Influence of Reducing the Load Level of Mission Profiles on the Remaining Useful Life of a TO220 Analyzed with a Surrogate Model

Tobias Daniel Horn[1], Jan Albrecht[2] and Sven Rzepka[3]

[1]*Fraunhofer ENAS, Technologie Campus 3, 09126 Chemnitz*
tobias.horn@enas.fraunhofer.de

[2,3]*Fraunhofer ENAS, Technologie Campus 3, 09126 Chemnitz and*
Chemnitz University of Thechnology, Center for Micro and Nano Technologies, Reichenhainer Str. 70, 09126 Chemnitz
jan.albrecht@enas.fraunhofer.de
sven.rzepka@enas.fraunhofer.de

ABSTRACT

A methodology for replacing finite element simulations with a fast-calculating surrogate model for fault tolerance in operating systems is presented. The study focuses on the TO220 rectifier system and explores methods to detect impending failures and calculate the resulting necessary load reduction. The finite element simulation model is described, highlighting the die attach as the relevant connection for failure. A surrogate model is developed using long-short-term-memory models to predict temperature and in-elastic strain. The surrogate model significantly reduces simulation time, allowing for the adjustment of load based on the system's current state of health. The rainflow counting algorithm is applied to calculate the number of cycles to failure, and the Palmgren-Miner linear damage accumulation relation is used to determine the damage and state-of-health. The dependency of the change in lifetime due to variations in scaling factor is evaluated and the results show that load reduction increases the lifetime of the system.

1. INTRODUCTION

The increased requirements for fault tolerance (e. g. for SAE level L3 and onward, defined by the Society of Automotive Engineers (SAE) in SAE International (2021)) requires, the operating system must continue to operate with reduced power until other measures are initiated. Therefore, the system must be able to detect the impending failure and start the fault handling. Furthermore, the result of the intervention must be predicted in order to apply right failure rectification. These requirements can be met by various methods, as mentioned by Moeller, Inamdar, van Driel, Bredberg, Hille, Knoll and Vandevelde (2024). For example, a system that regularly undergoes rest phases can run self-diagnoses processes by using standard load cycles during these rest phases. From the deviation of the resulting response to the response in the undamaged state, the damage and the resulting necessary reduction in load can be calculated (as shown in e.g. Chacko, Moeller, Kolas, Albrecht, and Rzepka (2024)). However, the disadvantage of this method is that the fault can be detected at the earliest in the first rest phase after the first measurable deviations have occurred. On the other hand, the advantage is that the calculation must not be carried out in the system itself, but can also be performed in the cloud, for example.

Alternatively, a digital twin of the system can be created and this representation can be digitally loaded in parallel with the real system. The digital twin then calculates the damage to the real system based on the real load. In order to achieve this, the digital twin must be capable of mapping the failure mechanism that occurs and calculating the damage from this. In addition, the calculation of the damage due to the load in the digital twin must be performed faster than the load is applied in reality (this depends on the available calculation resources). Only if both of these conditions are met the current state of health of the system can be mapped correctly and an appropriate regulation can be calculated.

In this work, the chosen system is a TO220 rectifier. The TO220 is a Silicon Carbide Schottky diode for ultra-high performance, low loss, high efficiency power conversion applications. For example, it can be used as a switched-mode power supply, AC-DC and DC-DC converter, in battery charging infrastructure, server and telecommunications power supply, uninterruptible power supply and as a photovoltaic inverter (Nexperia 2023). As already shown by Albrecht, Horn, Habenicht and Rzepka (2023), it is possible to generate a validated digital representation of this rectifier in the form of a combined multi-field FE simulation, from which the damage under real loads can be calculated. However, the calculation time of these FE simulations is much too long. This paper shows how the FE simulation model can be replaced by a fast-calculating surrogate model. After a brief introduction of the TO220 rectifier and the corresponding FE simulation in section 2, the training and setup of the surrogate model and the calculation of the state of health are presented. Subsequently, this surrogate model is used to calculate the change in lifetime due to the reduction of the load.

2. Finite Element Simulation

The structure of the TO220 rectifier can be seen in the FE simulation model in Figure 1 and Figure 2. The die attach is the relevant connection regarding the failure (as shown in Albrecht et al. (2023)), representing the connection between the chip and the lead frame. The current flows from the contact via the bond wire to the die and is then transferred to the lead frame via the die attach. The materials heat up due to the current flow (Joule heating). The temperature is dissipated via the heatsink.

