# Lawrence Berkeley National Laboratory

## Title

Analyzing the impact of design factors on solar module thermomechanical durability using interpretable machine learning techniques

## Permalink

https://escholarship.org/uc/item/2kp131d1

## Authors

Chen, Xin

Karin, Todd

Jain, Anubhav

## Publication Date

2025

## DOI

10.1016/j.apenergy.2024.124462

## Copyright Information

Peer reviewed
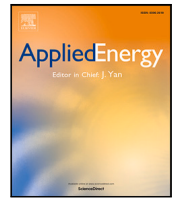
# Analyzing the impact of design factors on solar module thermomechanical durability using interpretable machine learning techniques☆

Xin Chen [a,b], Todd Karin [c], Anubhav Jain [a,*]

[a] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
[b] *University of California, Berkeley, Berkeley, CA, USA*
[c] *Kiwa PVEL, Member of Kiwa Group, Napa, CA, USA*

## ARTICLE INFO

## ABSTRACT

Solar modules in utility-scale systems are expected to maintain decades of lifetime to rival conventional energy sources. However, cyclic thermomechanical loading often degrades their long-term performance, highlighting the importance of effective design to mitigate thermal expansion mismatches between module materials. Given the complex composition of solar modules, isolating the impact of individual components on overall durability remains a challenging task. In this work, we analyze a comprehensive data set that comprises bill-of-materials (BOM) and thermal cycling power loss from 251 distinct module designs to identify the predominant design factors and their impacts on the thermomechanical durability of modules. The methodology of our analysis combines machine learning modeling (random forest) and Shapley additive explanation (SHAP) to correlate design factors with power loss and interpret the model's decision-making. The interpretation reveals that silicon type (monocrystalline or polycrystalline), encapsulant thickness, busbar numbers, and wafer thickness predominantly influence the degradation. With lower power loss of around 0.6% on average in the SHAP analysis, monocrystalline cells present better durability than polycrystalline cells. This finding is further substantiated by statistical testing on our raw data set. The SHAP analysis also demonstrates that while thicker encapsulants lead to reduced power loss, further increasing their thickness over around 0.6 to 0.7 mm does not yield additional benefits, particularly for the front side one. In addition, other important BOM features such as the number of busbars are analyzed. This study provides a blueprint for utilizing explainable machine learning techniques in a complex material system and can potentially guide future research on optimizing the design of solar modules.

## 1. Introduction

Utility-scale photovoltaic (PV) systems are expected to achieve an extended operating lifetime to be competitive with conventional energy sources [1]. However, solar modules installed in the field are subject to multiple environmental stresses such as ultraviolet light, temperature variation, mechanical loading induced by snow, wind, hail [2]. These factors introduce multiple pathways of degradation, reducing the durability of the module. One of the sources of long-term degradation is the cyclic thermomechanical deformation of solar modules caused by temperature variation. Over time, thermal cycling can cause the degradation of components within solar modules, such as interconnections, and lead to a decrease in module power [2,3]. Therefore, it is essential

to identify the potential problems in the current design of solar modules and optimize module robustness to thermal cycling degradation.

Fig. 1(a) illustrates an example of a typical glass/backsheet module. Its multi-layered construction comprises a front frame, a glass layer, polymer encapsulant layers commonly made from ethylene vinyl acetate (EVA) or polyolefin elastomer (POE), a solar cell layer composed of silicon solar cells and copper interconnections, and a backside polymer backsheet made from layers of Tedlar Polyester Tedlar (TPT). The copper ribbons are connected to the cell metallization with solder (SnPb). A defining characteristic of these components is their distinct thermal expansion coefficients (CTE, $\alpha$). This variance in CTEs means that they expand or contract at different rates in response to daily and seasonal temperature fluctuations in their operating environment.

## Nomenclature

| | |
|---|---|
| PV | Photovoltaic |
| BOM | Bill of materials |
| TC | Thermal cycling |
| CTE | Coefficient of thermal expansion |
| PQP | Product qualification program |
| Encaps | Encapsulant |
| Glass2 | Rear glass of glass-glass modules |
| Mono-c | Monocrystalline |
| Poly-c | Polycrystalline |
| EVA | Ethylene vinyl acetate |
| POE | Polyolefin elastomer |
| TPT | Tedlar Polyester Tedlar |
| ML | Machine learning |
| KNN | K-nearest neighbors |
| SVR | Support vector regression |
| RF | Random forest |
| RMSE | Root mean square error |
| SHAP | SHapley Additive exPlanation |
| $\phi$ | Shapley value |
| CI | Confidence interval |
| ANOVA | Analysis of variance |

Previous research [4] showed that the power loss of different models of PV modules after extended thermal cycling tests of up to 600 cycles varied from 0.8% to 14.5%. This degradation is partially due to the CTE mismatch between glass ($\alpha_{glass} \approx 8 \times 10^{-6}$ K$^{-1}$) and Si cells ($\alpha_{cell} \approx 2.5 \times 10^{-6}$ K$^{-1}$). The cells adhere to the glass by encapsulants, so the CTE mismatch causes a non-uniform in-plane displacement of cells and glass [5]. The gap between cells can vary with different temperatures, as shown in Fig. 1(b). Such temperature-driven variations over the lifetime of a module can result in cyclic deformation of the copper ribbons that connect cells and cause interconnection fatigue [5–7]. Additionally, fragments in cracked cells can shift due to temperature variation, leading to wear and tear at metal contacts [8]. Another prevalent degradation mode caused by temperature variation is solder disconnection. The CTE mismatch between the copper ($\alpha_{Cu} \approx 16.7 \times 10^{-6}$ K$^{-1}$) ribbon and Si cell ($\alpha_{Si} \approx 2.5 \times 10^{-6}$ K$^{-1}$) triggers non-uniform deformations in these layers [9]. As depicted in Fig. 1(c), at elevated temperatures, Si and Cu experience tensile and compressive stresses, and at low temperatures, the stresses are reversed, which causes a periodic change of shear stress within the solder layer [9]. This cyclic loading in the solder often culminates in solder disconnection [10–12]. Solder degradation is one of the failure modes that occurs in the early stage of module operation, which can be probed by 600 cycles of thermal cycling accelerated aging tests [4].

As shown above, thermal cycling degradation involves multiple modes, strongly impacted by the specific bill-of-materials (BOM) of solar modules. This encompasses the dimensions and material properties of each module component. Understanding which design factors predominantly affect thermal cycling power loss is of great significance in guiding future research on module optimization. Several previous studies have investigated the impact of various design factors. Bosco et al. [13] performed a regression analysis to identify design factors that may influence the degradation of the solder by simulating the accumulated damage to the solder joints during thermal cycles. The top sensitive factors are the thickness of the solder layer, Cu ribbon, and Si wafer layer. Park et al. [9] also confirmed that reducing the thickness of the Si cell and copper ribbon can increase the lifetime of the solder in thermal cycling by simulation. Zhu et al. [14] fabricated different mini-modules and did a simulation to investigate the effect of viscoelasticity of encapsulant materials on the solder joint fatigue. They found that modules with encapsulants of higher viscous properties presented more power loss after the thermal cycling test. Beinert et al. [15] found that increasing cell size and changing full cells into half-cut cells could decrease thermal stress on the cell layer using simulation. They further qualitatively concluded that cell thickness, encapsulant CTE and glass/backsheet CTE strongly influence the stress in the cell layer [16]. Hanifi et al. [12] also found that the higher rigidity of encapsulant materials and changing full cells into half-cut cells could mitigate ribbon fatigue during temperature variation. As exemplified by this past research, most sensitivity analysis work in this field relies on simulated data. Computational constraints often force researchers to model simplified structures, such as a single cell rather than a full-size solar module, or ignore the busbars, which may bypass some real-world effects.

Previous research [13] correlated module design with durability using a linear model to explore the effects of these design factors. Typically, the standard regression analysis with linear model imposes prerequisites on the data set, such as multivariate normality, homoskedasticity in the data. In addition, the linear model requires the linear relationship between the independent variables and the target variable; otherwise transformations of the independent variables such as logarithm and reciprocal transformation [13] are necessary to capture the non-linear relationship. On the other hand, machine learning (ML) models are mostly nonparamatric and present an outstanding ability to discover the underlying pattern in the data. While the enhanced predictive accuracy of advanced ML models is commendable compared to linear models, a significant challenge emerges in their interpretability. Several methodologies have been proposed to solve this trade-off between accuracy and interpretability. Noteworthy among these are Partial Dependence Plot (PDP) [17], Local Interpretable Model-agnostic Explanations (LIME) [18], and SHapley Additive exPlanations (SHAP) [19]. In particular, SHAP is able to do both global and local interpretation, and its adoption across diverse scientific disciplines is a testament to its efficacy [20–23]. These model-agnostic methodologies facilitate data inference with more predictive machine learning models beyond traditional linear models, whilst preserving interpretability.

This study identifies the predominant BOM features that affect the thermomechanical durability of modules and explores their impacts using real-world data. Our data set comprises full-size module data from the industry that includes both BOM features and the corresponding power loss after 600 standard thermal cycles [24]. Since these are existing records from the industry rather than from a well-designed data sampling process, we propose using nonparametric machine learning models to correlate the BOM features with the power loss and then applying the model-agnostic method to interpret the black-box models. By comparing various ML models, we develop a random forest (RF) model [25] with a testing root mean square error (RMSE) of 1.179%. Subsequently, we employ the SHAP method to interpret the developed model, thus identifying the predominant BOM features as Si type, encapsulant thickness, number of busbars, and wafer thickness. Also, with SHAP analysis, we elucidate the impacts of these top important factors on the thermomechanical durability. In the end, we apply statistical testing on the original data set to verify our conclusions.

