Organic Photovoltaic Efficiency Predictor: Data-Driven Predictions of Power Conversion Efficiencies of Non-Fullerene Acceptor Organic Solar Cells

Brianna L. Greenstein and Geoffrey R. Hutchison*

Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, Pennsylvania 15260

E-mail: geoffh@pitt.edu

Abstract

In the fabrication of organic solar cells, there has been a need for materials with high power conversion efficiencies (PCE). Scharber's model is commonly used to predict efficiency, however it exhibits poor performance with new non-fullerene acceptor (NFA) devices (RMSE=2.53%). In this work, an empirical model is proposed that can be a more accurate alternative for NFA organic solar cells. Additionally, many screening studies use computationally expensive methods. A model based on using the semi-empirical simplified time-dependent density functional theory (sTD-DFT) as an alternative method can accelerate the calculations and yields similar accuracy. The models presented in this paper, referred to as Organic Photovoltaic Efficiency Predictor (OPEP) models, have shown significantly lower errors than previous models, with OPEP/B3LYP yielding errors of 1.53% and OPEP/sTD-DFT of 1.55%. The proposed computational models can be utilized for fast and accurate screening of new high-efficiency NFAs and donor polymer pairs.

Photovoltaic research is a growing topic of both scientific and technological advancement as a mechanism to inexpensively convert abundant solar energy.[1] One of the most active areas of research for organic solar cells is the design of more efficient materials for the active layer. In recent years, there have been improvements in the design of non-fullerene acceptors (NFAs), with power conversion efficiencies (PCEs) of single-junction devices surpassing 18%.[2–4] To accelerate the discovery of materials with higher PCEs, computational exploration of new materials is vital. Screening methods such as high-throughput virtual screening (HTVS) and genetic algorithms have shown promise, due to their capabilities of efficiently searching a large chemical space and finding promising candidates.[5–11] The largest HTVS study for organic photovoltaics (OPVs) is the Harvard Clean Energy Project that screened 1.3 million donors[12] and 51,000 NFA materials.[6] These results, however, and similar research rely on Scharber's model[13] to evaluate the PCE of the materials.

While Scharber's model is commonly used, and has proven to be a useful guide, it makes many assumptions that lead to poor performance with non-fullerene acceptors.[13] The model was designed for fullerenes, which are poor absorbers of light, so there is no term to account for acceptor light absorption, as in NFA materials. Additionally, it treats the donors as inorganic solids, with the assumption that all photons above the band gap are absorbed, which does not consider varying polymer/molecule absorption profiles. Lastly, it sets the fill factor (FF) to a constant 65%; this parameter can drastically vary depending on device fabrication and material compatibility (Figure S1).

A few other physically-motivated models have been developed to improve the PCE prediction. Imamura's model[14] is a modified version of Scharber's model where the LUMO offset between the donor and acceptor is also considered. Additionally, the FF is set to 70%. Alharbi's model[15] considers the acceptor's absorption spectrum and exciton diffusion length, as well as setting the FF as a function of the open-circuit voltage (V_{OC}). A more detailed explanation of these models can be found in Note S1 in the Supporting Information (SI).

To predict the PCE, an understanding of the photoconversion process is required, since

many factors are relevant to overall device performance. Photons are converted to free charge carriers through a series of five main processes: (1) absorption of the photon by either the donor or acceptor, exciting an electron which results in an exciton, (2) this exciton diffuses to the interface between the donor and acceptor, (3) if the donor absorbs the photon, then the excited electron transfers from donor to n-type acceptor molecules; similarly if an acceptor absorbs, holes transfer from the acceptor to p-type donor, (4) if there is enough of a driving force between these energy levels, the exciton will split into free charges, (5) the free charges diffuse through the active layer to either the cathode or anode.[16,17] Throughout these processes, there are multiple mechanisms where the exciton can recombine, resulting in energy loss.[18]

In this work, we consider a data-driven model, drawing on real-world published experimental performance, considering 47 descriptors that can play a role in photovoltaic device performance to find trends in the PCE. While many factors are likely relevant to device efficiency, the development of an empirical data-driven model can help to resolve which factors are most relevant to predicting real-world device PCE. Figure 1 gives an overview of some of the descriptors examined in this study and their method of calculation. Some descriptors analyzed include the highest occupied molecular orbital, (HOMO), HOMO-1, lowest occupied molecular orbital (LUMO), LUMO+1 eigenvalues, energy level offsets between the donor and acceptor, optical band gap, transition energies, oscillator strengths, molecular planarity,[19–21] and conjugation length.[22] A complete list of descriptors and their explanations can be found in Note S2 in the SI.