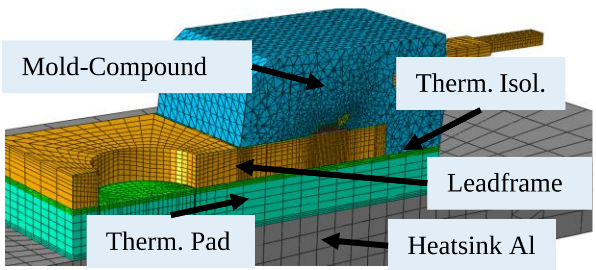

Figure 1: Finite Element model of the TO220 rectifier (cut view).

The FE simulation is based on a sequential approach, where the electric-thermal behavior is simulated first, followed by the thermal-mechanical behavior of the component. A current load profile is used as input for the electric-thermal simulation. The computed temperature field is then used as input for the thermal-mechanical simulation. From the thermal-mechanical simulation the in-elastic strain in the region of the die-attach corners (the relevant area for the failure) is extracted using an averaging approach.

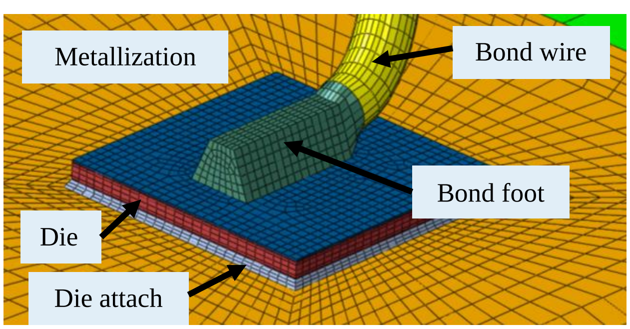

Figure 2: Finite Element model showing the die attach and the bond foot.

In order to calculate the state-of-health from the simulation result, the Coffin-Manson lifetime model

$$N_f = C_1 \Delta\varepsilon_{pl}^{C_2} \tag{1}$$

is used. With this model the number of cycles to failure N_f is calculated by using the in-elastic amplitude allocated to the cycle $\Delta\varepsilon_{pl}$ as well as two model parameters C_1 and C_2. In this calculation, the parameters identified by Darveaux and Banerji (1991) for Pb95Sn5 were used. Since real loads are used in this calculation rather than standard cycles, rainflow counting is carried out based on the temperature profile. The rainflow counting extracts cycles from the real load, and the corresponding change in the in-elastic strain is assigned to these cycles. From the number of cycles to failure N_f the state-of-health can be calculated (as shown in section 4). The methodology of calculating the state-of-health by using FE simulations is also shown in Figure 3.

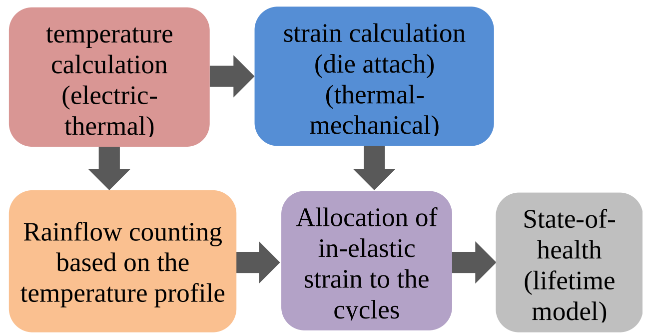

Figure 3: Applied methodology for the calculation of the state-of-health.

A crucial aspect of the FE simulation is the accurate description of the materials used in the component. Most of the materials exhibit strong temperature dependence, the bond wire and bond foot (both aluminum) as well as the lead frame (copper) are modeled by using linear elastic and bilinear kinematic hardening plasticity behavior and the solder in the die attach is modeled by using the Anand law. Especially the solder is highly non-linear and the material behavior depends strongly on the strain experienced in the past. Further details on the materials and the FE simulation in general are shown in Albrecht et al. (2023).

Using the calibrated model, a complete Worldwide harmonized Light vehicles Test Procedure (WLTP) mission profile was simulated by varying the electrical load over time. In order to obtain the current from the WLTP profile, an inverter module was added before the simulation. The temperature field from the electric-thermal simulation was used as input for the thermal-mechanical simulation and the stress as well as the strain in the die-attach corners is calculated and averaged. The results are shown in Figure 4.

