

PV Inverter Fault Classification using Machine Learning and Clarke Transformation

Louelson Costa, Ana Silva, Ricardo J. Bessa,
Centre for Power and Energy Systems (CPES)
INESC TEC
Porto, Portugal
{louelson.costa; ana.silva; ricardo.j.bessa}@inesctec.pt

Rui Esteves Araújo
Centre for Power and Energy Systems (CPES)
INESC TEC and Faculty of Engineering of the University of Porto
Porto, Portugal
raraujo@fe.up.pt

Abstract—In a photovoltaic power plant (PVPP), the DC-AC converter (inverter) is one of the components most prone to faults. Even though they are key equipment in such installations, their fault detection techniques are not as much explored as PV module-related issues, for instance. In that sense, this paper is motivated to find novel tools for detection focused on the inverter, employing machine learning (ML) algorithms trained using a hybrid dataset. The hybrid dataset is composed of real and synthetic data for fault-free and faulty conditions. A dataset is built based on fault-free data from the PVPP and faulty data generated by a digital twin (DT). The combination DT and ML is employed using a Clarke/space vector representation of the inverter electrical variables, thus resulting in a novel feature engineering method to extract the most relevant features that can properly represent the operating condition of the PVPP. The solution that was developed can classify multiple operation conditions of the inverter with high accuracy.

Index Terms—Machine learning, digital twin, Clarke transformation, photovoltaic, faults

I. INTRODUCTION

The photovoltaic (PV) power plant (PVPP) has multiple types of equipment working in cooperation to ensure proper operation, such as weather stations, electrical equipment (sensors, inverters, transformers, etc.), and supervisory control and data acquisition (SCADA) system [1]. Even though every piece of equipment is prone to fault or failure [2], the inverters are the type of equipment that is the target of most of the maintenance interventions [1], [3], closely followed by the PV modules. However, compared to PV modules, the inverter is not much used for studies of faults detection and classification [4], [5]. When detection techniques are applied to the inverter, they do not have a more detailed approach, seeing the inverter as a “black box” (or a simplified model of this component) [6].

The studies regarding the faults and failures [2] in PV systems focus on the DC side (PV modules and DC-DC converter). Some issues in the PV modules may be identified by using image processing [1], [7], as multiple issues are related to the surface of the PV modules (delamination, cracks,

etc.), while others may be identified using electrical measurements processing [3], [7]. This condition-based monitoring is particularly important to avoid critical issues that may cause a significant reduction in power generation or even stop it, whilst the early fault detection increase the safety and production of the PVPP.

Regarding the fault detection and classification of the inverters, some papers tackle these issues within the context of renewable and PV systems [3], [8], however, most of the research on inverter faults is for motor drive [9], [10]. As the inverter is a critical asset in the PVPP, this paper is motivated by the fault classification of such equipment using machine learning (ML) tools. The study was carried out having as a focus the Clarke transformation, i.e., the measurement of the output currents of the inverter, combined with other key measurements that are directly related to the inverter (inverter inputs and outputs measurements). For the scope of the present work, the AC/output currents of the inverter and DC/input currents of the inverter are taken into account.

The Parks-based approach is already applied to motor drives [10], where the behaviour of the space vector is directly related to some faulty conditions of the inverter and/or motor. Thus, a similar approach to the application of the Clarke transformation used to summarize the behaviour of the PV inverter into key features is proposed and validated. The results are always presented in normalized (per unit, p.u.) values to show that the discussion can be exported to any given three-phase PV system, regardless of its power level.

To do so, a real dataset of fault-free condition PVPP was taken as a basis. A digital twin (DT) was built so the faulty data could be generated. Thus, by using a typical approach due to fault-free and faulty data imbalance [3], [11], the dataset that is used to train, validate and test the proposed feature engineering and methodology is composed of fault-free real data and faulty simulated data.

