Expert Systems With Applications 296 (2026) 128890

Contents lists available at ScienceDirect

Expert Systems With Applications

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

General machine learning models for interpreting and predicting efficiency degradation in organic solar cells

David Valiente [a,b,*], Fernando Rodríguez-Mas [b], Juan V. Alegre-Requena [c], David Dalmau [c], María Flores [a], Juan C. Ferrer [a]

[a] University Institute for Engineering Research, Miguel Hernandez University, Avenida de la Universidad, s/n, Elche, 03202, Spain
[b] Communications Engineering Department, Miguel Hernandez University, Avenida de la Universidad, s/n, Elche, 03202, Spain
[c] Department of Inorganic Chemistry, Institute of Chemical Synthesis and Homogeneous Catalysis (ISQCH), CSIC, University of Zaragoza, Pedro Cerbuna 12, Zaragoza, 50009, Spain

ARTICLE INFO

Keywords:
Organic solar cells
Power efficiency degradation
Multilayer structure
Machine learning
Regression

ABSTRACT

Photovoltaic (PV) energy plays a key role in addressing the growing global energy demand. Organic solar cells (OSCs) represent a promising alternative to silicon-based PVs due to their low cost, lightweight, and sustainable production. Despite achieving power conversion efficiencies (PCEs) over 20 %, OSCs still face challenges in stability and efficiency. Recent advances in manufacturing, artificial intelligence and machine learning (ML) achieve optimized and screened OSCs for greater sustainability and commercial viability, thus potentially reducing costs while ensuring stable and long term performance. This work presents optimal ML models to represent the temporal degradation on the PCE of polymeric OSCs with structure ITO/PEDOT:PSS/P3HT:PCBM/Al. First, we generated a database with 166 entries with measurements of 5 OSCs, and up to 7 variables regarding the manufacturing and environmental conditions for more than 180 days. Then, we relied on a software framework that provides a conglomeration of automated ML protocols that execute sequentially against our database by simply command-line interface. This easily permits hyper-optimizing the ML models through exhaustive benchmarking so that optimal models are obtained. The accuracy for predicting PCE over time reaches values of the coefficient determination widely exceeding 0.90, whereas the root mean squared error, sum of squared error, and mean absolute error are significantly low. Additionally, we assessed the predictive ability of the models using an unseen OSC as an external set. For comparative purposes, classical Bayesian regression fitting are also presented, which only perform sufficiently for univariate cases of single OSCs.

1. Introduction

Electricity consumption has evolved exponentially over the last three decades, increasing from a total estimate of 15277 to 29479 TWh between 2000 and 2023, respectively (Ritchie and Roser, 2024). One of the immediate consequences has been reflected in the price of this energy. Simultaneously, in the technological realm, there has been a notable shift in electricity generation towards renewable sources. Particularly, the maturity of conventional photovoltaic (PV) cells and the emergence of new organic and hybrid materials have made it possible to get significant improvements in PV generation and its power conversion efficiency (PCE) (New Media Consortium, 2025). However, even though the global consumption of PV energy has increased from 1.08 to 1629.9 TWh between 2000 and 2023 (Ritchie and Roser, 2024), its generation still falls short of meeting current demands. Hence, there are many challenges and open lines for improving the generation and efficiency of PV cells.

In this context, organic solar cells (OSCs) have emerged as a promising alternative to silicon-based solar cells since, amongst other benefits, their production involves low-temperature manufacturing methods (Seo et al., 2015) and reduced carbon footprint. Their appeal also lies in the ease of processing, low cost, flexibility, and lightweight (Yeh & Yeh, 2013). Moreover, the evolution of their PCEs has been much superior to classical technologies and it has been on par with other considered emerging technologies, presenting efficiencies that exceed values of 20 % (Ayub et al., 2023). Even though the commercialization of OSCs still faces various challenges, including stability and efficiency, some cost studies (Chiappafreddo & Gagliardi, 2010; Rodriguez-Mas et al., 2023; Venkata et al., 2016) indicate that once these obstacles are

* Corresponding authors.
E-mail addresses: dvaliente@umh.es (D. Valiente), fernando.rodriguezm@umh.es (F. Rodríguez-Mas), jv.alegre@csic.es (J.V. Alegre-Requena), ddalmau@unizar.es (D. Dalmau), m.flores@umh.es (M. Flores), jc.ferrer@umh.es (J.C. Ferrer).

https://doi.org/10.1016/j.eswa.2025.128890
Received 15 February 2025; Received in revised form 13 May 2025; Accepted 1 July 2025
Available online 16 July 2025

overcome, OSCs can be manufactured at a cost lower than one dollar per peak Watt (Brabec et al., 2005).

The integration of machine learning (ML) before transferring to industrial implementation is being extensively tested these days. For instance, within the field of cheminformatics, it is of paramount importance in the design of new materials. Numerous studies have demonstrated the potential of deep learning algorithms to analyze complex chemical data, thereby accelerating the process of discovering, for example, drugs (Dalmau & Alegre-Requena, 2024a), as well as identifying promising compounds with improved properties (Lusci et al., 2013; Sanosa et al., 2024). Some early studies (Hansen et al., 2013; Rupp et al., 2012) laid the foundations for the application of ML models in predicting molecular properties, highlighting their great strengths. Others concentrated on examining chemical databases to identify potential materials (Unterthiner et al., 2015; Wallach et al., 2015), whose focus moved forward to the inclusion of PV features (Hachmann et al., 2011).

General advances in artificial intelligence (AI) and ML have demonstrated great improvements on the performance, reliability, and efficiency of PV. Please follow Table 1 to get the main insights of relevant studies which have focused on PV data processing with ML. It is worth highlighting some works, i.e. Serrano-Lujan et al. (2016) and Tang et al. (2024), where short-term memory networks have been implemented for timely evolving systems; data-intensive studies which identify influencing factors on the degradation of OSCs (Alsulami et al., 2024; Borazan, 2019; Zhang et al., 2017); and automated interpretable ML models that aid in the discovery of high PCE materials (Lee, 2023; Moore et al., 2022).

More recent works have significantly impacted the development of OSCs, offering data-driven tools for predicting performance and improving stability. Liu et al. (2024) demonstrate that ML can model the relationships between parameters, morphology, and long-term stability of OSCs; Du et al. (2024) have identified electronic structure with effective bandgap as key predictors for environmental resilience; Hußner et al. (2024) have introduced ML techniques to extract charge carrier dynamics from J–V curves, permitting efficient characterization without complex measurements; Zhao et al. (2022) applied ML to visualize and optimize the influence on PCE of fabrication parameters in OSCs, highlighting the benefits for experimental design; Osterrieder et al. (2023) further accelerated OSC optimization through an AI-driven autonomous platform using Bayesian optimization and spectral data; finally, Lin et al. (2025) addressed long-term stability by designing new copolymers that, when coupled with predictive modeling, showed enhanced durability. In sight of all these developments, a clear horizon for this work is to exploit ML in capturing the temporal degradation behavior of OSCs.

In particular, we explore different ML models to encode the PCE performance of polymeric OSCs based on multilayer ITO/PEDOT:PSS/P3HT:PCBM/Al (for definitions please check Sections 2.2 and 5). To that aim, we have used the ROBERT[1] program (Dalmau & Alegre-Requena, 2024b), which automates data curation, screening of ML models, assessment of predictive ability, and feature analysis (for reproducibility details, see Appendix C). Our dataset[2] consists of PCE measurements acquired for 180 days, of 5 OSC devices, resulting in a database with 166 entries. Although these devices are encapsulated, the database contains up to seven descriptors, some related to environmental conditions (temperature, hummidity, dew point and pressure) to inspect possible dependencies on their future performance and some others related to their composition and manufacturing (quantities of: solvent in the high transportation layer (HTL), i.e., PEDOT:PSS; P3HT, PCBM; and volume ratio of P3HT:PCBM). The accuracy metrics demonstrate the validity of these models to learn and represent the temporal behavior of our OSCs, with determination coefficient up to $R^2 = 0.96$, and low root mean squared error (RMSE), sum of squared error (SSE) and mean absolute error (MAE). Additionally, traditional Bayesian regression models sustained by non-linear least squares (LS) have been introduced to confirm the advantages of the ML-automated protocols against these classical approaches. While these conventional methodologies (Bozorg et al., 2020; Owolabi et al., 2017; Wolf & Benda, 2013) may yield satisfactory results in univariate regression domains, our approach demonstrates that computational learning modeling offers a much broader and comprehensive solution compared to those based on classical statistical models. ML models show increased robustness and reliability for:

- Generating multivariate models to capture PCE behavior over time while ensuring the interpretability of variable influences.
- Learning from an entire dataset of multiple OSCs, not solely from a specific OSC device as in classical regression.
- Screening new OSC devices not used during the learning phase, thereby predicting its behavior, as an unknown device.
- Detecting optimal variables to fabricate the OSCs with optimal PCE and/or its stability over time.

In summary, the key contributions of this research are as follows:

1. Assessment of optimal ML models that learn the PCE behavior over time for polymer-based OSC devices with ITO/PEDOT:PSS/P3HT:PCBM/Al structures.
2. Comparative benchmarking among different ML models and classical statistical regression approaches.
3. Identification of an optimal ML model capable of predicting the PCE behavior of unseen OSC devices.
4. Feature analysis to establish the influence of variables on the performance of the OSC devices.
5. Reproducibility and transparency of the obtained ML models using a standardized framework for command line replication.

The rest of the paper is organized as follows: Section 2 describes the specific OSCs manufactured in our laboratory, their electrical parameters, and the periodic measurements that comprise the database used by the ML methods. Subsequently, these methods are defined, starting from a preliminary scope of regression problems, along with the software framework that enables their benchmarking and, consequently, the extraction of models with optimal hyper-parameters. Then, Section 3 outlines the experiments and results. Finally, Section 4 draws conclusions from this work.