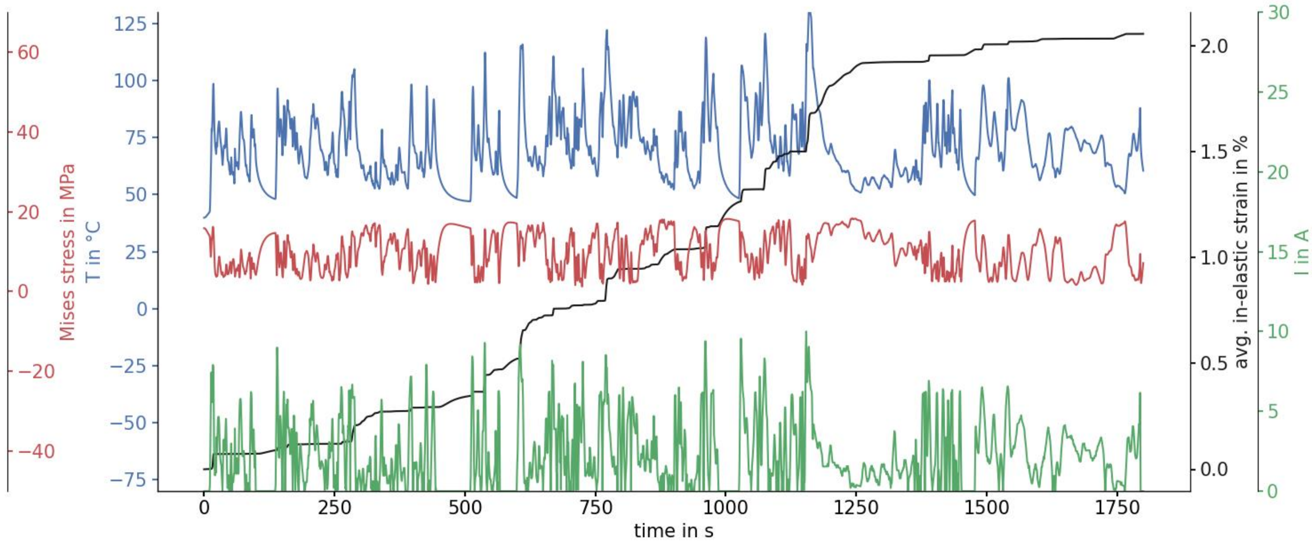

Figure 4: WLTP (current) mission profile (green) as well as some of the calculated results from the FE simulation: temperature (blue), von Mises stress (red) and the averaged total in-elastic strain (black).

3. Surrogate Model

As mentioned before, the WLTP cycle can be simulated by using FE simulation and the results fit to experimental measurements. So, the FE simulation model is a digital representation of the TO220 rectifier. However, the simulation time for calculation the WLTP cycle of 1800s is round about two days. In order to reduce the simulation time, the finite element simulation must be replaced by a surrogate model.

Therefore, the restrictions are: The surrogate model must take the current as input and the temperature as well as the in-elastic strain as outputs. Additionally, the surrogate model must be able to store all information of the reality – and because the FE simulation fits to the reality also all information of the FE simulation model. Due to this, the type of the surrogate model cannot be chosen randomly.

As described before, the materials of the TO220 rectifier are highly non-linear and also strongly dependent to the history. Due to this, as model type the long-short-term-memory (LSTM) model is used (Hochreiter & Schmidhuber (1997)). LSTMs are effective in capturing long-range dependencies in sequential data and have the ability to remember information over long periods of time, as for example shown in Zheng, Ristovski, Farahat and Gupta (2017). Analogous to the FE simulation, two different LSTM models were trained: one model to predict the temperature and one model to predict the in-elastic strain. The training data were produced by the FE simulation model and as the LSTM model is to be applied to real loads, the simulation data from the WLTP cycle is used for training. The training/validation split is 80/20 % and the Adam optimizer (Kingma & Ba 2014) is used. In order to generate additional data that the model had not seen in training, the mission profile was varied using different methods and a total of six variations were calculated using FE simulation. The shown mean absolute error / mean absolute percentage error (MAE/MAPE) is calculated on all seven mission profiles. The models are trained without considering the time. So, for the data a constant time step of 10 Hz is used.