This paper is organized as follows: The PVPP and its factors of influence, including a general overview of the most common electrical equipment and their faults, are presented in Section II; Section III describes the dataset and the feature engineering method based on the Clarke transformation; Section IV presents the ML algorithms for classification of inverter faults; Section V describes ML models training and testing; Section

This work is financed by the ERDF - European Regional Development Fund, through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme under the Portugal 2020 Partnership Agreement, within project A14PV, with reference POCI-01-0247-FEDER-111936.

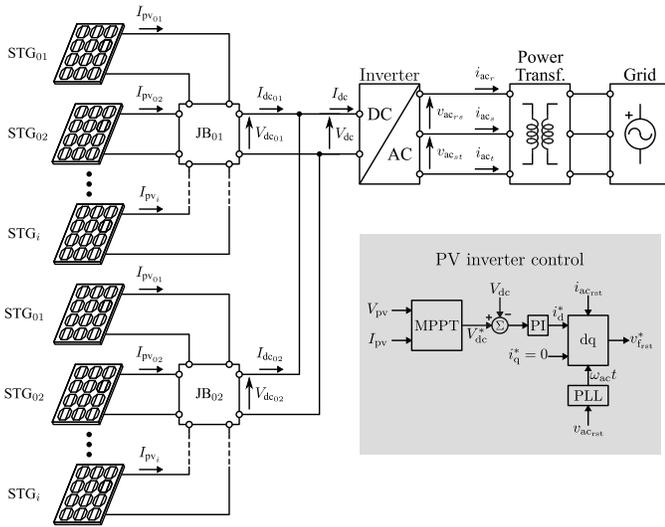


Fig. 1. PVPP electrical circuit with the junction/combiner boxes connection to the DC-link of the inverter, and the inverter connection to the grid through a power transformer. In detail, the PV inverter control.

VI presents the conclusions and future work.

II. PVPP PLANT FAULT CHARACTERIZATION

The PVPP are large structures that can be composed of hundreds to thousands of PV modules, feeding multiple inverters for the grid connection. Other systems, such as weather stations, and SCADAs, are also part of PVPP. Thus, a study of the most common faults is a key step to developing proper detection tools. A simplified representation of the PVPP electrical circuit is presented in Fig. 1. It is worth noting that the transformer is not modelled in this paper.

There are multiple fault detection techniques for each part of a PVPP. Besides the specific object of study within a PVPP (PV modules, DC cables and combiner/junction boxes, inverters, power transformer, etc.), there are multiple strategies (AI-based, model-based, etc.) that can be used to do such a task. Regardless of the technique or strategy, they have some clear steps to be followed [7], [11], [12].

Face the size of the PVPP is very important to understand the range of a fault to develop an effective detection algorithm. For example, a problem that originates in one of the sub-systems may be observed on multiple measurement points. At the same time, it is also important to understand how a problem at a given point may affect the adjacent sub-systems (forward-or backward-wise) [13].

A. PV modules

The mismatch faults of PV modules happen when a PV module, or a set of them, for some reason, are not operating in the same conditions as the other sets connected in series or parallel. In a series connection, it is expected that all of the PV modules work at the same current. Similarly, in a parallel connection, it is expected that all of the PV modules (or PV strings) work at the same voltage. The PV cell degradation faults are related to multiple issues that may lead to the

sunlight blocking (partial or total) of specific areas of PV modules. They may also cause a series resistance increase, resulting in a mismatch as well. The PV module faults can be detected by electrical and meteorological measurements (real vs. expected out power, for instance).

B. DC connection and combiner boxes

In a PVPP, the PV modules are connected in series (to increase their output voltage, resulting in PV strings) and, subsequently, connected in parallel at the combiner boxes (increasing the current). Among the problems associated with the DC connection and combiner boxes, and even though the occurrence of problems in these types of equipment is negligible compared to the PV modules and power electronics in terms of production loss [1], one of the most common alongside short circuits is the cable degradation. The cable degradation is the gradual resistance increase of the DC cables caused by soldering issues, ageing or electrical stress, and the open circuit connections, which is the disconnection between the PV modules or junction boxes and inverters, or fuse tripping.