To sum up, the schematic workflow of this project is demonstrated in Fig. 2(a)–(e). The principal contributions of this study are summarized as follows:

(1) This work identifies that Si type, encapsulant thickness, busbar numbers and wafer thickness are the predominant BOM features that influence module's thermomechanical durability.

(2) We find that using mono-c Si cells presents lower TC power loss (0.6%) than poly-c Si ones, and thicker encapsulant over 0.6–0.7 mm does not bring additional benefits to the durability.
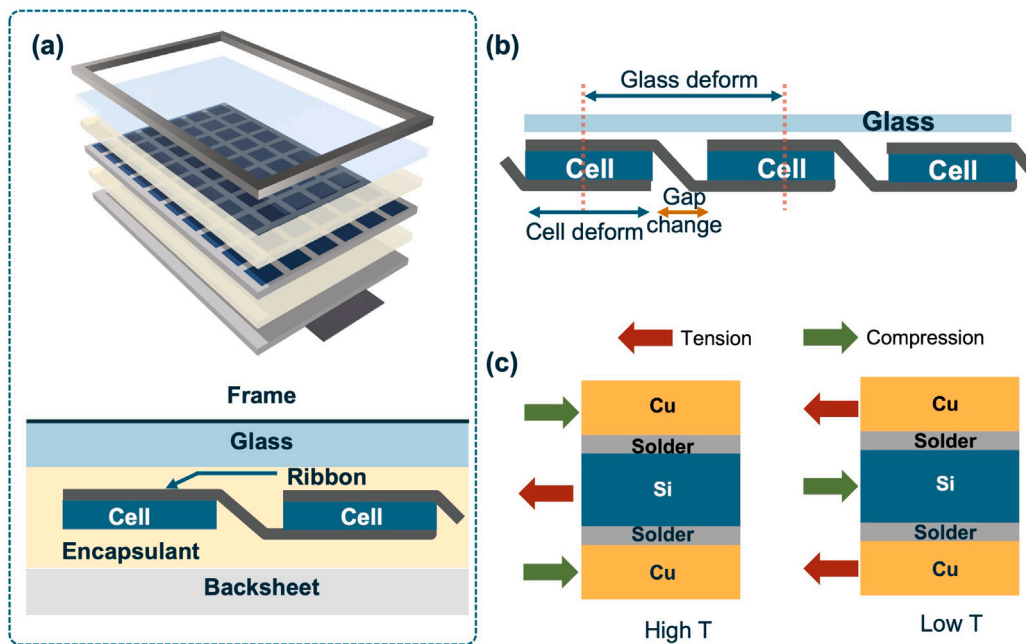
**Fig. 1.** Schematic structure of a glass/backsheet PV module and the mechanism of its thermomechanical degradation. (a) The layered structure of a PV module. (b) Mechanism of the cyclic deformation of ribbons. (c) Mechanism of cyclic thermal stress in solder layer.
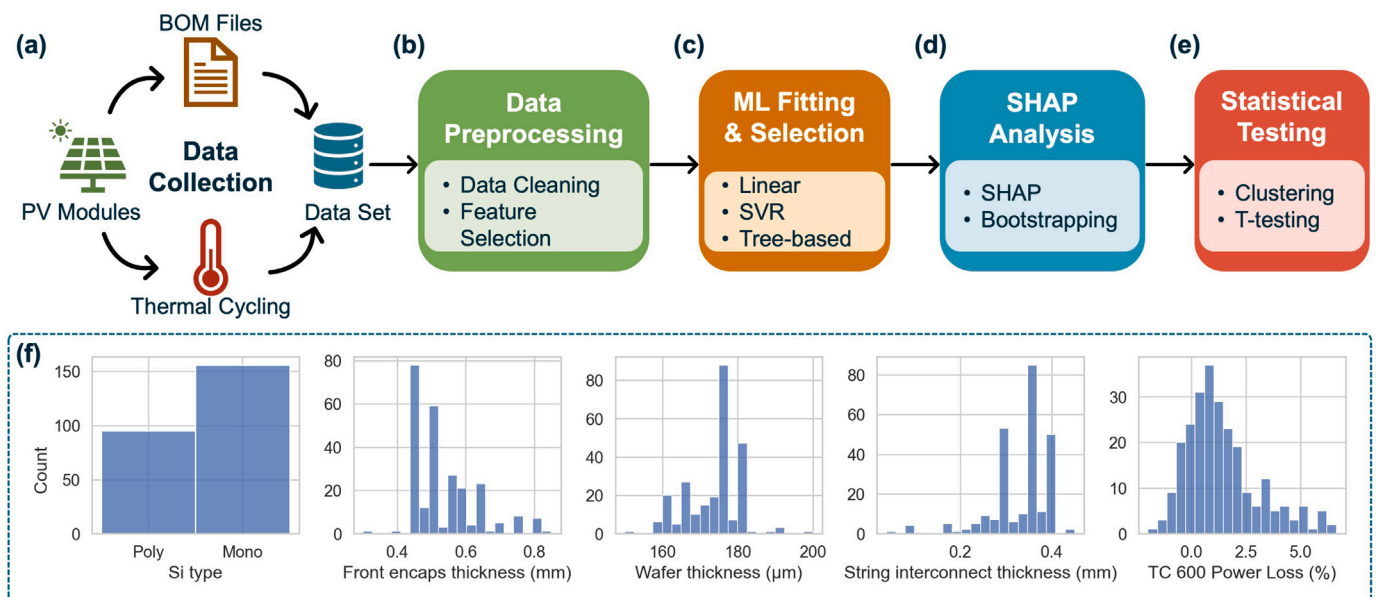


**Fig. 2.** The workflow of the study and data distribution. (a)–(e) illustrate the steps performed in this study. (a) Data collection. The data set is constructed by manually extracting information from BOM files and measuring the power loss under thermal cycling tests. (b) Data preprocessing. The data set is then cleaned, and the corresponding features are selected to build machine learning models. (c) ML model selection. Multiple models such as linear models, support vector regressor (SVR), and tree-based models are fitted and compared to correlate the BOM features with the TC 600 power loss. The best-performing model is selected for interpretation. (d) SHAP analysis. SHAP method is then used to interpret this optimal model to determine the predominant design factors and their impacts on the power loss. The uncertainty of Shapley values is quantified using bootstrapping. (e) Validation of the interpretation. In the end, post hoc statistical testing is used to validate the interpretation. Clustering method is used to regroup the data set and t-testing is used to test the statistical significance. (f) Distribution of the selected feature values and the power loss of each module. The distribution of other features is shown in the SI.

(3) We propose a novel workflow to the field of understanding the complex correlations between BOM data and module durability. The methodologies combine machine learning modeling (RF model), model-agnostic interpretation method (SHAP), and post hoc validation with statistical testing.

(4) We publish the processed BOM data and codes in our work. This data set can be beneficial for future research on improving

the durability of PV modules considering the difficulty in tracking the BOM features. Some values in this public data set are redacted due to non-disclosure agreement.

The rest of this article is organized as follows: Section 2 illustrates the details of thermal cycling setup and computation methodologies, including data collection, the feature matrix constructed by feature

selection, ML modeling, SHAP interpretation and statistical testing. Section 3 reports the results derived from the application of these methods. We first compare the performance of ML models and report the generalization of the optimal RF model. After that we rank the importance of each BOM features and demonstrate the impact of predominant features on the durability by interpreting the aforementioned RF model with SHAP analysis. We further discuss the underlying mechanisms of the connections between BOM and the module durability. In the end, we validate the interpretation of the ML model using statistical testing. Section 4 summarizes this paper and provides further insight into the current module design and the durability of modules.

## 2. Methodology

All experimental data were collected at PVEL [26]. All codes and models compiled in this research were written in Python3 [27] and executed on a MacBook Pro (Apple M2 Pro chip, 16 GB memory).

### 2.1. BOM data collection

The data comes from PVEL's product qualification program (PQP) [28]. 275 full-size commercial solar modules of different designs were collected from 47 different manufacturers by PVEL. These modules are estimated to cover around 70% various types of designs in the market by the date of collection. The BOM data were extracted manually from the document files provided by these manufacturers and yielded over 100 features, including supplier information, dimensions and configurations, materials of PV components such as cells, connections, encapsulants, backsheets and glasses. In this study, we select certain features that are essential for machine learning modeling based on the steps described in the following Section 2.3.

### 2.2. Thermal cycling test

The thermal cycling test is carried out at PVEL according to the standard outlined in IEC 61215-2:2016 [24]. Each module is placed in an environmental chamber and subjected to temperature cycles from $-40\,°C$ to $+85\,°C$. There are 600 thermal cycles, and each cycle contains a maximum $100\,°C/hr$ ramp rate and a minimum dwell time of 10 min. Each complete cycle takes around 6 h. A current–voltage (IV) flash test (Pasan SunSim 3B) is performed under standard test conditions following IEC 60904-1:2006 [29] before the aging test and after every 200 cycles to measure the IV curve of the modules.

