An Application of Feature Engineering and Machine Learning Algorithms on Condition Monitoring of SiC Converters

Loghmani Moghaddam Toussi, Afshin; Bahman, Amir Sajjad; Iannuzzo, Francesco; Blaabjerg, Frede

Published in:
Proceedings of the 2021 IEEE Energy Conversion Congress and Exposition (ECCE)

DOI (link to publication from Publisher):
10.1109/ECCE47101.2021.9595152

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

An Application of Feature Engineering and Machine Learning Algorithms on Condition Monitoring of SiC Converters

Afshin Loghmani Moghaddam Toussi
dept. of Energy Technology
Aalborg University
Aalborg, Denmark
almt@et.aau.dk

Amir Sajjad Bahman
dept. of Energy Technology
Aalborg University
Aalborg, Denmark
amir@et.aau.dk

Francesco Iannuzzo
dept. of Energy Technology
Aalborg University
Aalborg, Denmark
fia@et.aau.dk

Frede Blaabjerg
dept. of Energy Technology
Aalborg University
Aalborg, Denmark
fbl@et.aau.dk

***Abstract*— This paper proposes an approach for non-invasively fault detection in a wide range of power electronics applications. Simulations were used at this stage for gathering the large amount of data needed, but the conclusions can be applied as-is to experimental data sets. These data were fed to the machine-learning algorithm for training and validation. A buck converter based on silicon-carbide MOSFETs is selected as the case study. The beauty of the proposed approach is that it only uses the output terminal signals for diagnostics. Our approach starts with extracting features from the output voltage signal, then inferring criteria to evaluate and rank these features in terms of relevance (feature engineering phase). The final step in this phase is reducing the space dimension by selecting sets of meaningful features. Finally, several artificial neural network structures demonstrate the effectiveness of this approach.**

Keywords*— condition monitoring, fault detection, feature engineering, power electronic converter, Silicon Carbide (SiC), machine learning, artificial neural network, radial basis function networks (RBF networks), principal component analysis (PCA)

I. Introduction

Silicon carbide (SiC) devices have become more popular recently. This is due to their superior characteristics in comparison to Silicon (Si) devices. Higher power density, better on-state losses/blocking voltage trade-off, and higher thermal conductivity are some of those key characteristics [1], [2].

Although SiC technology is progressing fast, the reliability of such devices becomes better and better. However, until it reaches a mature stage, accurate condition monitoring and early identification of faulty components can help in a safer adoption of this technology. This study focuses more on faults that occur due to the components' degradation, such as gate threshold voltage drift occurring in MOSFET devices [3]. Moreover, the effects of gate oxide breakdown are considered, which might happen due to overcurrent or high electric field and temperature [2], [4].

Machine learning (ML) algorithms have become some of the most sought-after approaches for fault detection. Artificial Neural Networks (ANNs), Hidden Markov Models (HMMs), and Fuzzy systems are some of the most employed ML algorithms [5], [6]. However, some researchers have shown their interest in using metaheuristic approaches [7], though those methods are relatively slow. The performance of machine learning algorithms heavily depends on the datasets. Therefore, an important phase called feature engineering should be implemented, which plays a critical role in enabling and improving the dataset's accuracy.

This research aims to investigate the possibility of detecting the parametric drift of each component from the output signal alone and determining convenient health indicator sets. ML algorithms are trained to detect faults using these feature sets, see Fig. 1. We achieved this goal by developing an integrated MATLAB - PSpice tool to generate accurate simulation results for feature engineering and later on for training and testing ML algorithms.

We investigate some of the most common faults and their symptoms in power electronic converters, which occur due to degradation, including variations in the switch parameters (drain-source on-resistance ($R_{DS(on)}$), threshold voltage, junction temperature), and output capacitor's ESR. At the same time, the input voltage and load can also vary.

After data generated via simulations, we implement a scoring method to select the most appropriate features. To reduce the dimensionality and to make our artificial neural network more stable and require reducing the training data set size, we use principal component analysis (PCA) [8], [9].

Finally, we use machine learning algorithms to detect faults. We used two artificial neural network structures, Feed-forward artificial neural networks (also known as multilayer perceptrons) and radial basis function (RBF) networks, to evaluate our approach [10]–[12]. These networks are trained and tested using the health indicators obtained in the previous steps.

