and thus, only combinations of constant and linear covariance functions need to be considered for this model.

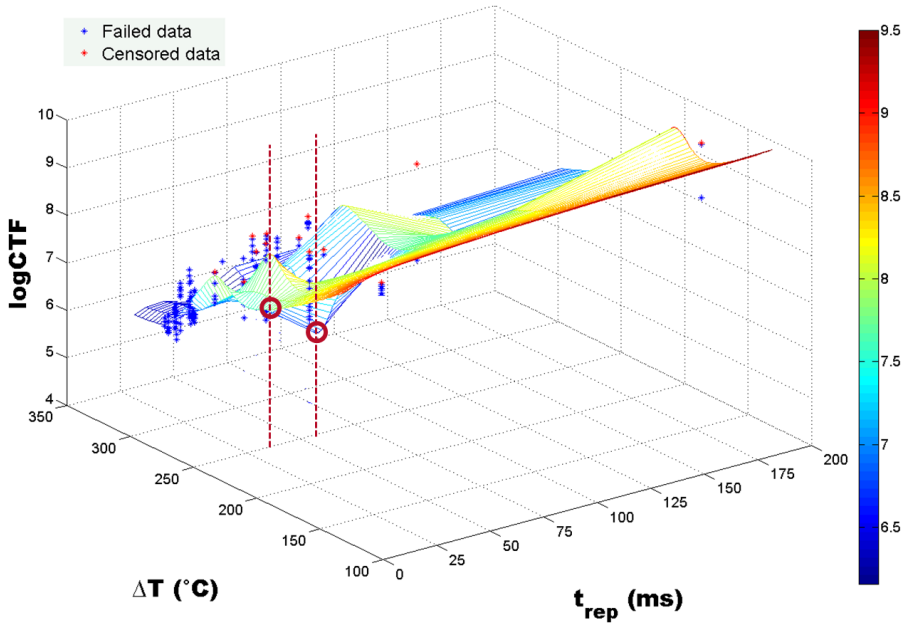


Figure 4: Lifetime prediction for the second mixture component at $T_{case} = 150°C$ using a standard GP with squared exponential covariance function.

For the investigation of appropriate combinations of covariance functions (single, sum or product) the BIC is again used. Results are listed in Table 2, where BIC_1 and BIC_2 denote the BIC value for the first and second mixture component, respectively. It turns out that for the first mixture component, the application of a linear covariance function provides the best fit, whereas for the second mixture component the sum of a constant and a linear covariance function is more appropriate. For the purpose of a common model assumption, the sum of a constant and a linear covariance function is used for both mixture components.

Table 2: BIC values for combinations of constant (c) and linear (l) covariance functions.

covariance function	c	l	sum(c, l)	prod(c, l)
BIC_1	182.2	**172.6**	177.7	177.7
BIC_2	453.0	356.8	**355.9**	361.9

With this, $k(\boldsymbol{x},\boldsymbol{x}'|\boldsymbol{\theta_k})$ is defined by

$$k(\boldsymbol{x},\boldsymbol{x}'|\boldsymbol{\theta_k}) = I\sigma^2_{const,k} + \boldsymbol{x}^T\Sigma_{lin,k}\boldsymbol{x}',$$

where $\boldsymbol{\theta_k} = (\sigma_{const,k}, \sigma_{\log t_{rep},k}, \sigma_{\log \Delta T,k}, \sigma_{T^{-1}_{peak},k})$, I denotes the unity matrix with dimension three and $\Sigma_{lin,k} = diag(\sigma^2_{\log t_{rep},k}, \sigma^2_{\log \Delta T,k}, \sigma^2_{T^{-1}_{peak},k})$ is the diagonal matrix containing the prior variances of the linear model coefficients. Based on missing prior knowledge, standard priors following log-uniform distributions for the data variance (σ^2_k) and function parameters ($\sigma^2_{const,k}, \sigma^2_{\log t_{rep},k}, \sigma^2_{\log \Delta T,k}, \sigma^2_{T^{-1}_{peak},k}$) of each mixture component are applied.

Bayesian inference is performed in MATLAB® using the statistical toolbox *GPstuff* (Vanhatalo et al. 2013). This toolbox was extended to handle mixtures of log-normally distributed components where data can be censored. For parameter estimation a combination of surrogate slice sampler for the latent variable $\boldsymbol{f}$ (Murray, Adams, and MacKay 2010) and elliptical slice sampler for the parameter vector $(\boldsymbol{\theta}_k, \boldsymbol{\sigma}^2_k)$ (Murray and Adams 2010) is applied.