### 2.3. Data preprocessing and feature selection

The final feature matrix **X** we constructed to train the machine learning models contains 251 modules with 22 BOM feature columns and the target variable **Y** is the power loss (%) after 600 thermal cycles. The distribution of some selected BOM features, along with the power loss after 600 cycles of thermal cycling test (TC 600), is plotted in Fig. 2(b). The complete distribution of the 22 features can be found in the SI Section "Feature Selection". The power loss is defined in Eq. (1).

$$\Delta P(\%) = \frac{P_{init}^{max} - P_{after}^{max}}{P_{init}^{max}} \times 100\% \tag{1}$$

where $P_{init}^{max}$ is the module's maximum output power before the aging test and $P_{after}^{max}$ is the maximum output power after 600 thermal cycles.

To construct this feature matrix (*i.e.,* the input X), we preprocess the raw data by data cleaning and feature selection. Data cleaning includes outlier detection using residual analysis, data type casting, and missing data handling using a K-nearest neighbors (KNN) imputer [30]. The outlier detection step removes 24 modules that do not have TC 600 power loss data or have large residuals (over 97% quantile) in the residual analysis, as demonstrated in the Supporting Information

(SI) section titled "Data Preprocessing". Detailed descriptions of these preprocessing steps are also included in this SI section.

Subsequent to the data cleaning process, we select features from more than 100 BOM attributes in the raw data set. The goal of feature selection is to identify features with potential correlations to the target variable among other BOM features in the original data set and reduce inter-feature dependencies because high correlation among features makes the interpretation of feature importance less accurate and less stable. This selection is primarily guided by domain-specific knowledge. According to previous research [13,15,16], the features that influence the stress distribution in solar modules include the thickness, width, and length of each layer and the interconnection thickness, so these features are included in the feature matrix. Mechanical properties such as viscoelasticity of the polymer encapsulant may also influence the durability of the module, but these properties are difficult to track in the BOM data and are not discussed in this work. We also perform exploratory data analysis using a correlation matrix to further test the association among features and between features and power loss. These associations are calculated using the Python package Pingouin [31] and Scipy [32].

We employ different statistics to measure the associations for different types of variables. We use Spearman correlation to test the correlation between two numerical variables $X$ and $Y$. For instance, Fig. 3(a) is a correlation matrix that shows some highly dependent features related to module dimensions, so these features apart from glass length and width are excluded from the feature matrix to mitigate multicollinearity. We explain the impact of dependent features on feature importance in the SI section "Feature Selection". Furthermore, we select features that may not be explored in previous regression analyses [9,13,16] but still show correlation with thermal cycling degradation based on the correlation matrix in Fig. 3(b). For example, the number of busbars was not considered in the previous simulation [13] to save computation time, but we still include this feature in our models because this feature shows a high Spearman correlation in this data exploration process. It should be noted that the Spearman correlation is applicable to numerical variables. Regarding other types of variables, we use the Chi-square test [33] to test the statistical association among categorical features, and Analysis of Variance (ANOVA) [33] to test the statistical association between numerical features and categorical features as detailed in the SI Section "Feature Selection". We note that the dependence computed here is only used to help construct the feature matrix and does not necessarily reflect the true correlation between features and power loss due to the interaction of features. Therefore, further analysis like statistical analysis or machine learning modeling is required to determine correlations or feature importance.

### 2.4. ML models

In this study we compare multiple regression models, including linear (Lasso and ridge regression), support vector regression, and tree-based (RF and XGBoost) models to find the model that minimizes prediction error. Linear models act as a baseline for model comparison. We use k-fold ($k = 5$) cross-validation to test the robustness. The hyperparameters of these models are tuned by using grid search, as described in the SI. In all cases, the model performance is assessed using root mean square error (RMSE) between the measured values and the predicted values. We build ML models using toolkits from the Scikit-learn package [34] to model the correlation between BOM features and the power loss. The XGBoost model is not originally included in Scikit-learn and is constructed using the open source XGBoost package [35]. The split of the development set, the model evaluation metric, and the description of model architectures are illustrated as follows:

#### 2.4.1. Data splitting and model evaluation

The data set is split into the training set and the testing set with a splitting ratio of 8:2. The training set is further divided into five
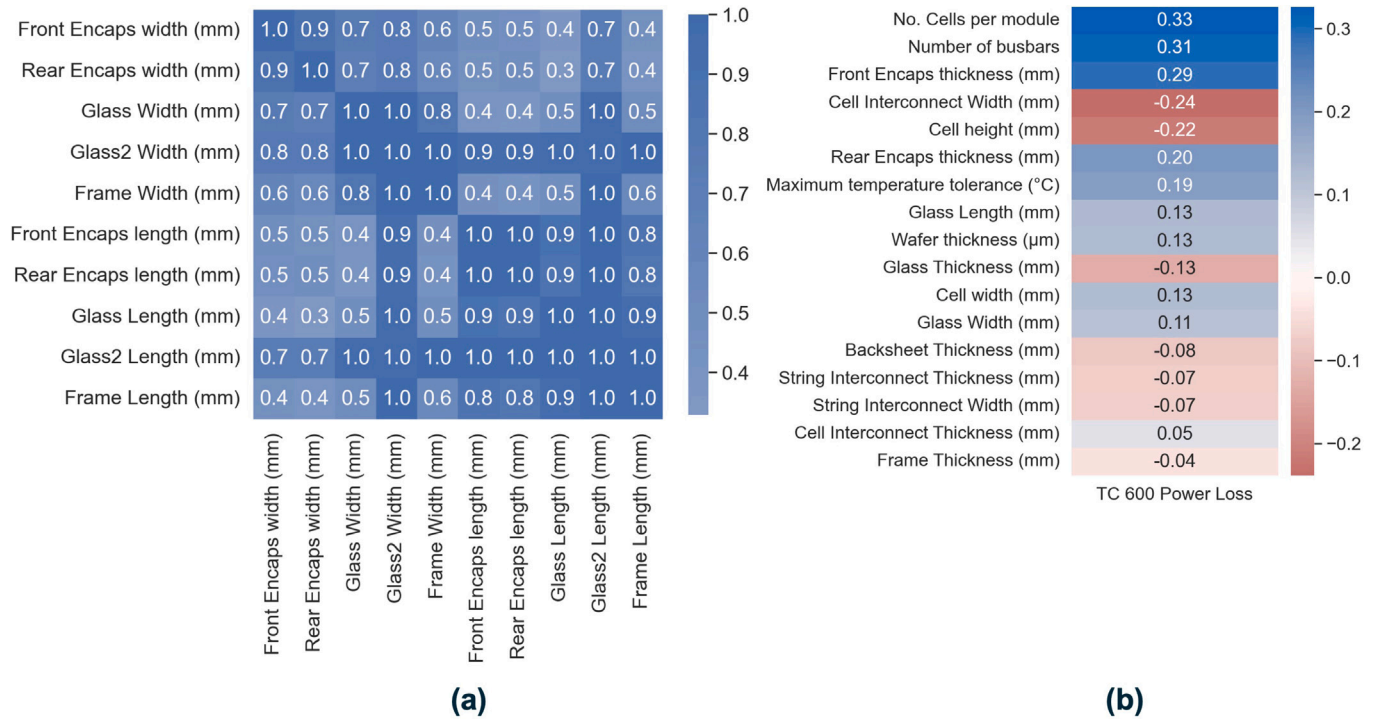
**Fig. 3.** The Spearman correlation between numerical features and the target variable is calculated to assist the construction of the feature matrix in combination with domain knowledge. The correlations shown here are solely used as a reference for feature selection and do not necessarily reflect the true correlations or importance since multiple variables may be entangled. (a) Correlation matrix of numerical features related to module dimensions. (b) Spearman correlation score between selected features and power loss. Features with high correlation coefficient such as the number of busbars are used to construct the feature matrix, although they were not explored in some previous analyses [13,16]. The full correlation matrices of numerical features as well as categorical features are shown in the SI.

parts for 5-fold cross-validation to compare models and do hyperparameter tuning. In each validation iteration, four parts are used for training and the rest part is used for validation. The performance after cross-validation is the average performance from each iteration. Furthermore, the hold-out testing set is used to evaluate the generalization of the optimal model picked from the cross-validation process. The root mean squared error (RMSE) used to evaluate model performance is defined in Eq. (2).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \hat{y}_i - y_i \right)^2} \tag{2}$$

where $n$ is the number of data points, $\hat{y}_i$ is the prediction for the data point $i$ and $y_i$ is the ground truth of $i$. The RMSE shares the same unit (%) as the power loss.

### 2.4.2. Linear model

The generalized formula of a linear model follows Eq. (3):

$$\mathbf{y} = \mathbf{Xw} \tag{3}$$

where $\mathbf{y} \in \mathbb{R}^n$ is the target vector, $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the feature matrix, and $\mathbf{w} \in \mathbb{R}^d$ is the model weight. $n$ is the number of measurements and $d$ is the number of features. The training aims to find the $\mathbf{w}$ that minimizes the loss between the measured values and the predicted values, as shown in Eq. (4).

$$\hat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \|\mathbf{y} - \mathbf{Xw}\|_2^2 + \lambda R \tag{4}$$

where $\lambda R$ is the regularization term with the hyperparameter $\lambda$ to prevent overfitting. Lasso regression [36] with $L1$ regularization ($\lambda \|\mathbf{w}\|_1$) and Ridge regression [37] with $L2$ regularization ($\lambda \|\mathbf{w}\|_2^2$) were trained. The equations for these regularization terms are shown in the SI Section "ML Modeling".

### 2.4.3. SVR

Support vector regression [38] finds a hyperplane with margins that minimize the error between the true value and predicted values. The equations for SVR is shown in the SI Section "ML Modeling".