The remainder of this paper is organized as follows. In section II, we discuss the data acquisition methodology and the initial data processing. Section III explains feature engineering techniques, including feature extraction, the scoring criteria, and feature set selections. Section IV evaluates the effectiveness of the proposed approach by a few ML algorithms. Section V presents the results of simulations, feature extraction, feature scoring, and ML performance in fault detection. Section VI discusses the results and

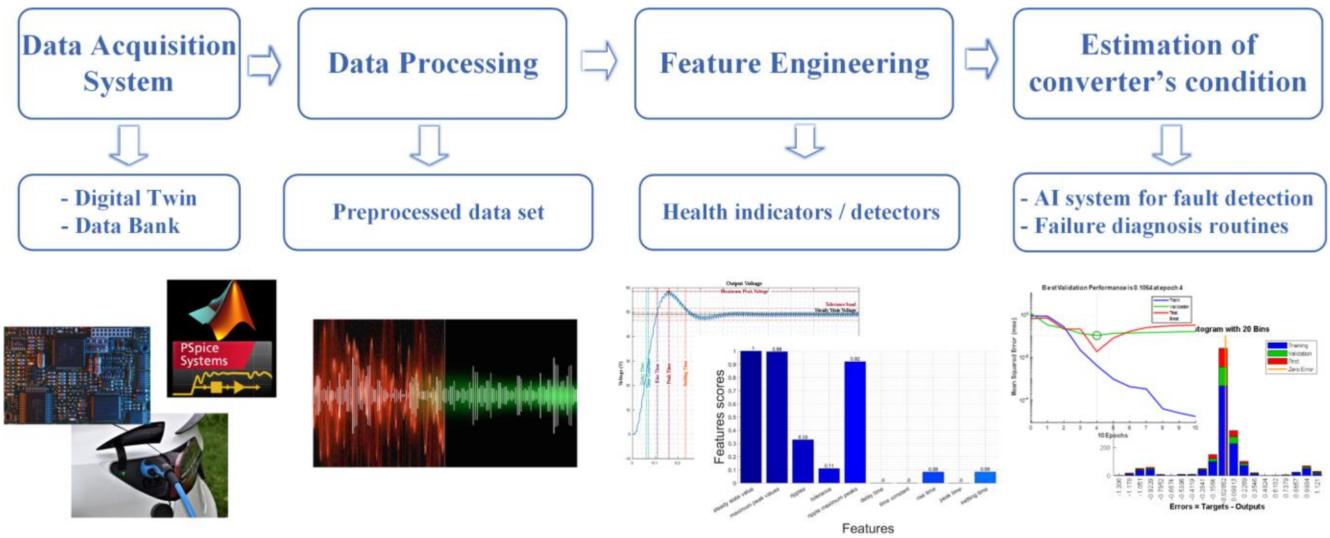

Fig. 1. Flow chart of the proposed algorithm

investigates how this approach has helped boost machine learning algorithms' accuracy and why they have failed to deliver sound results in some cases. The final section provides some concluding remarks.

II. Simulations, Data Acquisition, and Initial Data Processing

Our approach consists of four steps, as depicted in Fig 1. The first step is to gather data by simulations. A detailed model of the converter creates a digital twin of the system in PSpice to achieve higher accuracy results. MATLAB governs the variation of components' parameters and executes the simulations in Pspice to simulate the degradation of components and their effects on the system's output response. Fig. 2 shows the buck converter structure investigated in this study.

Table I shows the specifications of these components for this 400V to 48V buck converter.

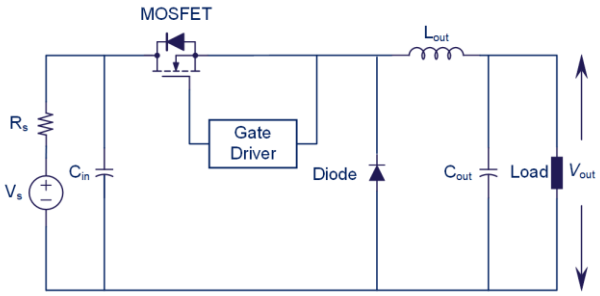

Fig. 2. The Buck converter schematic

It is also necessary to probe the simulations' data, as there is a chance of wrong results data due to convergence problems in PSpice simulations, leading to missing or erroneous data.

III. Feature Engineering

Although this article proposes a general approach that is useful for a wide variety of power electronic devices, the focus is on electric vehicles converters. In such converters, it is worth considering the transient part of output voltage response alongside their steady-state behaviors. Fig. 3 illustrates the transient and steady-state features of the described buck converter's output voltage signal.

Two critical methods in feature engineering are feature selection and feature extraction. The first one focuses on finding a subset of original data, while the latter tries to find a mapping of multidimensional space into a space of (usually) fewer dimensions.

The next step in our feature engineering approaching is assessing each parameter's correlation and sensitivity to these features. Ideally, we want to diagnose all components' health status, which is linked to their corresponding parameters simultaneously. However, some of these parameters might not have a high correlation or sensitivity, and consequently, they might not be detectable by monitoring the output signal and its features.

