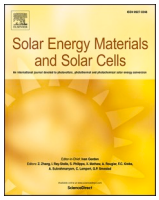
Solar Energy Materials & Solar Cells 257 (2023) 112398

Contents lists available at ScienceDirect

Solar Energy Materials and Solar Cells

journal homepage: www.elsevier.com/locate/solmat



Dynamics of backsheet-driven insulation issues

Claudia Buerhop Lutz [a,*], Larry Lüer [b], Oleksandr Stroyuk [a], Jens Hauch [a], Ian Marius Peters [a]

[a] Forschungszentrum Jülich GmbH, Helmholtz-Institute Erlangen-Nürnberg for Renewable Energies (HI ERN), D-91058, Erlangen, Immerwahrstraße 2, Germany
[b] Friedrich-Alexander Universität Erlangen-Nürnberg, Material Science, iMEET, D-91058, Erlangen, Martensstraße 5, Germany

ARTICLE INFO

Keywords:
Backsheet
Degradation rate
Insulation resistance
Machine learning
Gaussian Process Regression

ABSTRACT

To reap in the full benefit of high-efficiency solar cells and modules, the quality of the polymer encapsulant and backsheet (BS) materials is essential. Recent field studies show that early degrading backsheets cause safety issues and inverter shutdowns, resulting in yield losses. The dynamic development of degradation is important but is rarely studied because of the lack of proper datasets with meaningful and sufficient data. We studied inverter data of solar panels with a combined capacity of about 1 MWp for more than nine years including ground impedance (*GI*), and we labelled BS-types using our in-house identification method. Every inverter was connected to modules with exclusively a single BS-type, namely three-multilayer BSs denoted by their outer air layer: PA (polyamide), FC (fluorinated coating), and PVDF (polyvinylidene fluoride). We present, for the first time, an analysis of the degradation dynamics of *GI*, using a trained Gaussian Process Regression model. Using this model, we derived key degradation parameters, namely *GI* loss rates and time points for the onset of degradation. We found that PVDF BSs are associated with very low *GI* loss rates, while loss rates for PA are threefold higher with an onset at 6.6 year of operation for PA. FC BS-related loss rates were humidity dependent and two times higher those of PA; the onset was determined at 4.9 year of operation.

1. Introduction

The choice of polymer materials in PV modules is important for persistence and yield. Improvements in cell efficiency can only be exploited sustainably if the cells are well encapsulated and protected against adverse climatic effects. Early ageing of the polymers, especially the backsheets [1,2], results in safety risks jeopardizing the reliable operation and a high yield of the solar parks. The BS degradation results in insufficient insulation resistance, pushing inverters into fault mode [3].

It is becoming increasingly clear that damaged BSs constitute a serious issue. Little research has been done on the modes of BS field-degradation since proper and sufficient data are usually not available. To fill this gap, we profited from an access to historical, BS-labelled inverter data and applied a machine learning (ML) method, namely Gaussian Process Regression (GPR), to analyse the dynamics of the historic evolution of the BS-related ground impedance (*GI*) in correlation with ambient conditions. If *GI* becomes too low, the inverter does not start operation, resulting in a loss of yield and revenue. The present study focuses on the *GI* threshold defined by standard [4] as $GI > 1$ MΩ or specified by the manufacturer, e.g. as $GI_{th} > 0.4$ MΩ for the inverters installed in the PV power station of interest. Low *GI* values can be due to e.g. internal inverter issues or low insulation resistance of PV modules (at high humidity). At that, we expect internal inverter issues to be statistically evenly distributed for all inverters, regardless of the BS types of the PV modules connected to the inverter. An insight into BS-driven insulation issues requires a unique dataset with different known BS types for the inverters with specific BS characteristics.

Therefore, the goal of this study is to go beyond sporadic spot inspections, i. e. documentation of degradation phenomena in BS, e.g. cracks, and random insulation measurements of strings or modules. Since we cannot analyse the cracking retrospectively, we aim to address the change in electrical properties to better understand the dynamics of aging of the BS and its impact on performance. To achieve this goal, a suitable dataset consisting of monitoring data time series of inverters with identified BS of all modules plus weather data was created. With these data, we have trained a ML tool, in particular GPR, which is capable of learning non-linear relationships and calculating confidence intervals. After successful training and testing, the learned kernel is used to calculate key points based on the derivatives and slopes. BS specific *GI* loss rates and characteristic time points can be derived from the results.

* Corresponding author.
E-mail address: c.buerhop-lutz@fz-juelich.de (C. Buerhop Lutz).

https://doi.org/10.1016/j.solmat.2023.112398
Received 24 April 2023; Received in revised form 16 May 2023; Accepted 18 May 2023
Available online 2 June 2023

2. Experimental procedure

For studying the dynamics of insulation issues in PV power stations, the available dataset as well as the knowledge of the bill of material (BOM) are both crucial for applying a ML method with meaningful output. Both, dataset and the used GPR, will be described in detail.

2.1. Description of dataset

We analyzed a PV power station with a capacity of 6.8 MWp (28,030 PV modules, belonging to 1,245 strings, which are connected to 423 inverters) commissioned in Germany in 2012. The available historic monitoring data span almost ten years of operation and include *GI*, which describes the insulation resistance measured by the inverters. At that, there is a single *GI* data point per day and inverter. For studying the BS-dependent insulation evolution, we identified the BSs of all modules on-site using near-infrared absorption (NIRA) analysis [5,6]. At that, a huge variety of BSs was revealed on inverter level [3,7]. In order to find a clear correlation between BS material and insulation issues, we reduced the dataset to inverters that are connected exclusively to modules with one type of BS, considering exclusively cases with more than two inverters for statistical purposes. The further analysis, therefore, contains 68 inverters with 4,692 PV modules and a capacity of 1.2 MWp.