### 2.4.4. Tree-based model

Random forest and XGBoost are trained in this research. Random forest regression builds an ensemble of decision trees on different subsets of the original data using bootstrapping and splits nodes on a random subset of features, helping to increase robustness and prevent overfitting. The final prediction result is the average of the predictions of each decision tree. XGBoost uses boosting method by building decision trees sequentially, each trained to correct its predecessor's errors. Both tree-based models can be regularized by limiting the depth of trees or the number of nodes. XGBoost can also be regularized using $L1$ or $L2$ regularization in gradient boosting.

### 2.5. SHAP analysis

To interpret ML models and understand the correlation between BOM features and power loss, we use SHAP method to explain the ML models and apply bootstrapping to quantify the uncertainty of this method. SHAP analysis reveals the impact of a certain feature on model prediction while marginalizing the effect of other features. A prediction (*i.e.,* power loss) of the model can be considered as the sum of the contributions (*i.e.,* Shapley values) of each BOM feature. The average contributions of each BOM feature represent the feature importance, and the dependence between feature values and Shapley values reflects the feature impacts on the power loss.

### 2.5.1. SHapley additive explanations

SHAP is a model-agnostic method that interprets each feature's marginal contribution to a specific prediction of the machine learning model. It can interpret both local prediction and global contributions

of each feature via the additive method. SHAP measures the marginal contribution by computing the Shapley value of each feature ($\phi_j$), defined as:

$$\phi_j = \sum_{S \subseteq \{1,\ldots,p\} \backslash \{j\}} \frac{|S|!(p - |S| - 1)!}{p!} (\text{Val}(S \cup \{j\}) - \text{Val}(S)) \quad (5)$$

where $S$ is a subset of the features used in the model, and $p$ is the number of features. $\{1, \ldots, p\} \backslash \{j\}$ represents the set without feature $j$. $\text{Val}(S)$ is the prediction for feature values in set $S$ that are marginalized over features not included in set $S$.

A prediction can be interpreted as the sum of Shapley values of each feature, as shown in Eq. (6), where $\hat{f}(\boldsymbol{x})$ is the prediction, $E_X[\hat{f}(X)]$ is the expectation of the prediction, $\phi_j$ is the Shapley value (*i.e.,* contribution) of the feature $j$, $\boldsymbol{x}$ is a feature vector that contains feature values in this prediction. Eq. (6) illustrates that each Shapley value is correlated with the prediction, so the impact of a feature on its Shapley value also reflects its impact on the target variable.

$$\hat{f}(\boldsymbol{x}) = E_X[\hat{f}(X)] + \sum_{j=1}^{p} \phi_j \quad (6)$$

### 2.5.2. Bootstrapping

The uncertainty of the Shapley value is determined by using Bootstrapping [39], which randomly sampled with replacement from the original data set 1,000 times and Shapley values of each feature are computed in each iteration. This results in the distribution of the Shapley value of each feature. The confidence interval (CI) of the Shapley value is quantified by using a 95% confidence level.

### 2.6. Post hoc statistical analysis

Utilizing SHAP analysis allows us to understand the most influential BOM features and their impacts on power loss. However, the validity of SHAP interpretation is contingent upon the performance of the machine learning models, necessitating supplementary investigations. The straightforward method of validation is to conduct an independent post hoc statistical testing on the power loss in the original data set. For instance, we can directly compare the power loss of poly-c Si modules and mono-c Si modules to test which Si type is more durable. However, since our data set comprises modules from various manufacturers, other BOM features are not controlled when we compare different Si types. Therefore, we need to reorganize the original data set. The steps include regrouping the original data set to generate control groups using the clustering method and subsequent statistical testing.

### 2.6.1. Clustering

K-means [40] implemented with Scikit-learn is used to do the clustering, which partitions the data into a predefined number of clusters ($K$) by iteratively assigning each data point to the nearest centroid, and then recalculating the centroids as the mean of all the points in the cluster until the centroids stabilize or a maximum number of iterations is reached. The "elbow" method is used to determine the optimal number of clusters to be used. Before clustering, the feature matrix is standardized following Eq. (7) and weights are assigned to each feature based on their importance following Eq. (8)

$$Z = \frac{X - \mu}{\sigma} \quad (7)$$

where $Z$ is the standardized feature matrix, $X$ is the training feature matrix, $\mu$ and $\sigma$ are the training samples' mean and standard deviation.

$$\text{feature}_i^* = \text{feature}_i \times \frac{\text{SHAP}_i}{\sum_j \text{SHAP}_j} \quad (8)$$

where $\text{feature}_i$ is the standardized value of feature $i$ and $\text{SHAP}_i$ is the mean absolute Shapley value of feature $i$.

**Table 1**

Mean value and standard deviation of RMSE of each model during the 5-fold cross validation.

| Model | RMSE_mean (%) | RMSE_std (%) |
|---|---|---|
| Lasso (Linear) | 1.614 | 0.156 |
| Ridge (Linear) | 1.631 | 0.159 |
| SVR | 1.657 | 0.195 |
| RF (Tree-based) | 1.427 | 0.247 |
| XGBoost (Tree-based) | 1.491 | 0.133 |

### 2.6.2. T-test

T-test is conducted using the Pingouin package [31]. The *t*-value for the group *a* and *b* is defined in the SI Section "Post hoc Statistical Analysis".

## 3. Results and discussion

### 3.1. ML modeling

In this work, we train a random forest model by taking the constructed BOM feature matrix as input and the TC 600 power loss as target variable. This model shows a validation RMSE of 1.427% which outperforms other models in our study and a testing RMSE of 1.179%, which illustrates the generalization of this model. It should be noted that the unit (%) of RMSE is the same as the power loss. Before selecting this model, we compare several ML models as illustrated in Section 2. Table 1 and Fig. 4(a) compare the mean value and standard deviation of RMSE of each model during the 5-fold cross-validation. In general, tree-based models outperform other models, with RMSE validation of 1.427%(±0.247%) for RF and 1.491%(±0.133%) for XGBoost. This is unsurprising because linear models do not capture the nonlinear relationships between the BOM features and the power loss, and SVR tends to develop severe overfitting as shown in Fig. 4(b). Considering the low mean error and simpler implementation, we select RF as the optimal model for subsequent SHAP analysis. It should be noted that although the difference in RMSE values between tree-based models and other models may seem marginal in Table 1, given the low target variable values, even minor RMSE variations can signify substantial prediction performance disparities. This significant difference in performance among models can be visualized in the scatter plots of predicted values *vs.* measured values. For example, the plot for Lasso model in Fig. 4(b) clearly shows underfitting since the scatter points are not distributing along the "Measured=Predicted" line (the diagonal dash line). The scatter plots for all the models are demonstrated in the SI Section "ML Modeling". In addition, to mitigate the influence of large differences in ranges of features in the linear model, especially when these features have different units, we also train linear models on a standardized data set following the standardization equation (Eq. (7)). However, the performance of the linear model is not varied (details included in SI). We note that tree-based models are not influenced by this difference between feature ranges and do not require standardized data set.

After comparing the RF model with other models, we further test its generalization on the hold-out 20% of the data set (*i.e.,* testing set), as shown in Fig. 4(c). The RMSE (1.179%) of the hold-out testing data is not drastically higher than the training error, as opposed to the overfitted SVR model. This indicates that the fitted RF model is able to capture the underlying correlation between the BOM data and the degradation and has generalization to new data. Following model selection, we retrain the optimal RF model on the combined data set of both training and testing set, achieving a RMSE of 0.558%, and proceed to interpret this model using the SHAP method.
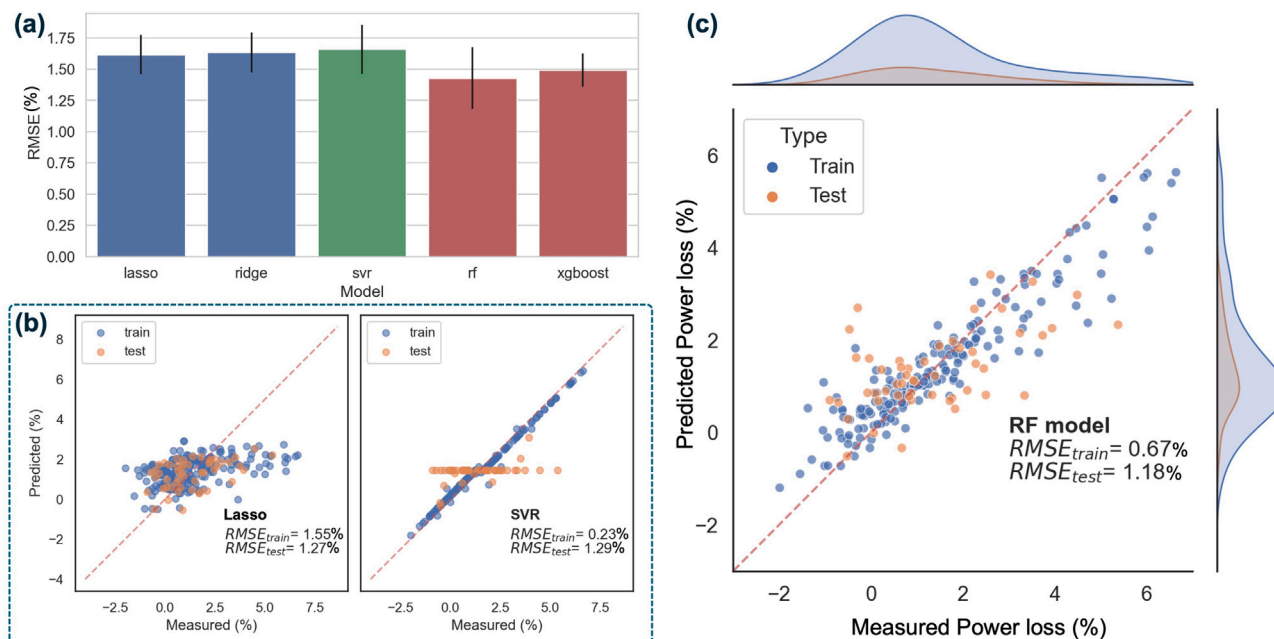
**Fig. 4.** Comparison of machine learning models and generalization test of the optimal model. (a) The 5-fold cross-validation RMSE score for each model. The error bar shows the standard deviation of scores in each fold. Tree-based models (RF and XGBoost) present the best performance. (b) The performance of the Lasso and SVR on the training and testing set, which present underfitting and overfitting. (c) Performance of the selected RF model in the hold-out testing set. Relatively low testing RMSE compared to the training score shows that the RF model is not severely overfitted and generalized well on the new data set.