Moreover, considering concurrent variations in all parameters results in an exponential increase in the simulation numbers, we can avoid unnecessary simulations by executing

Table I- Key Components of The Buck Converter

Components	*Manufacturer / part number*	*Values*
MOSFET	Wolfspeed N channel SiC Power MOSFET E3M0065090D	V_{DSmax} 900 V, I_D (@ 25˚C) 35 A, $R_{DS(on)}$ 65 mΩ
Diode	Wolfspeed SiC Schottky Diode Z-Rec Rectifier C5D25170H	Blocking Voltage 1700 V, Current Rating 25 A, VF 1.5 V, IF 25 A
Output filter inductor (L_{out})	PA5141.105NL	110 μH
Output filter capacitor (C_{out})	Würth Elektronik WCAP-ATLI Aluminum Electrolytic Capacitor	15 μF, ESR = 0.85Ω, ESL, Par_res @100khz

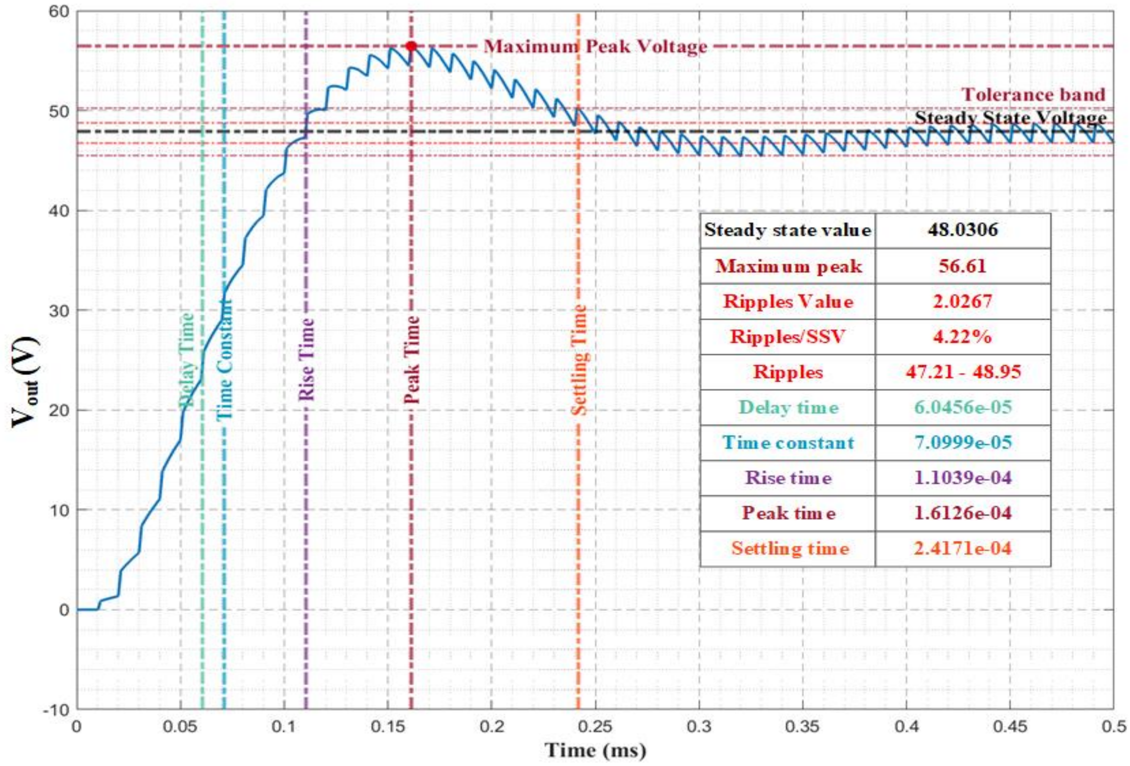

Fig. 3. Output signal and its time-domain features

this assessment for each parameter separately. This method reduces the simulation numbers from N^p simulations, for N parameters and p possibilities for each of them, to $N \times p$ simulations.

NB, multi-parameter variation simulations are done for the final step of feature engineering when principal component analyses (PCA) are used to reduce the dimensionality.

Fig. 4 shows a variation of the output signal when a parameter changes from its nominal value to an extreme value so that the component is considered faulty. Fig. 5 shows that in this case, among various time features of the output signal, the most notable variation from the initial condition occurs in settling time, then ripples, and the maximum peak of the signal. This figure might give us some inspiration about which features are better indexes for monitoring the status of this component. However, it might also be misleading as the variation of components parameters on a specific range might have a non-linear effect on the output signal. Therefore, we are going to introduce a criterion that helps us select health indicators that are useful over the desired monitoring range of parameters variations.