The consequence of those faults is, at the least, a disconnection of the affected string or combiner boxes. A short-circuit may lead to fire and result in cable and/or circuit disconnection. DC cables and combiners boxes are reliable and hardly prone to failure, but in case of a problem, as they connect the multiple parts of the PVPP, the production associated with them will be stopped. This can go from a single PV module, through entire strings that are feeding the inverter.

C. Power electronics and reactive components

Some works are focused on the PV modules and “reduce” the power electronics and reactive components to a simplistic black box labelled as “inverter”. However, the inverter is composed of multiple interconnected systems, and a single problem in those components may lead to a cascade effect that will result in poor performance of the PVPP [14].

Besides the three-phase inverters and PV module faults already being researched, there is a lack of research on ML techniques applied to fault analysis in the inverter of PV power plants. Mostly, the effort of this research towards failure detection in DC-AC converters is on the open or short circuit analysis, as those conditions can lead to power loss increase, THD increase, current, voltage, or thermal stress over the components (semiconductors and reactive), etc. Regarding the IGBTs and MOSFETs, a switch failure will result in a non-balanced output (in the case of three-phase systems) [6].

D. Fault Operation Modes

For proof of concept of the proposed methodology, five operation modes (one for fault-free conditions and four for faulty conditions) directly related to the inverter are considered:

- Fault-free condition (*noFault*): regular operation of the PVPP without any fault;

- DC cable degradation (*dcCabDeg*): this type of fault is very common due to the ageing of multiple components. In this case, this operation mode is representing issues with the DC cable, or in the connection in the junction boxes or strings, etc.;
- DC cable open-circuit (*dcCabOC*): this type of fault can be caused by several degradations that may lead to the cable disconnection or fuse tripping during sunny and cold days;
- switch degradation (*switchDeg*): this type of fault is caused by stressful currents and/or voltages profiles applied to the switches (IGBTs or MOSFETs) of the inverter, it can lead to a switch fault;
- switch open-circuit (*switchOC*): one of the most common faults that can happen to an inverter, not only in PV systems. This can be caused by stressful current and/or voltage profiles, over temperature, mechanical damage, bad assembling, etc.

It is noteworthy that the proposed solutions in this paper are not dependent on the control, inverter topology, or power level, and the signals that are being processed by the algorithms are electrical and weather data, that can be retrieved from PVPP SCADA systems.

III. DIGITAL TWIN AND FEATURE ENGINEERING

To train and validate the proposed methodology, it is necessary to build the dataset and to organize/process it so the data can be intelligible for the algorithms. In that sense, due to the lack of real faulty data, a DT (implemented in Simulink/MATLAB[®]) is developed so that the hybrid dataset can be assembled. After that, the Clarke transformation and feature extraction is carried out.

A. PVPP model validation and faulty dataset

This paper is based on the PVPP Monte das Flores (approximately 38° N, 8° W), owned and operated by EDP Renewable. The meteorological data (i.e., weather measurements such as irradiance and ambient temperature) and SCADA data (i.e., electrical measurements such as currents, voltages, power, etc.) were used for building the DT model.

Once the DT is assembled, it is validated across multiple days during a year. Some of these results can be seen in Fig. 2 for the RMS value of one of the AC output currents of the inverter. It can be noticed that the DT has a similar output when compared to the real data. However, under low irradiance (usually below 100 W/m²), the DT performance is compromised. Also, during cloudy scenarios, the performance is not as good as under sunny days, even though it is a satisfactory result for the objectives of this paper.

After validating the DT model, the faulty data is generated for the four faulty conditions prior selected. Both DC cable-related problems are located between one of the junction boxes and the DC link of the inverter. For the degradation, a series resistance is added to the circuit. For the open circuit, a sudden disconnection manoeuvre is done. Both switch-related problems are applied directly to a switch of the inverter. For

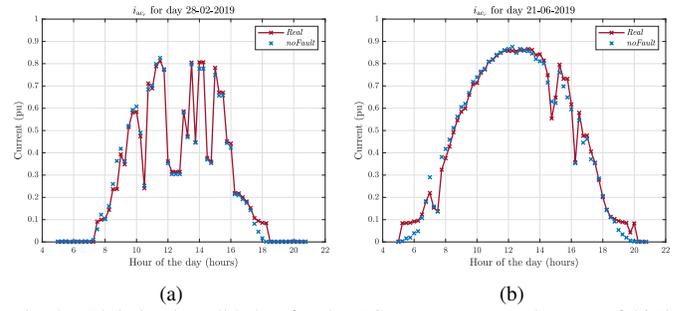


Fig. 2. Digital twin validation for the AC currents across the year of 2019: (a) February 28th; and (b) June 21st.