### 3.2. SHAP analysis for ranking BOM feature importance

To investigate the importance of features related to power loss, we use the SHAP method to interpret the RF model demonstrated in Section 3.1. With this method, we find that BOM features such as Si type, encapsulant thickness, busbar numbers and wafer thickness remain the top important factors, as shown in Fig. 5(a) which ranks the average impact of each feature on the prediction in the order of decreasing mean absolute Shapley values. In general, Fig. 5 demonstrates that "Si type (poly=1)", which represents whether the solar cell is mono-c (denoted as 0) or poly-c (denoted as 1) Si, is listed as the most important feature. Also, this feature still remains the most important even when the lower bound of the confidence interval (shown as the error bars) is considered. This indicates that the Si type is the predominant design factor that impacts thermal cycling power loss. The subsequent influential features are "Number of busbars", "Wafer thickness", "Front encaps thickness" and "Rear encaps thickness". If we consider the thickness of the front and rear encapsulant together due to their similarity and add up their Shapley values, the encapsulant thickness is the next important BOM feature other than the Si type as shown in the red dashed line in Fig. 5(a).

To quantify the uncertainty of Shapley values, the 95% confidence intervals are shown as the error bars in Fig. 5(a). The narrow interval indicates the robustness of this method to random selection in the training data. It should be noted that Si type and encapsulant thickness are the two top important features with robustness, but the importance rank of other features can vary according to the confidence interval in Fig. 5(a). Fig. 5(b) shows an example of the bootstrapping distribution of the Shapley value of the "Number of busbars", which follows a normal distribution consistent with the central limit theorem [33].

### 3.3. SHAP analysis for interpreting impacts of BOM features on power loss

To understand the impact of varying a specific BOM feature on the power loss, we also examine the Shapley value of each feature in each measurement, as shown in Fig. 5(c). We will separately discuss the detailed interpretation of the impacts of each top important feature and

possible mechanisms in the next few paragraphs. Due to the additive property of the SHAP formula shown in Section 2, the power loss of a module can be considered as the sum of Shapley values of each feature, and thus the dependence between the feature value and the Shapley value reflects the relation between the BOM feature and the power loss. For instance, the "Si type" row in Fig. 5(c) illustrates that poly-c Si modules (red clusters) possess positive Shapley values, which implies a tendency towards higher power loss for such modules.

The dependence between BOM features and power loss is clear for categorical variables such as "Si type" in Fig. 5(c), but it becomes challenging to understand this dependence for other numerical features in this plot. To better visualize this dependence between BOM features and power loss for numerical features, we also construct Fig. 6 based on Fig. 5(c) to illustrate the relationship between the feature values and their corresponding Shapley values. The Shapley values ($< 0.1$) of other features are not significant compared to these primary ones and are not discussed in detail in this paper. Fig. 6(a)–(d) show the dependence plot of the top five important features besides the Si type and Fig. 6(e)(f) demonstrate two features that were investigated in previous studies [9,13].

#### 3.3.1. Impact of Si type

Based on the SHAP interpretation, Si type is the predominant design factor regarding the module's thermomechanical durability and poly-c Si modules present higher power loss. Fig. 5(a) illustrates that using poly-c Si instead of mono-c Si increases the Shapley value by around 0.6 on average, which indicates an increase of 0.6% in the power loss. We note that this value only quantifies the impact of using poly-c Si, but because of the noises in our data set and model, the change in power loss after replacing mono-c Si with poly-c Si may vary in reality. To the best of our knowledge, this difference in durability between mono-c and poly-c Si modules was not reported previously. Since mono-c and poly-c Si have similar CTE [41], the variance of power loss may not be caused by the difference in thermal expansion. One plausible reason involves the grain boundary in poly-c Si that serves as the initiation sites of micro-cracks [42]. These micro-cracks can propagate with successive thermal cycles, leading to a degradation
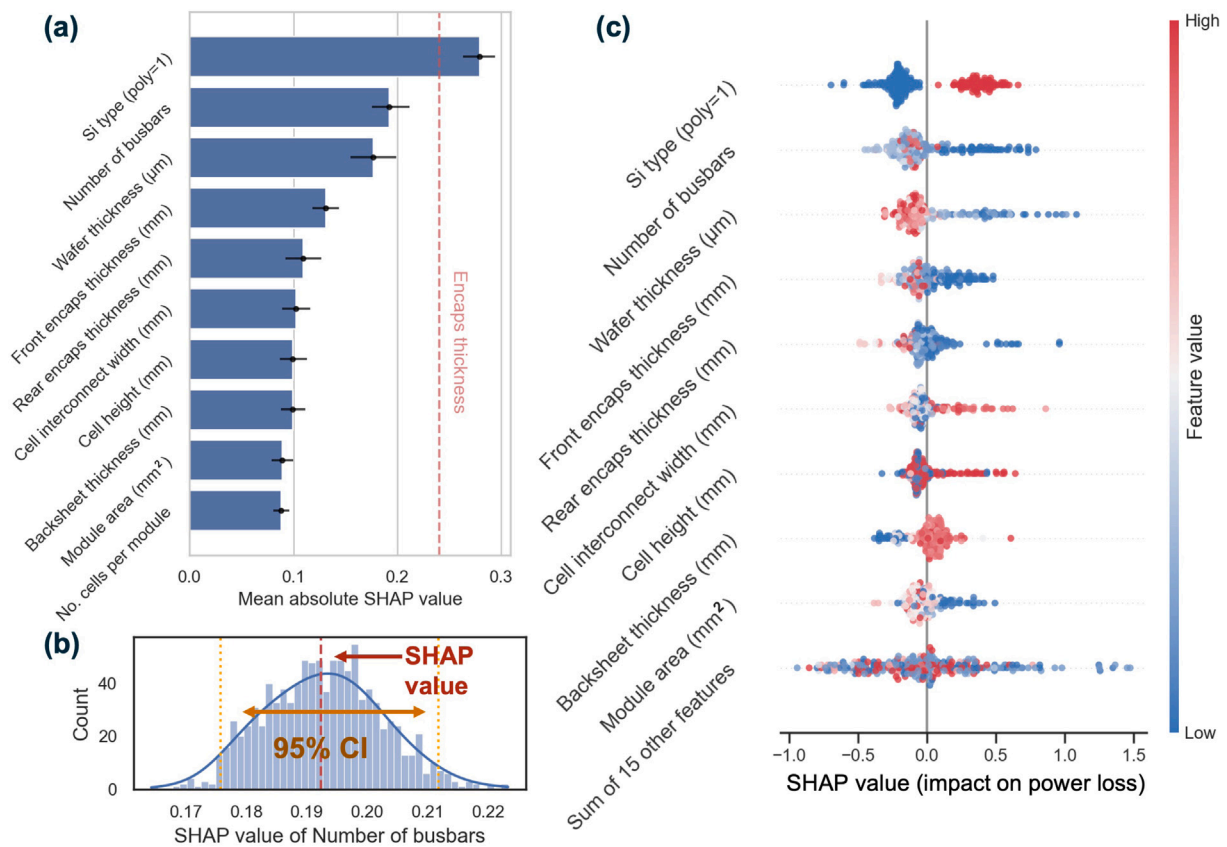
**Fig. 5.** SHAP interpretation of the machine learning model and uncertainty determination. (a) The rank of importance of the BOM features. The mean absolute Shapley value is calculated over all the modules in the original data set. The error bar shows the uncertainty with 95% confidence level using bootstrapping (random sampling from the original data set). The red dashed line shows the sum of Shapley values of both front and rear encapsulant thickness, indicating that the encapsulant thickness is also a very important feature other than the Si type. (b) An example of the bootstrapping distribution of the number of busbars using 1000 sampling. The Shapley value computed from the original data is indicated by the red dashed line, and the boundaries of the 95% CI computed from the bootstrapping are indicated by the orange dashed line. (c) The beeswarm plot of Shapley values. It shows the dependent relation between each feature and the power loss. The color of each point denotes the feature values. The red color denotes a larger feature value, and the blue color denotes a smaller feature. The *x*-axis is the Shapley value (*i.e.*, impact on degradation) of each feature. For example, the blue dots (lower feature value) in the row "Wafer thickness" with positive Shapley values (higher power loss) suggest that a lower wafer thickness tends to increase power loss.

in the performance of the solar module over time even before the crack length reaches the threshold of a catastrophic breakage. In addition, different cutting methods for processing the silicon ingot can influence the fracture toughness of Si wafers [43], so this manufacturing step can be a potential confounding variable that causes this phenomenon. We note that the impact of micro-cracks on module electric power and the evolution of micro-cracks in long-term cyclic loading remain a topic of ongoing research [2,44,45], so other factors may also cause the difference in power loss here. Another possible reason is that the current data set may be influenced by other confounding variables. The mitigation of the influence of confounding variables will be illustrated in the following subsection "Statistical Validation of SHAP Interpretation". We note that the solar industry is phasing down the usage of poly-c Si cells primarily because of the lowered cost and higher efficiency of mono-c Si cells. The findings from this analysis can provide another motivation for the switch from poly-c to mono-c Si cells.