One of the main applications for predictive descriptor models is in screening studies, in which fast and accurate property predictions of molecules are desired. Typical evaluation of excitation energies and oscillator strengths of electronic excitations use time-dependent density functional theory (TD-DFT). An alternative method is simplified TD-DFT (sTD-DFT),[23] a semi-empirical method that can speed up calculations by 2-3 orders of magnitude and with comparable accuracy.[24] The calculated properties from DFT/TD-DFT and sTD-

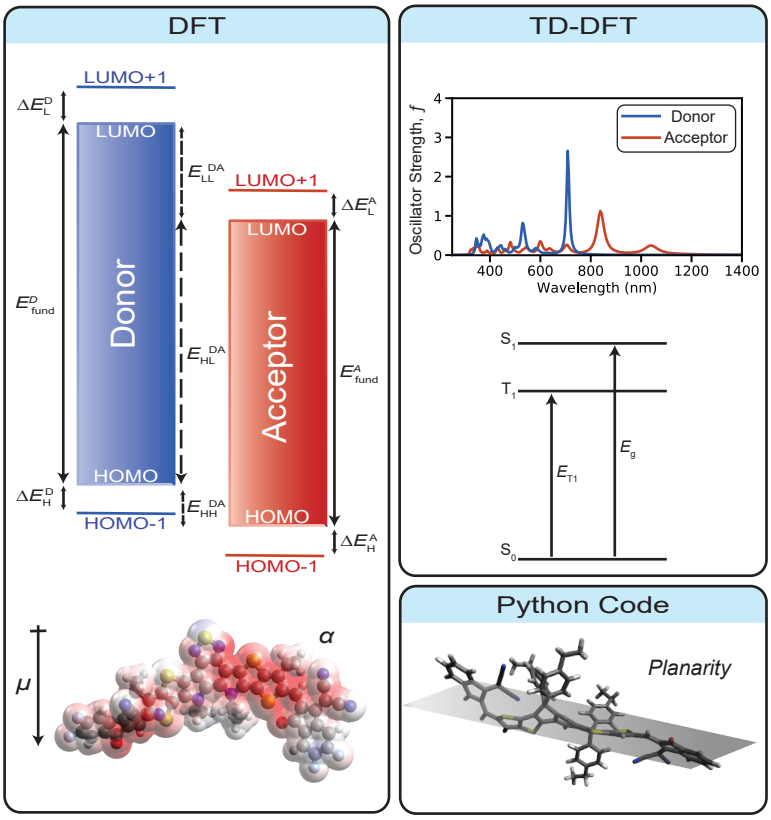

Figure 1: Diagrams labeling some descriptors used in this study and method of calculation. Using density functional theory (DFT), the difference between orbital eigenvalues is calculated for the acceptor and donor molecules, as well as the polarizability, α, and dipole moment, μ. TD-DFT calculated the absorption spectrum with oscillator strengths and energies of optical transitions, including to first excited singlet or triplet states. A python script was written to calculate the planarity of the molecule using the distance of each atom to a plane of best fit.

DFT are found to be in good agreement, as seen in correlation plots in Figure S2.

In this work, empirical descriptor models are proposed to predict FF, short-circuit current (J_{SC}), V_{OC}, and PCE as alternatives to current models for NFA-based OPVs. The first organic photovoltaic efficiency predictor (OPEP) model called OPEP/B3LYP is based on ground-state DFT and TD-DFT B3LYP calculations, while the OPEP/sTD-DFT models are based on sTD-DFT CAM-B3LYP calculations. OPEP/sTD-DFT was developed using the same dataset as OPEP/B3LYP, while the related OPEP/sTD-DFT Expanded was trained across a larger dataset containing molecules too large to compute with TD-DFT methods in a reasonable time frame. These proposed models are then compared to existing models such

as Scharber, Imamura, and a modified version of Alharbi.

A dataset was prepared using 84 donor-acceptor pairs found in recent literature, comprising of sixty-six NFAs, two fullerene acceptors, and six polymer donors (Table S1). Since the main goal of screening studies is to find materials with high PCE, the models were trained on experimental pairs with PCE greater than 9%, of which there are 59 pairs.

The authors note that experiments with FF below 40% were not included, since FF is highly dependent on the manufacturing process and treatments, and low FFs may be an indication that the processing conditions were not fully optimized for the donor/acceptor materials and thus the measured PCE would be lower than in an optimized device.

Since many descriptors were calculated for each material, the least absolute shrinkage and selection operator (LASSO) was utilized to eliminate descriptors and perform regression analysis. The resulting models were 5-fold cross-validated.

Using LASSO and manual selection of descriptors to avoid overfitting, models were prepared to predict FF, J_{SC}, V_{OC}, and PCE with B3LYP and with sTD-DFT CAM-B3LYP. Two versions were prepared for each model, one with a wide range of experimental PCE and another with experimental PCE above 9% (Eqs S13-S36). The equation to predict PCE using B3LYP (OPEP/B3LYP) for high experimental PCE is:

$$PCE = -136.5 + 6.42E_{T_1} + 0.917\Delta\mu_{ge} + 8.31\omega_D^- + 0.00389E^D_{oscs-10} - 0.00047E^A_{oscs} - 0.0036\alpha^A \quad (1)$$

where E_{T_1} is the electronic transition energy from the singlet ground state to the first triplet excited state of the acceptor, in units of eV, $\Delta\mu_{ge}$ is the change in dipole moment from going from the ground state to the first excited state of the acceptor, ω_D^- is the electro-donating power of the donor, $E^D_{oscs-10}$ is the energy of transition with the largest oscillator strength within the first 10 transitions from the ground state to the excited state of the donor, in units of cm^{-1}, E^A_{oscs} is the lowest energy of electronic transition of the acceptor with an oscillator

strength greater than 0.1, in units of cm^{-1}, and α^A is the polarizability of the acceptor in units of a.u.