2. Materials and methods

2.1. Organic solar cells

The OSCs characterized in this work are manufactured through spin-coating technique, which involves the deposition of overlapped polymeric thin films. Polymers are deposited in solution, then a process of rotation followed by heating aids in removing the solvent, hence forming the layer. The structure of the devices was as follows: ITO/PEDOT:PSS/P3HT:PCBM/Al, where each layer refers to:

- ITO: Indium tin oxide.
- PEDOT:PSS: poly(3,4-ethylenedioxythiophene) polystyrene sulfonate.
- P3HT:PCBM: (poli(3-hexiltiofeno-2,5-diil):[6,6]-phenyl-C61-butyric acid methyl ester.
- Al: Aluminium.

As a substrate, a glass with a 60 nm thick semi-transparent ITO layer was used. The substrates were placed in the spinner with the ITO layer facing upwards so that the thin layers could be deposited on it. The first layer was the PEDOT:PSS film. It was deposited at room temperature at a rotation speed of 6000 rpm for 60 s. The remaining solvent was removed by heating at 150 °C for 10 min. Once the first layer was deposited, the

[1] https://github.com//jvalegre/robert
[2] https://github.com/POLI-NANO/OSCs

Table 1
Literature review.

Ref.	Topic	Material	Method(s)	Data	Application	Contribution
Eibeck et al. (2021)	Efficiency prediction	Organic PV	SVR, RF, GB, ANN, KNN	Experimental	Performance prediction	ML comparison for predicting PCE
Serrano-Lujan et al. (2016)	Prediction with ensemble learning	General PV	LSTM + GRU + Attention	Meteorological	Forecasting	Prediction via ensemble deep learning methods
Toledo et al. (2019)	Parameter estimation	PV panels	ANN	Simulated	Modeling	Improved accuracy for PV parameter estimation
Lopez et al. (2016)	Efficiency prediction	Organic PV	ANN, RF	Experimental	Efficiency Estimation	ML comparison for performance prediction
Moore et al. (2022)	Energy levels prediction	Organic PV	Transfer Learning	Material datasets	Material Discovery	Transfer learning for predicting efficient HOMO/LUMO
Malhotra et al. (2021)	Defect classification	PV panels	CNN + transfer learning	Images	Fault Detection	Transfer-based detection with pretrained networks
Ryu et al. (2019)	ML model comparison	General PV	XGBoost, SVM, ANN	Real-field	Performance Evaluation	Systematic comparison of ML models
Sun et al. (2019)	Fault prediction	General PV	SVM, ANN	Real-field	Fault Detection	Early fault prediction using statistical features
Miyake and Saeki (2021)	ML screening	Perovskite PV	RF, KNN, SVM	Laboratory	Materials Discovery	ML for discovering stable compositions
Mahmood et al. (2022)	Optimizing ML pipelines	Organic PV	AutoML	Benchmark	Forecasting	Automated pipeline for model selection
Mahmood and Wang (2021)	Parameter identification	Organic PV	Hybrid DL + evolutionary algorithm	Simulated + experimental	Modeling	Hybrid optimization for parameter extraction
Ju et al. (2018)	Material properties	Perovskite PV	XGBoost, DNN	Material datasets	Materials Discovery	Prediction of perovskite bandgaps using ML
Mahmood et al. (2023)	Bandgap prediction	Organic PV	SVR + GSA optimization	Material datasets	Materials Discovery	SVR tuned via metaheuristics for prediction
Jobayer et al. (2023)	Overview of ML	Various PV	Multiple ML models	Literature	Forecasting, modeling	Overview of ML for PV prediction and modeling
Tang et al. (2024)	Dual-attention forecasting	PV panels	GRU + Attention + clustering	Real-field	Forecasting	Robust predicting model for weather with transfer learning
Nguyen and Ishikawa (2023)	Output prediction	General PV	XGBoost, DNN	Simulated	Forecasting	High-accuracy output estimation
Dwivedi et al. (2024)	Defects detection	PV panels	Deep CNN + attention	Images	Fault Detection	Cross-system fault detection with attention DL
Yildirim et al. (2023)	Material screening	Perovskite PV	AutoML	Material datasets	Materials Discovery	AutoML for HSL discovery in perovskites
Mammeri et al. (2023)	Feature prediction	Perovskite PV	RF, SVM	Experimental	Efficiency Estimation	Multi-feature analysis for optimal PV design
Lee (2023)	Ternary OSC efficiency	Organic PV	Interpretable ML + descriptors	Experimental	Efficiency Estimation	Accurate prediction using molecular descriptors
Zhang et al. (2022)	Degradation and stability	Perovskite PV	Statistical + ML analysis	Stability dataset	Stability Analysis	Data-driven insight into degradation trends
Alsulami et al. (2024)	Large dataset stability	Perovskite PV	ML model (unspecified)	Literature + experimental	Stability Analysis	Stability trends across multiple studies
Wolf and Benda (2013)	Parameter extraction	General PV	Statistical + analytical models	Experimental	Parameter Estimation	Combined analytical/statistical estimation
Zhang et al. (2017)	Stability Analysis	Organic PV	Statistical	Experimental	Stability Analysis	Indoor and outdoor stability with environmental impact on degradation
Borazan (2019)	Lifetime Study of Photovoltaic Fibers	Organic PV	Statistical + analytical	Laboratory	Lifetime Prediction	Lifetime estimation under different stress conditions.

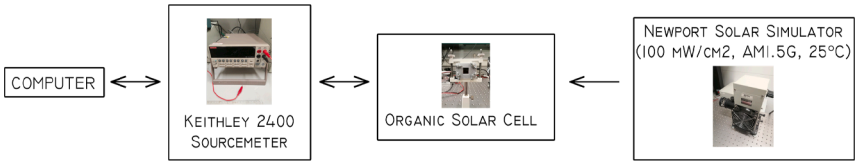

Fig. 1. Block diagram of the acquisition system used to acquire the OSC database.

active layer composed by the P3HT:PCBM polymer blend was applied. This was done at 300 rpm for 3 min. It was then dried at 80 °C for 1 h.

The final layer deposited on the devices was an aluminium film. Aluminium was not deposited using the spin-coating technique but rather by metal evaporation in a vacuum chamber. The equipment consists of a high vacuum chamber, two vacuum pumps, one of which is a high vacuum pump, and a power supply to provide the necessary current to evaporate the aluminium. Once the samples were placed inside the chamber, vacuum was achieved, reaching pressures of 10^{-6} mbar. When this pressure was reached, the aluminium evaporated by Joule effect.

2.2. OSC Database

As initially mentioned in Section 1, our dataset consists of PCE measurements of 5 OSC devices, where up to 7 variables were registered for more than 180 days (accesible on Github [3] and Appendix B). Electrical parameters (current-density vs voltage, J-V curves) of these OSCs were acquired along with climate conditions. The acquisition process comprises a Keithley-2400 equipment that acts as source generator for the voltage sweep as well as recorder of the generated PV current. The J-V curves were measured under light conditions; i.e. 100 mW/cm^2, Air Mass (AM) 1.5 Global (G), and 25 °C, as generated by a solar simulator Newport xenon arc lamp and an AM 1.5G filter. The electrical characterization was completed by determining the characteristic electrical parameters, including the short-circuit current density (J_{sc}), the open-circuit voltage (V_{oc}), the maximum power point (P_{mpp}), the fill factor (FF), and the PCE. Obtaining these parameters allowed to extract the PCE value as follows:

$$PCE = \frac{P_{mpp}}{P_{inc}} = \frac{J_{sc} V_{oc} FF}{G} \tag{1}$$

being P_{inc} the incident solar power on the OSCs, which derives from the incident irradiance of the solar simulator, G.

The block diagram of the equipment for the acquisition system is presented in Fig. 1. This permitted obtaining a database with 166 entries. The variables and their units are presented in Table 2. Notice that, in order to get further insights on the future performance of the OSCs, environmental conditions at laboratory have also been measured, despite the devices were encapsulated.

As an example, Fig. 2 presents an OSC device contained in the database, for which its current density (J) data have been periodically acquired against voltage (V) over time. Specifically, Fig. 2(a) displays the voltage range on the X-axis from 0 to 0.5 V, while Fig. 2(b) adjusts the scale to show the range from 0 to 0.25 V, allowing for a clearer observation of the evolution of the J-V curve when days go by (different colors of the legend). The degradation principles of these organic devices suggest that over time, these curves, from which maximum power and energy conversion efficiency are obtained, should consistently decay.

Likewise, Fig. 3 compares the temporal evolution of three different OSCs in terms of their normalized PCE values, ranging from 0 to 1, for more than 180 days. Please note that normalization is used for visualization purposes, as typically used in this field (Würfel & Würfel, 2016) when analyzing PCE decay. Once again, it is confirmed that time leads to the degradation of the device's conversion efficiency. However, it also emerges that differences in the fabrication of the OSCs, as well as the environmental conditions during their measuring, may influence the trend followed by the PCE over time.

[3] https://github.com/POLI-NANO/OSCs

Table 2
Detail of the variables in the dataset: Manufacturing variables and environmental conditions at laboratory. It contains 166 entries.

Variables	Values [min-max]	std
PCE	[0.27-1.32] %	0.25
quantity DS HTL	[250-1000] μl	0.19
P3HT	[1-1.2] mg	0.07
PCBM	[0.8-1] mg	0.07
ratio P3HT:PCBM	[1-1.25] -	0.09
Temperature	[12-23] °C	2.81
Hummidity	[33-88] %	17.78
Dew point	[3-19] °C	3.99
Pressure	[997-1022] hPa	7.25
Time	[0-181] days	60.21

2.3. Regression problem with classical approaches

This study aims to automatically model the degradative behavior exhibited by manufactured OSCs using ML approaches, which is a regression problem. In order to establish further comparisons that highlight and reinforce the abilities of ML-based protocols, a preliminary step was carried out to present regression fitting using traditional LS-supported Bayesian methods.