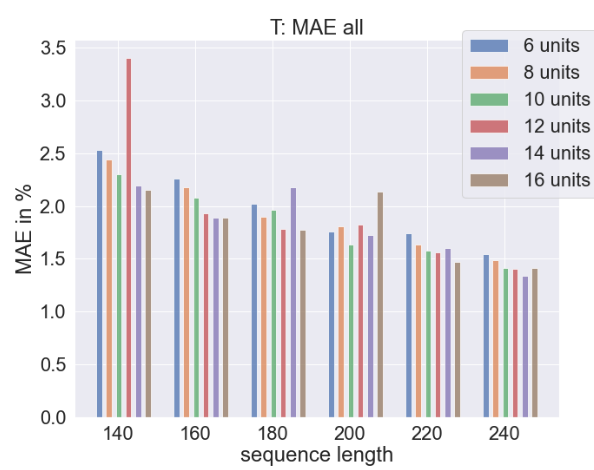

Figure 5: MAPE for variation of the sequence length and the units for the LSTM predicting the temperature.

For the LSTM there are many parameters, such as sequence length, number of unit layers, number of units per layer, features, predictions, learning rate etc. These parameters are optimized by a combination of a variation study and a hyperparameter variation. Exemplarily for the LSTM predicting the temperature, which only uses one unit layer, in Figure 5 the MAPE for the variation of the sequence length and the number of units in the unit layer is shown. The increase of the sequence length (the history taken into account) significantly increases the prediction quality. Simultaneously, increasing the sequence length reduces the difference between the models with different units.

For predicting the temperature, the LSTM model with the current as feature, a sequence length of 240, one unit layer with 14 units is chosen. The prediction quality (also for all seven mission profiles) is shown in Figure 6.

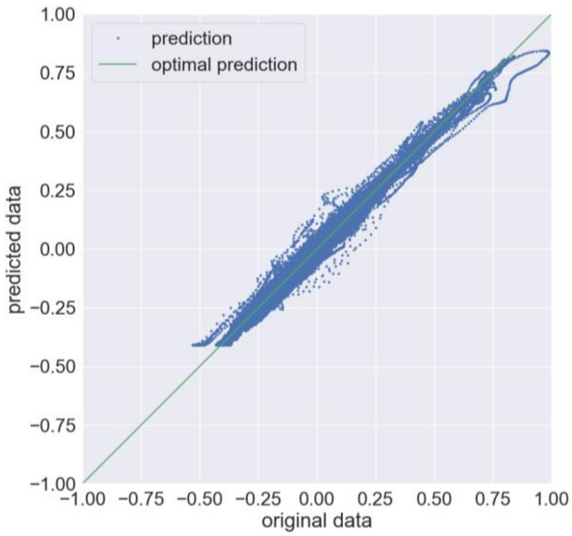

Figure 6: Prediction quality of the LSTM for temperature prediction. The green line indicates the optimal prediction.

As mentioned before, for the prediction of the in-elastic strain a second LSTM model is trained. Therefore, the features are the current together with its first and second derivative and also the temperature (predicted by the other LSTM) together with its first derivative. The result of the parameter variation and hyperparameter variation gives a model with two unit layers with 16 units and 10 units respectively. The sequence length is identified to 200.

The in-elastic strain is a continuously increasing quantity where the changes are constantly added up. Therefore, not the in-elastic strain itself was predicted, but the incremental in-elastic strain. In post-processing after the prediction itself, the in-elastic strain is calculated from the incremental in-elastic strain via integration. Due to this in Figure 7, where the predictions quality for the (total) in-elastic strain is shown for the seven mission profiles, seven connected lines of prediction points are visible.

The application of integration also means the integration of errors. This has both advantages and disadvantages. Assume that only one error occurs at a specific point in time. Then, from this data point onwards, a deviation can be seen in all subsequent data points in Figure 7. This deviation will also be included in the MAPE calculation. On the other hand, a further, opposing error can cancel out the original error. Nevertheless, the prediction is sufficiently accurate, as also can be seen in Figure 8.

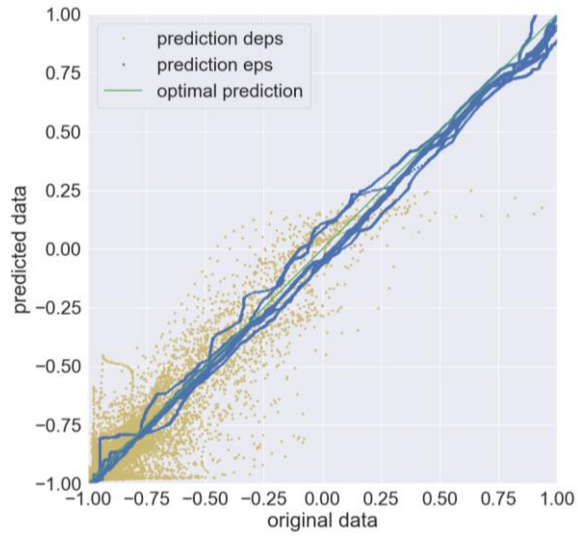

Figure 7: Prediction quality of the LSTM for in-elastic strain prediction. The yellow points indicate the prediction quality for the incremental strain, the blue for total strain.