The distribution of errors of the predicted values for FF, V_{OC}, J_{SC}, and PCE for the proposed OPEP / B3LYP and OPEP/sTD-DFT models, as well as Scharber, Imamura, and Alharbi's models are shown in Figure 2. Examination of the predicted FF in Figure 2a shows large deviations and underestimation of FF for Scharber's and Imamura's models, resulting from the selection of constant FFs of 65% and 70%, respectively. While these values are close to the median FF of 67% from the sampling of 107 OPV devices (Figure S1), they can not adequately describe the complex FF parameter. As data-driven regressions, our models have small deviations close to zero for FF while Alharbi's model consistently overestimates the FF, with a median deviation of approximately 12%. The OPEP models for both B3LYP and sTD-DFT have large distributions close to zero deviation, however, the height of the distribution suggests that there are frequent outliers.

As illustrated in Figure 2b, Scharber and Imamura have the same distributions, since they use the same equation to predict V_{OC}. Alharbi's model again significantly overestimates this parameter, with a high positive deviation. The proposed OPEP models have many deviations close to zero, therefore outperforming the previous models.

In Figure 2c, Scharber and Imamura's predictions have large negative deviations, with some more than 20 mA/cm^2, and a wide distribution of these deviations, which is understandable since these models consider only one or two parameters for Scharber's and Imamura's models, respectively. Alharbi's full model for J_{SC} was not implemented in this work, since it treats optical effects in the films.

Examination of the PCE in Figure 2d demonstrates that Scharber, Imamura, and a modified version of Alharbi's model (Scharber's J_{SC} was substituted in the calculations) mostly underestimate the PCE. The OPEP models perform significantly better with fewer outlier points. The OPEP models trained on all experimental ranges of PCE had slightly larger errors, however performed better than the other models (Figure S3)