Three criteria are taken into account to assess the possibility of each feature becoming a health indicator. These

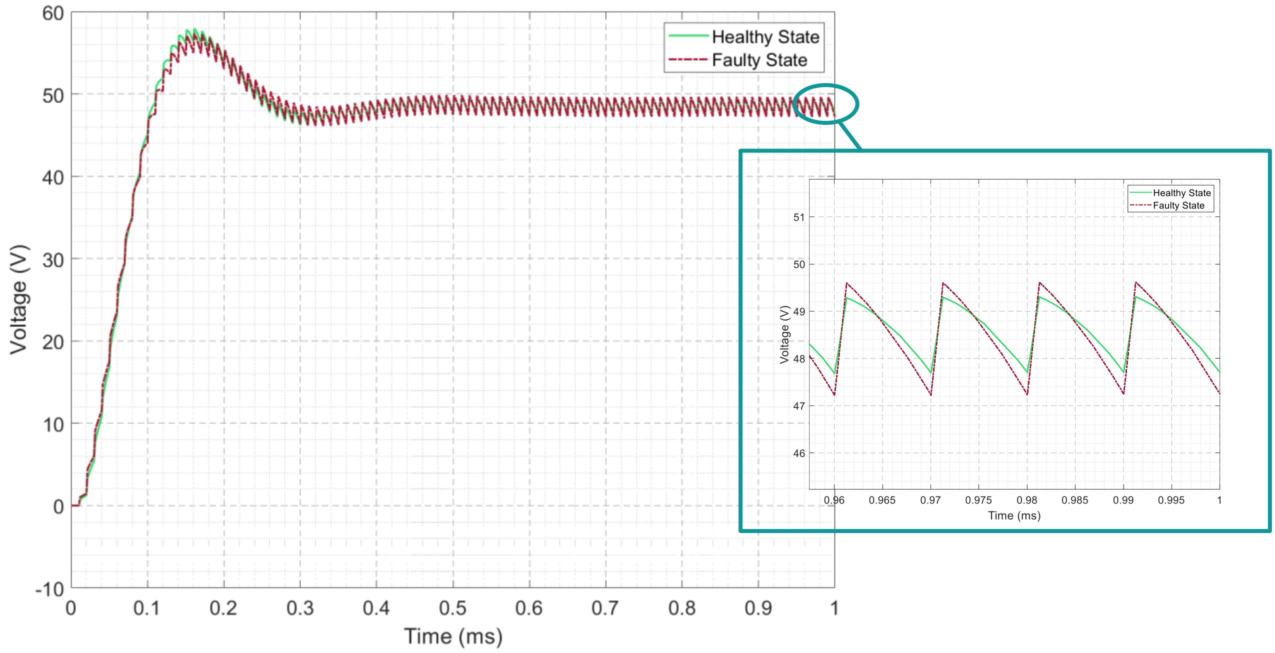

Fig. 4. Variation of the output signal by drift in parameters of components

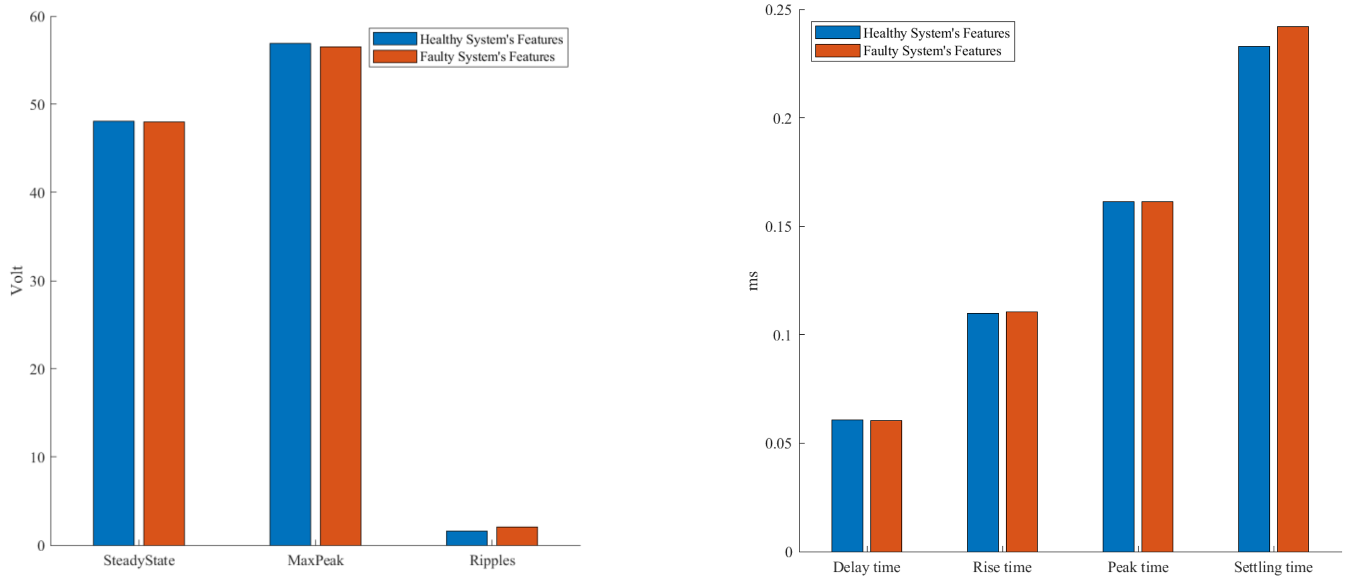

Fig. 5. Time domain features variation for 40% drift in ESR value

criteria are their sensitivity and correlation with the drifting parameter and also if the changes in these features are detectable by measuring devices, i.e., the practicality criterion:

$$score = practicality(w_1 sensitivity + w_2 correlation) \quad (1)$$

Correlation and the modified sensitivity ($\bar{S}_{\alpha}^{T}$) of a parameter (T) to a signal feature (α) are calculated by equations (2) and (3) [13], respectively.