Regression is a method for estimating the relationship between a response or output variable and one or more predictor or input variables. Linear and non-linear regression serve as estimators of values between observed data points. From that starting point, a regression model relates response data to predictor data through one or more coefficients. Then a parametric fitting algorithm is needed to calculate some model coefficients from a set of input data. Therefore, the parametric algorithm estimates deterministic components, whereas random components are typically described as the error. Considering the model as a coefficient function of the independent variable, the error encodes random fluctuations around a Gaussian probability distribution. In many scopes, the goal is to minimize that error, classically addressed by means of LS fitting approaches, such as: linear mean LS, weighted LS, robust LS, non-linear LS, etc.

$$y = f(X, \beta) + \varepsilon \tag{2}$$

where y is the output vector data of n×1, corresponding to the input data in X of n×m, after being applied f as a non-linear function of the coefficient vector β (m×1), being ε the vector of unknown errors of n×1. Afterwords, the SSE is minimized, understood as the residual sum of squares, given a set of n data values, the residual value of the i-th value r_i is calculated as:

$$r_i = y_i - \hat{y}_i \tag{3}$$

where y_i represents the i-th observed value and $\hat{y}_i$ represents the i-th estimated value, and accordingly:

$$SSE = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{4}$$

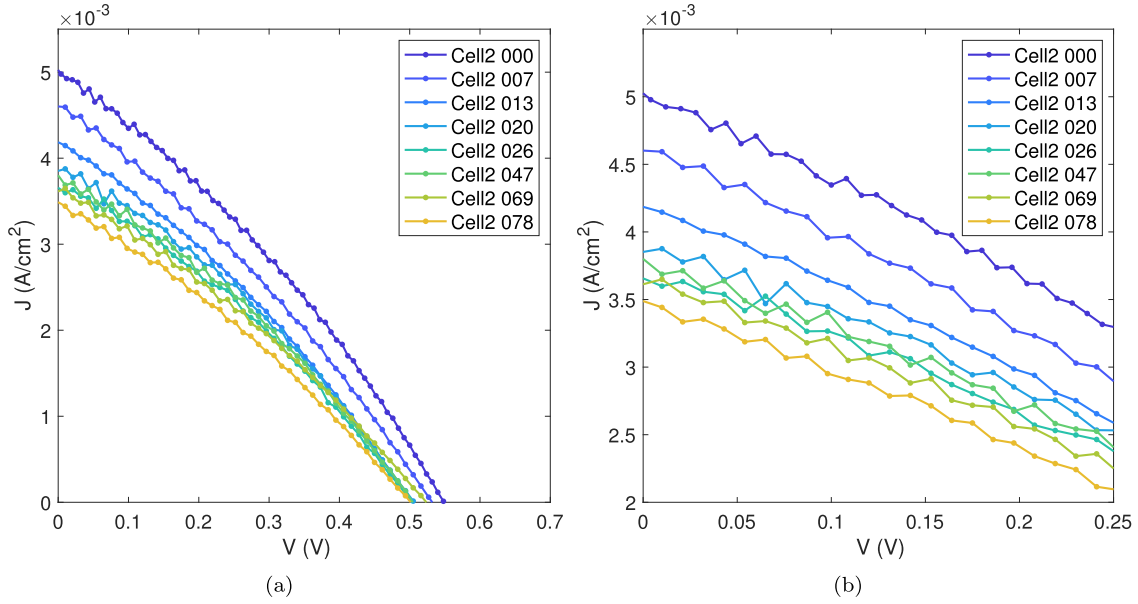

Fig. 2. Evolution of current density (J) vs. voltage (V) over time (days) for the OSC *Cell*2. (**a**) Curve J-V with V ∈[0-0.7] V. (**b**) Same curve J-V with V ∈[0-0.25] V.

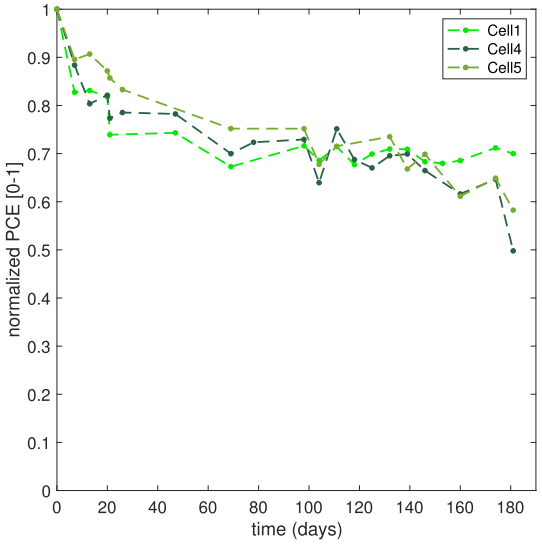

Fig. 3. Evolution of the normalized PCE over time for three different OSCs.

Subsequently, the algorithm proceeds iteratively calculating the coefficients from an initial seed. Sometimes non-linear models trust on heuristic schemes to calculate initial values. For others models, coefficients randomly initialized in ranges from [0-1]. Then the response value is given as $\hat{y} = f(X, \beta)$, computed using the jacobian matrix of $f(X, \beta)$, as the matrix that contains the partial derivatives with respect to the coefficients of β. Finally, the adjustment of the coefficients for the next iteration lies on some non-linear LS algorithms, such as Levenberg-Marquardt, Gradient descent or Gauss-Newton (Holland & Welsch, 1977). Whenever the fitting meets the specified convergence criteria, the final solution is assumed as valid.

After observing in Section 2.2 the behavior of PCE over time for the various OSCs contained in our database, it suggests that the LS regression fitting models yielding the best results are those with non-linear characterization. Table 3 displays the selected models along with their expressions as a function of time, dependent on the adjustment coefficients. It should be noted that the capability of these classical models lies solely in modeling the univariate behavior of the time effect on the PCE values. While it is possible to explore other LS regression fitting in the multivariate domain, they only allow for establishing linear relationships, which do not adequately capture the behavior of our devices.

Table 3
LS Bayesian regression fitting models to estimate PCE, denoted as $f(x)$, where x represents time by means of parametric models.

Parametric model	Coefficient expression of $f(x)$
$exp1$	ae^{bx}
$exp2$	$ae^{bx} + ce^{dx}$
$gauss1$	$a_1 e^{-[(x-b_1)/c_1]^2}$
$gauss2$	$a_1 e^{-[(x-b_1)/c_1]^2} + a_2 e^{-[(x-b_2)/c_2]^2}$
$poly3$	$p_1 x^3 + p_2 x^2 + p_3 x + p_4$

2.4. ML Framework

In contrast to the previous classical approaches, ML moves forward to produce non-parametric regression models that adjust more complex behaviors. In this work, we exploit the advantages of a software framework developed under Python, ROBERT[4] (Dalmau & Alegre-Requena, 2024b) (v1.0.6; see Appendix C for reproducibility details), that facilitates hyper-optimization and benchmarking over well-recognized ML regression models by single command line instruction. This automated framework consists of the following modules (for further details please check the online documentation (Dalmau and Alegre Requena, 2025)):

- CURATE (data curation): It processes the input dataset in order to filter correlated descriptors, noise, duplicates, as well as to identify and to convert categorical variables into one-hot descriptors.
- GENERATE (model selection): It iterates through multiple hyper-optimized models from scikit-learn (Pedregosa et al., 2011), including Random Forest (RF) (Breiman, 2001), Multivariate Linear Model (MVL) (Draper & Smith, 1998), Gradient Boosting (GB) (Friedman, 2001), Gaussian Process (GP) (Rasmussen & Williams, 2005), AdaBoost Regressor (AdaB) (Freund & Schapire, 1997), MLP Regressor Neural Network (NN) (Hornik et al., 1989), and Voting Regressor (VR) (Rokach, 2010). The algorithms are combined with different

[4] https://github.com//jvalegre/robert

training-validation split sizes, from 60-40 % to 90-10 % using random data splitting. For each combination of algorithm and training size, two models are generated: one with all the descriptors and another with only the most important variables detected by permutation feature importance (PFI) analysis. Among all the possibilities, the program selects two optimal models based on RMSE error.

- PREDICT (external predictions selection): the framework is able to predict new target values. Moreover, it provides feature importances using SHapley Additive exPlanations (SHAP) and Permutance Feature Importance (PFI) analysis, and outlier detection.
- VERIFY (assesing predictive ability): It assesses the predictive ability of the models, considering tests such as y-shuffle, y-mean, k-fold cross-validation, and prediction with one-hot encoding.
- REPORT (generation of PDF reports): With the aim to enhance reproducibility and transparency, this module offers a detailed report containing comprehensive information about the ML models utilized and replication instructions through command line executions.

3. Results

This section introduces the results obtained using both classical LS regression fitting and ML models to estimate the temporal behavior of our OSCs. The selected error metrics (de Azevedo Takara et al., 2024; Khan & Choi, 2025) for analysis are briefly presented below:

- Coefficient of determination R^2: It quantifies how well the independent variables explain the variability of the dependent variable. Higher values indicate that the model fits the data well and captures a larger proportion of the variability in the dependent variable.

$$R^2 = 1 - \frac{SSE}{SS_{tot}} \tag{5}$$

where SSE was defined in (4) and SS_{tot} is the total sum of squares.

- RMSE: It provides a measure of the average magnitude of the errors made by the model in its predictions. Minimizing this error is often a goal when training regression models in ML.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

where n is the number of samples, y_i is the real observed value and $\hat{y}_i$ is the predicted value.

- MAE: Another common objective error that measures the average magnitude of the errors between the predicted values and the actual values of the target variable. Unlike RMSE, which penalizes large errors more heavily, MAE treats all errors equally by taking the average of their absolute values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{7}$$

3.1. Classical LS regression fitting results

Fig. 4 presents the accuracy metrics of the five Bayesian regression models introduced in Table 3 to estimate the performance of PCE over 180 days, for each of the OSCs available in the database (Table 2). It is worth noting that these results correspond to the mean values, along with the standard deviation for each OSC. In general terms, it can be observed that these models tend to perform slightly better over short time periods compared to the longer experiments.