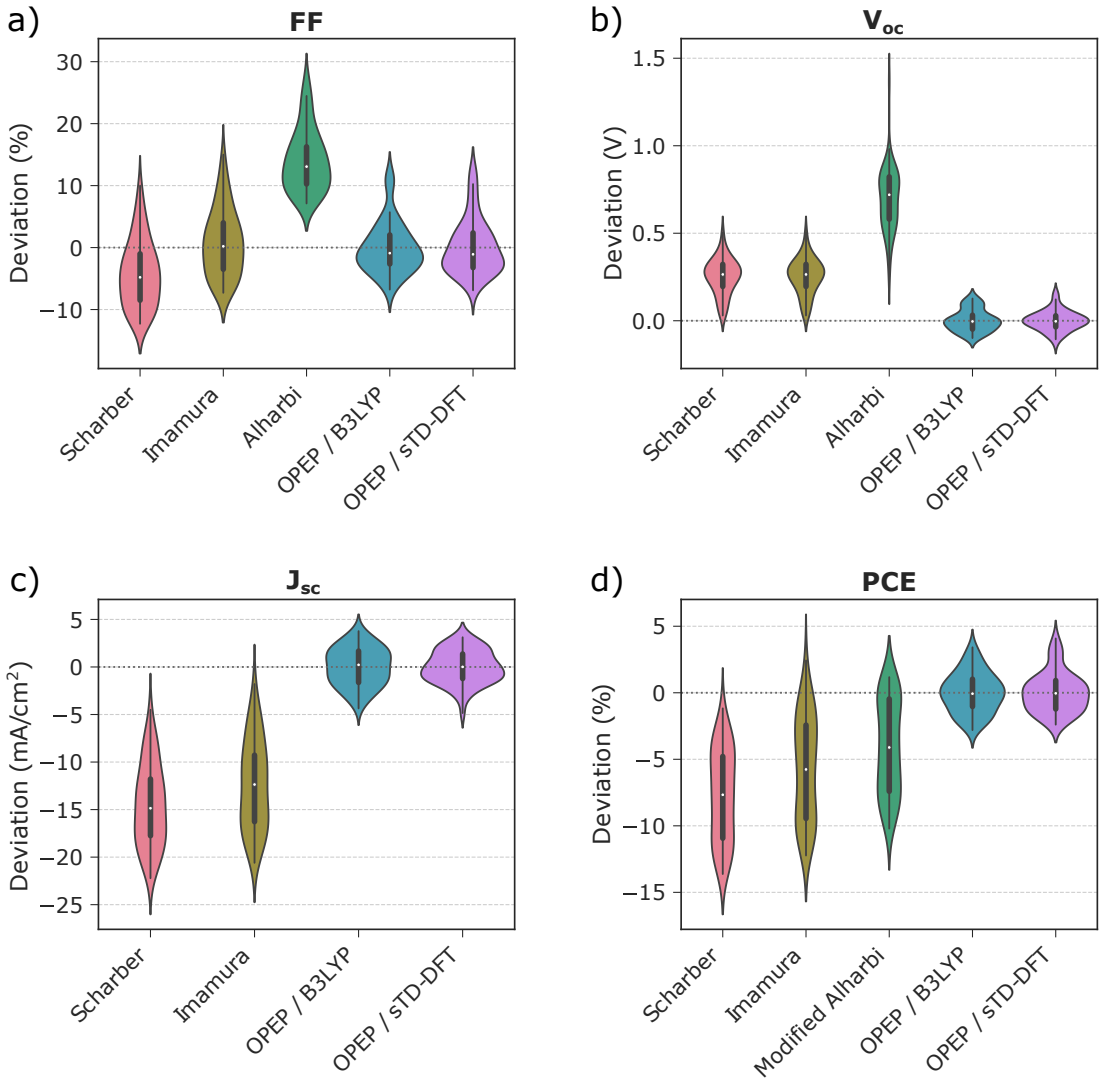

Figure 2: Violin plots showing the distribution of deviation (predicted-experimental values) for (a) FF, (b) V_{OC}, (c) J_{SC}, and (d) PCE. Experimental values were used for donor-acceptor pairs with PCE greater than 9%. In 2d, the modified Albarbi model refers to using Scharber's J_{SC} for the calculation of PCE.

Comparisons of the predicted PCE from each model against the experimentally reported PCE can be seen in Figure 3. R^2 is the correlation coefficient and compares the predicted data points to a horizontal line. R^2 is typically in the range 0 to 1, but can be below zero when the fit of the data points is worse than the average. In Figure 3, the 5-fold cross validated R^2 is *negative* for all of the previous models, suggesting that these models are inferior to just setting

the average PCE of the experimental values as the predictions. Both OPEP models show significant improvements in the correlation of predicted PCE versus the experimental PCE. The OPEP/B3LYP model has R^2 of 0.640 ± 0.006, the highest of all the models examined, while OPEP/sTD-DFT has R^2 of 0.630 ± 0.005, essentially indistinguishable. The OPEP models trained on high experimental PCE were used to predict the PCE for low performing devices (Figure S4) and performed worse with R^2 of 0.36 ± 0.08, 0.30 ± 0.06, and 0.35 ± 0.08 for OPEP/B3LYP, OPEP/sTD-DFT, and OPEP/sTD-DFT Expanded, respectively. While higher correlations would be desirable, the performance of the OPEP models are notable since they do not directly include predictors of film morphology, crystal packing, or charge transport, among other properties. Future work to include some or all will undoubtedly improve the accuracy of the OPEP models.

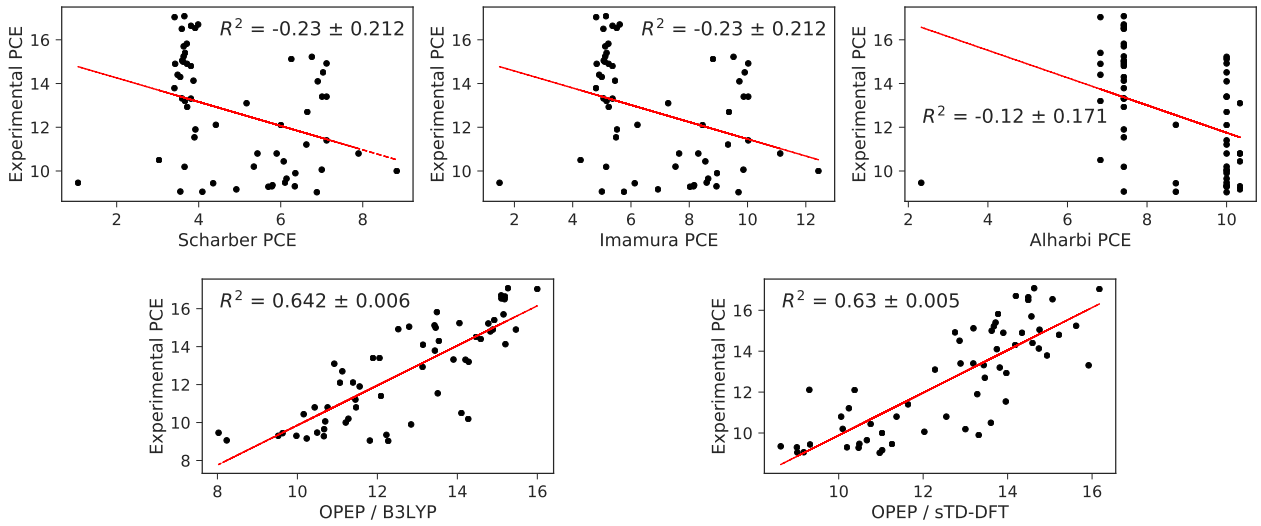

Figure 3: Correlation plots comparing predicted PCE of models versus the experimental PCE for pairs with experimental PCE greater than 9%.