The PV modules of these inverters are all from the same manufacturer, but they feature different BSs. Using NIRA, we identified three multi-layer BS types on-site (see Table 1); 14 inverters with PA-PA-PA BSs, 34 inverters with FC-PET-PP BSs, and 20 inverters with PVDF-PET-PE BSs. Throughout the paper, the inverters will be addressed by their BS air layers: FC- (fluorinated coating), PA- (polyamide), and PVDF- (polyvinylidene fluoride) BSs for easier reading.

BS anomalies were observed for modules with PA BSs and FC BSs. 60–80% of modules with PA show macroscopic cracks and "chalking" (rutile titania deposits, also described and studied by other groups [1, 8–10]). 100% of modules with FC BSs show inner micro cracks, from those 10% having additional strong corrosion (rarely found in literature yet [2,7,11]). Modules with PVDF-BS show no apparent anomalies, compare Fig. 1.

All inverters have been operated at the same site in Germany for approximately ten years (3,437 days). The weather at the installation site is evaluated on an annual basis for 2%- and 98%-percentiles for daily values of ambient temperature *T*, relative humidity *RH*, and irradiance *I*: $I = 0.3–8.4$ kWh/m^2, $T = -2.4 – 25.6$ °C, $RH = 45.7–94.3\%$) [12]. Humidity and temperature are of high relevance for this study. Days with temperatures of $T = 0–5$ °C with humidity of $RH = 0.9–0.95$ are most frequent with a kernel density above 0.3, shown for bivariate relationship between *T* and *RH* in Fig. 2. Other combinations, e.g. $T > 20$ °C with $RH \approx 0.5$ are seldom with occurrence probabilities below 0.04.

GI was measured by inverters every morning before connecting to the grid. For data reduction, we aggregated inverter data in terms of daily median *GI* sorted by BS type, as shown in Fig. 3. We note that single *GI* values for particular inverters and days might deviate from the median *GI*. These time series are analyzed and linked with the weather parameters, namely *T* and *RH*, which are averaged for a period of 8 h prior to grid connection.

Table 1
Summary of the identified BSs, PA: polyamide, PET: polyethylene terephthalate, PVDF: polyvinylidene fluoride, PE: polyethylene, PP: polypropylene, FC: fluorinated coating, R: rutile TiO_2.

BS-type	Thickness [µm]	air layer	core layer	inner layer	anomalies	# of inverters
PA	400	PA + R	PA + R	PA + R	Open cracks	14
FC	215–220	FC + R	PET	PP + R	Closed inner cracks	34
PVDF	470–480	PVDF + R	PET	PE + R	–	20

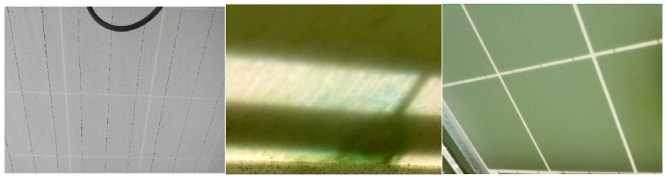


Fig. 1. Photographs of BSs of field-aged PV modules, a) open cracks in PA, b) green-coloured corrosion and closed inner cracks in FC, and c) no signs of degradation for PVDF.

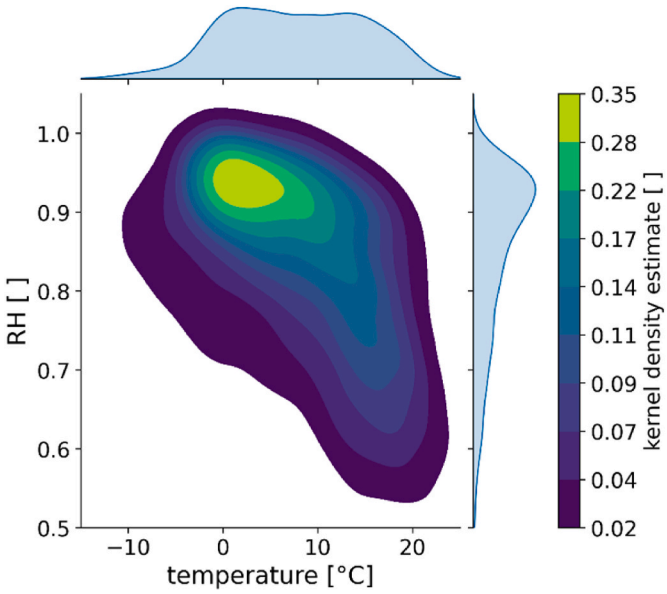


Fig. 2. Kernel density distribution for humidity and temperature at the site of the PV power station in Germany for the studied period.

The original data (Fig. 3 left) show a seasonal influence, with low *GIs* in summer when the ambient temperature is high and high *GIs* in winter when the ambient temperature is low. While the mean daily temperature changes from 0 °C to 20 °C, *GI* drops by roughly 5 MΩ. We associate the seasonal *GI* fluctuation with a temperature-driven inverter drift. At that, we calculated ratios of GI_{PA}/GI_{PVDF} and GI_{FC}/GI_{PVDF} (Fig. 3 right) in order to demonstrate this effect assuming no degradation for inverters with PVDF. The resulting ratios levelized close to a ratio of one. The ratio of GI_{FC}/GI_{PVDF} developed increasingly characteristic seasonal pattern with low values in winter not observed for GI_{PA}/GI_{PVDF}.