The maximum errors for all temporal fittings are bounded to values ~0.06. According to the results of Fig. 4, it is observed that on average, the *gauss*2 model, consisting of two Gaussian terms, is the one that best estimates the performance of the OSCs in terms of PCE over time. However, it should be noted that for a 30-day fit, this model requires more points to achieve a valid R^2. It is worth mentioning that these methods cannot characterize the behavior of all devices contained in

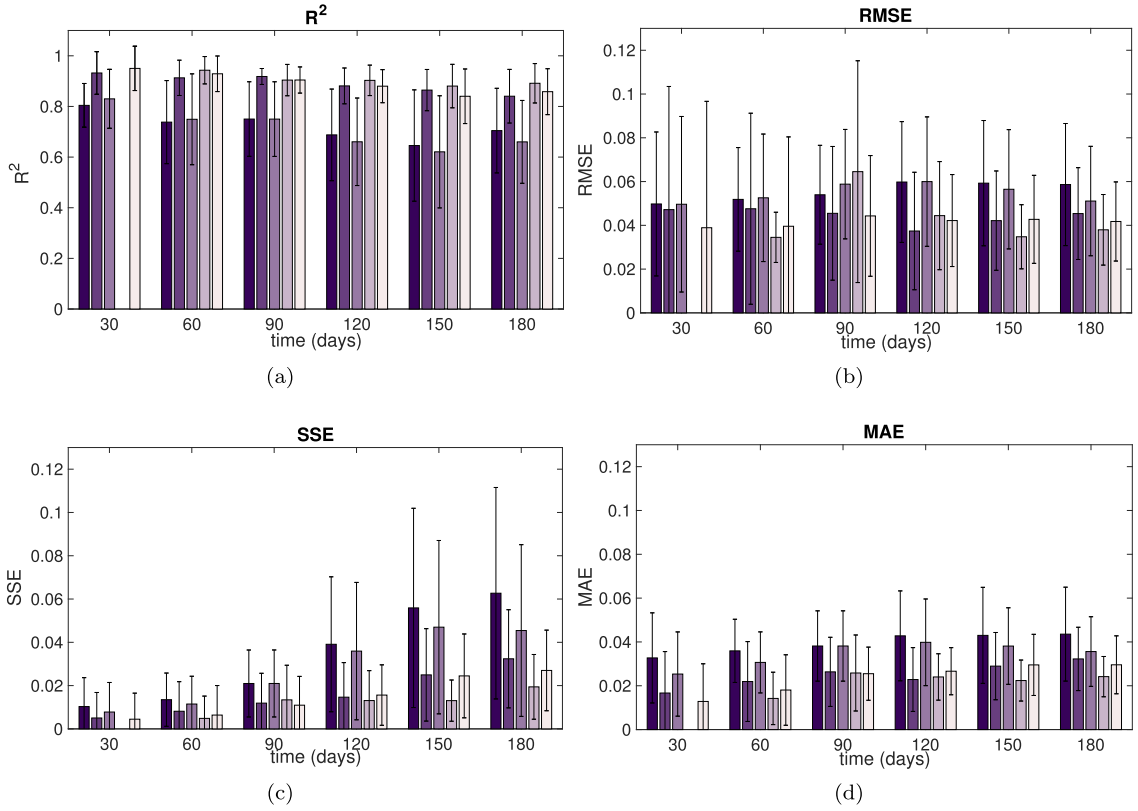

Fig. 4. Accuracy metrics of the LS-supported Bayesian regression fitting over time. Five different models are evaluated: *exp*1 ■, *exp*2 ■, *gauss*1 ■, *gauss*2 ■, *poly*3 ■. (**a**) Coefficient of determination, R^2. (**b**) Root mean squared error, RMSE. (**c**) Sum of squared errors, SSE. (**d**) Mean absolute error, MAE.

our database in a global manner, and the results obtained reflect solely the mean performance of fitting for individual OSCs.

3.1.1. PCE Prediction with LS

Once preliminary fittings have been made with traditional LS regression methods, it has been found that the best fitting corresponds to a double-term Gaussian function. Now we analyze the robustness of this method to predict the behavior of PCE against the temporal variable. For this purpose, four regression fittings have been obtained with PCE data acquired up to: 30, 60, 90, and 120 days, respectively. Then the behavior of PCE at 180 days has been predicted for each one (Fig. 5 and Table 4). It can be confirmed that only the fitting with data up to 120 days is able to predict reliably future PCE values. This result demonstrates that even the best parametric LS fitting regression model requires more than half of the temporal data to reliably model the PCE evolution at future time values.

3.2. ML Regression results

In this section, we generate ML models to study the temporal behavior of PCE. Notice that, unlike the previous LS-based regression fittings presented in Section 3.1, these ML models allow us to characterize the performance of all the OSC devices under a single model, while also using the multiple variables in the dataset. These include manufacturing parameters and environmental conditions at each measurement. To comparatively assess the ML models, Table 5 presents detailed accuracy results for the algorithms GB, NN, MVL, and RF, with training data up to 180 days. Please note that all these models have PFI enabled, so that the most important features are considered. For further details about these features see Section 3.2.2 and Appendix C.

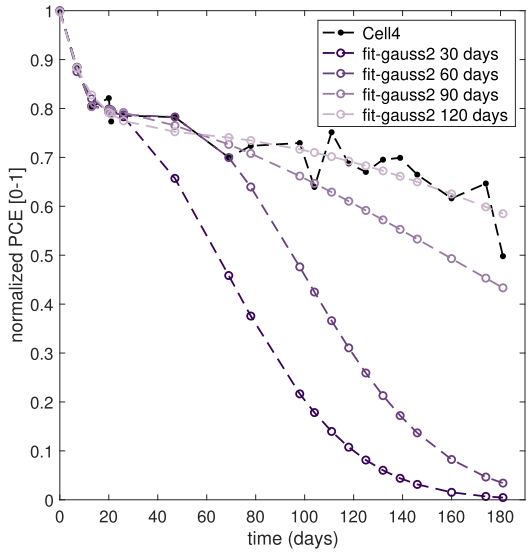

Fig. 5. PCE predition over time with the best LS-supported Bayesian regression fitting model (*gauss2*). Four temporal datasets are used to compute the fittings: 30 days –○–; 60 days –○–; 90 days –○– and 120 days –○–.

Table 4
Accuracy metrics of the best LS-supported Bayesian regression fitting model (*gauss2*) to predict PCE over time, presented in Fig. 5.

Fitting data	Prediction	RMSE	SSE	MAE
30 days	60 days	0.0516	0.0213	0.0277
	90 days	0.1141	0.1302	0.0680
	120 days	0.1872	0.4905	0.1339
	150 days	0.2224	0.8405	0.1705
	180 days	0.2442	1.3117	0.1979
60 days	90 days	0.0454	0.0206	0.0220
	120 days	0.1083	0.1641	0.0685
	150 days	0.1511	0.3879	0.1046
	180 days	0.1911	0.8038	0.1445
90 days	120 days	**0.0150**	**0.0032**	**0.0103**
	150 days	0.0160	0.0043	0.0119
	180 days	0.0190	0.0079	0.0118
120 days	150 days	0.0158	0.0042	0.0127
	180 days	0.0138	0.0038	0.0105

Table 5
Accuracy metrics of the ML regression models computed with training data up to 180 days. Run on an Apple M1 8-core 3,2 Ghz; Python 3.10.12.

Training data	ML model	$\mathbf{R}^2$	RMSE	SSE	MAE	time (s)
180 days	GB-90-10	0.96	0.0640	0.0595	0.0520	8.01
	GB-80-20	0.96	0.0720	0.0600	0.0560	7.73
	GB-70-30	0.94	0.0840	0.0672	0.0650	7.70
	GB-60-40	0.95	0.0700	0.0431	0.0500	7.77
	NN-90-10	0.66	0.1400	0.3732	0.1100	23.46
	NN-80-20	0.76	0.1100	0.1138	0.0670	22.99
	NN-70-30	0.60	0.1600	0.3555	0.1200	17.70
	NN-60-40	0.43	0.2500	0.7265	0.1800	15.65
	MVL-90-10	0.81	0.1000	0.2929	0.0810	**4.06**
	MVL-80-20	0.81	0.1000	0.2783	0.0800	4.07
	MVL-70-30	0.74	0.1400	0.2696	0.1000	4.10
	MVL-60-40	0.74	0.1300	0.2802	0.0920	4.09
	RF-90-10	**0.96**	**0.0480**	**0.0246**	**0.0390**	10.01
	RF-80-20	0.97	0.0520	0.0309	0.0370	10.66
	RF-70-30	0.94	0.0670	0.0398	0.0460	10.65
	RF-60-40	0.93	0.0850	0.0644	0.0580	10.39

The nomenclature for the models is: ALGORITHM-TRAINING-VALID (e.g. RF-90-10). Results with different training-validation ratios are presented: from 90-10 % to 60-40 %. In this study, the optimal ML models used RF algorithms with a 90-10 training-validation partitioning. It can be observed that, in general, all MAE and RMSE values are bounded within ranges ~[0.02-0.03], for training data of 120 days onwards.

The most robust model is highlighted in bold, considering the overall performance across all presented metrics, with priority given to the R^2 value. Please note that the supplementary material (Appendix A) contains these same metrics for further inspection of models obtained with other temporal ranges. Likewise, the detailed parameters for each model can be consulted in Appendix C. It is worth noting that errors of the ML methods are clearly bounded, regardless they operate on a multivariate database, in contrast to univariate Bayesian fitting presented in Section 3.1.

3.2.1. PCE Prediction with ML

After comparatively evaluating the performance of ML models, this section validates their ability to predict the temporal behavior of the PCE for an OSC device never seen by the models. Firstly, the models are trained and validated without data of *Cell*4. Standing out above the others are the GB-90-10 and RF-90-10 models. In line with previous results (Table 5), and thus considering its robustness to model the entire dataset, RF-90-10 has been selected to get predictive inferences up to 180 days for the unseen OSC *Cell*4. Fig. 6 provides such results. In particular, Fig. 6(a) presents the results of the model when the dataset is splitted into traning and validation, at 90 %-10 % (blue and yellow points), whereas Fig. 6(b) shows the prediction for *Cell*4 when its data are used as external test.

Next, Fig. 7 presents validation tests. Fig. 7(a) shows the RMSE error of the model compared to those of the tests: y-mean, y-shuffle and onehot. Fig. 7(b) produces a 5-fold cross-validation test (Kohavi, 1995), with satisfactory results. In consequence, neither data leakage nor overfitting is evidenced.