### 3.3.2. Impact of encapsulant thickness

The encapsulant thickness remains another important design factor. From Fig. 6(a)(b), both front and rear encapsulant thickness exhibit negative relationship with the Shapley value when the thickness is lower than around 0.6–0.7 mm, suggesting that thinner encapsulant might result in higher power loss. Particularly, when the encapsulant thickness is reduced to around 0.4 mm, the influence becomes pronounced, with the front encapsulant increasing the power loss by approximately 0.2% and the rear encapsulant by approximately 0.6%.

This corresponds to previously simulated results [13] that thicker encapsulant decreases the accumulated thermal stress in the solder layer. Also, as a soft embedding of silicon solar cells, the encapsulant can compensate for the strain coming from the glass during deflection [46]. Interestingly, when the front encapsulant thickness is greater than 0.7 mm, the power loss increases, especially for the front encapsulant. This opposite trend was not reported by previous regression analysis [9,13]. A possible cause for this trend is a change in the mechanical properties of the polymer at lower temperature. At high temperature, the encapsulant layer is soft and acts as a compensation layer for the strain difference between the glass and Si layer. However, lower temperature, especially approaching or below the glass transition temperature ($\sim -33$ °C [47] for EVA and $\sim -25$ °C [48] for POE), limits the mobility of polymer chains and thus increases the stiffness of the encapsulant. This transition transforms the laminated structure of solar module into so-called glass-encapsulant-Si "sandwich" structure and more strain is conducted to the cell layer by the stiff encapsulant as the module bends during thermal cycling [46,49]. In this structure at low temperature, thicker encapsulant between the glass and Si layer, which is the front encapsulant, can increase the tensile stress in the Si layer and lead to higher Si fracture probability. Such a phenomenon was previously reported by Dietrich et al. [46] that, at −40 °C, thicker encapsulant leads to higher probability of failure. Therefore, in our BOM data set, the power loss first decreased with thicker encapsulant and then increased as the failure at lower temperature dominates. The hypothesis here can also explain that front encapsulant presents steeper increase at tail region (> 0.7 mm) in Fig. 6(a) than the rear
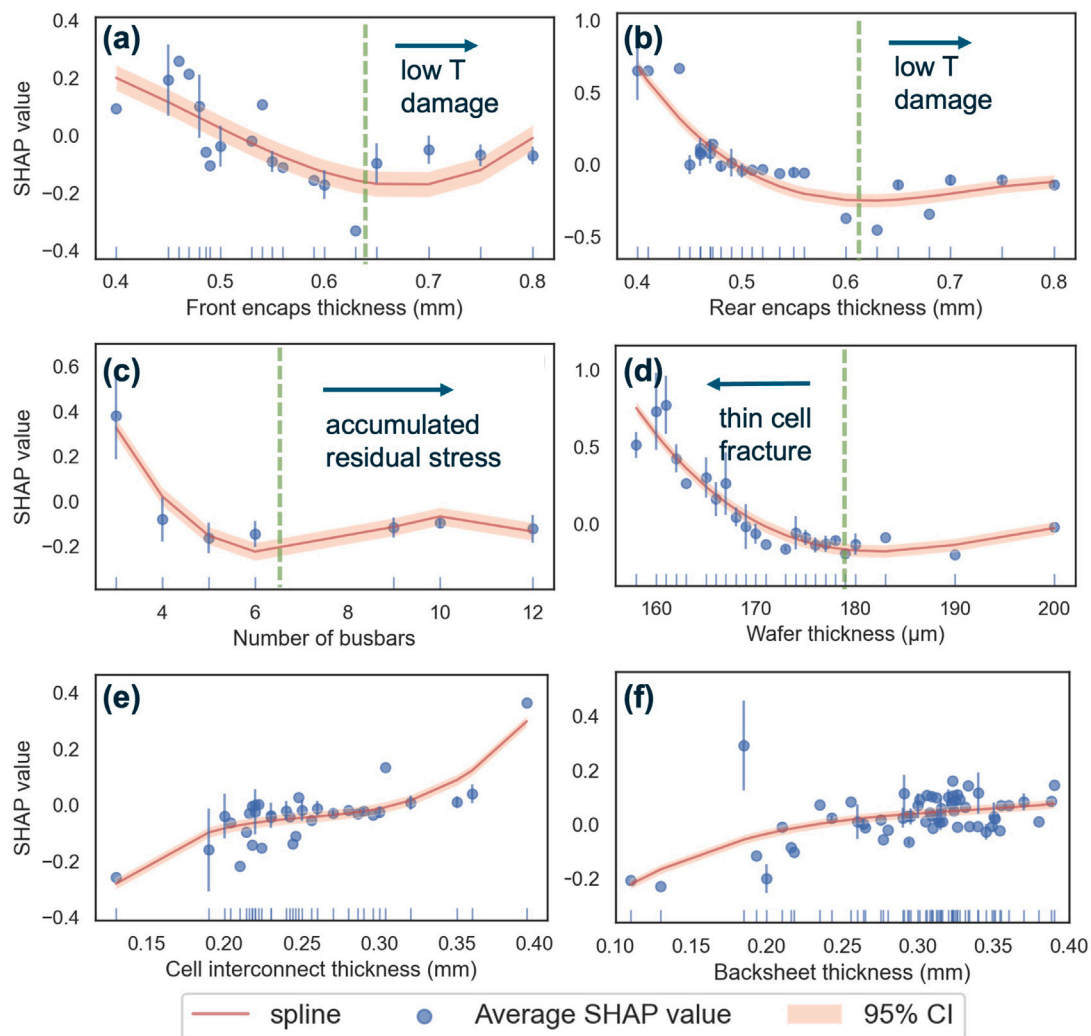
**Fig. 6.** Dependence plot of the relation between numerical features and Shapley value. In these dependence plots, the mean Shapley value corresponding to each feature value is displayed as one dot and the standard deviation is used for the error bar. The bottom rug plot in each sub-figure shows the distribution of the data. A smooth spline interpretation with 95% prediction interval is used to illustrate the trend of the relationship. Due to the additive property of the SHAP formula, the Shapley value is related to the power loss so the dependence plot also reflects the relation between the power loss and the BOM features. (a)–(d) shows the impacts of the top five features apart from the Si cell type. (e) and (f) examine the impact of the two features that were considered in the previous literature.

encapsulant in Fig. 6(b) since the front encapsulant locates between the glass and Si layer conduct more strain. However, we note that since the measurements at the tail region are sparse, this trend of increase may also be influenced by noise. Further testing with more controlled data sets may be needed to validate the hypothesis.

*3.3.3. Impact of busbar number*

Fig. 6(c) demonstrates that the power loss decreases about 0.6% as the number of busbars increases from 4 to around 7 and then increases slightly or remains unchanged with more busbars soldered on solar cells. This is reasonable because increasing the number of busbars can increase the probability of current connection between cells and external circuits even when the solder layer beneath some busbars got cracked due to thermal stress. However, too many busbars soldered on the Si wafer can yield more residual thermal stress during the fabrication process, which causes solder disconnection and even wafer fracture during operation [50,51]. Therefore, the accumulated residual stress counteracts the benefit of increasing busbar number. We also notice the drop of Shapley value at 12 busbars but since the measurements are sparse, it cannot represent the general trend of impacts. The relationship between busbar number and power loss in this analysis should attract attention because multi-busbar design is

becoming prevalent in recent module manufacturing, but its long-term durability needs further study [52].

*3.3.4. Impact of wafer thickness*

The wafer thickness was considered as a top important factor in previous studies [9,13] and also remains an important feature in our data set. However, the impact of wafer thickness in our real-world data set shows a different trend to previously simulated results. Fig. 6(d) indicates that a thinner wafer thickness (from 180 μm to 160 μm) can increase power loss by over 0.7%, contrary to previous studies that recommended thinner wafers. However, when the thickness is greater than 180 μm, thinner wafer is beneficial for durability. This discrepancy may arise because earlier simulations primarily focused on damage within the solder layer. However, as a brittle material, silicon wafers are more likely to experience catastrophic fractures if the thickness is excessively reduced. This potentially offsets the advantages of reduced damage in the solder layer. This trend aligns with another study showing that higher wafer thickness is important in module design [16].