2.2. Methodology - GPR

To study the dynamics of *GI* changes with time, we used GPR, which is a non-parametric ML model that can be used for modelling spatial and time series data. GPR is a robust supervised ML algorithm used to solve regression and classification problems and assuming that the underlying data is normally distributed and normally jointly distributed. The premise of GPR is that a function is modelled derived from a finite number of points but allowing values at unobserved locations as well as unseen data points to be revealed with high confidence. The core of GPR is finding a suitable kernel function for modelling similarities between such data points distributed in space and time.

$$\textit{learned kernel} = \textit{amplitude} \bullet \textit{RBF}(\text{length-scale}[\textit{days}, T, \textit{RH}]) + \text{WhiteKernel}(\textit{noise})$$

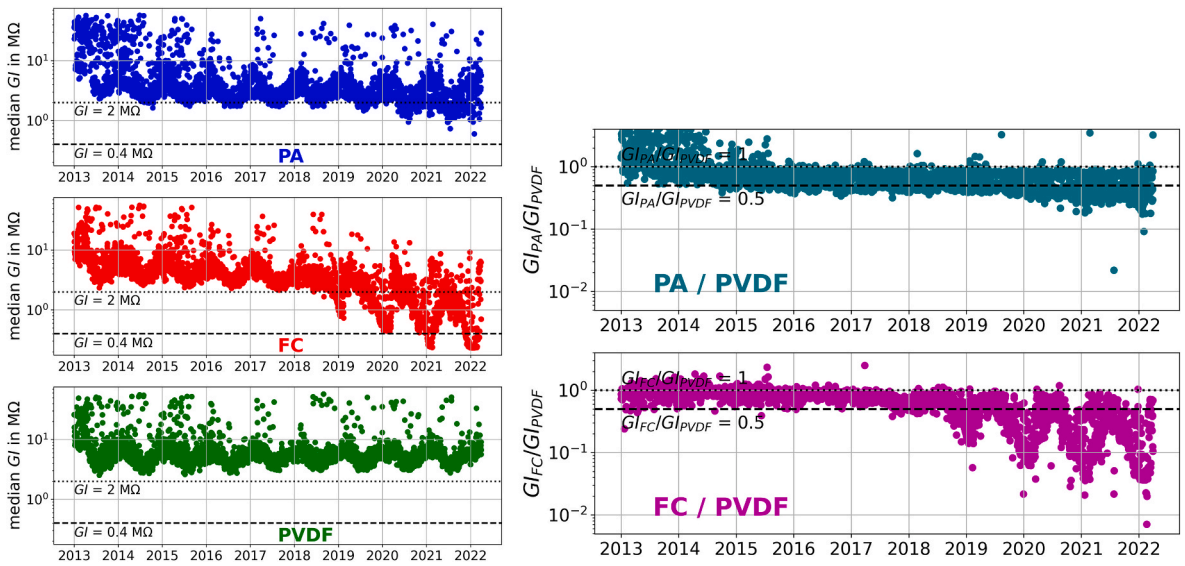


Fig. 3. Original monitoring data of GI measured by the inverter every day sorted by BS-type, PA, FC, and PVDF (left). For eye guidance, horizontal lines mark GI = 2 MΩ (dotted line) and GI = 0.4 MΩ (dashed line). Ratios of GI_{PA}/GI_{PVDF} and GI_{FC}/GI_{PVDF} demonstrating the impact of seasonal temperature fluctuations on GI, which coexists with GI variations due to potential BS insulation issues. For eye guidance, horizontal lines mark $GI_{BS}/GI_{PVDF} = 1$ (dotted line) and $GI_{BS}/GI_{PVDF} = 0.5$ (dashed line), BS is either PA or FC, (right).

The (non-linear) function, the learned kernel, if trained by given training data points with two input parameters: namely the mean (m) and the kernel function (K) to ensure smoothness. To design the kernel to use with our GPR, we can make assumptions regarding the available data. In particular, a long-term falling trend could be fitted using a Radial basis function (RBF) kernel with a large length-scale parameter (forcing this component to be smooth) describing a pronounced seasonal variation and some minor irregularities. The noise of the dataset can be accounted for with a kernel consisting of an RBF kernel contribution, which aims to explain the correlated noise components such as local weather phenomena, and a WhiteKernel contribution for the white noise. The dataset is split into a train dataset and a test dataset. During fitting the kernel, hyper-parameters are optimized, which results in the learned kernel. Once the kernel function parameters are optimized using a training set, one can make predictions on the test set using the predictive equations. A large length-scale (close to 100) describes an almost linear correlation and a smooth behaviour, while a short length-scale (close to 1) describes non-linear behaviour and captures local irregularities (local maxima or minima may result).

3. Results and discussion

This section is divided into two parts. In the first part, we describe the validation of the learned kernel for Gaussian Process Regression method. The second part is focused on the discussion of the application of the learned kernel for calculation of characteristic values for *GI* time series and comparison of different BSs.

3.1. Gaussian Process Regression – learned kernel and validated model

As the *GI* data vary by an order of magnitude during the studied operation time, we found that GPR performed better with logarithmically transformed data, modelling log *GI* instead *GI*. Furthermore, the dataset was reduced to one quarter of the original data points (3,437 days) to 845 daily values avoiding overstressing the GPR modelling. The hyper-parameters of the learned kernels are listed in Table 2. We interpret the values as follows: the pre-factor (amplitude) explains most of the target signal, by a long-term trend for 5–7.5 *GI*/kΩ with a daily length-scale of 2–5.5 days, a length-scale for *T* of 2–21 °C and a length-scale for *RH* of 2–95%. The white-noise contribution is between 0.03 and 0.15 *GI*/kΩ. Thus, the overall noise level is very small, indicating that the data can be very well explained by the models.

Comparing the different BSs, FC has the smallest values regarding pre-factor and length-scales compared to PA and PVDF. Thus, we expect stronger non-linearity for GI_{FC} than for GI_{PA} and GI_{PVDF}. The longer the length-scale of a component the smoother it is. At that, there is a clear ranking among the length-scales of the components: length-scale (days) < length-scale (*T*) < length-scale (*RH*).

Resulting plots in Fig. 4 show predicted versus experimental data (plotting predicted against ground truth for each data point). The overlaying data points for test and train data sets demonstrate the well trained GPR model, also underpinned by low RMSE values (root mean square error) in Table 2.