Finally, Fig. 8 compares the temporal evolution of the PCE for *Cell*4 with the predicted data obtained by the model.

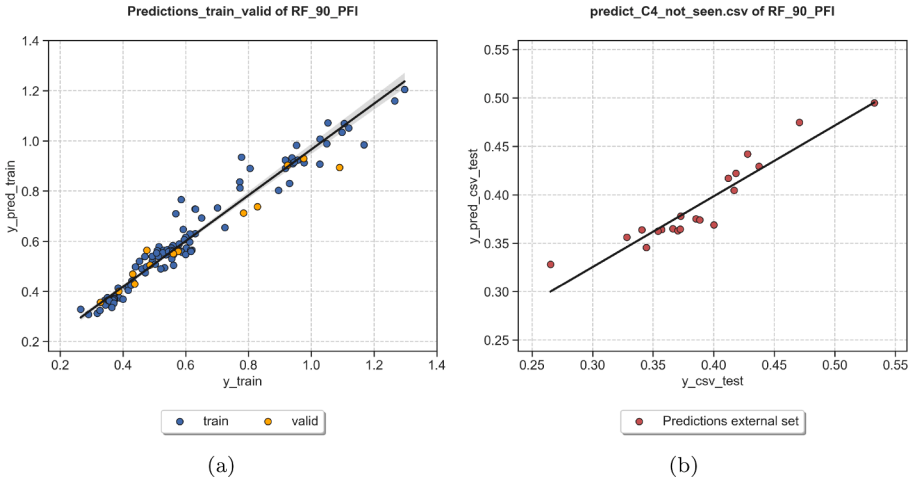

Fig. 6. Results of the best ML model, RF-90-10: (**a**) training-validation without data of *Cell*4 ($R^2 = 0.96$, MAE = 0.05, RMSE = 0.071); (**b**) prediction of *Cell*4 as external test ($R^2 = 0.88$, MAE = 0.015, RMSE = 0.021).

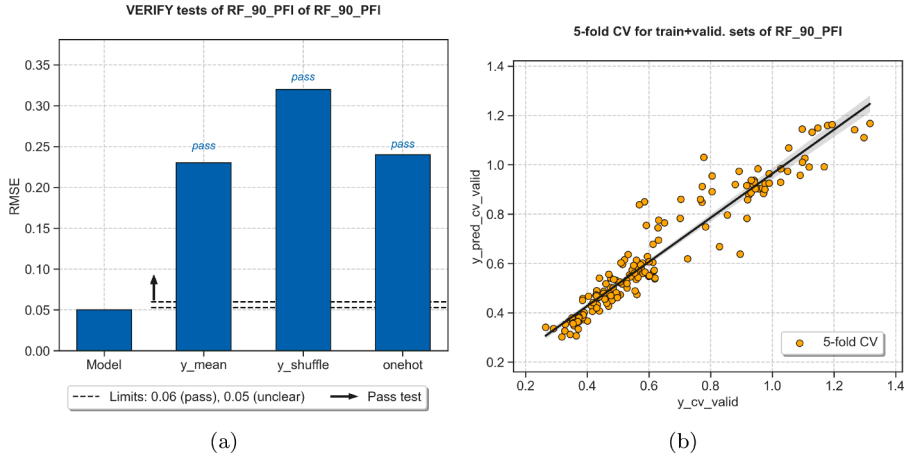

Fig. 7. Validation of the best ML model, RF-90-10: (**a**) RMSE values for validation tests y-mean, y-shuffle and onehot; (**b**) 5-fold cross-validation test ($R^2 = 0.89$, MAE = 0.06, RMSE = 0.09).

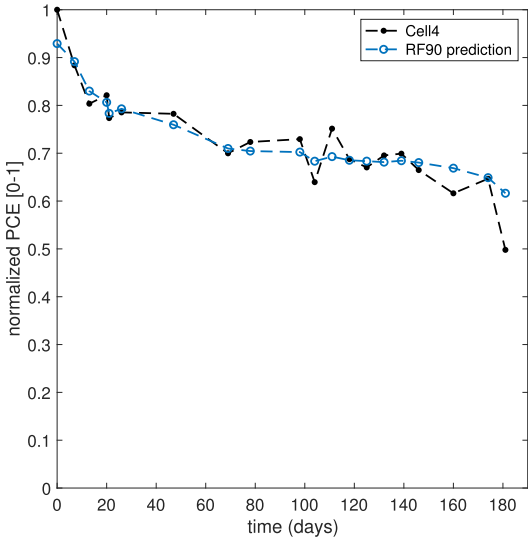

Fig. 8. PCE prediction over time with the best ML regression model (RF-90-10, blue) for an OSC not seen during the training-validation process (*Cell*4, black).

3.2.2. Feature analysis

Considering the results obtained in the previous section, it is worthwhile to study the top-performing ML model that characterizes the behavior of OSCs (i.e. RF-90-10). ML models trained with a limited number of data points are unable to characterize and predict with sufficient accuracy. For this reason, we focused on analyzing the complete dataset, with measures up to 180 days. In this regard, PFI and SHAP analyses are presented below, by means of Figs. 9 and 10, respectively. PFI evaluates the significance of individual features in a ML model by measuring the increase in the model's prediction error after permuting the values of a specific feature. The resulting increase in error indicates the dependency on that feature (Breiman, 2001; Fisher et al., 2019). Besides, SHAP permits interpreting individual predictions of ML models by computing the contribution of each feature to the model's as an additive feature attribution method (Lundberg & Lee, 2017), i.e., it permits showing how much a feature biases the model's prediction up or down.

Then Fig. 9 presents the influence of the most relevant feature for the model, after applying PFI filtering, that is, removing variables with low effect on the R^2. It is confirmed that the amount of solvent (PEDOT:PSS) in the HTL layer has the most significant influence. Likewise, the P3HT:PCBM ratio also demonstrates certain relevance. Additionally, dependencies with the value of PCBM are also observed, since it constraints the P3HT:PCBM ratio, given its non-linear influence in the denominator.

As for the environmental conditions, it is known that low humidity proves to be beneficial for these OSCs, which in our geographical location is normally correlated with high temperatures and atmospheric pressures. In a similar manner it acts the dew point, which directly

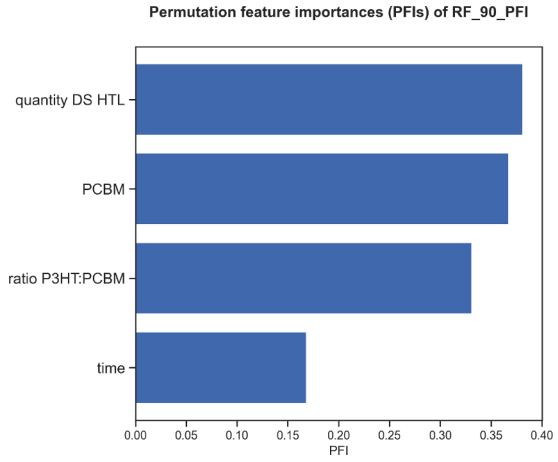

Fig. 9. PFI results for the best ML model generated with data up to 180 days.

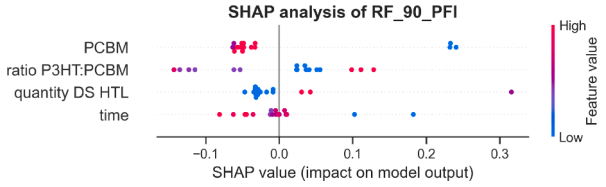

Fig. 10. SHAP results for the best ML model generated with data up to 180 days.

correlates with humidity. Nonetheless, the device encapsulation demonstrates that the effect of these variables is minimized, as PFI filtering proves that they are not relevant enough to influence substantially the model.

Regarding the SHAP analysis, Fig. 10 validates the previous insights: high values of the amount of solvent PEDOT:PSS in the HTL layer and high ratios of P3HT:PCBM contribute positively to get higher and stable PCE values. Besides this, the effect of the PCBM value on the higher P3HT:PCBM ratio is again confirmed.

Overall, the importance of the quantity of PEDOT:PSS in the HTL layer has been demonstrated as it plays an essential role in the multilayer structure of the OSCs. It confirms its relevance as the second layer to cover sufficiently the substrate and thus increasing the PCE and its stability. Moreover, the ratio P3HT:PCBM also demonstrates its positive role in dealing as charge carrier in the HTL layer and therefore assuring higher PCE values.

4. Discussion and conclusions

This paper has presented the application of optimal ML frameworks to characterize the degradation behavior, in terms of PCE, of OSCs with multilayer structure ITO/PEDOT:PSS/P3HT:PCBM/Al. A dataset with 166 entries entries was created, containing PCE values of various OSC devices measured over 180 days. This dataset was supplemented with seven variables describing environmental conditions of the experiments and manufacturing parameters of the devices. Through hyperoptimization of a set of ML models we have presented an accuracy analysis of different methods, which were fed with OSC data periodized into sets from 30 to 180 days. The benchmarking has confirmed the validity of models like RF or GB to confer R^2 values over 0.90, reaching in some extents R^2~0.96–0.97 and error metrics (RMSE, SSE and MAE) significantly low when long term data is used for training. To reinforce the suitability of these ML models, classical LS regression fitting methods have been compared. These proved not to be suitable for a multivariable dataset like ours, especially when dealing with long term data of the OSCs. Consequently, their ability to predict PCE values is highly unreliable.

ML models proved to offer high feasibility to predict the behavior of unknown OSCs. Feature analysis suggests that the most influential variable is the solvent in the HTL layer, i.e., the amount of PEDOT:PSS. This is explained as the multilayer structure of the OSCs needs a minimum value of PEDOT:PSS to ensure a layer that completely covers the substrate. Moreover, the ratio P3HT:PCBM also exhibits significant importance, being higher values representative of greater impact on the model. Finally, it has been observed that variables such as temperature, humidity, dew point and pressure have lesser impact on the models, explained by the encapsulation made to the OSCs during their manufacturing.