5 MODELING & PREDICTION RESULTS

In this section the modeling and prediction results given by the GP prior are compared with the currently applied MoE Norris-Landzberg model, also called modified Coffin-Manson (NIST 2010), where the expectation of (2) can be modelled by

$$E[f_k|\boldsymbol{x}_i, \boldsymbol{\theta}_k] = \log A_k + \alpha_k \log t_{rep,i} + \beta_k \log \Delta T_i + \frac{\Delta H_k}{BT_{peak,i}} \qquad k = 1,2, \tag{4}$$

with $\boldsymbol{\theta}_k = (\log A_k, \alpha_k, \beta_k, \Delta H_k)$. Thereby, ΔH and B denote the activation energy and the Bolzmann constant, respectively. For this model prior information is available (Plankensteiner 2011), which is included via a Normal-Inverse-Gamma prior (Frühwirth-Schnatter 2006). To provide a direct comparison between the two models, the mixture components are modeled independently from each other and then they are mixed by the estimated mixture weight η_i, where η_i is modeled by a cumulative Beta distribution function (Plankensteiner 2011). To evaluate the interpolation and extrapolation quality, only a subset of device type A (data of 20 test conditions) is used to train the model. The remaining data is used for model and prediction validation.

For the Goodness-of-Fit (GoF) evaluation, the BIC for each mixture component and model assumption is determined and listed in Table 3. Based on the BIC values, the GP model provides the better fit for both mixture components. Using the BIC approximation for the log marginal likelihood (Held 2008), additionally the Bayes Factor (BF) (Held 2008) can be listed. For the first and second mixture component BFs of 1.9 and 10.6 are achieved, respectively. Based on the fact that a BF > 3 indicates a significant improvement in model quality, the improvement by the GP model is barely worth mentioning for the first mixture component, but substantial for the second one. Therewith, it is hypothesized that the GP prior leads to a significant improvement in prediction quality.

Table 3: BIC and BF values for both models.

GoF	MoE	GP	BF
BIC_1	179.02	**177.7**	1.9
BIC_2	360.62	**355.9**	10.6

To evaluate the prediction quality, the Sum of Squared Errors of Prediction (SSEP) for each test setting is determined. Since this evaluation is only of importance for failed data, data sets with no fails are excluded at this point. With this the amount of available data sets reduces to 51. For the training set, leave-one-out

Cross-Validation (CV) (James et al. 2013) is performed. Then the whole training set is used to develop the lifetime model, which is used for prediction of the validation data sets. The means of SSEPs are listed in Table 4. For device types A, B, C, and E the results are almost identical, whatever model is taken. With respect to device type D, the MoE model gains slightly better results. On average, a mean SSEP of 1.95 for the MoE and 2.08 for the GP model are achieved.

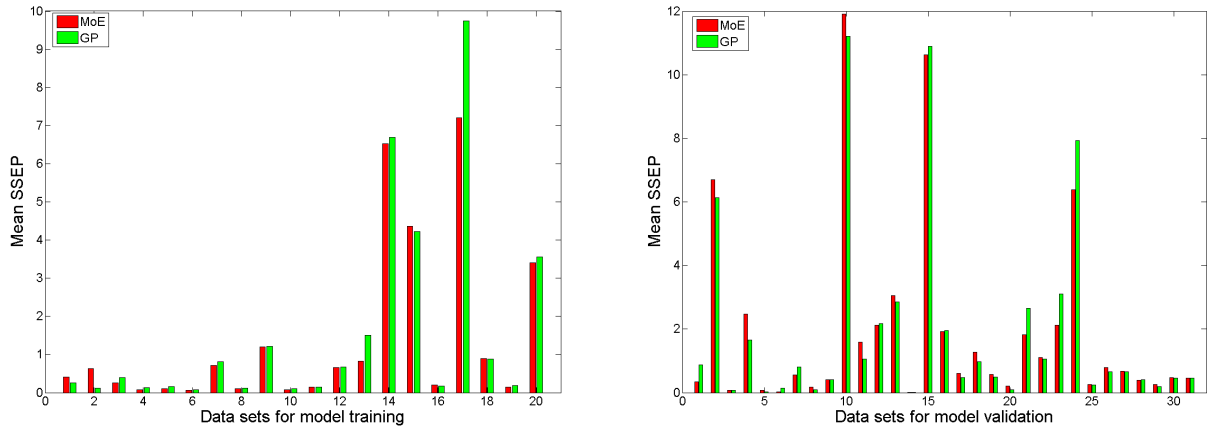
Table 4: Means of SSEP for device types A, B, C, D, and E.