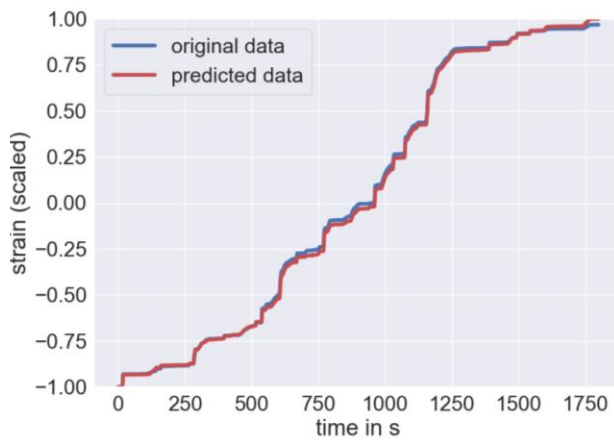

Figure 8: Comparison of the result of the FE simulation (blue) and the LSTM model prediction (red) for the WLTP cycle. The strain is scaled to (-1, 1).

So, with this surrogate model a reduction of the calculation time from two days to 19 seconds (7 seconds for the prediction of the temperature and 12 seconds for the prediction of the strain) for the WLTP cycle of 1800 seconds is achieved.

4. State-of-Health and Remaining Useful Life

As mentioned before for the FE simulations, the rainflow counting algorithm is applied in order to transfer the load profile and the resulting continuously increasing in-elastic strain into separate cycles. By using Coffin-Manson lifetime model (equation (1)), which describes the shape of the strain

Wöhler curve in the low cycle fatigue range, the number of cycles to failure $N_{f,i}$ for each cycle i is calculated. The inelastic strain $\Delta\varepsilon_{cr,i}$ for each sub-cycle per cycle (closed cycles within a larger cycle) is subtracted. From the number of cycles to failure the damage per cycle

$$D_i = 1/N_{f,i} \tag{2}$$

is calculated. Then Palmgren-Miner linear damage accumulation relation (proposed by Palmgren (1924) and further developed by Miner (1945)), is used to sum up all damage contributions

$$D = \sum_i D_i = \sum_i \frac{1}{N_{f,i}} \tag{3}$$

From this, the damage can be transferred into the state-of-health

$$SoH = 1 - D \tag{4}$$

Subsequently, the state of health was calculated for a series of WLTP cycles, whereby initially the WLTP cycles were not changed for the entire period. With this a lifetime of 143 days is calculated. In addition, when a health status of 50% was reached, the load level of the WLTP cycle was reduced. This is used to simulate a reduction in power in response to the damage reached. As shown in Figure 9, this reduction in power significantly increases the lifetime.

Consequently, this model can be used to adjust the load of the TO220 rectifier to the current state of health. Due to the short calculation time of the surrogate model, the influence of the load reduction on the lifetime can be predicted. This allows to adjust the load in a targeted manner, which is necessary for a control.

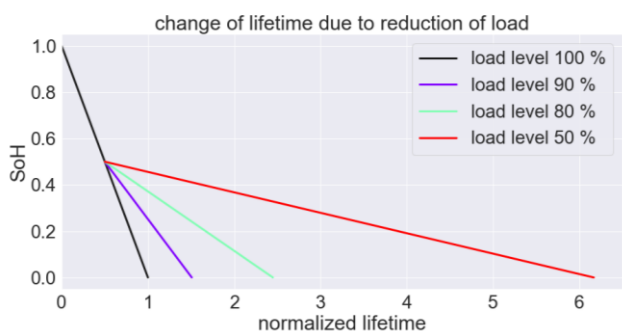

Figure 9: Change of the lifetime due to the reduction of the load level of the WLTP cycle.

5. Conclusion

In this work the methodology of replacing the FE simulation model by a surrogate model is shown. The amount of calculation time is massively reduced and due to this the surrogate model will be implemented on a micro controller in order to finalize the digital twin.