As the main limitations and challenges faced during this implementation of ML with OSCs, we observed that the accuracy of the measurement equipments is a key aspect so as to have valid data to feed the models. Valid and wide ranges for the data are highly needed, according to feasible OSC manufacturing process. Otherwise, models might predict unfeasible OSCs, whose manufacture is not possible.

5. Glossary

- Al: Aluminium.
- ITO: Indium tin oxide.
- PEDOT:PSS: poly(3,4-ethylenedioxythiophene) polystyrene sulfonate.
- P3HT:PCBM: (poli(3-hexiltiofeno-2,5-diil):[6,6]-phenyl-C61-butyric acid methyl ester.
- PEDOT:PSS: poly(3,4-ethylenedioxythiophene) polystyrene sulfonate.

CRediT authorship contribution statement

Conceptualization D.V. and J.V.A.-R.; Methodology D.V., J.V.A.-R., D.D. and F.R.-M.; Software D.V., J.V.A.-R., D.D. and M.F.; Data curation D.V. and M.F.; Formal analysis D.V., F.R.-M.; Funding acquisition D.V., J.C.F. and J.V.A.-R.; Investigation D.V., F.R.-M. and J.V.A.-R.; Resources F.R.-M. and J.C.F.; Supervision J.C.F. and J.V.A.-R.; Validation D.D. and M.F.; Visualization D.V. and M.F; Writing – original draft D.V. and J.V.A.-R.; Writing – review & editing J.V.A.-R., D.D. and F.R.-M.

Data availability

The program ROBERT (v1.0.6) to model the database is publicly available at: https://github.com/jvalegre/robert/releases. Its online documentation (Dalmau and Alegre Requena, 2025) at: https://robert.readthedocs.io/en/latest/index.html. The datasets are available in Appendix B and at: https://github.com/POLI-NANO/OSCs

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.

Acknowledgements

This work is part of the project CIAICO/2023/193 funded by Generalitat Valenciana. The authors acknowledge that support. J.V.A.-R. and D.D. acknowledge Gobierno de Aragón Fondo Social Europeo (Research Group E07_23R), the State Research Agency of Spain (MCIN/ AEI/ 10.13039/501100011033/ FEDER, UE) for financial support (PID2022-140159NA-I00) and the European Union's Recovery and Resilience Facility-Next Generation in the framework of the General Invitation of the Spanish Government's public business entity Red.es to participate in talent attraction and retention programmes within Investment 4 of Component 19 of the Recovery, Transformation and Resilience Plan (MOMENTUM, MMT24-ISQCH-01).

Appendix A. Extended accuracy metrics of ML models

This section comprises extended accuracy metrics of the ML models used in this work, when the temporal scope of the training-validation data is varied: 30, 60, 90, 120 and 150 days, respectively. In particular, the accuracy study of the best ML model (RF-90-10) over time, is presented in Fig. A.1 (Tables A.1–A.5).

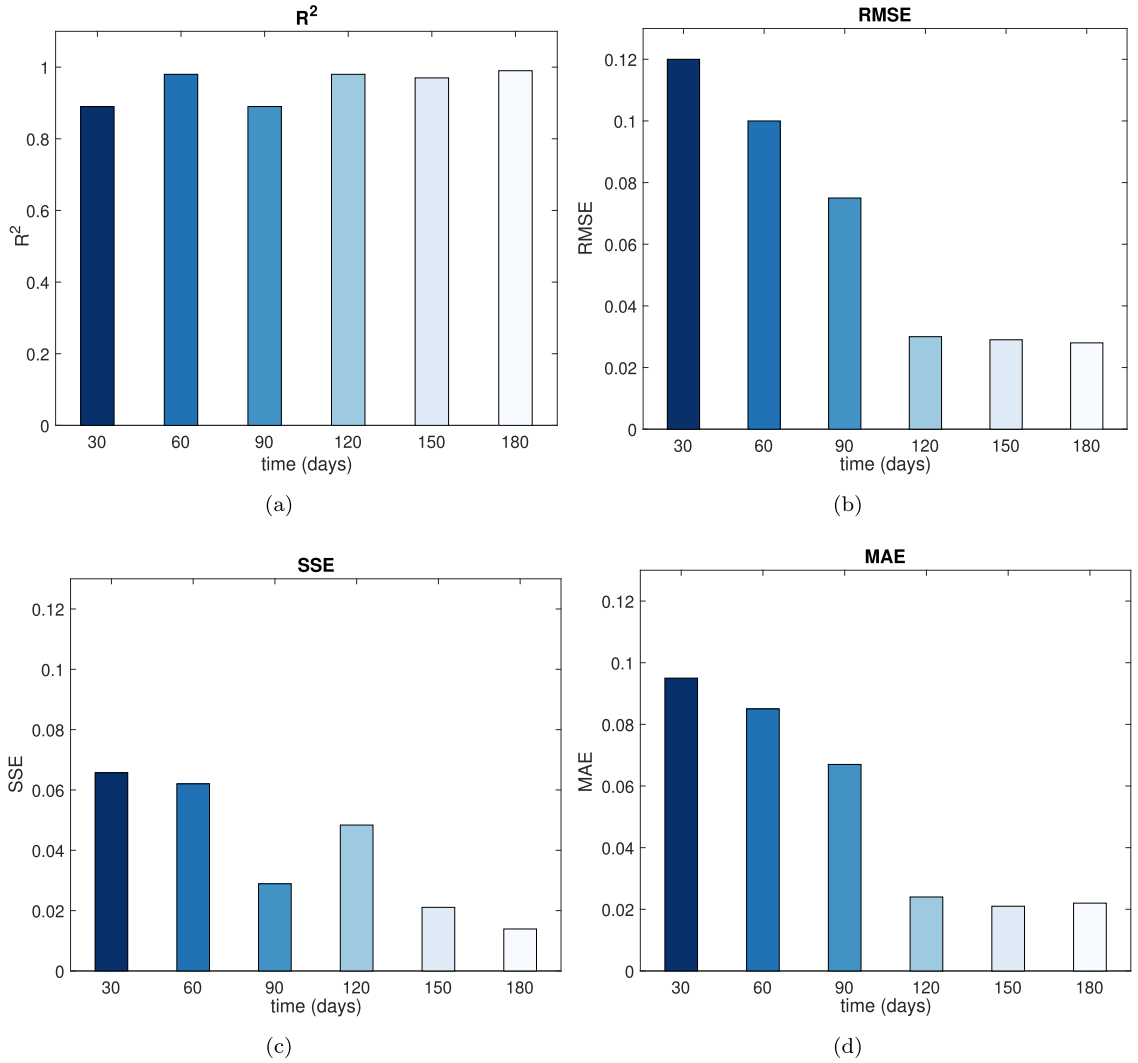

Fig. A.1. Accuracy of the best ML regression models over time: 30 days RF-60-40 ■; 60 days RF-90-10 ■; 90 days RF-90-10 ■; 120 days RF-90-10 ■; 150 days RF-90-10 ■ and 180 days RB-90-10 ■. (**a**) Coefficient of determination, R^2. (**b**) Root mean squared error, RMSE. (**c**) Sum of squared errors, SSE. (**d**) Mean absolute error, MAE.

Table A.1
Accuracy metrics of the ML regression models computed with training data up to 30 days.

Training data	ML model	$\mathbf{R^2}$	RMSE	SSE	MAE
30 days	GB-80-20	0.86	0.140	0.0647	0.1200
	GB-70-30	0.73	0.140	0.0622	0.1200
	GB-60-40	0.84	0.160	0.1103	0.1500
	NN-80-20	0.76	0.180	0.1655	0.1800
	NN-70-30	0.51	0.170	0.0832	0.1600
	NN-60-40	0.62	0.170	0.0931	0.1400
	MVL-80-20	0.71	0.200	0.0984	0.1500
	MVL-70-30	0.62	0.220	0.1006	0.1400
	MVL-60-40	0.66	0.170	0.1067	0.1400
	RF-80-20	0.86	0.130	0.0573	0.1000
	RF-70-30	0.78	0.130	**0.0558**	0.1000
	RF-60-40	**0.89**	**0.120**	0.0725	**0.0950**

Table A.2
Accuracy metrics of the ML regression models computed with training data up to 60 days.

Training data	ML model	$\mathbf{R}^2$	RMSE	SSE	MAE
60 days	GB-90-10	0.85	0.1000	**0.0155**	**0.0800**
	GB-80-20	0.76	0.1500	0.0433	0.1300
	GB-70-30	0.84	0.1500	0.0597	0.1300
	GB-60-40	0.78	0.1700	0.1275	0.1500
	NN-90-10	0.89	0.1300	0.1239	0.1000
	NN-80-20	0.86	0.1700	0.1677	0.1500
	NN-70-30	0.88	0.1500	0.0506	0.1300
	NN-60-40	0.79	0.1600	0.0677	0.1400
	MVL-90-10	0.71	0.1500	0.1059	0.1300
	MVL-80-20	0.80	0.1700	0.1120	0.1500
	MVL-70-30	0.63	0.1700	0.1117	0.1500
	MVL-60-40	0.61	0.1900	0.1290	0.1500
	RF-90-10	**0.98**	**0.1000**	0.0702	0.0850
	RF-80-20	0.92	0.1400	0.0832	0.1300
	RF-70-30	0.78	0.1600	0.0747	0.1400
	RF-60-40	0.81	0.1300	0.0578	0.1000

Table A.3
Accuracy metrics of the ML regression models computed with training data up to 90 days.