The approximate objective functions for predictor pairs provide an insight into bilateral correlations. Fig. 5 shows the distribution of the data points (lightly shaded corridor) and confidence interval for predictions (dark shaded corridor), which is narrow due to the large number of data points. The three materials show different behaviour

Table 2
Learned kernel hyper parameters and metrics.

kernel	Log-marginal-likelihood	RBF				White Kernel	RSME	
		amplitude log*GI* [*GI*/kΩ]	Length-scale time [d]	Length-scale T [°C]	Length-scale RH [%]	White-noise [*GI*/kΩ]	Test-data	Train-data
PA	−97.637	7.47	4.60	20.7	21.6	0.073	0.27	0.27
FC	−464.809	5.16	1.94	2.31	1.8	0.148	0.37	0.40
PVDF	195.708	6.38	5.45	13.5	95.3	0.033	0.18	0.18

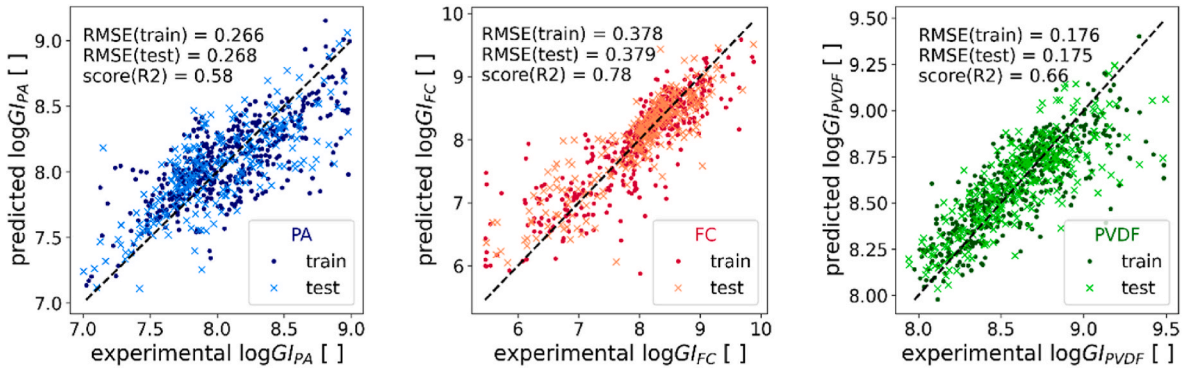


Fig. 4. Result plots for inverters with modules with a single BS-type a) PA, b) FC, and c) PVDF.

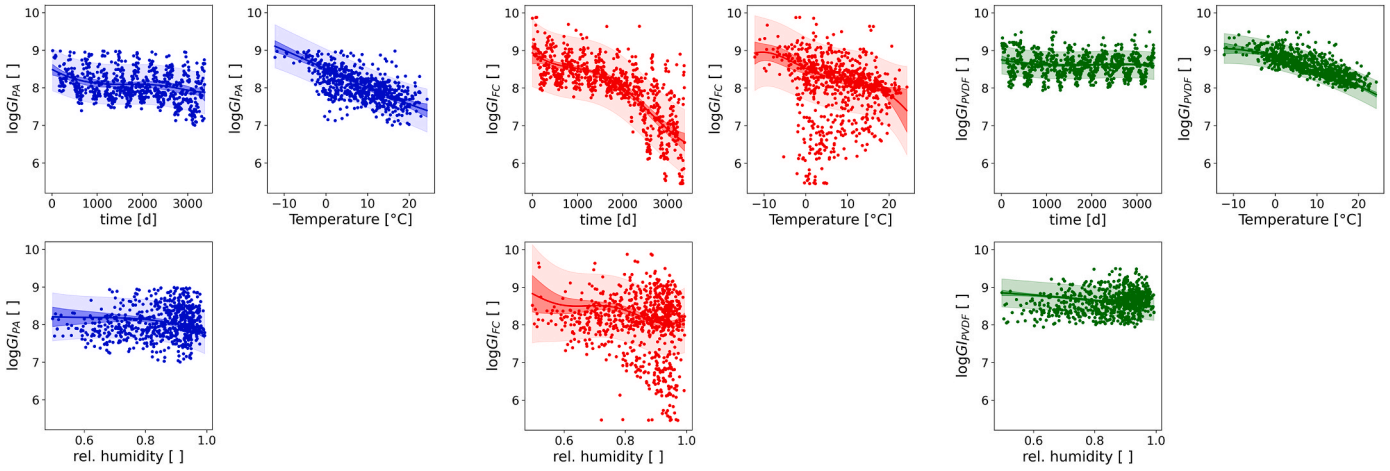


Fig. 5. Approximate objective function for predictor pairs, left) PA, middle) FC, and right) PVDF, showing also a spread of data points (lightly shaded area) and a confidence interval for predictions (dark shaded area).

depending on factors: time, temperature or relative humidity. Log *GI* fluctuates seasonally up and down with time. Temperature dependence shows a negative trend, while at varied humidity log *GI* remains almost constant. Meanwhile PA and PVDF show narrow distributions of the data points, for FC stronger and different dependencies are observed; the amplitude of log *GI* increased with time, the spread of log *GI* values is increased at temperatures between 0 °C and 15 °C and humidity above 80%.

For visualizing the prediction quality, Fig. 6 shows the prediction curves versus the observed data points for specific ambient conditions. Such conditions have been previously identified as anomalous/interesting [3], e.g. at around $T = 5$ °C with $RH = 0.85$. PA and PVDF inverters show almost no changes, stay at a high *GI* level. At the same time, log *GI* decays strongly for FC inverters for all tested conditions after fifth year of operation. Generally, log *GI* of FC inverters are lower than those of PA or PVDF inverters are.