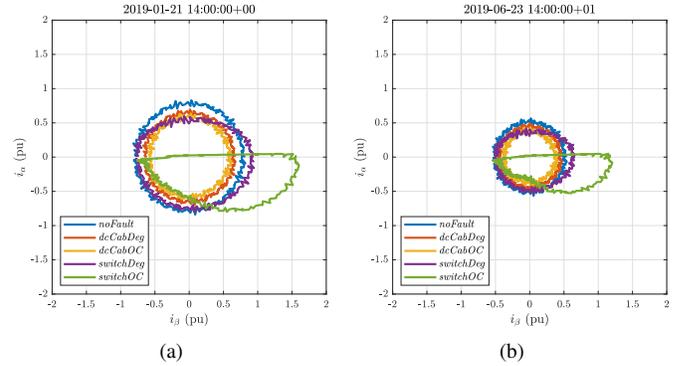


Fig. 3. Clarke space vector dataset without center dislocation samples from 2019 seasons at 2 PM: (a) Winter; and (b) Summer.

the degradation, an internal series resistance is added to the on-resistance of the switch. For the open circuit, the switch stops switching, always at a tuned-off state.

B. Clarke transformation and feature engineering

The Clarke transformation, or alpha-beta transformation, is a matrix multiplication that can be applied to three-phase signals to achieve another representation of the three-phase currents. Thus, it starts from a three-phase system into an orthogonal reference frame:

$$i_\alpha(t) = \frac{2}{3}i_r(t) - \frac{1}{3}(i_s(t) - i_t(t)), \quad (1)$$

$$i_\beta(t) = \frac{2}{\sqrt{3}}(i_s(t) - i_t(t)), \quad (2)$$

$$i_\gamma(t) = \frac{2}{3}(i_r(t) + i_s(t) + i_t(t)). \quad (3)$$

The i_γ , or homopolar component, is not being used in this project, thus a simplified version of the transform can be written using only i_α, i_β . For a perfectly balanced and fault-free operating inverter, the three-phase currents are represented by i_α, i_β as a perfect circle in the space vector, as shown in Fig. 3 by the blue (*noFault*) waveform.

However, under faulty conditions, the space vector will suffer deformations in its shape, radius, etc. For instance, Fig. 3 shows an overlap of the multiple conditions under analysis. It can be seen that in the *noFault*, *switchDeg* and *switchOC* the deformation/difference on the space vector is

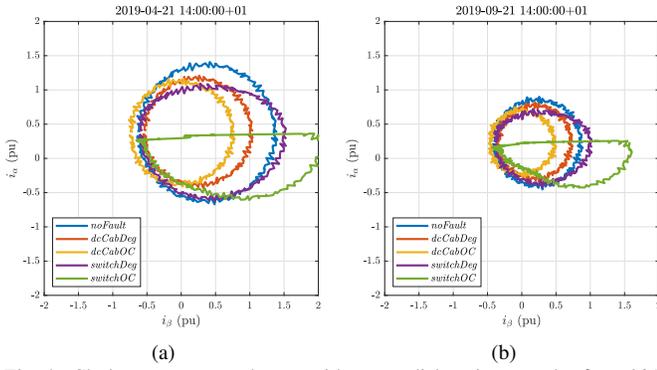


Fig. 4. Clarke space vector dataset with center dislocation samples from 2019 seasons at 2 PM: (a) Spring; and (b) Autumn.

more recognizable when compared to the *dcCabDeg* and *dcCabOC*. The switch-related problems cause the space vector to change its shape, resulting in something closer to an ellipse or even a half-circle.