Training data	ML model	$\mathbf{R}^2$	RMSE	SSE	MAE
90 days	GB-90-10	0.87	**0.0730**	**0.0255**	**0.0660**
	GB-80-20	0.87	0.1100	0.0736	0.0860
	GB-70-30	0.88	0.0940	0.0775	0.0770
	GB-60-40	0.75	0.1400	0.0755	0.1100
	NN-90-10	0.86	0.0790	0.0332	0.0630
	NN-80-20	0.82	0.1300	0.1595	0.1200
	NN-70-30	0.46	0.1900	0.1071	0.1700
	NN-60-40	0.52	0.1900	0.1611	0.1500
	MVL-90-10	0.61	0.1100	0.3161	0.1000
	MVL-80-20	0.63	0.1700	0.1468	0.1400
	MVL-70-30	0.54	0.1700	0.1378	0.1400
	MVL-60-40	0.55	0.1800	0.1569	0.1500
	RF-90-10	**0.89**	0.0750	0.0300	0.0670
	RF-80-20	0.87	0.0940	0.0909	0.0700
	RF-70-30	0.80	0.1000	0.0613	0.0780
	RF-60-40	0.74	0.1400	0.0860	0.1100

Table A.4
Accuracy metrics of the ML models computed with training data up to 120 days.

Training data	ML model	$\mathbf{R}^2$	RMSE	SSE	MAE
120 days	GB-90-10	0.95	0.0670	**0.0312**	0.0540
	GB-80-20	0.88	0.0700	0.0324	0.0510
	GB-70-30	0.89	0.0810	0.0453	0.0620
	GB-60-40	0.81	0.1300	0.1447	0.1100
	NN-90-10	0.88	0.0600	0.2140	0.0450
	NN-80-20	0.83	0.0890	0.0576	0.0660
	NN-70-30	0.71	0.1300	0.0767	0.0960
	NN-60-40	0.59	0.1700	0.1743	0.1300
	MVL-90-10	0.70	0.0990	0.1729	0.0850
	MVL-80-20	0.65	0.1400	0.1820	0.1100
	MVL-70-30	0.60	0.1600	0.2311	0.1400
	MVL-60-40	0.51	0.1800	0.2370	0.1500
	RF-90-10	**0.98**	**0.0300**	0.0507	**0.0240**
	RF-80-20	0.88	0.0680	0.0651	0.0480
	RF-70-30	0.86	0.0920	0.0609	0.0690
	RF-60-40	0.84	0.1000	0.0574	0.0770

Table A.5
Accuracy metrics of the ML regression models computed with training data up to 150 days.

Training data	ML model	R^2	RMSE	SSE	MAE
150 days	GB-90-10	0.98	0.0300	0.0510	0.0230
	GB-80-20	0.93	0.0660	0.0242	0.0420
	GB-70-30	0.94	0.0730	0.0351	0.0580
	GB-60-40	0.91	0.0820	0.0591	0.0640
	NN-90-10	0.90	0.0580	0.0760	0.0540
	NN-80-20	0.82	0.0110	0.1277	0.0630
	NN-70-30	0.88	0.0960	0.0871	0.0740
	NN-60-40	0.78	0.1200	0.1533	0.0910
	MVL-90-10	0.84	0.0920	0.2113	0.0800
	MVL-80-20	0.87	0.1100	0.2139	0.0840
	MVL-70-30	0.88	0.1200	0.2163	0.0900
	MVL-60-40	0.82	0.1100	0.2184	0.0910
	RF-90-10	**0.97**	**0.0290**	**0.0219**	**0.0210**
	RF-80-20	0.96	0.0510	0.0394	0.0324
	RF-70-30	0.94	0.0780	0.0417	0.0570
	RF-60-40	0.90	0.0930	0.1069	0.0943

Appendix B. OSC Database

Database in *csv* file consisting of 166 entries which includes up to seven variables regarding both the manufacturing process and environmental conditions for more than 180 days (Table 2). PCE values of several polymeric OSCs with a multilayer structure ITO/PEDOT:PSS/P3HT:PCBM/Al were measured.

In addition, to ensure reproducibility, another *csv* file provides a dataset without data of *Cell*4. Finally, another *csv* file can be used to test the predictive ability of the model to predict data of such cell (see Section 3.2.1.)

Appendix C. ROBERT Report

Detailed report obtained with ROBERT (v1.0.6) after modeling the entire database. Results and reproducibility details are contained in this file.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.eswa.2025.128890

References

Alsulami, B.N.N., David, T.W., Essien, A., Kazim, S., Ahmad, S., Jacobsson, T.J., Feeney, A., & Kettle, J. (2024). Application of large datasets to assess trends in the stability of perovskite photovoltaics through machine learning. *Journal of Materials Chemistry A*, *12*, 3122–3132. https://doi.org/10.1039/D3TA05966A

Ayub, A.R., Akram, W., Khan, M.A., Zeshan, M., Ateeb, M., Yaqoob, U., Dad, M.U., Elmushyakhi, A., Shawky, A.M., & Iqbal, J. (2023). Designing of asymmetric non-fullerene based acceptor materials by re-modification of spacers with PCE for organic solar cell. *Optik*, *278*, 170602. https://doi.org/10.1016/j.ijleo.2023.170602

Borazan, I. (2019). A study about lifetime of photovoltaic fibers. *Solar Energy Materials and Solar Cells*, *192*, 52–56. https://doi.org/10.1016/j.solmat.2018.12.003

Bozorg, M., Bracale, A., Caramia, P., Carpinelli, G., Carpita, M., & De Falco, P. (2020). Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting. *Protection and Control of Modern Power Systems*, *5*(1), 21. https://doi.org/10.1186/s41601-020-00167-7

Brabec, C.J., Hauch, J.A., Schilinsky, P., & Waldauf, C. (2005). Production aspects of organic photovoltaics and their impact on the commercialization of devices. *MRS Bulletin*, *30*(1), 50–52. https://doi.org/10.1557/mrs2005.10

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chiappafreddo, P., & Gagliardi, A. (2010). The photovoltaic market facing the challenge of organic solar cells: Economic and technical perspectives. *Transition Studies Review*, *17*(2), 346–355. https://doi.org/10.1007/s11300-010-0148-0

Dalmau, D., & Alegre-Requena, J.V. (2024a). Integrating digital chemistry within the broader chemistry community. *Trends in Chemistry*, *6*(8), 459–469. Special Issue: Emerging Leaders in Chemistry - Part II. https://doi.org/10.1016/j.trechm.2024.06.008

Dalmau, D., & Alegre-Requena, J.V. (2024b). ROBERT: Bridging the gap between machine learning and chemistry. *WIREs Computational Molecular Science*, *14*(5), e1733. https://doi.org/10.1002/wcms.1733

Dalmau, D., Alegre Requena, J.V., (2025). ROBERT's documentation. accessed on 2025-02-03. https://robert.readthedocs.io/en/latest/index.html.

de Azevedo Takara, L., Teixeira, A.C., Yazdanpanah, H., Mariani, V.C., & dos Santos Coelho, L. (2024). Optimizing multi-step wind power forecasting: Integrating advanced deep neural networks with stacking-based probabilistic learning. *Applied Energy*, *369*, 123487. https://doi.org/10.1016/j.apenergy.2024.123487

Draper, N.R., & Smith, H. (1998). Applied regression analysis. (3rd ed.). New York: Wiley. https://doi.org/10.1002/9781118625590

Du, X., Zhang, L., Zhao, P. et al. (2024). Machine learning and automated screening of organic solar cells for air and light resilience. *InfoMat*, *6*(4), e12554. https://doi.org/10.1002/inf2.12554

Dwivedi, D., Babu, K.V.S.M., Yemula, P.K., Chakraborty, P., & Pal, M. (2024). Identification of surface defects on solar PV panels and wind turbine blades using attention based deep learning model. *Engineering Applications of Artificial Intelligence*, *131*, 107836. https://doi.org/10.1016/j.engappai.2023.107836

Eibeck, A., Nurkowski, D., Menon, A., Bai, J., Wu, J., Zhou, L., Mosbach, S., Akroyd, J., & Kraft, M. (2021). Predicting power conversion efficiency of organic photovoltaics: Models and data analysis. *ACS Omega*, *6*(37), 23764–23775. https://doi.org/10.1021/acsomega.1c02156

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, *20*(177), 1–81. https://doi.org/10.48550/arXiv.1801.01489

Freund, Y., & Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/jcss.1997.1504

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Hachmann, J., Olivares-Amaya, R., Atahan-Evrenk, S., Amador-Bedolla, C., Sánchez-Carrera, R.S., Gold-Parker, A., Vogt, L., Brockway, A.M., & Aspuru-Guzik, A. (2011). The Harvard clean energy project: Large-scale computational screening and design of organic photovoltaics on the world community grid. *The Journal of Physical Chemistry Letters*, (17), 2241–2251. https://doi.org/10.1021/jz200866s

Hansen, K., Montavon, G., Biegler, F., Fazli, S., Rupp, M., Scheffler, M., von Lilienfeld, O.A., Tkatchenko, A., & Müller, K.R. (2013). Assessment and validation of machine learning methods for predicting molecular atomization energies. *Journal of Chemical Theory and Computation*, *9*(8), 3404–3419. https://doi.org/10.1021/ct400195d

Holland, P., & Welsch, R.E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics-theory and Methods*, *6*, 813–827. https://api.semanticscholar.org/CorpusID:122208291.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Hußner, M., Machui, F., & Berny, S. (2024). Machine learning extraction of recombination and mobility parameters from current? Voltage curves. *Advanced Energy Materials*, *14*(5), 2303000. https://doi.org/10.1002/aenm.202303000

Jobayer, M., Shaikat, M.A.H., Naimur Rashid, M., & Hasan, M.R. (2023). A systematic review on predicting PV system parameters using machine learning. *Heliyon*, *9*. https://doi.org/10.1016/j.heliyon.2023.e16815

Ju, M.G., Chen, M., Zhou, Y., Dai, J., Ma, L., Padture, N.P., & Zeng, X.C. (2018). Toward eco-friendly and stable perovskite materials for photovoltaics. *Joule*, *2*(7), 1231–1241. https://doi.org/10.1016/j.joule.2018.04.026

Khan, T., & Choi, C. (2025). Attention enhanced dual stream network with advanced feature selection for power forecasting. *Applied Energy*, *377*, 124564. https://doi.org/10.1016/j.apenergy.2024.124564

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on artificial intelligence* (pp. 1137–1143). Morgan Kaufmann Publishers.