A statistical analysis of all the models to predict FF, J_{SC}, V_{OC}, and PCE was performed. The R^2, mean absolute error (MAE), root mean square error (RMSE), and mean absolute percent error (MAPE) were calculated and 5-fold cross-validated for each model. MAE, RMSE, and MAPE are measures of the average magnitude of the error, while RMSE is commonly used to give larger weight to the larger errors. Figure 4 shows the comparison

of MAE and RMSE for each model with every parameter. For all parameters, the OPEP models have much lower MAE and RMSE, with RMSE reaching around 1.5% for predicting PCE.

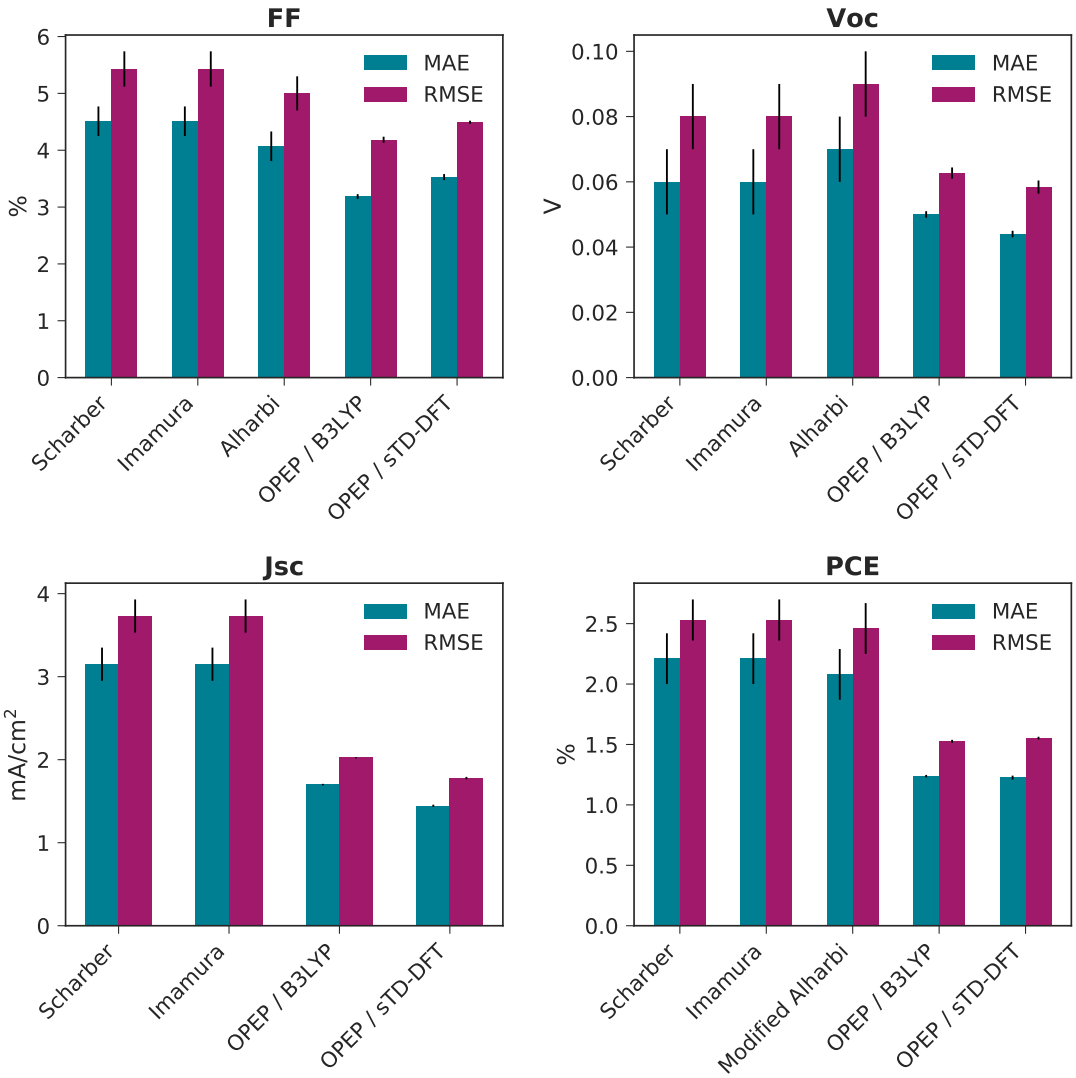

Figure 4: Statistical analysis of models for predicting OPV parameters with experimental PCE above 9%. The error bars are from the standard error of 5-fold cross-validation.