$$corr(\alpha,T) = \frac{\operatorname{cov}(\alpha,T)}{\sigma_{\alpha}\sigma_{T}} \quad (2)$$

$$\bar{S}_{\alpha}^{T} = \frac{1}{n}\sum_{\alpha\in D}\left|\frac{\Delta T}{T}\frac{\alpha}{\Delta\alpha}\right| \quad (3)$$

Fig. 6 shows the outcome of this scoring method for the MOSFET's junction temperature and Capacitor's ESR.

In the final step of feature engineering, multi-parameter variation simulation is executed by the preceding step's remaining parameters. The acquired data from these simulations will also be used for training and testing machine learning algorithms.

PCA is the most popular instance of projection methods. Employing PCA prevents machine learning algorithms from fitting themselves to features that are not deterministic in detecting system status. Also, it decreases the number of simulations required for training them. Obtaining the feature vector is according to the following five steps [9]:

0. Acquire data: the data collected from the simulations
1. Standardize the data: a simple way is to subtract the mean
2. Compute the covariance matrix
3. Extract the eigenvectors and eigenvalues of the covariance matrix. It is also possible to use singular value decomposition (SVD) instead [8]
4. Construct the feature vector: Reorder eigenvectors and determine the ones to retain
5. Obtain a new data set by applying the eigenvector to the data set

Bear in mind that our feature selection based on scoring falls in the feature selection category and PCA in the feature extraction category.

IV. Machine Learning

The effectiveness of the aforementioned feature engineering steps on machine learning algorithms' performance was evaluated by two artificial neural networks (ANN) structures employed to diagnose faults in the components. These two structures are feed-forward artificial

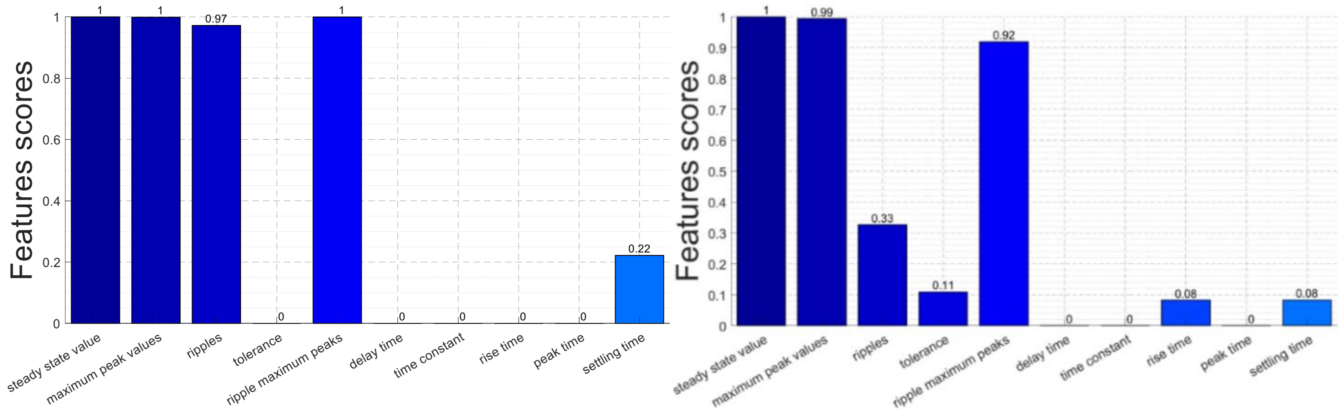

Fig. 6. Output signal features scored for: (a) MOSFET junction temperature (b) Capacitor's ESR.

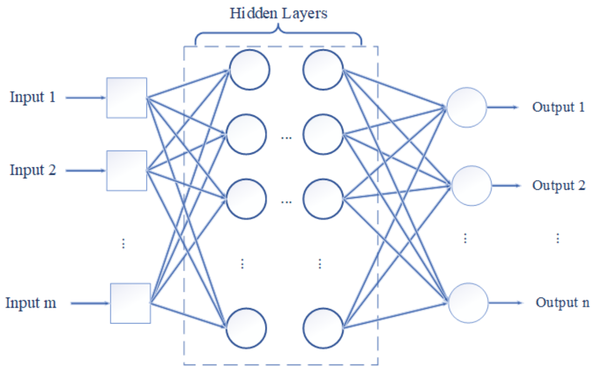

Fig. 7. FF-ANN architecture

neural networks (FF-ANNs) and radial basis functions networks (RBF networks). The networks' inputs were the extracted features, and the desired outcome is an appropriate diagnose of the components' states (pristine, degraded, and faulty).