3.2. Predictions with GPR model

The high quality of log *GI* predictions achieved using GPR encouraged and enabled us to study gradients and derivatives for characteristic points. The first derivative describes the velocity or rate *v* of the log *GI* loss (eq (1)). The second derivative (eq (2)) gives the acceleration *a* of the log *GI* loss.

$$v = \frac{d\ log\ GI}{dt} \qquad \text{eq 1}$$

$$a = \frac{dv}{dt} = \frac{d^2\ log\ GI}{dt^2} \qquad \text{eq 2}$$

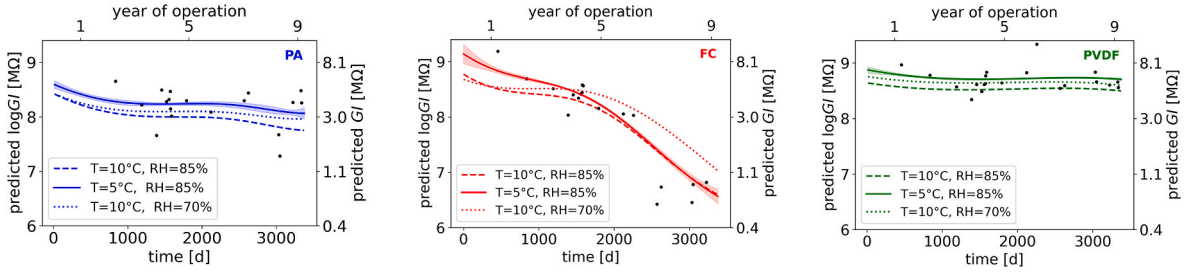


Fig. 6. Dynamics of predicted log *GI* for three specified ambient conditions for inverters with modules with a single BS-type a) PA, b) FC, and c) PVDF. For FC in b) the mean error is marked. Data points are included for the case T = 4–6 °C and RH = 80–90%.

The arrays of curves in Fig. 7, Fig. 8, and Fig. 9 visualize the predicted log *GI* with confidence of predictions and their derivatives (v and a) for selected *RH* and *T* values sorted by BS type. Please note, that the critical *GI* corresponds to log *GI* = 6 [*GI*/kΩ]. Predicted log *GI* differs depending on the BS type. For PA, log *GI* drops smoothly with time, more strongly for increasing humidity and increasing temperature. The temperature effect can be associated mainly with a temperature-determined drift of inverter electronics, more details are still to be collected. The log *GI* loss rate is always negative, $v < 0$, meaning a reduction of log *GI*. Critical values for inverter operation have not been reached, yet.

The curve of predicted log *GI* for FC does not appear as smooth and homogeneous as for PA. As expected from the calculated hyperparameters, there are more local extrema. Furthermore, the confidences are significantly larger for low humidity, most likely due to the low number and at the same time strongly scattered data points for FC. The derivative is always negative and a pronounced minimum is recognizable for $RH > 0.7$. For the first years, there are only small losses in log *GI*. However, after some years the log *GI* loss accelerates strongly and within short time (top graph in Fig. 8) showing a minimum for a (bottom graph in Fig. 8). It should be emphasized that inverters with FC BSs typically fall below the critical threshold for inverter operation.

Finally, for PVDF the arrays of curves are parallel and constantly on the same level. No log *GI* loss is observed in Fig. 9, which is in agreement with v almost zero and a negligible small.

From these results some key values which classify the degradation dynamics can be identified. We calculated three key points which describe certain time points: 1) the turning point t_1 at $v = 0$ and $a = 0$ (after an initial log *GI* loss, log *GI* stabilized on a high level); 2) the onset of log *GI* loss t_2 at $a = a_{\min}$ (the loss rate increases the most, which forms the "knee"-shaped dependence for log *GI*); 3) the point of maximum log *GI* loss t_3 at $v = v_{\min}$, where log *GI* drops at its highest rate.

These times were evaluated for the three BSs and compiled in Table 3. For FC, the shortest times were always calculated. While log *GI* of FC already starts to decrease after 4.9 years, PVDF remains constant and shows first signs of decreasing values only after $t_1 = 5.5$ years. PA follows FC approx. 1.5 years later, whereby the log *GI* loss rate is significantly lower. Only for FC large log *GI* loss rates have been observed so far, reaching their maximum after 7.1 years. For PA maximum loss rate have been reached in the eighth year, but one a much lower level than for FC. Similar results have not yet been observed for PVDF BSs.

Besides the temporal key points, the loss rates are of importance. They determine how fast critical values are reached. Therefore, we evaluated the maximum log *GI* loss rates by calculating the 10% percentiles including only values with originally more than three data points per *RH*-*T*-bins. Fig. 10 shows the log *GI* loss rates with their confidence intervals for three tested BS types. Irrespective of humidity, the log *GI* loss rate of PVDF is small with $v_{\mathrm{PVDF}} = -0.17$ e−3 *GI*/kΩ/d. For PA, it is higher than for PVDF with a maximum at $v_{\mathrm{PA}} = -0.6$ e−3 *GI*/kΩ/d, while FC shows the largest log *GI* loss rates. The value of v depends strongly on humidity and reaches a maximum of $v_{\mathrm{FC}} = -1.6$ e−3 *GI*/kΩ/d at high humidity. In summary, v_{FC} is many times larger than v_{PVDF} and twice as large as v_{PA}, which matches the rates reported using linear regression approach for binned data [13].

High *GI* loss rates and early onset of *GI* loss result in exponentially increasing instances of critical low *GI* values, as reported in [7,14]. Analysing individual inverters in year eight, statistically every FC inverter drops below 400 kΩ for 35 days, with this trend increasing. At the same time, PA inverters show drops only for two days, while PVDF inverters reveal no such instances at all.

In summary, inverters with FC drop and trip alerts, while PA and PVDF inverters show no such behaviour. This is also confirmed by confidential warning reports, indicating that 103 out of 423 inverters fail operation. These 103 inverters include 34 inverters with exclusively FC BSs and 69 inverters including at least one FC module. Therefore, preventive actions are required to ensure the availability of the inverters.

3.3. Discussion

Having the two necessary fundamental prerequisites: 1) a suitable dataset of interest, here the collection of historic ground impedance data, and 2) BS identification performed for all modules connected to these inverters, gives insight into the dynamics of the development of BS-related insulation issues.