Mean SSEP	CV A	A	B	C	D	E
MoE	**1.40**	2.27	**4.19**	**0.89**	**2.47**	0.50
GP	1.56	**2.12**	4.28	0.94	3.08	**0.47**

For a detailed evaluation of SSEPs, Figure 5 shows the SSEP for each data set and model. Precisely, Figure 5(a) shows the results of the CV and Figure 5(b) illustrates the evaluation of the validation sets. In total, both models lead to comparable results, regardless which device type is considered. All data sets with high SSEPs for the GP model, e.g. data set 17 in Figure 5(a) or data set 10 in Figure 5(b), also show high SSEPs for the MoE model. The high SSEPs indicate that either

(1) the model for the mean lifetime of the mixture component or/and
(2) the model for the mixture weight

lacks accuracy and thus, data is predicted insufficiently. Since the effect of changing test settings on the mixture weights is well understood (Plankensteiner 2011), the main focus lies on an accurate prediction for the mean lifetime of each mixture component.



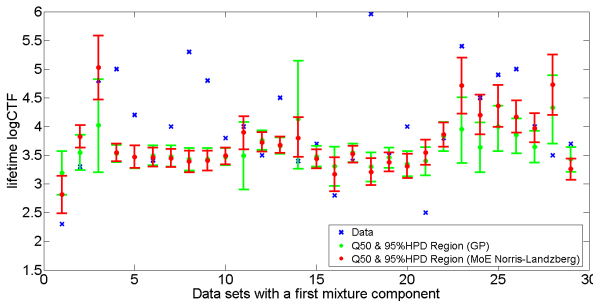
(a) Means of SSEP evaluation for the training sets. Sets 1 - 20 contain data from device type A.

(b) Means of SSEP evaluation for the validation sets. Sets 1 - 13, 14 - 16, 17 - 21, 22 - 25, and 26 - 31 contain data from device type A, B, C, D, and E, respectively.

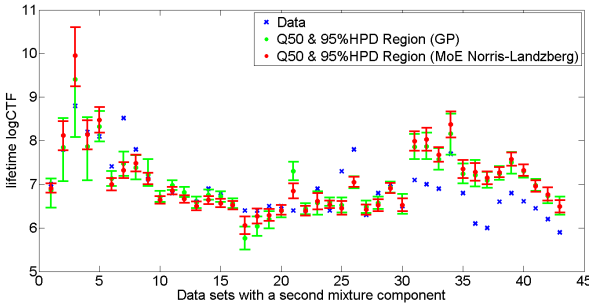
Figure 5: Means of SSEP for training and validation data sets.

To evaluate the prediction quality, the observed mean lifetime is compared with the corresponding 95% HPD regions (Robert 2001), illustrated in Figure 6. The broad 95% HPD regions for the first mixture component are thereby eye-catching, shown in Figure 6(a). This indicates a high uncertainty in both models. Moreover, five data sets (1, 12, 16, 21, and 28) lie below the predicted 95% HPD regions of both models. It is suspicious that even 3 of them (1, 12, and 16) belong to device type A, although data from this device type has been used for model development and thus, this data should give accurate results. Even more

surprising is the fact, that almost the half of the observed data lie above the predicted 95% HPD regions, where again a lot of data sets of device type A are included. The widths of the intervals as well as the lacking accuracy leads to the conclusion that the first mixture component is itself a mixture of different effects which cannot be captured by a single model representing only one physical failure mechanism. It might be more appropriate to use a mixture of overlapping distributions characterizing different failure mechanisms. These results also show that at this high stress level, close to the destruction, different mechanisms start to interact and make an accurate lifetime predict extremely challenging.



(a) Comparison of empirical and predicted data using GP and MoE model for the first mixture component. Sets 1 - 20, 21 - 22, 23 - 24, 25 - 28, and 29 contain data from device type A, B, C, D, and E, respectively. Sets 3 - 14 have been used for model development.



(b) Comparison of empirical and predicted data using GP and MoE model for the second mixture component. Sets 1 - 28, 29 - 30, 31 - 35, 36 - 38, and 39 - 43 contain data from device type A, B, C, D, and E, respectively. Sets 6 - 24 have been used for model development.

Figure 6: Observed data and predicted 95% HPD regions for GP & MoE model.