A. Feed-Forward ANNs

FF-ANNs are some of the most used ANN structures. These ANNs can be used for solving simple problems and easily extended to deep learning systems to solve very complex problems. It is imperative to design their architectures and select their parameters to suit their purpose. Fig. 7 shows a sample of these ANNs.

The architecture, which is the most popular nowadays, incorporates Rectified Linear Unit (ReLU) activation function for hidden layers and a Softmax layer as the output layer, see Fig. 9.

We used the "Scaled Conjugate Gradient Backpropagation Algorithm" for training the FF-ANN and "cross-entropy" as the performance evaluation criterion.

B. Radial Basis Function Networks (RBF Networks):

Another popular architecture for classification is the RBF network structure. These networks comprise three layers, an input layer, a hidden layer with RBF activation function, and a linear output layer, see Fig.8.

One of the characteristics of RBF networks is that their value only depends on the distance of its input to a fixed point. On the other hand, in the RBF networks, all features have the same importance as they have exactly the same weight in distance computation.

To sum up, all the steps in the proposed approach are as follow:

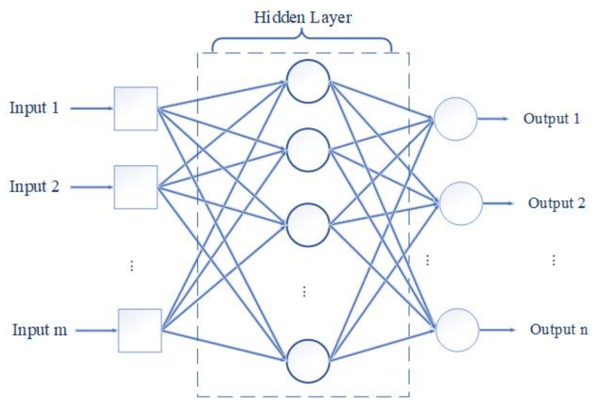

Fig. 8. RBF network architecture

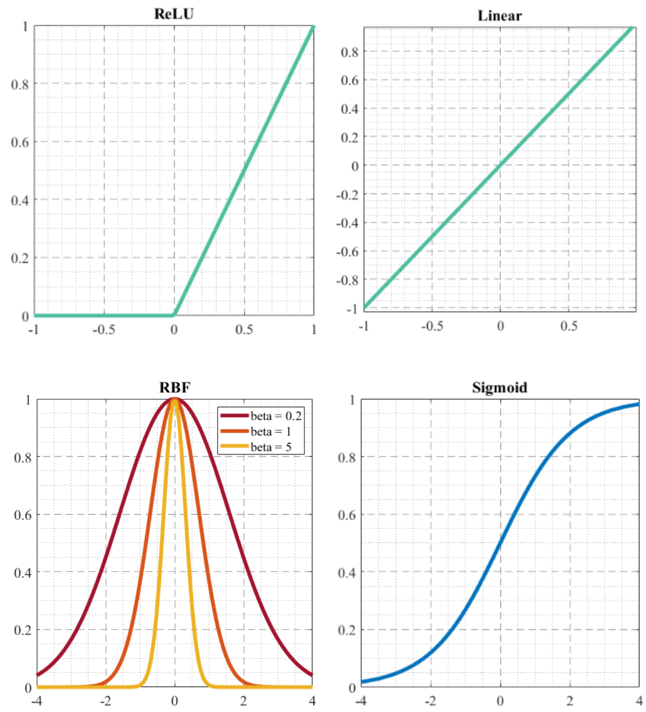

Fig. 9. Some of the most used activation functions in ANNs

1- Gather data from the simulations and sieve the unwanted data points

2- Nominate some features of the signal as health indicators (in our example, we considered some time-domain features)

3- Select the most appropriate features for monitoring

4- Apply PCA to reduce dimensionality and, therefore, the need for a large amount of data, also increasing the performance of machine learning algorithms

5- Select and customize a machine learning algorithm based on the characteristics of the problem

6- Train and evaluate the machine learning algorithm using the data gathered and processed in steps 1 to 4

V. RESULTS

Fig. 10 demonstrates how PCA projects all features to a 3D space. The eigenvalue magnitude signifies how much