*3.3.5. Impact of other features*

Furthermore, Fig. 6(e) and (f) demonstrate the impacts of the cell interconnection (*i.e.,* copper ribbon) thickness and backsheet thickness,
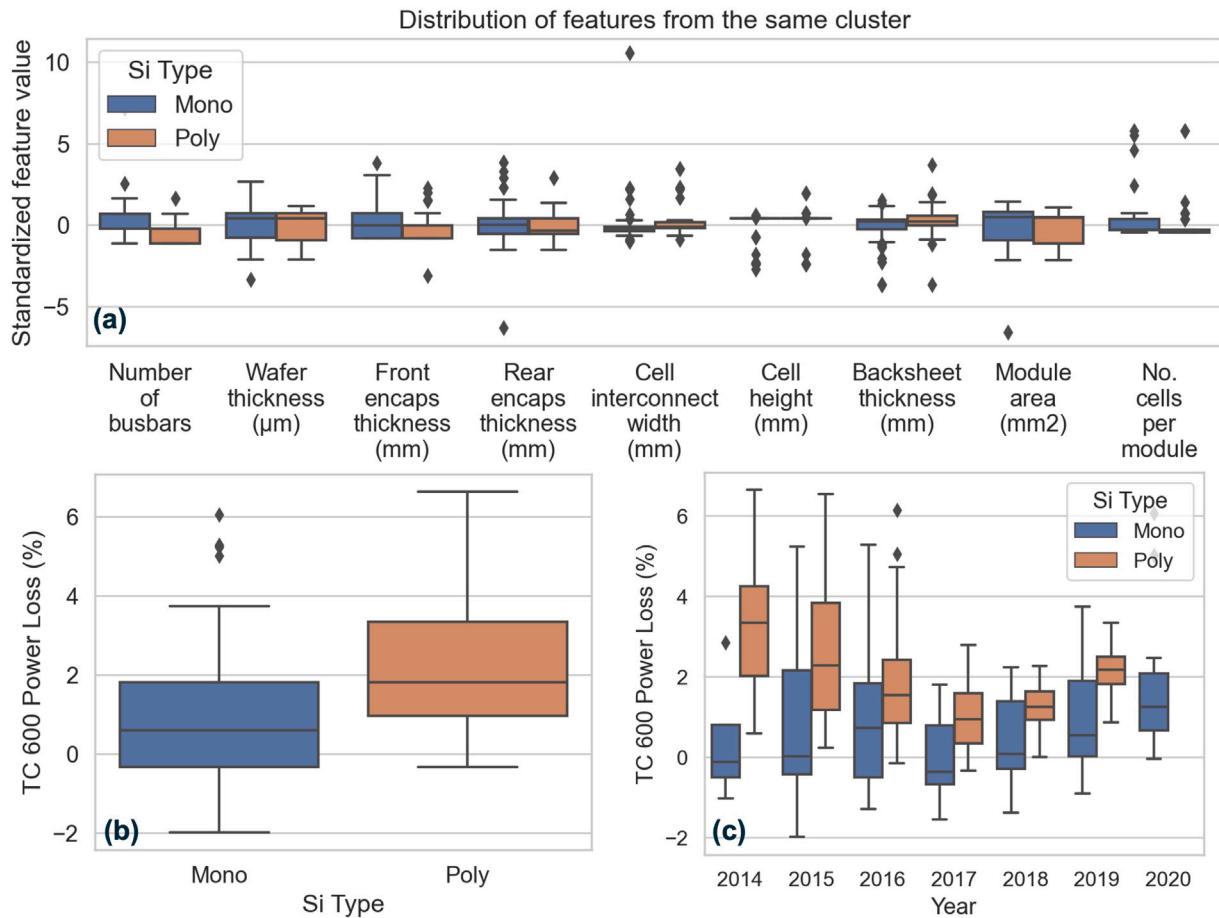
**Fig. 7.** Statistical testing to verify the impact of the Si cell type. Sample data is first extracted from the raw data set using the clustering method to guarantee that features other than Si cell type are randomized. Then the t-test is used to test the statistical significance of the impact of the Si cell type. (a) The distribution of several top important features in the same cluster. The similar distribution comparison between mono-c Si cells and poly-c Si cells indicates the randomization of these features. (b) The distribution of power loss between mono-c Si cells and poly-c Si cells in the same cluster. The *p*-value smaller than 5% indicates that poly-c Si modules have a statistically significant increase in the thermal cycling power loss. (c) The distribution of Si type in each year. The higher power loss of poly-c Si modules indicates that the manufacturing year is not the confounding variable.

which were explored in previous simulations [9,13]. From Fig. 6(e), it can be seen that modules with lower ribbon thickness have lower power loss because of lower thermal stress in the solder layer [9]. However, we consider this trend to be unreliable because the data points in the region over 0.35 mm and lower than 0.20 mm are sparse and may skew this trend. Fig. 6(f) suggests that the power loss is expected to increase as the backsheet thickness increases. This increase corresponds to the conclusion of the previous investigation [13] that thicker backsheet leads to more accumulated damage in the solder layer.

### 3.4. Statistical validation of SHAP interpretation

To validate the results of the aforementioned interpretation, we directly examine the power loss in the original data set. As explained in Section 2, the original data set is not well controlled, so we re-sample a controlled subset derived from the original one. To determine whether Si type is indeed impactful, we construct this subset using k-means clustering method so that feature values in this subset are similar apart from the Si type. Then we conduct a t-test on this subset and illustrate that poly-c Si modules have a statistically significant increase in the power loss. In this section we select the Si type for this analysis because not only is it the most impactful factor, but it is a categorical variable which makes the control process easier. The example of clustering other numerical features is included in the SI Section "Post hoc Statistical Analysis", which shows that the control group is hard to obtain for other variables.

In this controlled subset to test the impact of the Si type, the distribution of the top ten important BOM features other than Si type is shown in Fig. 7(a). It can be seen that the boxplots of most features for these cells are similar but contain differences in Si type, which means that these features are randomized between mono-c and poly-c Si. We notice that poly-c Si cells tend to have fewer busbars, which may cause a slight correlation between "Si type" and "number of busbars". This potential correlation may influence the importance of these two features. However, after checking the original data set, we find that the difference of median value of the number of busbars is not large between mono-c (median = 5) and poly-c Si cells (median=4), so the influence of this difference may not be significant since the difference of Shapley value in 4 and 5 busbars is around 0.1. More details of constructing this cluster is shown in the SI section "Post hoc Statistical Analysis". We also check the normality of the power loss using the quantile–quantile (Q-Q) plots, which is a prerequisite for further parametric statistical testing. The details are included in the SI.

With all prerequisite processing of the data set, we conduct a t-test to compare the mean power loss of poly-c Si modules and mono-c Si modules in this controlled subset. Fig. 7(b) compares the distribution of the power loss of mono-c and poly-c Si modules from the selected cluster, which displays that poly-c Si modules have more power loss on average. The null hypothesis of our t-test is:

$H_0$ : *Poly-c Si modules don't exhibit higher thermal cycling power loss than mono-c Si modules.*

The test result illustrates that the power loss of poly-c Si modules is greater than that of mono-c Si modules and this result is statistically significant (p-value $= 1.4 \times 10^{-8} < 5\%$). This validates the interpretation of the SHAP analysis.

Although the clustering can mitigate the influence of uncontrolled variables, some other variables that are not included in the ML modeling may still impact our interpretation. One of the most potential confounding variables is the manufacturing process. To mitigate the bias towards a specific supplier, this data set collects modules from various manufacturers. Also in Fig. 7(c), we compare the power loss of the two types of modules from the same cluster in each manufacturing year. It reveals that poly-c Si modules still have a higher power loss each year, which means that the manufacturing year is not the underlying reason for the difference. Despite these efforts, there is still a chance that some unrecorded manufacturing conditions during the fabrication of poly-c and mono-c Si modules may vary and influence the thermomechanical durability rather than the Si type itself. As in all such cases, data analysis can point the direction to highly plausible hypotheses, but careful and dedicated experiments would be needed to confirm each trend.

## 4. Conclusion

In this work, we identify that the predominant BOM features related to module's thermomechanical durability in the current module design are Si type, encapsulant thickness, busbar numbers, and wafer thickness. Our analysis is based on a unique data set that includes the BOM features of full-size modules from various manufacturers and their power loss after 600 thermal cycles. To analyze this data set, we first correlate the BOM features with the power loss using machine learning modeling. By comparing with other models, we develop a RF model with a testing RMSE of 1.179% as the optimal model and a RMSE of 0.558% on the whole data set for further SHAP interpretation. We subsequently apply SHAP analysis to the fitted RF model to interpret the whole data set and used statistical testing to verify the conclusion.

Overall, we find that the Si type (*i.e.,* whether the module is composed of poly-c or mono-c Si cells) is the most influential factor. The replacement of poly-c Si with mono-c Si decreases the power loss by approximately 0.6% in the SHAP analysis. The next most important feature is the thickness of the front and rear encapsulant; a higher thickness can reduce power loss, but further increases in thickness (over around $0.6 - 0.7$ mm) present an opposite impact, particularly for the front encapsulant. We also analyze the impacts of other features. In particular, we find that increasing the number of busbars initially reduces power loss but subsequently increases it after the number exceeds around 7. The wafer thickness displays an opposite trend to the previous findings [9,13]. In our data set, too low wafer thickness ($< 180$ μm) is not beneficial for the durability. We further find that the power loss is reduced when the ribbon thickness is decreased and thinner backsheet layer is beneficial for durability.

Finally, it is important to acknowledge the limitations of our study. Although our data set covers a breadth of BOM features, certain areas, such as the solder thickness or various CTEs of encapsulants, remained uncharted due to the difficulty of tracking this information from manufacturers. Furthermore, the use of commercially available modules as opposed to carefully controlled modules means that conclusions might still be influenced by confounding variables, underscoring the need for further investigation. Despite these constraints, we hope this study can reveal the potential application of machine learning and model-agnostic interpretation methods in examining BOM effects on reliability and provide insight into the direction of future module optimization.

## Code availability

Our codes for ML modeling and SHAP interpretation are available at Github repository(https://github.com/DuraMAT/bom_analysis).