Lee, M.H. (2023). Interpretable machine learning model for the highly accurate prediction of efficiency of ternary organic solar cells based on nonfullerene acceptor using effective molecular descriptors. *Solar RRL*, *7*(14), 2300307. https://doi.org/10.1002/solr.202300307

Lin, W., He, Z., Zhang, S. et al. (2025). Multi-component copolymerization for mechanically durable and stable organic solar cells. *Angewandte Chemie International Edition*, *64*(19), e202420121. https://doi.org/10.1002/anie.202420121

Lopez, S.A., Pyzer-Knapp, E.O., Simm, G.N., Lutzow, T., Li, K., Seress, L.R., Hachmann, J., & Aspuru-Guzik, A. (2016). The Harvard organic photovoltaic dataset. *Scientific Data*, *3*(1), 160086. https://doi.org/10.1038/sdata.2016.86

Liu, H., Kim, M., Karki, A., (2024). Data-Driven Insights into Organic Solar Cell Stability through Morphological Analysis and Machine Learning. *Advanced Materials 36* (3), 2300259. https://doi.org/10.1002/adma.202300259.

Lundberg, S.M., & Lee, S.I. (2017). A unified approach to interpreting model predictions. In *Advances in neural information processing systems*. (vol. *30*). https://arxiv.org/abs/1705.07874.

Lusci, A., Pollastri, G., & Baldi, P. (2013). Deep architectures and deep learning in chemoinformatics: The prediction of aqueous solubility for drug-like molecules. *Journal of Chemical Information and Modeling*, *53*(7), 1563–1575. https://doi.org/10.1021/ci400187y

Mahmood, A., Irfan, A., & Wang, J.L. (2022). Machine learning and molecular dynamics simulation-assisted evolutionary design and discovery pipeline to screen efficient small molecule acceptors for PTB7-Th-based organic solar cells with over 15% efficiency. *Journal of Materials Chemistry A*, *10*, 4170–4180. https://doi.org/10.1039/D1TA09762H

Mahmood, A., Sandali, Y., & Wang, J.L. (2023). Easy and fast prediction of green solvents for small molecule donor-based organic solar cells through machine learning. *Physical Chemistry Chemical Physics*, *25*, 10417–10426. https://doi.org/10.1039/D3CP00177F

Mahmood, A., & Wang, J.L. (2021). Machine learning for high performance organic solar cells: Current scenario and future prospects. *Energy & Environmental Science*, *14*, 90–105. https://doi.org/10.1039/D0EE02838J

Malhotra, P., Biswas, S., Chen, F.C., & Sharma, G.D. (2021). Prediction of non-radiative voltage losses in organic solar cells using machine learning. *Solar Energy*, *228*, 175–186. https://doi.org/10.1016/j.solener.2021.09.056

Mammeri, M., Dehimi, L., Bencherif, H., Amami, M., Ezzine, S., Pandey, R., & Hossain, M.K. (2023). Targeting high performance of perovskite solar cells by combining electronic, manufacturing and environmental features in machine learning techniques. *Heliyon*, *9*(11). https://doi.org/10.1016/j.heliyon.2023.e21498

Miyake, Y., & Saeki, A. (2021). Machine learning-assisted development of organic solar cell materials: Issues, analyses, and outlooks. *The Journal of Physical Chemistry Letters*, *12*(51), 12391–12401. https://doi.org/10.1021/acs.jpclett.1c03526

Moore, G.J., Bardagot, O., & Banerji, N. (2022). Deep transfer learning: A fast and accurate tool to predict the energy levels of donor molecules for organic photovoltaics. *Advanced Theory and Simulations*, *5*(5), 2100511. https://doi.org/10.1002/adts.202100511

New Media Consortium, (2025). New Media Consortium Horizon Report. accessed on 2025-03-02. https://www.nrel.gov/pv/cell-efficiency.html.

Nguyen, D.C., & Ishikawa, Y. (2023). On predicting annual output energy of 4-terminal perovskite/silicon tandem PV cells for building integrated photovoltaic application using machine learning. *Heliyon*, *9*(7). https://doi.org/10.1016/j.heliyon.2023.e18097

Osterrieder, B., Gao, K., Lewis, N., & Aspuru-Guzik, A. (2023). AI-guided autonomous optimization of organic solar cells via spectral data and bayesian models. arXiv preprint arXiv:2305.08248 https://arxiv.org/abs/2305.08248

Owolabi, T.O., Akande, K.O., Olatunji, S.O., Aldhafferi, N., & Alqahtani, A. (2017). Modeling energy band gap of doped TiO_2 semiconductor using homogeneously hybridized support vector regression with gravitational search algorithm hyper-parameter optimization. *AIP Advances*, *7*(11), 115225. https://doi.org/10.1063/1.5009693

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html.

Rasmussen, C.E., & Williams, C.K.I. (2005). Gaussian Processes for Machine Learning. The MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001

Ritchie, H., Roser, M., (2024). Energy. accessed on 2024-21-12. https://ourworldindata.org/energy.

Rodriguez-Mas, F., Valiente, D., Ferrer, J.C., Alonso, J.L., & Fernandez de Avila, S. (2023). Towards a greener photovoltaic industry: Enhancing efficiency environmental sustainability and manufacturing costs through solvent optimization in organic solar cells. *Heliyon*, *9*(12), e23099. https://doi.org/10.1016/j.heliyon.2023.e23099

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, *33*(1), 1–39. https://doi.org/10.1007/s10462-009-9124-7

Rupp, M., Tkatchenko, A., Müller, K.R., & von Lilienfeld, O.A. (2012). Fast and accurate modeling of molecular atomization energies with machine learning. *Physical Review Letters*, *108*, 058301. https://doi.org/10.1103/PhysRevLett.108.058301

Ryu, S., Kwon, Y., & Kim, W.Y. (2019). A Bayesian graph convolutional network for reliable prediction of molecular properties with uncertainty quantification. *Chemical Science*, *10*, 8438–8446. https://doi.org/10.1039/C9SC01992H

Sanosa, N., Dalmau, D., Sampedro, D., Alegre-Requena, J.V., & Funes-Ardoiz, I. (2024). Recent advances of machine learning applications in the development of experimental homogeneous catalysis. *Artificial Intelligence Chemistry*, *2*(1), 100068. https://doi.org/10.1016/j.aichem.2024.100068

Seo, J.H., Um, H.D., Shukla, A., Hwang, I., Park, J., Kang, Y.C., Kim, C.S., Song, M., & Seo, K. (2015). Low-temperature solution-processed flexible organic solar cells with PFN/agNWs cathode. *Nano Energy*, *16*, 122–129. https://doi.org/10.1016/j.nanoen.2015.06.013

Serrano-Lujan, L., Cadenas, J.M., Faxas-Guzman, J., & Urbina, A. (2016). Case of study: Photovoltaic faults recognition method based on data mining techniques. *Journal of Renewable and Sustainable Energy*, *8*(4), 043506. https://doi.org/10.1063/1.4960410

Sun, W., Li, M., Li, Y., Wu, Z., Sun, Y., Lu, S., Xiao, Z., Zhao, B., & Sun, K. (2019). The use of deep learning to fast evaluate organic photovoltaic materials. *Advanced Theory and Simulations*, *2*(1), 1800116. https://doi.org/10.1002/adts.201800116

Tang, Y., Yang, K., Zhang, S., & Zhang, Z. (2024). Photovoltaic power forecasting: A dual-attention gated recurrent unit framework incorporating weather clustering and transfer learning strategy. *Engineering Applications of Artificial Intelligence*, *130*, 107691. https://doi.org/10.1016/j.engappai.2023.107691

Toledo, C., Serrano-Lujan, L., Abad, J., Lampitelli, A., & Urbina, A. (2019). Measurement of thermal and electrical parameters in photovoltaic systems for predictive and cross-correlated monitorization. *Energies*, *12*(4). https://doi.org/10.3390/en12040668

Unterthiner, T., Mayr, A., & Wegner, J.K. (2015). Deep learning as an opportunity in virtual screening. In *Deep learning workshop at NIPS*. https://api.semanticscholar.org/CorpusID:38235267.

Venkata, M., Agrios, A., & Ahmad, S. (2016). Current and emerging technologies and materials for solar power conversion (vol. 1). CRC Press. https://doi.org/10.1201/9781315369853

Wallach, I., Dzamba, M., & Heifets, A. (2015). Atomnet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv:1510.02855 https://doi.org/10.48550/arXiv.1510.02855

Wolf, P., & Benda, V. (2013). Identification of PV solar cells and modules parameters by combining statistical and analytical methods. *Solar Energy*, *93*, 151–157. https://doi.org/10.1016/j.solener.2013.03.018

Würfel, P., & Würfel, U. (2016). Physics of solar cells: From basic principles to advanced concepts. Wiley-VCH. https://doi.org/10.1201/9781315369853

Yeh, N., & Yeh, P. (2013). Organic solar cells: Their developments and potentials. *Renewable and Sustainable Energy Reviews*, *21*, 421–431. https://doi.org/10.1016/j.rser.2012.12.046

Yildirim, M.O., Gok Yildirim, E.C., Eren, E., Huang, P., Haris, M.P.U., Kazim, S., Vanschoren, J., Uygun Oksuz, A., & Ahmad, S. (2023). Automated machine learning approach in material discovery of hole selective layers for perovskite solar cells. *Energy Technology*, *11*(1), 2200980. https://doi.org/10.1002/ente.202200980

Zhang, Y., Yi, H., Iraqi, A., Kingsley, J., Buckley, A., Wang, T., & Lidzey, D.G. (2017). Comparative indoor and outdoor stability measurements of polymer based solar cells. *Scientific Reports*, *7*(1), 1305. https://doi.org/10.1038/s41598-017-01505-w

Zhang, Z., Wang, H., Jacobsson, T.J., & Luo, J. (2022). Big data driven perovskite solar cell stability analysis. *Nature Communications*, *13*(1), 7639. https://doi.org/10.1038/s41467-022-35400-4

Zhao, Y., Xu, S., Li, X. et al. (2022). Machine learning optimization of organic photovoltaics: Visualization of the process?performance landscape. *Advanced Intelligent Systems*, *4*(3), 2100261. https://doi.org/10.1002/aisy.202100261