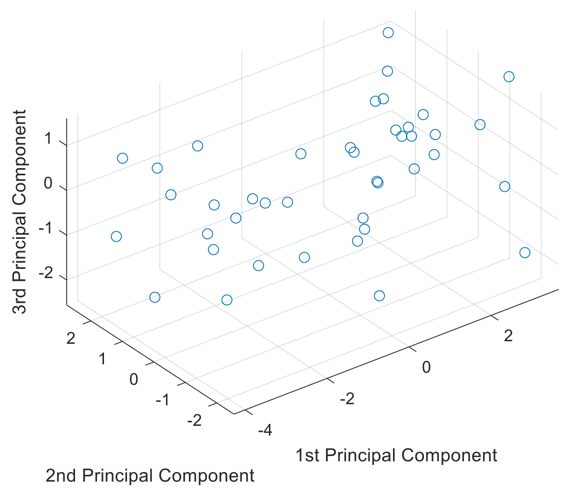

Fig. 10. PCA reduces dimensionality by projecting data set to a 3D space

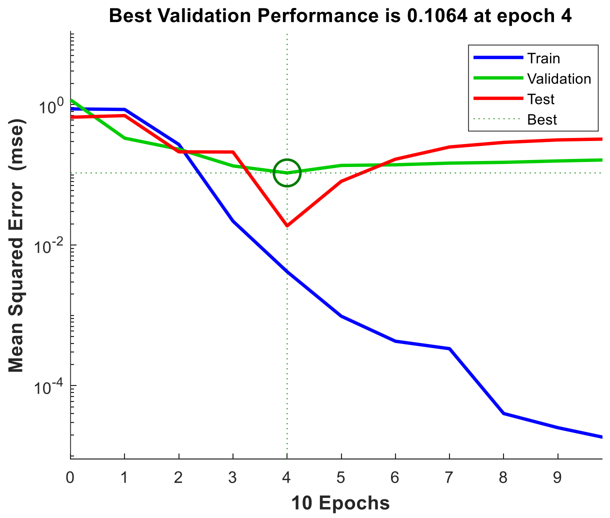

Fig.11. FF-ANN convergence and performance

variance each component explains. In this instance, the first component describes 93.1%, the second 6.2%, and the third one 0.6 of the data set.

Fig. 11 illustrates an FF-ANN's performance during training, validation, and testing phases to detect a component's state. The testing results for different components varied drastically from close to perfect detection to an unacceptable precision. In simple cases, there is only one varying parameter FF-ANN and RBF's performance are almost the same. As the number of varying parameters increases, the performances of ANNs diverge. For instance, FF-ANN successfully diagnosed 85% of components state correctly for three varying parameters, while RBF could achieve above 92% accuracy.

VI. Discussion

Selecting the most appropriate features for condition monitoring and promoting them as the health indicators might be the most important and also challenging part of condition monitoring, which is usually overshadowed by the importance of machine learning algorithms that are used for fault detection. As discussed in section III, the variations in the raw output might not satisfy the requirements for decent monitoring. Moreover, screening the features extracted at a single point of operation might be misleading, that is the main reason we needed to develop our scoring method, which assigned a quantitative value to each feature and made it possible to have a criterion to compare and select the most appropriate features.

As it is shown by Fig. 10, PCA can effectively map features into space with much lower dimensions. In this instance, these three features cover 99.9% of data set characteristics, and by studying them instead of initial features, we will reduce the dimensionality drastically without losing precision.

Fig. 11 shows that overtraining does not always improve performance. Therefore, during the training phase, we use a validation set to prevent overfitting.In the following, this section discusses why ML algorithms are successful in some cases and fail to deliver good results in others.

Parallel coordination plots reveal a cause of ML algorithms' failure to successful detection in those cases; One of such cases is the MOSFET's junction temperature, as shown in Fig. 12, where