## CRediT authorship contribution statement

**Xin Chen:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Todd Karin:** Writing – review & editing, Funding acquisition. **Anubhav Jain:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The data set is available at DuraMat Datahub (https://datahub.duramat.org/dataset/bom_thermal_cycling_degradation). Some values in the published data set are redacted according to non-disclosure agreement.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Writefull & Chatgpt in order to check grammar and improve readability. The author(s) did not use this tool/service to directly generate the manuscript but only to polish the manuscript written by the author(s). After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.apenergy.2024.124462.

## References

[1] Ardani K, Denholm P, Mai T, Margolis R, O'Shaughnessy E, Silverman T, et al. Solar futures study. Tech. rep., U.S. Department of Energy; 2021.

[2] Aghaei M, Fairbrother A, Gok A, Ahmad S, Kazim S, Lobato K, et al. Review of degradation and failure phenomena in photovoltaic modules. Renew Sustain Energy Rev 2022;159:112160.

[3] Jordan DC, Silverman TJ, Wohlgemuth JH, Kurtz SR, VanSant KT. Photovoltaic failure and degradation modes. Prog Photovolt, Res Appl 2017;25(4):318–26.

[4] Kawai S, Tanahashi T, Fukumoto Y, Tamai F, Masuda A, Kondo M. Causes of degradation identified by the extended thermal cycling test on commercially available crystalline silicon photovoltaic modules. IEEE J Photovolt 2017;7(6):1511–8.

[5] Eitner U, Köntges M, Brendel R. Use of digital image correlation technique to determine thermomechanical deformations in photovoltaic laminates: Measurements and accuracy. Sol Energy Mater Sol Cells 2010;94(8):1346–51.

[6] Borri C, Gagliardi M, Paggi M. Fatigue crack growth in silicon solar cells and hysteretic behaviour of busbars. Sol Energy Mater Sol Cells 2018;181:21–9.

[7] Wiese S, Meier R, Kraemer F. Mechanical behaviour and fatigue of copper ribbons used as solar cell interconnectors. In: 2010 11th international thermal, mechanical & multi-physics simulation, and experiments in microelectronics and microsystems (euroSimE). IEEE; 2010, p. 1–5.

[8] Silverman TJ, Bliss M, Abbas A, Betts T, Walls M, Repins I. Movement of cracked silicon solar cells during module temperature changes. In: 2019 IEEE 46th photovoltaic specialists conference. PVSC, IEEE; 2019, p. 1517–20.

[9] Park S, Han C. Reliability-driven design optimization of si solar module under thermal cycling. J Mech Sci Technol 2022;36(8):4099–114.

[10] Jeong J-S, Park N, Han C. Field failure mechanism study of solder interconnection for crystalline silicon photovoltaic module. Microelectron Reliab 2012;52(9–10):2326–30.

[11] Itoh U, Yoshida M, Tokuhisa H, Takeuchi K, Takemura Y. Solder joint failure modes in the conventional crystalline Si module. Energy Procedia 2014;55:464–8.

[12] Hanifi H, Pander M, Zeller U, Ilse K, Dassler D, Mirza M, Bahattab MA, Jaeckel B, Hagendorf C, Ebert M, et al. Loss analysis and optimization of PV module components and design to achieve higher energy yield and longer service life in desert regions. Appl Energy 2020;280:116028.

[13] Bosco N, Silverman TJ, Kurtz S. The influence of PV module materials and design on solder joint thermal fatigue durability. IEEE J Photovolt 2016;6(6):1407–12.

[14] Zhu J, Owen-Bellini M, Montiel-Chicharro D, Betts TR, Gottschalg R. Effect of viscoelasticity of ethylene vinyl acetate encapsulants on photovoltaic module solder joint degradation due to thermomechanical fatigue. Japan J Appl Phys 2018;57(8S3):08RG03.

[15] Beinert AJ, Romer P, Heinrich M, Mittag M, Aktaa J, Neuhaus DH. The effect of cell and module dimensions on thermomechanical stress in PV modules. IEEE J Photovolt 2019;10(1):70–7.

[16] Beinert AJ, Romer P, Heinrich M, Aktaa J, Neuhaus H. Thermomechanical design rules for photovoltaic modules. Prog Photovolt, Res Appl 2022.

[17] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001;1189–232.

[18] Ribeiro MT, Singh S, Guestrin C. " Why should i trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016, p. 1135–44.

[19] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017;30.

[20] Bannigan P, Bao Z, Hickman RJ, Aldeghi M, Häse F, Aspuru-Guzik A, et al. Machine learning models to accelerate the design of polymeric long-acting injectables. Nature Commun 2023;14(1):35.

[21] Giles SA, Sengupta D, Broderick SR, Rajan K. Machine-learning-based intelligent framework for discovering refractory high-entropy alloys with improved high-temperature yield strength. npj Comput Mater 2022;8(1):235.

[22] Kumar R, Singh AK. Chemical hardness-driven interpretable machine learning approach for rapid search of photocatalysts. npj Comput Mater 2021;7(1):197.

[23] Hartono NTP, Thapa J, Tiihonen A, Oviedo F, Batali C, Yoo JJ, et al. How machine learning can help select capping layers to suppress perovskite degradation. Nature Commun 2020;11(1):4172.

[24] Commission IE. Terrestrial photovoltaic (PV) modules - Design qualification and type approval - Part 2: Test procedures. Tech. Rep. 61215–2, International Electrotechnical Commission; 2016.

[25] Ho TK. Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, vol. 1. IEEE; 1995, p. 278–82.

[26] PV Evolution Labs [Online]. Available: https://www.pvel.com/.

[27] Van Rossum G, Drake FL. Python 3 reference manual. Scotts Valley, CA: CreateSpace; 2009.

[28] PVEL. The 2023 PV module reliability scorecard. 2023, [Online]. Available: https://scorecard.pvel.com/.

[29] Commission IE. Photovoltaic devices – Part 1: Measurement of photovoltaic current-voltage characteristics. Tech. Rep. 60904–1, International Electrotechnical Commission; 2006.

[30] Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics 2001;17(6):520–5.

[31] Vallat R. Pingouin: statistics in Python. J Open Source Softw 2018;3(31):1026.

[32] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 2020;17:261–72.

[33] Bruce P, Bruce A, Gedeck P. Practical statistics for data scientists: 50+ essential concepts using R and Python. O'Reilly Media; 2020.

[34] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825–30.

[35] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016, p. 785–94.

[36] Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Stat Methodol 1996;58(1):267–88.

[37] Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 1970;12(1):55–67.

[38] Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V. Support vector regression machines. Adv Neural Inf Process Syst 1996;9.

[39] Efron B. Bootstrap methods: another look at the jackknife. In: Breakthroughs in statistics: methodology and distribution. Springer; 1992, p. 569–93.

[40] Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. J R Stat Soc Ser C 1979;28(1):100–8.

[41] Watanabe H, Yamada N, Okaji M. Linear thermal expansion coefficient of silicon from 293 to 1000 K. Int J Thermophys 2004;25(1):221–36.

[42] Zavattieri PD, Espinosa HD. Grain level analysis of crack initiation and propagation in brittle materials. Acta Mater 2001;49(20):4291–311.

[43] Wu H, Melkote SN. Effect of crystal defects on mechanical properties relevant to cutting of multicrystalline solar silicon. Mater Sci Semicond Process 2013;16(6):1416–21.

[44] Silverman TJ, Deceglie MG, Owen-Bellini M, Hobbs WB, Libby C. Cracked solar cell performance depends on module temperature. In: 2021 IEEE 48th photovoltaic specialists conference. PVSC, IEEE; 2021, p. 1691–2.

[45] Libby C, Paudyal B, Chen X, Hobbs WB, Fregosi D, Jain A. Analysis of PV module power loss and cell crack effects due to accelerated aging tests and field exposure. IEEE J Photovolt 2022;1–9.

[46] Dietrich S, Sander M, Pander M, Ebert M. Interdependency of mechanical failure rate of encapsulated solar cells and module design parameters. In: Reliability of photovoltaic cells, modules, components, and systems v. 8472, SPIE; 2012, p. 123–31.

[47] Agroui K, Collins G, Farenc J. Measurement of glass transition temperature of crosslinked eva encapsulant by thermal analysis for photovoltaic application. Renew Energy 2012;43:218–23.

[48] Baiamonte M, Colletti C, Ragonesi A, Gerardi C, Dintcheva NT. Durability and performance of encapsulant films for bifacial heterojunction photovoltaic modules. Polymers 2022;14(5):1052.

[49] Dietrich S, Pander M, Sander M, Schulze SH, Ebert M. Mechanical and thermomechanical assessment of encapsulated solar cells by finite-element-simulation. In: Reliability of photovoltaic cells, modules, components, and systems III. 7773, SPIE; 2010, p. 117–26.

[50] Rendler LC, Kraft A, Ebert C, Eitner U, Wiese S. Mechanical stress in solar cells with multi busbar interconnection—Parameter study by FEM simulation. In: 2016 17th international conference on thermal, mechanical and multi-physics simulation and experiments in microelectronics and microsystems. IEEE; 2016, p. 1–5.

[51] Shin H, Han E, Park N, Kim D. Thermal residual stress analysis of soldering and lamination processes for fabrication of crystalline silicon photovoltaic modules. Energies 2018;11(12):3256.

[52] Zuboy J, Springer M, Palmiotti E, Barnes T, Karas J, Smith B, Woodhouse M. Technology scouting report: Reliability implications of recent pv module technology trends. 2022, [Online]. Available: https://www.nrel.gov/docs/fy22osti/82871.pdf.