First, BSs degrade differently which can be described by two key factors: time point and *GI* loss rate. The onset for *GI* loss is late and the *GI* loss rate is close to zero for PVDF inverters compared to PA and FC inverters. For FC inverters noticeable *GI* loss is already found in 5th year of operation as well as extremely high *GI* loss rate which is much earlier and 2.5-fold stronger than for PA inverters (6.6 a). Thus, strings and inverters including modules with obvious signs of degradation, e. g. open macroscopic cracks in PA and inner micro cracks in FC, can be associated with low insulation resistance. Visually, modules with FC BSs are characterized by crack structures in the inner PP layer. Our

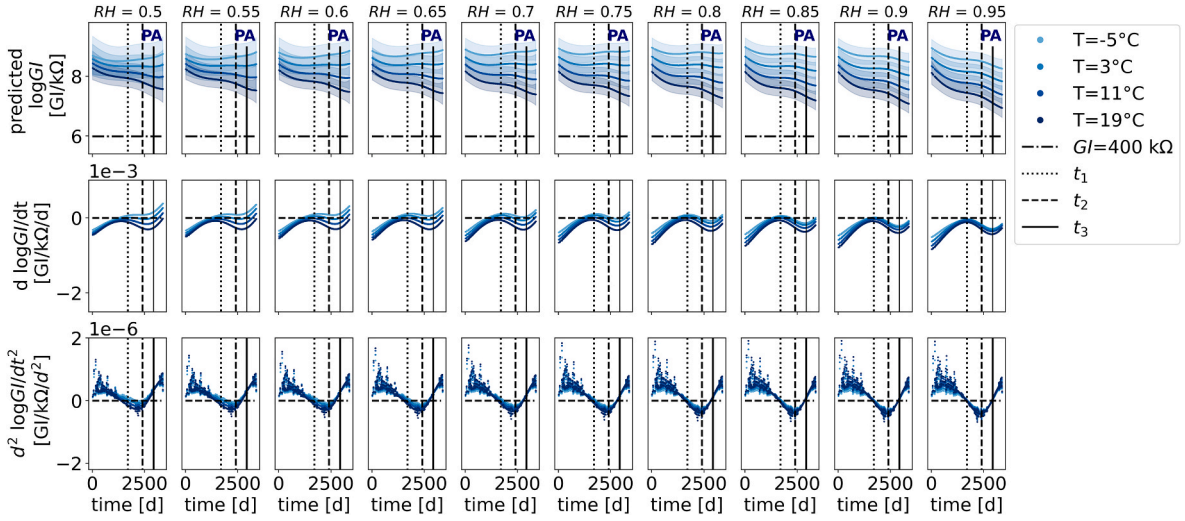


Fig. 7. Predicted log *GI dynamics as a function of T and RH for inverters with modules having PA BS, top) predicted log GI, middle) loss rate v, and bottom) loss acceleration a.*

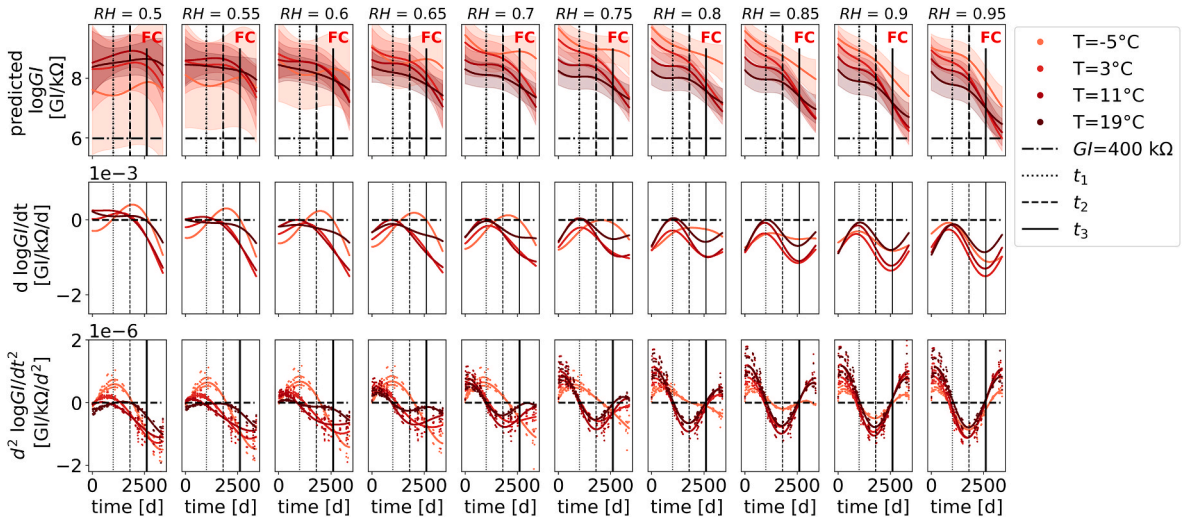


Fig. 8. Predicted log *GI dynamics as a function of T and RH for inverters with modules having FC BS, top) predicted* log *GI, middle) loss rate v, and bottom) loss acceleration a.*