For the prediction of the temperature the final trained model is a LSTM model with just one unit layer and six units therein. Due to this low complexity, in future work a change to a less complicated model type will be taken under consideration.

Additionally, the surrogate model is currently being trained with a complete WLTP cycle. This has the disadvantage that the generation of the training data requires a relatively large amount of resources. However, it can be assumed that it contains multiple pieces of information that are not necessarily required for training the surrogate model. For this reason, the training of the surrogate model is to be simplified in future work. The WLTP cycle and other realistic load profiles will be analyzed using methods from time series analysis (e.g., the matrix profile, what is presented in Imani, Madrid, Ding, Crouter, and Keogh (2018) or Mercer, Alaee, Abdoli, Singh, Murillo, and Keogh (2021)), the relevant patterns and anomalies will be determined and calculated as separate profiles, weighted and specified in the training. This reduces the effort required to generate the training data.

Acknowledgement

This work was partially supported by HiEFFICIENT project. This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement no. 101007281. The JU receives support from the European Union's Horizon 2020 research and innovation program and Austria, Germany, Slovenia, Netherlands, Belgium, Slovakia, France, Italy, and Turkey. Further, it was partially supported by Trust-E project. This project has received funding from PENTA - EUROPIDES[2] and BMBF (16ME0324).

Furtheron, Wolfgang Wondrak and Leonhard Hertenstein from Mercedes Benz AG are highly acknowledged for providing the current profile of the WLTP driven cycle.

References

Albrecht, J., Horn, T., Habenicht, S., & Rzepka, S. (2023, December). Mission Profile related Design for Reliability for Power Electronics based on Finite Element Simulation. In *2023 IEEE 25th Electronics Packaging Technology Conference (EPTC)* (pp. 722-726). IEEE. doi: 10.1109/EPTC59621.2023.10457766.

Chacko, J., Moeller, H., Kolas, K. A., Albrecht, J., & Rzepka, S. (2024, April). Investigation of an approach for the determination of the current State-of-Health and improvement of service life in power modules. In *2024 25th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)* (pp. 1-9). IEEE. *doi: 10.1109/EuroSimE60745.2024.10491510*

Darveaux, R., & Banerji, K. (1991, May). Fatigue analysis of flip chip assemblies using thermal stress simulations and a Coffin-Manson relation. In *1991 Proceedings 41st*

Electronic Components & Technology Conference (pp. 797-805). IEEE. doi: 10.1109/ECTC.1991.163971

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780, doi: 10.1162/neco.1997.9.8.1735

Imani, S., Madrid, F., Ding, W., Crouter, S., & Keogh, E. (2018). Matrix profile XIII: Time series snippets: A new primitive for time series data mining. In *2018 IEEE international conference on big knowledge (ICBK)* (pp. 382-389). IEEE. doi: 10.1109/ICBK.2018.00058

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Mercer, R., Alaee, S., Abdoli, A., Singh, S., Murillo, A., & Keogh, E. (2021). Matrix profile xxiii: Contrast profile: A novel time series primitive that allows real world classification. In *2021 IEEE International Conference on Data Mining (ICDM)* (pp. 1240-1245). IEEE. doi: 10.1109/ICDM51629.2021.00151

Miner, M. A. (1945). Cumulative Damage In Fatigue. *J. Appl. Mech.*, 12(3), A159-A164 Doi: https://doi.org/10.1115/1.4009458

Moeller, H., Inamdar, A., van Driel, W.D., Bredberg, J., Hille, P., Knoll, H. & Vandevelde, B. (2024). Digital Twin Technology in Electronics. In van Driel, W.D., Pressel, K. & Soytürk, M. (Eds.), *Recent Advances in the Reliability Assessment of Electronic Devices* (287-328). Location: Publisher. *(in press)*

Nexperia (2023). PSC1065 650 V, 10A SiC Schottky diode in TO-220-2 R2P. https://www.nexperia.com/products/diodes/silicon-carbide-sic-schottky-diodes/PSC1065K.html

Palmgren, A. (1924). Durability of ball bearings. *ZVDI*, *68*(14), 339-341.

SAE International (2021) *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles* (SAE J3016).

Zheng, S., Ristovski, K., Farahat, A., & Gupta, C. (2017, June). Long short-term memory network for remaining useful life estimation. In *2017 IEEE international conference on prognostics and health management (ICPHM) (pp. 88-95).* IEEE.