The full statistical analysis is displayed in Table 1. For FF, our models have significantly lower RMSE than the previous models, indicating OPEP can predict the FF more appropriately than previous models. The lowest RMSE achieved was $4.19 \pm 0.05\%$ for the OPEP/B3LYP model. However, this FF RMSE is still relatively high. This is most likely

due to the lack of efficient descriptors in the model, such as charge transport properties, dielectric constants, and solubility, which can be challenging to predict.

The OPEP J_{SC} models have RMSE almost 2 mA/cm^2 less than the other published models, with the best RMSE of 1.78 ± 0.02 mA/cm^2 for OPEP/sTD-DFT. Since J_{SC} is known to be primarily controlled by charge generation and charge transport processes, incorporating descriptors such as the sum of the oscillator strengths and an absorption figure of merit that gives higher weight to transitions in the regions with higher solar radiation intensity increases the accuracy of these proposed new models over treatments such as Scharber.

For the V_{OC} models, Scharber and Imamura have RMSE of 0.08 ± 0.01 V, and Alharbi has RMSE of 0.09 ± 0.01 V. The OPEP models have lowered this error to 0.058 ± 0.002 V and 0.063 ± 0.002 V for OPEP/sTD-DFT and OPEP/B3LYP, respectively. These models include descriptors such as electrodonating power and polarizability. What is interesting to note is that none of the OPEP models include the energy between the donor HOMO and acceptor LUMO eigenvalues, the only descriptor in Scharber's model. Development of the OPEP models shows that while this descriptor can contribute, other descriptors can be more predictive of the V_{OC} (i.e., the LASSO technique includes them preferentially).

Lastly, for comparing the PCE, both OPEP models have RMSE around 1.5%, and RMSE as low as 1.53 ± 0.01% for OPEP/B3LYP. Results of a T-test indicate that the performance of the OPEP models are indistinguishable (p=0.98). For every statistic in Table 1, the standard error of the mean (SEM) for the OPEP models is much lower than that of other models. This is further evidence that these proposed models can serve as replacements for Scharber's model for NFA screening studies.

While the proposed models are a useful alternative to Scharber's model for NFA OPVs, they can be improved as more experimental data becomes available. Low-performing device efficiencies are rarely reported, yet are required to develop more accurate models and establish design rules. Moreover, many NFA experimental studies use the same donor materials. A larger dataset containing 77 acceptor-donor pairs with PCE over 9% and FF below 40%

Table 1: Statistical analysis of all models by comparing R^2, mean absolute error (MAE), root mean squared error (RMSE). All models are cross-validated with experimental PCE greater than 9%. The format is the predicted value ± standard error of the mean (SEM). Units for MAE and RMSE for each parameter are: FF is %, J_{SC} is mA/cm^2, V_{OC} is V, and PCE is %.

	Model	R^2	MAE	RMSE
FF	OPEP/B3LYP	0.399 ± 0.016	3.186 ± 0.041	4.185 ± 0.053
	OPEP/sTD-DFT	0.307 ± 0.008	3.528 ± 0.053	4.492 ± 0.027
	Alharbi	0.050 ± 0.056	4.070 ± 0.260	5.000 ± 0.300
	Imamura	-0.120 ± 0.063	4.510 ± 0.260	5.430 ± 0.310
	Scharber	-0.120 ± 0.063	4.510 ± 0.260	5.430 ± 0.310
J_{SC}	OPEP/sTD-DFT	0.739 ± 0.004	1.444 ± 0.014	1.778 ± 0.015
	OPEP/B3LYP	0.662 ± 0.003	1.700 ± 0.009	2.024 ± 0.008
	Imamura	-0.920 ± 0.618	3.150 ± 0.200	3.730 ± 0.200
	Scharber	-0.920 ± 0.618	3.150 ± 0.200	3.730 ± 0.200
V_{OC}	OPEP/sTD-DFT	0.610 ± 0.027	0.044 ± 0.001	0.058 ± 0.002
	OPEP/B3LYP	0.552 ± 0.025	0.050 ± 0.001	0.063 ± 0.002
	Imamura	-0.010 ± 0.224	0.060 ± 0.010	0.080 ± 0.010
	Scharber	-0.010 ± 0.224	0.060 ± 0.010	0.080 ± 0.010
	Alharbi	-0.230 ± 0.089	0.070 ± 0.010	0.090 ± 0.010
PCE	OPEP/B3LYP	0.642 ± 0.006	1.240 ± 0.010	1.526 ± 0.012
	OPEP/sTD-DFT	0.630 ± 0.005	1.226 ± 0.016	1.553 ± 0.011
	Alharbi	-0.120 ± 0.171	2.080 ± 0.210	2.460 ± 0.210
	Scharber	-0.230 ± 0.212	2.210 ± 0.210	2.530 ± 0.170
	Imamura	-0.230 ± 0.212	2.210 ± 0.210	2.530 ± 0.170

was created. This dataset includes 20 donor oligomers, all as hexamers, as opposed to the six donors in the original dataset. This larger dataset was not calculated with DFT and TD-DFT methods due to slow convergence of the larger oligomers. With this expanded dataset, sTD-DFT was used to see if the model is improved by adding more data. The resulting model, known as OPEP/sTD-DFT Expanded, is:

$$PCE = -33.08 + 4.26\Delta E_H^A + 1.377\Sigma f_A - 0.459\Sigma f_D + 0.174 Abs_{FOM}^D + 2.45\omega_D^- + 9.5 \times 10^{-4} E_g^D \quad (2)$$

where ΔE_H^A is the energy difference between the HOMO and HOMO-1 of the acceptor, Σf_A and Σf_D are the sums of the oscillator strengths of the acceptor and donor, respectively, Abs_{FOM}^D is a new descriptor proposed that takes into account the oscillator strengths of the donor and the solar spectrum, ω_D^- is the electro-donating power of the acceptor, in units of eV, and E_g^D is the optical band gap of the donor in units of cm^{-1}.

The performance of OPEP/sTD-DFT Expanded can be seen in Table 2. Comparisons between OPEP/sTD-DFT Expanded and the original OPEP/sTD-DFT show the performance is statistically indistinguishable (p=0.88) and either model can be used for the prediction of PCE with similar accuracy.

Table 2: Statistical analysis of OPEP/sTD-DFT Expanded. The model was cross-validated and based on experimental PCE greater than 9%. The format is predicted value ± standard error of the mean (SEM). Units for MAE and RMSE for each parameter are: FF is %, J_{SC} is mA/cm^2, V_{OC} is V, and PCE is %.

	R^2	MAE	RMSE
FF	0.310 ± 0.003	3.531 ± 0.020	4.446 ± 0.010
J_{SC}	0.682 ± 0.011	1.557 ± 0.021	1.924 ± 0.032
V_{OC}	0.519 ± 0.014	0.047 ± 0.001	0.061 ± 0.001
PCE	0.552 ± 0.021	1.297 ± 0.025	1.601 ± 0.036

Further improvement of PCE prediction models can come from the inclusion of relevant factors to FF and J_{SC} such as solubilities and charge transport properties such as Marcus

reorganization energies. Solubility is crucial to the domain size, which in turn is directly related to the exciton recombination rate and overall PCE of the device. If the solubility between conjugated polymers and NFAs can be predicted, better FF models can be established. Improvements in these predictions can be achieved if more experimental solubility data between small molecules and conjugated polymers is available, as well as easily accessible computational methods to calculate interaction parameters. According to Marcus theory, reorganization energies are related to the driving force for exciton dissociation and can influence the electron and hole mobility throughout the active layer,[22,25] yet calculating this parameter requires lengthy calculations. A machine learning method recently developed by Abarbanel and Hutchison shows the potential for efficient screening of reorganization energies of polythiophenes[26] and is promising for future application towards NFAs.

In summary, new data-driven models to predict the FF, J_{SC}, V_{OC}, and PCE for OPV devices incorporating NFAs are proposed. The use of experimental data can help to reflect the critical balance of factors relevant in determining device efficiency. A model based on DFT and TD-DFT with B3LYP (OPEP/B3LYP) and a model with a larger dataset using sTD-DFT CAM-B3LYP (OPEP/sTD-DFT Expanded) are presented, both with higher accuracy than previous models. OPEP/B3LYP can predict the PCE with an RMSE of $1.52 \pm 0.01\%$, while the sTD-DFT model with the expanded dataset can predict the PCE with RMSE of $1.60 \pm 0.04\%$. Using sTD-DFT is a useful alternative to standard DFT and TD-DFT calculations and shows similar accuracy. The OPEP models provide a simple route to pre-screen NFAs before synthesis or to evaluate candidates in computational material searches, and we intend to update them with reliable community data.

Acknowledgement

We acknowledge support from Department of Energy-Basic Energy Sciences Computational and Theoretical Chemistry (Award DESC0019335) and the University of Pittsburgh Center

for Research Computing through the computational resources provided.

Supporting Information Available

Details on Scharber's, Imamura's, and Alhabri's models, description of descriptors analyzed, experimental dataset, OPEP models for B3LYP, sTD-DFT, and sTD-DFT Expanded trained on experimental PCE above 9% and on all ranges of PCE to predict FF, J_{SC}, Voc, and PCE, figures: histogram of experimental FF, correlations between B3LYP and CAM-B3LYP, violin plots for all models examined across all ranges of PCE, performance of OPEP models trained on high PCE on all ranges of experimental PCE.

The data that supports the findings of this study are openly available at: `https://github.com/hutchisonlab/OPEP`

References

(1) U.S. Energy Information Administration - EIA - Independent Statistics and Analysis. `https://www.eia.gov/outlooks/aeo/`.

(2) Ma, R.; Liu, T.; Luo, Z.; Guo, Q.; Xiao, Y.; Chen, Y.; Li, X.; Luo, S.; Lu, X.; Zhang, M.; Li, Y.; Yan, H. Improving open-circuit voltage by a chlorinated polymer donor endows binary organic solar cells efficiencies over 17%. *Science China Chemistry* **2020**, *63*, 325–330.