For the DC cable-related problems, the difference from the fault-free condition is the radius, as the circle shape is maintained. This can make it harder to distinguish between a *dcCabDeg* and *dcCabOC* by only exploring the i_α, i_β representation.

Based on this difficulty to distinguish between the *dcCabDeg* and *dcCabOC* problems, the study proposed to consider more features to be processed by the ML algorithms or even altering the space vector to add some elements that could help to tell the difference between *dcCabDeg* and *dcCabOC*. One simple but effective addition was the centre dislocation, which consists of using the input DC-link currents of the inverter (I_{dc01}, I_{dc02} as seen in Fig. 1) as the centre of the space vector, generating the new space vectors presented in Fig. 4. With the centre dislocation, it is easier to notice the difference between *dcCabDeg* and *dcCabOC*, thus it is expected that the ML algorithms have a better discrimination capability between them as well.

The Figs. 3 and 4 show that depending on the irradiance and temperature level, which vary along the day, months, and seasons, there will be multiple space vectors. The proposed features extracted from the space vector and some weather-related features to be processed by the ML algorithm are presented in Table I.

It is worth noting that the normalization of the data is done based on the nominal ratings (current, voltage, power) of the inverter, which can be found on their datasheet. When necessary, the weather-related data (irradiance and temperature) are normalized based on the standard test conditions (STC). This approach aims to provide a generalized overview of the methodology, which can be exported to any other PV system.

IV. ML ALGORITHMS FOR FAULT DETECTION

The alpha-beta classification is responsible for indicating the most likely fault that occurred. Therefore, four supervised machine learning methods will be considered [5], [15]:

TABLE I
PROPOSED FEATURE ENGINEERING AND BRIEF CONTEXTUALIZATION

Feature	Explanation/formula
α_0, β_0	Space vector center
$\alpha_{\max}, \beta_{\max}$	Maximum value within a cycle of i_α, i_β
$\alpha_{\text{mean}}, \beta_{\text{mean}}$	Mean value within a cycle of i_α, i_β
$\alpha_{\min}, \beta_{\min}$	Minimum value within a cycle of i_α, i_β
$\alpha_{\text{rms}}, \beta_{\text{rms}}$	RMS value within a cycle of i_α, i_β
$\Delta\alpha_{\max}, \Delta\beta_{\max}$	Difference between a given
$\Delta\alpha_{\text{mean}}, \Delta\beta_{\text{mean}}$	α, β characteristic and
$\Delta\alpha_{\min}, \Delta\beta_{\min}$	the time-wise previous characteristic
I_{dc01}, I_{dc02}	DC-link input currents
G	Global irradiance
Temp	Ambient temperature
season	Season of a given subset of the dataset
skytype	Indicator from clouded to clear-sky

- Decision Tree (DTr): Although decision trees require little pre-processing, they have some disadvantages, such as the creation of high-depth trees does not allow a reasonable generalization of data (overfitting) and the instability due to low variations of training samples.
- Random Forest (RF): an ensemble of decision trees that corrects the overfitting of the previous method during training. For classification problems, the result is the class selected by most trees.
- K-Nearest Neighbors (kNN): this algorithm aims to restrict the decision space by choosing the k-nearest neighbors in the training dataset of a given testing value.
- Artificial Neural Network (ANN): a biological-inspired computational network that includes the supervised learning algorithm Multi-Layer Perceptron (MLP). It is worth mentioning the adoption of a feedforward architecture with backpropagation for the alpha-beta classification, which consists of training the dataset using some gradient descent-based method by propagating the error back into the nodes (layers) and updating the parameters (weights and biases) to minimize the loss.

V. NUMERICAL RESULTS

Combining the dataset and its features (alpha-beta currents and weather data) with the select ML algorithms, the training, validation, and testing of each model is carried out. The general pipeline is depicted in Fig. 5. The training and testing are done with different datasets from different years.