With respect to the second mixture component, visualized in Figure 6(b), the relationship between the empirical and predicted data shows an evident trend. For device type A, only the lifetime of data set 21 is on average significantly shorter than predicted by both models. It turns out that this data set has a high T_{case} with a long t_p - a fatal combination of test settings leading to a significantly shorter lifetime on average. For data sets 7, 25 and 26 the mean lifetime is underestimated by both models. For all other data sets of device A and all data sets of device B the predicted 95% HPD regions overlaps the empirical data, which indicates a good prediction quality. For device types C, D and E the predicted lifetime is for all data sets, except data set 34, significantly higher than observed. One reason for this inaccuracy might be the fact that the shape of the active DMOS area differs between device types A, B and device types C, D, and E. For the first two device types, the active DMOS area is L-shaped, whereas the active DMOS area for the latter ones are rectangle-shaped. With this, it can be concluded that the information about the active DMOS area needs to be considered as well for an accurate lifetime prediction, especially in case of extrapolation to other device designs.

Summarizing the results, the MoE Norris-Landzberg and GP models are comparable and show the same level of accuracy regarding the predicted 95% HPD regions and the SSEPs.

6 SUMMARY & OUTLOOK

In this paper a GP prior is used to model complex lifetime data from five different device types. The data follows a mixture of two log-normally distributed heteroscedastic components and contain censored observations.

For model development MCMC simulations are exploited. For this purpose, the statistical toolbox *GPstuff* in MATLAB® has been extended. Based on the BIC evaluation, the GP model indicates a promising improvement to the currently applied MoE Norris-Landzberg model.

To evaluate the prediction quality, posterior predictive distributions for the validation sets were determined. Additionally a CV for each training set was performed. Although the BIC of the GP model

indicated better results than applying the MoE model, SSEPs and 95% HPD regions show that the models are comparable and that the GP model does not provide a significant improvement in prediction quality. Thus, the GP model serves as a reliable alternative to currently applied methods.

Summarizing the research work of the last years, the model and prediction accuracy gained by modeling smart power semiconductor lifetime data with MoE, BN and GP priors fall below expectations. Thus, it is assumed that, in general, regression based models can capture only a part of the complex lifetime behavior of the observed data. Consequently, it is unavoidable to consider alternative hypotheses:

- Only modeling the mean lifetime is insufficient because the observed data indicates test setting dependent and DUT specific variances. Moreover, the mixture weights also differ for the investigated device types. Thus, including a model for the data variances and improving the model for the mixture weights is indispensable.
- The currently used input parameters cannot capture the whole lifetime behavior, meaning that a better representation of the geometric design (e.g. size/ shape of the active DMOS area), electro-thermal and thermo-mechanical effects is needed. To include these substantial parameters into the lifetime model, improvements in simulation models as well as used material properties are obligatory.
- The lifetime cannot be explained by a monotonic and continuous function. There are different, superposed and interacting degradation effects, that make a regression tenuous. More realistic approaches like modeling the internal damage and degradation with (interacting) stochastic processes need to be considered, e.g. change point for failure mechanisms. It is assumed, that this leads to more accurate results than modeling the observed lifetime data directly (Meeker, Escobar, and Lu 1998).
- What if the investigated data suffers from inconsistencies that are not evident? Consequently, the model and prediction inaccuracy can not be given by used methods' properties, but depends on devices and observed data. This might call the existence of an appropriate and sufficiently precise model, e.g. for extrapolation tasks, into question. Especially this consideration should be kept in mind and requires a detailed investigation in the future.

ACKNOWLEDGMENTS

The authors would like to thank Roland Sleik and Michael Ebner for the measurement support, Robert Illing and Mirko Bernardoni for the temperature simulations, Balamurugan Karunamurthy and Sebastian Eiser for the support on simulation topics and thermo-mechanical analytics as well as Michael Glavanovics, Michael Nelhiebel, Walther Heinz and Christoph Schreiber for valuable discussions.

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

Bluder, O. 2011. *Prediction of Smart Power Device Lifetime based on Bayesian Modeling*. Ph. D. thesis, Alpen-Adria-Universität of Klagenfurt.

Bluder, O., M. Glavanovics, J. Pilz, and K. Plankensteiner. 2012. "A Bayesian Mixture Coffin-Manson Approach to Predict Semiconductor Lifetime". In *Proceedings of Stochastic Modeling Techniques and Data Analysis*, 45–52.

Chen, W., and C. Neson. 1979. "Thermal Stress in Bonded Joints". *IBM Journal of Research and Development* 23(2):179–188.