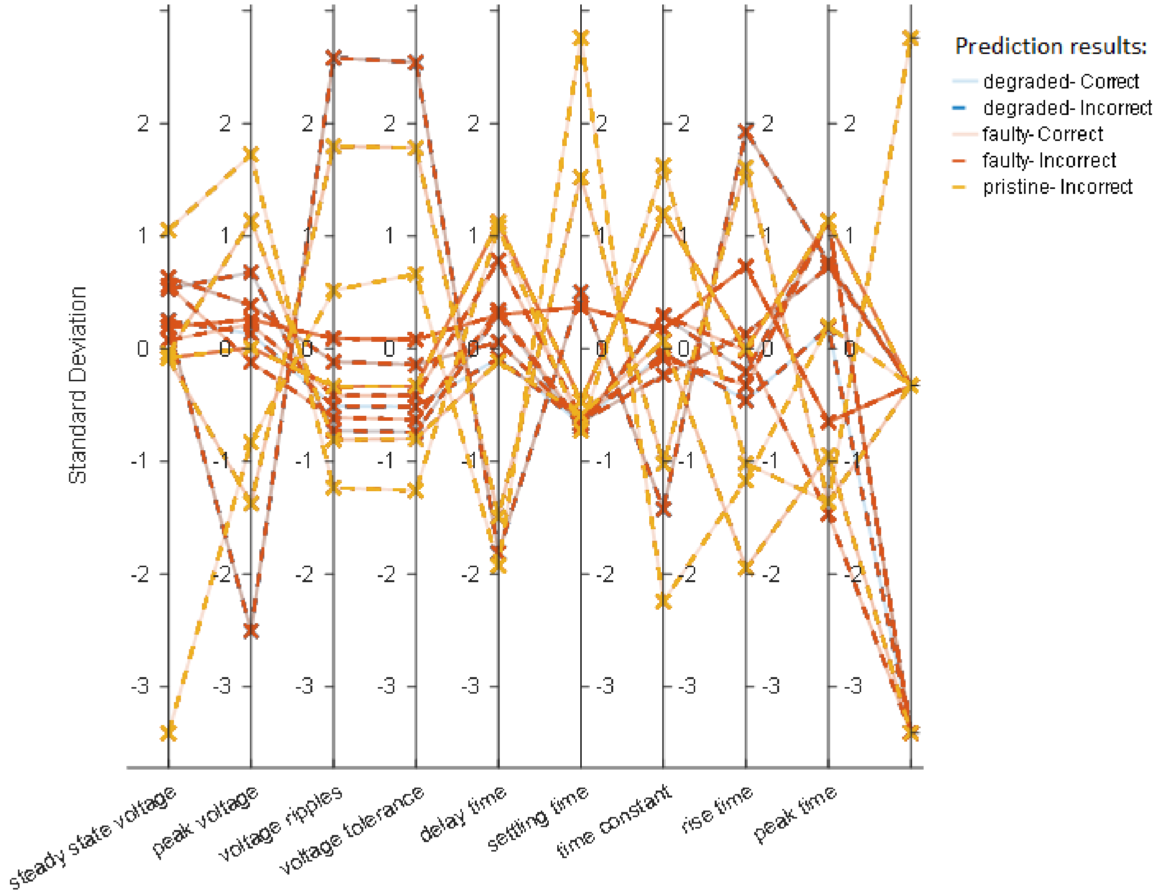

Fig. 12. Parallel coordination plot for junction temperature

the features of different states coincide. To overcome this problem, we need to increase the resolution of variations in parameter drifts.

VII. Conclusions

This paper proposed an approach that eliminates the experimental data collection burden by using simulations. This approach also helps to achieve real-time monitoring of systems by shifting the computationally intensive signal processing and training of machine learning algorithms to an offline procedure, allowing the implementation of fast machine learning algorithms.

A SiC buck converter is selected as the case study since they have wide application in electric vehicles. We simulated the degradation effects of the component by a digital twin of the system. A scoring criterion is introduced to select the best features extracted from the raw output signal of the system. Then we reduced the dimensionality using PCA. Finally, the processed data was used to train and evaluate the performance of different structures of neural networks.

References

[1] J. Rabkowski, D. Peftitsis, and H.-P. Nee, "Silicon carbide power transistors: A new era in power electronics is initiated," *IEEE Ind. Electron. Mag.*, vol. 6, no. 2, pp. 17–26, 2012.

[2] J. A. Schrock *et al.*, "Failure Analysis of 1200-V/150-A SiC MOSFET Under Repetitive Pulsed Overcurrent Conditions," *IEEE Trans. Power Electron.*, vol. 31, no. 3, pp. 1816–1821, 2016.

[3] B. J. Nel and S. Perinpanayagam, "A Brief Overview of SiC MOSFET Failure Modes and Design Reliability," in *Procedia CIRP*, 2017.

[4] R. Ouaida *et al.*, "Gate oxide degradation of SiC MOSFET in switching conditions," *IEEE Electron Device Lett.*, vol. 35, no. 12, pp. 1284–1286, 2014.

[5] C. Bhargava *et al.*, "Review of Health Prognostics and Condition Monitoring of Electronic Components," *IEEE Access*, vol. 8, pp. 75163–75183, 2020.

[6] A. Loghmani Moghaddam Toussi and A. A. Mashayekhi, "Prognosis of Failure of Bearings of Rey Gas Power Plant Using Hidden Semi- Markov Models," in *6th International Conference on Electrical Power Generation*, 2014.

[7] Y. Peng, S. Zhao, and H. Wang, "A Digital Twin Based Estimation Method for Health Indicators of DC-DC Converters," *IEEE Trans. Power Electron.*, vol. 36, no. 2, pp. 2105–2118, Feb. 2021.

[8] J. Shlens, "A tutorial on principal component analysis," *arXiv Prepr. arXiv1404.1100*, 2014.

[9] L. I. Smith, "A tutorial on principal components analysis," 2002.

[10] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT Press, 2012.

[11] C. M. Bishop, "Pattern recognition," *Mach. Learn.*, vol. 128, no. 9, 2006.

[12] E. Alpaydin, *Introduction to machine learning*. MIT Press, 2020.

[13] A. Loghmani Moghaddam Toussi, A. S. Bahman, F. Iannuzzo, and F. Blaabjerg, "Parameters sensitivity analysis of silicon carbide buck converters to extract features for condition monitoring," *Microelectron. Reliab.*, vol. 114, p. 113910, Nov. 2020.