A. ML models dataset

In the alpha-beta classification, for the training and validation (cross-validation) is used 80% of the hybrid dataset from 2021. Afterwards, the models described in Table II are tested using the remaining 20%. Nevertheless, as the purpose of the model is the classification of faults, regardless of the year under study, the hybrid dataset from 2019 will be used as a second and final testing dataset.

At first, we considered features related to date and weather (*skytype* and *season*). However, it is worth remembering that the faulty dataset was generated through a DT developed in Simulink/MATLAB[®] modelling tool, which used the same

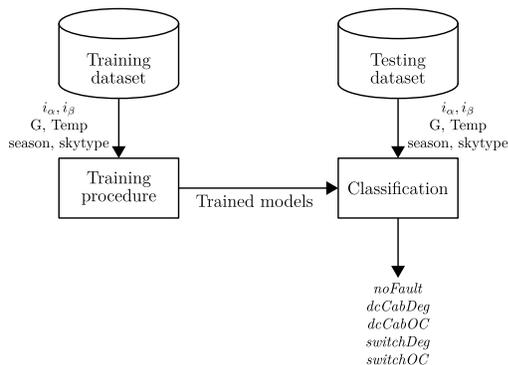


Fig. 5. Fault detection pipeline. Based on the diagram from [11].

TABLE II
ML ALGORITHMS HYPER-PARAMETERS AND TRAIN ACCURACY

Algorithm	Hyper-parameters	Train accuracy (without; with dislocation)
ANN	activation = tanh, hidden layers = 55, iterations = 1000, random state = 42 or 0, tolerance = 0.001, learning rate = constant	97.9%; 97.7%
DTr	depth = 3, random state = 0	93.5%; 89.5%
kNN	k = 5, weights = distance	100.0%; 100.0%
RF	depth = 3, random state = 42	95.1%; 94.5%

values of the fault-free condition for the previous features. This assumption created redundancy in considering features related to date and weather as inputs, leading to their removal.

A similar conclusion was reached regarding the features $\sum \alpha$ and $\sum \beta$. After an investigation based on the Pearson Correlation Coefficients of the comparison of pairs that compose the set $C = \{(\alpha_{\text{mean}}, \sum \alpha), (\beta_{\text{mean}}, \sum \beta)\}$, it was found a coefficient value of 1.00 ($\rho = 1.00$). Thus, the features $\sum \alpha$ and $\sum \beta$ were disregarded.

B. Hyper-parameters tuning

One of the most common methods for hyper-parameter tuning is the Grid Search. Its purpose is to determine the estimator with the most accurate predictions. The optimization of the models is implemented by the function GridSearchCV of the Scikit-Learn/Python library, which includes the cross-validation of the training dataset, considering five subsamples ($cv = 5$).

This allows finding the best (or as close as possible to the optimal) hyper-parameters for each model. After applying the Grid Search technique to both dataset approaches (with and without centre dislocation), the hyper-parameters for the ML models are listed in Table II.

Their train accuracy shows that ANN and kNN are the most promising solutions so far. Although, will be tested a different dataset from another year for a better comparison between the ML models.

TABLE III
ALGORITHMS METRICS FOR TESTING, WITHOUT CENTER DISLOCATION

Algorithm	Accuracy	Precision	Sensitivity	Specificity
ANN	97.3%	97.3%	97.3%	99.3%
DTr	92.1%	93.2%	92.1%	98.1%
kNN	97.2%	97.2%	97.2%	99.3%
RF	93.9%	95.1%	93.9%	98.6%

TABLE IV
ALGORITHMS METRICS FOR TESTING, WITH CENTER DISLOCATION

Algorithm	Accuracy	Precision	Sensitivity	Specificity
ANN	97.0%	97.1%	97.0%	99.3%
DTr	87.1%	90.8%	87.1%	97.0%
kNN	97.3%	97.3%	97.3%	99.3%
RF	93.5%	93.8%	93.5%	98.4%

C. ML models comparison

The comparison between the ML models is done using the metric of accuracy, precision, sensitivity, and specificity, which are functions of the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The listed metrics for the dataset without center dislocation are presented in Table III, whilst the metrics for center dislocation are presented in Table IV.