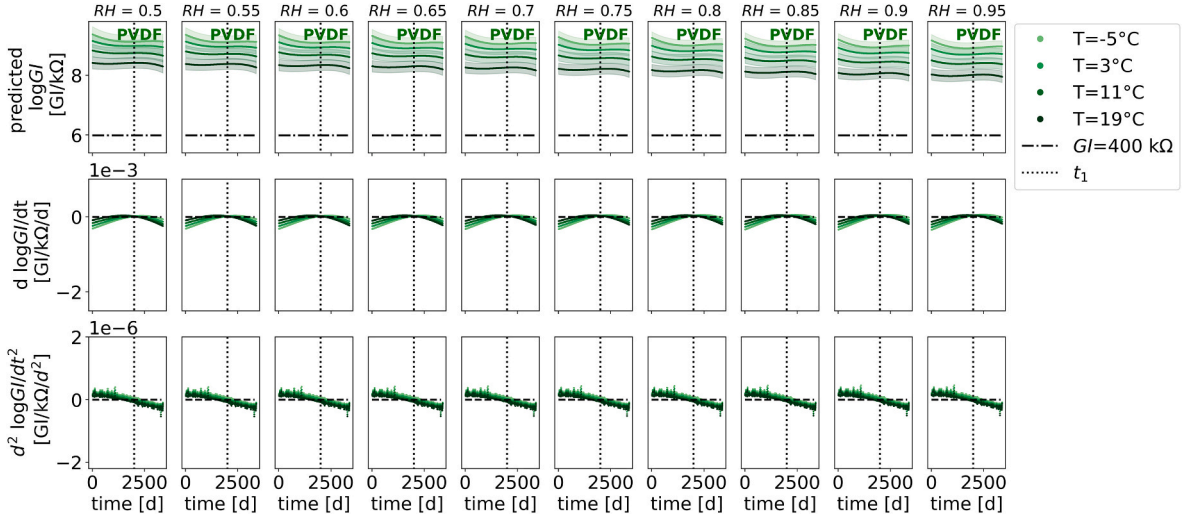


Fig. 9. Predicted log *GI dynamics as a function of T and RH for inverters with modules having PVDF BS, top) predicted log GI, middle) loss rate v, and bottom) loss acceleration a.*

Table 3
Characteristic points of *GI* loss deduced from predicted log *GI* values using GPR.

Time [a]	Turning point	Onset of log *GI* loss	Point of maximum log *GI* loss
	$t_1 = t\,(v = 0, a = 0)$	$t_2 = t\,(a = a_{\min})$	$t_3 = t\,(v = v_{\min})$
PA	4.6	6.6	8.0
FC	2.7	4.9	7.1
PVDF	5.5	Not yet	Not yet

hypothesis is that these cracks support the water transport and leakage current. Especially at low temperatures ($T \approx 5$ °C) and high humidity, there are optimal conditions for water transport; for $T << 5$ °C the mobility of water starts to decrease and approaches zero when water turns to ice. For $T \gg 5$ °C moisture evaporates. Thus, the electrical conductivity is high when the mobility of water inside the BS is high.

Thus, an early onset of *GI* loss and strong *GI* loss rate can result in severe operational inverter issues. The weakest modules in terms of insulation resistance of an inverter determines the inverter availability, connection to the grid, feeding-in and income. Using field-suitable

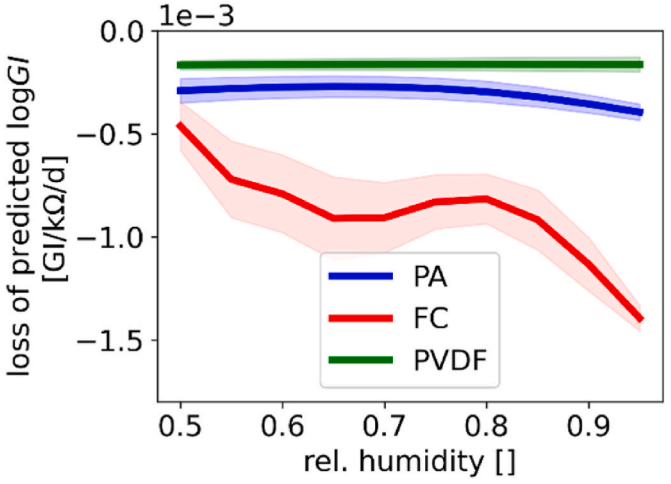


Fig. 10. Strongest log *GI* loss rates determined by 10%-percentiles of loss rate v *for inverters with modules with one BS-type (PA, FC, and PVDF).*

methods, like NIRA, identify the BSs onsite and give insight into the huge variety of used BS materials. Careful, regular visual and optical inspection can be a potential strategy to detect signs of degradation early. Proper inspection allows inverters with more than one module with a critical BS to be identified and flagged in advance, before insulation issues begin and lead to everyday failures. Open is the question of a representative sample (size) to estimate the risk for BS-driven inverter issues reliably. To answer this question, more systematic scientific research of BS degradation at field conditions is needed. This includes also the investigation of minor components, e.g. additives, stabilizers, promoters, which are not identifiable using non-destructive spectroscopic testing, as well as processing parameters. The role of such additives for degradation is not completely understood, yet, because of lack of knowledge of the original composition and the tedious, destructive analysis to investigate reactions of unspecified, minor components in polymers.

Second, the differing impact of different BS materials on the *GI* loss with time can be extracted from monitored daily *GI* values using e.g. GPR. There are key points, as time points and *GI* loss rates, which indicate changes. The three time points can be viewed as three levels of warning that become more severe over time. If operators and analysts watch out for these time points, they have an early warning system and can initiate maintenance actions and prepare adequate countermeasures. Otherwise, troubleshooting is necessary to keep the inverter operating, like finding and disconnecting modules with low insulations resistance electrically from the string until the string becomes too short for inverters operation. Thus, with the right and existing data, here historic *GI* time series, and a trained and tested ML algorithms, investors and operators have a databased method for preventive decision-making.

Third, the study showed that the weather conditions, namely humidity and temperature, are of importance for the *GI* loss. PVDF performs on a constantly high level whereas PA and PVDF show significant *GI* drop at the same operation conditions. Thus, polymers degrade differently in the same climate. As a result, investors should select proper polymers for long-lasting PV modules considering the climate conditions at the installation site in order to fulfil the expected lifetime of 20 years and longer, avoiding the risk of failing within the first half.

Finally, we want to emphasize that the BS types are of statistical relevance for the inverter availability. The fact, that FC inverters drop much more often and stronger than PA or PVDF inverters cannot be considered as a random effect. Being solely an inverter related issue, the *GI* drop should be statistically equally distributed between the different BS types, which is not supported by the above-discussed experimental observations.