Frühwirth-Schnatter, S. 2006. *Finite Mixture and Markov Switching Models*. 1st ed. Berlin: Springer.

Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin. 2014. *Bayesian Data Analysis*. 3rd ed. Boca Raton, USA: Chapman & Hall/ CRC Press.

Glavanovics, M., H. Köck, V. Kosel, and T. Smorodin. 2007. "A New Cycle Test System Emulating Inductive Switching Waveforms". In *Proceedings of the 12th European Conference on Power Electronics and Applications 2007*, 1–9.

Glavanovics, M., and H. Zitta. 2001. "Thermal Destruction Testing: An Indirect Approach to a Simple Dynamic Thermal Model of Smart Power Switches". In *Proceedings of the 27th European Solid-State Circuit Conference*, 221 – 224.

Held, L. 2008. *Methoden der statistischen Inferenz: Likelihood und Bayes*. Heidelberg: Springer.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning with Application in R*. New York: Springer.

Kosel, K., R. Sleik, and M. Glavanovics. 2007. "Transient Non-linear Thermal FEM Simulation of Smart Power Switch and Verification by Measurements". In *Proceedings of Thermal Investigations of ICs and Systems*, 110–114.

Kravchenko, G., B. Karunamurthy, M. Nelhiebel, and H. Pettermann. 2014. "FEM Study of Fatigue Crack Growth in a Power Semiconductor Chip Subjected to Transient Thermomechanical Loading". In *20th European Conference on Fracture, submitted*.

Meeker, W., L. Escobar, and C. Lu. 1998. "Accelerated Degradation Tests: Modeling and Analysis". *Technometrics* 40(2):89–99.

Murray, I., and R. Adams. 2010. "Slice Sampling Covariance Hyperparameters of Latent Gaussian Models". In *Proceedings of Advances in Neural Information Processing Systems 23*, 1732–1741.

Murray, I., R. Adams, and D. MacKay. 2010. "Elliptical Slice Sampling". In *The Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, 541–548.

Nelhiebel, M., R. Illing, C. Schreiber, A. 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, and S. Fraissè. 2011. "A Reliable Technology Concept for Active Power Cycling to Extreme Temperatures". In *Proceedings of the 28th European Solid-State Circuits Conference*, 295–298.

NIST 2010, July. "NIST Handbook". http://www.itl.nist.gov/div898/handbook/.

Plankensteiner, K. 2011. "Application of Bayesian Models to Predict Smart Power Switch Lifetime". Master's thesis, Alpen-Adria-Universität of Klagenfurt.

Plankensteiner, K., O. Bluder, and J. Pilz. 2013. "Application of Bayesian Networks to Predict Smart Power Semiconductor Lifetime". In *Proceedings of the 9th Conference on Ph. D. Research in Microelectronics and Electronics*, 281–284.

Riccio, M., A. Irace, G. Breglio, S. Spirito, V. Kosel, M. Glavanovics, and A. Satka. 2009. "Thermal Simulation and Ultrafast IR Temperature Mapping of a Smart Power Switch for Automotive Applications". In *Proceeding of 21st International Symposium on Power Semiconductor Devices & IC's*, 200–203.

Robert, C. P. 2001. *The Bayesian choice*. 2nd ed. New York: Springer.

Shi, J., and T. Choi. 2011. *Gaussian Process Regression Analysis for Functional Data*. Boca Raton, USA: Chapman & Hall/CRC Press.

Suresh, S. 1998. *Fatigue of Materials*. Cambridge Solid State Science Series, Cambridge University Press.

Vanhatalo, J., J. Riihimäki, J. Hartikainen, P. Jylänki, V. Tolvanen, and A. Vehtari. 2013, April. "GPstuff: Bayesian Modeling with Gaussian Processes". *Journal of Machine Learning Research* 14:1175–1179.

AUTHOR BIOGRAPHIES

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.

OLIVIA BLUDER is a Junior Researcher at KAI- Kompetenzzentrum für Automobil- und Industrieelektronik GmbH, Austria. She received her M.S and Ph.D. 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-ail.at.

JÜRGEN PILZ is a full Professor and the Head of the Department of Statistics at the Alpen-Adria-Universität of Klagenfurt, Austria. He holds M.S. and Ph.D. degrees in Mathematical Statistics at the Technical University Bergakademie Freiberg, Germany. His research interests are among others Bayesian modeling, computer science, spatial statistics and design of experiments. His email address is Juergen.Pilz@uni-klu.ac.at.