Considering the results without center dislocation, the kNN and ANN models presented the best results for all of the metrics, with ANN slightly ahead. This shows that even though the ANN are more complex solutions, a kNN model is enough to properly classify the multiple conditions under study.

A similar analysis can be made for the result with center dislocation. Whilst the DTr and RF presented a worse result, the kNN model presented the best metrics. Overall, the accuracy of the ANN without center dislocation and the kNN with center dislocation was the same: 97.3%. It is worth noting that both models do not overfit, which is verified through the proximity of training and testing metrics, such as accuracy. The accuracy during the cross-validation is slightly higher, but this is an expected outcome.

The confusion matrices of the best ML models for this study are presented in Fig. 6. It can be noticed that the conditions *noFault*, *switchDeg* and *switchOC* are the most easily classified, which is supported by the space vector plots. On the other hand, by analyzing the results for *dcCabDeg* and *dcCabOC*, their classification stills pose some challenges when compared to the other conditions. This is also visually noticeable in the space vectors, whereas those two conditions are the ones with the most similar space vector plots. However, it is worth noting that both solutions for the ML models represented in Fig. 6 achieve an accuracy higher than 90%.

VI. CONCLUSIONS AND FUTURE WORK

The paper describes the Alpha-Beta/Clarke representation built through a DT model using the Simulink/MATLAB[®] tool. Subsequently, by applying supervised classification methods, the models implemented allowed the identification of four

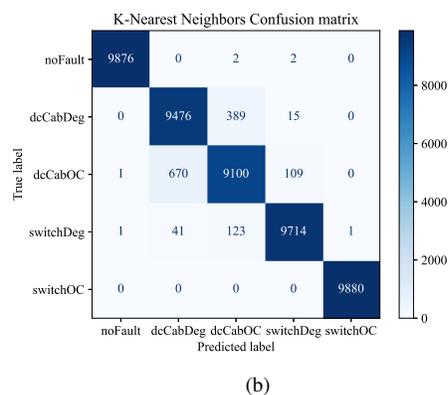
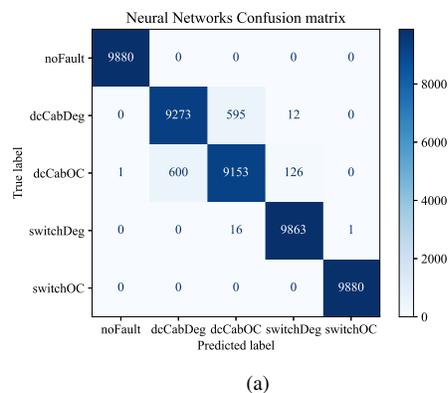


Fig. 6. Confusion matrices for the testing of the alpha-beta classification for the ML algorithms that resulted in the highest accuracy: (a) Without center dislocation for ANN (Table III); and (b) With center dislocation for kNN (Table IV).

faults related to DC cables and switches. The following achievements were presented:

- For fault detection, it was possible to model the alpha-beta representation through supervised machine learning methods (whereas ANN and kNN presented the best results);
- For the switch-related faults, it was possible to distinguish them from the other conditions, on the space vector representation, by their shape change (an ellipse for *switchDeg* and a semicircle in the *switchOC*);
- The radius deviation caused by the DC cable-related faults in the space vector representation was more noticeable due to the center dislocation.

For future work, the authors are investigating the possibility of including more electrical measurements related to the inverter and its possibility to classify failures in the junction boxes, PV modules, and grid. In summary, future improvement will pursue the next steps:

- The inclusion of more electrical readings that can lead to more classification options;
- Increase the number of conditions to be classified, from both the DC-side and AC-side of the inverter;
- Further investigation of the inclusion of data for date and

weather-related features.

ACKNOWLEDGMENT

The authors would like to thank the partners from the AI4PV (<https://ai4pv.eu/>) project, EDP Centre for New Energy Technologies and Isotrol.