4. Conclusion

With a proper dataset, ML techniques become powerful tools for generating meaningful predictive outcome relevant for decision-making. We linked chemical information on polymers in PV module BSs with electrical ground impedance of inverters. The created dataset includes three different BS types with data and time series collected for ca. ten years of operation. The *GI* values predicted using a GPR enabled us to evaluate the characteristic evolution of *GI* with respect to different BSs. The learned kernel allowed deduction of key points important for inverter availability, *GI* loss rates and time points with the resulting ranking $GI(\mathrm{PVDF}) > GI(\mathrm{PA}) > GI(\mathrm{FC})$. For FC, the earliest onset of *GI* loss in year 2.7 years and the highest log *GI* loss rates of -1.6 e-3 *GI*/kΩ/d were predicted from the monitoring data. At that, the onset for GI loss is two years earlier than for PA and three years earlier than for PVDF. The log *GI* loss rate of FC is 2.5 times higher than for PA and 10 times larger than for PVDF. As early signs of degradation can be found in the data, early warnings should have been possible based on the ML predictions. The data analysis points out that GI values drop over time for PA and strongly for FC. Inverter tripping will result. Without countermeasures inverters will fail operation, do not connect to the grid. If inverters do not feed in, income losses are the consequence. This ML approach highlights that key points for degradation exist and that they can be quantified. In addition to degradation rates, we could determine specific time points for degradation. These time points can function as stepwise warning levels to initiate actions in the future. In general, long-lasting BS and durable polymers need to be specified for inverter availability at specific conditions for long-term operation.

CRediT authorship contribution statement

Claudia Buerhop Lutz: Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Larry Lüer:** Methodology. **Oleksandr Stroyuk:** Writing – review & editing. **Jens Hauch:** Funding acquisition, Supervision. **Ian Marius Peters:** Writing – review & editing, Supervision.

Data availability

The data that has been used is confidential.

Acknowledgement

We gratefully thank the German Federal Ministry for Economic Affairs and Climate Action (BMWK) for financial funding of the project and “dig4morE” (FKZ: 03EE1090B) as well as by the Bavarian State Government (project “PVTera—Reliable and cost-efficient photovoltaic power generation on the Terawatt scale” no. 44-6521a/20/5).

References

[1] G.C. Eder, Y. Voronko, W. Mühleisen, C. Hirschl, G. Oreski, H. Sonnleitner, Error analysis of aged modules with cracked backsheets, in: 36th EU-PVSEC, Marseille, France, 2019, https://doi.org/10.4229/EUPVSEC20192019-4BO.13.3.
[2] J. Markert, S. Kotterer, D.E. Mansour, D. Philipp, P. Gebhardt, Advanced analysis of backsheet failures from 26 power plants, in: 38th European PV Solar Energy Conference and Exhibition, Lisboa, Portugal, 2021.
[3] C. Buerhop, O. Stroyuk, J. Zöcklein, T. Pickel, J. Hauch, I.M. Peters, Wet leakage resistance development of modules with various backsheet types, Prog. Photovoltaics Res. Appl. (2021), https://doi.org/10.1002/pip.3481.
[4] DIN/VDE, Selbsttätige Schaltstelle zwischen einer netzparallelen Eigenerzeugungsanlage und dem öffentlichen Niederspannungsnetz, 2013.
[5] O. Stroyuk, C. Buerhop, A. Vetter, J. Hepp, J. Hauch, I.M. Peters, C.J. Brabec, Distinguishing between different types of multi-layered PET-based backsheets of PV modules with near-infrared spectroscopy, Prog. Photovoltaics Res. Appl. (2021), https://doi.org/10.1002/pip.3465.
[6] O. Stroyuk, T. Pickel, T. Winkler, C. Buerhop, J.A. Hauch, I.M. Peters, Field characterization of silicon solar module backsheets by near-infrared absorption (NIRA) spectroscopy, in: 38th EU-PVSEC, Lisboa, Portugal, 2021, 3-936338-78-7.
[7] C. Buerhop-Lutz, T. Pickel, O. Stroyuk, J. Hauch, I.M. Peters, Insulation resistance in relation to distribution of backsheet types in strings and inverters, Sol. Energy Mater. Sol. Cell. 246 (2022), 111913, https://doi.org/10.1016/j.solmat.2022.111913.
[8] A. Omazic, G. Oreski, M. Halwachs, G.C. Eder, C. Hirschl, L. Neumaier, G. Pinter, M. Erceg, Relation between degradation of polymeric components in crystalline silicon PV module and climatic conditions: a literature review, Sol. Energy Mater. Sol. Cell. 192 (2019) 123–133, https://doi.org/10.1016/j.solmat.2018.12.027.
[9] G. Eder, Y. Voronko, W. Mühleisen, C. Hirschl, G. Oreski, K. Knöbl, H. Sonnleitner, Aging induced cracking of polymeric backsheets: analytical approach to identify the drivers, in: IEEE PVSC-46, Chicago, IL, USA, 2019.
[10] P. Gebhardt, L.P. Bauermann, D. Philipp, Bachsheet Chalking - theroretical background and relation to backsheet cracking and insulation failures, in: 35th EU-PVSEC, Brussels, Belgium, 2018, pp. 1097–1100. https://10.4229/35thEUPVSEC20182018-5DO.7.5.
[11] C. Buerhop, O. Stroyuk, T. Pickel, T. Winkler, J. Hauch, I.M. Peters, PV-Modules and their Backsheets - a case study of a multi-MW PV power station, SOLMAT (2021) 231, https://doi.org/10.1016/j.solmat.2021.111295.
[12] Global Modeling and Assimilation Office, MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly, Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed: 2021 April 1st, DOI:10.5067/VJAFPLI1CSIV.
[13] C. Buerhop, O. Stroyuk, T. Pickel, J. Hauch, I.M. Peters, Evolution of inverter ground impedances for PV modules with various backsheet types, Prog. Photovoltaics Res. Appl. (2022) 1–9, https://doi.org/10.1002/pip.3621.
[14] C. Buerhop, O. Stroyuk, J. Zöcklein, T. Pickel, J. Hauch, I.M. Peters, Wet leakage resistance development of modules with various backsheet types, Prog. Photovoltaics Res. Appl. 30 (2022) 938–947, https://doi.org/10.1002/pip.3481.