REFERENCES

- [1] G. D. Lorenzo, R. Araneo, M. Mitolo, A. Niccolai, and F. Grimaccia, "Review of o&m practices in pv plants: Failures, solutions, remote control, and monitoring tools," *IEEE Journal of Photovoltaics*, vol. 10, no. 4, pp. 914–926, 2020.
- [2] R. Isermann, *Fault-Diagnosis Applications: Model-Based Condition Monitoring Actuators, Drives, Machinery, Plants, Sensors, and Fault-Tolerant Systems*, 1st ed. Springer Publishing Company, Incorporated, 2011.
- [3] P. Jain, J. Poon, J. P. Singh, C. Spanos, S. R. Sanders, and S. K. Panda, "A digital twin approach for fault diagnosis in distributed photovoltaic systems," *IEEE Transactions on Power Electronics*, vol. 35, no. 1, pp. 940–956, 2020.
- [4] L. M. de Oliveira, F. E. Mendes da Silva, G. M. da Cunha Pereira, D. S. Oliveira, and F. L. Marcelo Antunes, "Siamese neural network architecture for fault detection in a voltage source inverter," in *2021 Brazilian Power Electronics Conference (COBEP)*, 2021, pp. 1–4.
- [5] A. Malik, A. Haque, V. B. Kurukuru, M. A. Khan, and F. Blaabjerg, "Overview of fault detection approaches for grid connected photovoltaic inverters," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 2, p. 100035, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2772671122000079>
- [6] J. A. Dhanraj, A. Mostafaeipour, K. Velmurugan, K. Techato, P. K. Chaurasiya, J. M. Solomon, A. Gopalan, and K. Phoungthong, "An effective evaluation on fault detection in solar panels," *Energies*, vol. 14, no. 22, 2021. [Online]. Available: <https://www.mdpi.com/1996-1073/14/22/7770>
- [7] G. M. Tina, C. Ventura, S. Ferlito, and S. De Vito, "A state-of-art-review on machine-learning based methods for pv," *Applied Sciences*, vol. 11, no. 16, 2021. [Online]. Available: <https://www.mdpi.com/2076-3417/11/16/7550>
- [8] Q. Navid, A. Hassan, A. A. Fardoun, R. Ramzan, and A. Alraeesi, "Fault diagnostic methodologies for utility-scale photovoltaic power plants: A state of the art review," *Sustainability*, vol. 13, no. 4, 2021. [Online]. Available: <https://www.mdpi.com/2071-1050/13/4/1629>
- [9] G. A. Orcajo, J. M. Cano, M. G. Melero, M. F. Cabanas, C. H. Rojas, J. F. Pedrayes, and J. G. Normiella, "Diagnosis of electrical distribution network short circuits based on voltage park's vector," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1964–1972, 2012.
- [10] M. B. Abadi, A. M. S. Mendes, and S. M. A. Cruz, "Three-level npc inverter fault diagnosis by the average current park's vector approach," in *2012 XXth International Conference on Electrical Machines*, 2012, pp. 1893–1898.
- [11] A. E. Lazzaretti, C. H. d. Costa, M. P. Rodrigues, G. D. Yamada, G. Lexinoski, G. L. Moritz, E. Oroski, R. E. d. Goes, R. R. Linhares, P. C. Stadzisz, J. S. Omori, and R. B. d. Santos, "A monitoring system for online fault detection and classification in photovoltaic plants," *Sensors*, vol. 20, no. 17, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/17/4688>
- [12] A. Mellit and S. Kalogirou, "Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: Challenges, recommendations and future directions," *Renewable and Sustainable Energy Reviews*, vol. 143, p. 110889, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032121001830>
- [13] J. Keller and B. Kroposki, "Understanding fault characteristics of inverter-based distributed energy resources," 1 2010. [Online]. Available: <https://www.osti.gov/biblio/971441>
- [14] Y. Xia and Y. Xu, "A transferrable data-driven method for igbt open-circuit fault diagnosis in three-phase inverters," *IEEE Transactions on Power Electronics*, vol. 36, no. 12, pp. 13 478–13 488, 2021.
- [15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.