(3) Lin, F.; Jiang, K.; Kaminsky, W.; Zhu, Z.; Jen, A. K.-Y. A Non-fullerene Acceptor with Enhanced Intermolecular π-Core Interaction for High-Performance Organic Solar Cells. *Journal of the American Chemical Society* **2020**, *142*, 15246–15251.

(4) Liu, Q.; Jiang, Y.; Jin, K.; Qin, J.; Xu, J.; Li, W.; Xiong, J.; Liu, J.; Xiao, Z.; Sun, K.; Yang, S.; Zhang, X.; Ding, L. 18% Efficiency organic solar cells. *Science Bulletin* **2020**, *65*, 272–275.

(5) Kanal, I. Y.; Owens, S. G.; Bechtel, J. S.; Hutchison, G. R. Efficient Computational Screening of Organic Polymer Photovoltaics. *The Journal of Physical Chemistry Letters* **2013**, *4*, 1613–1623.

(6) Lopez, S. A.; Sanchez-Lengeling, B.; de Goes Soares, J.; Aspuru-Guzik, A. Design Principles and Top Non-Fullerene Acceptor Candidates for Organic Photovoltaics. *Joule* **2017**, *1*, 857–870.

(7) Meftahi, N.; Klymenko, M.; Christofferson, A. J.; Bach, U.; Winkler, D. A.; Russo, S. P. Machine learning property prediction for organic photovoltaic devices. *npj Computational Materials* **2020**, *6*, 1–8.

(8) Sun, W.; Zheng, Y.; Yang, K.; Zhang, Q.; Shah, A. A.; Wu, Z.; Sun, Y.; Feng, L.; Chen, D.; Xiao, Z.; Lu, S.; Li, Y.; Sun, K. Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Science Advances* **2019**, *5*, eaay4275.

(9) Sahu, H.; Yang, F.; Ye, X.; Ma, J.; Fang, W.; Ma, H. Designing promising molecules for organic solar cells *via* machine learning assisted virtual screening. *Journal of Materials Chemistry A* **2019**, *7*, 17480–17488.

(10) Wilbraham, L.; Berardo, E.; Turcani, L.; Jelfs, K. E.; Zwijnenburg, M. A. High-Throughput Screening Approach for the Optoelectronic Properties of Conjugated Polymers. *Journal of Chemical Information and Modeling* **2018**, *58*, 2450–2459.

(11) Nagasawa, S.; Al-Naamani, E.; Saeki, A. Computer-Aided Screening of Conjugated Polymers for Organic Solar Cell: Classification by Random Forest. *The Journal of Physical Chemistry Letters* **2018**, *9*, 2639–2646.

(12) Hachmann, J.; Olivares-Amaya, R.; Jinich, A.; Appleton, A. L.; Blood-Forsythe, M. A.; Seress, L. R.; Román-Salgado, C.; Trepte, K.; Atahan-Evrenk, S.; Er, S.; Shrestha, S.;

Mondal, R.; Sokolov, A.; Bao, Z.; Aspuru-Guzik, A. Lead candidates for high-performance organic photovoltaics from high-throughput quantum chemistry–the Harvard Clean Energy Project. *Energy & Environmental Science* **2014**, *7*, 698–704.

(13) Scharber, M.; Mühlbacher, D.; Koppe, M.; Denk, P.; Waldauf, C.; Heeger, A.; Brabec, C. Design Rules for Donors in Bulk-Heterojunction Solar Cells-Towards 10 % Energy-Conversion Efficiency. *Advanced Materials* **2006**, *18*, 789–794.

(14) Imamura, Y.; Suganuma, M.; Hada, M. Computational Study on the Search for Non-Fullerene Acceptors, Examination of Interface Geometry, and Investigation of Electron Transfer. *The Journal of Physical Chemistry C* **2019**, *123*, 17678–17685.

(15) Alharbi, F. H.; Rashkeev, S. N.; El-Mellouhi, F.; Lüthi, H. P.; Tabet, N.; Kais, S. An efficient descriptor model for designing materials for solar cells. *npj Computational Materials* **2015**, *1*, 15003.

(16) Clarke, T. M. How to split an exciton. *Nature Energy* **2020**, *5*, 644–645.

(17) Brabec, C. J.; Heeney, M.; McCulloch, I.; Nelson, J. Influence of blend microstructure on bulk heterojunction organic photovoltaic performance. *Chem. Soc. Rev.* **2011**, *40*, 1185–1199.

(18) Ji, Y.; Xu, L.; Hao, X.; Gao, K. Energy Loss in Organic Solar Cells: Mechanisms, Strategies, and Prospects. *Solar RRL* **2020**, *4*, 2000130.

(19) Kuzmich, A.; Padula, D.; Ma, H.; Troisi, A. Trends in the electronic and geometric structure of non-fullerene based acceptors for organic solar cells. *Energy & Environmental Science* **2017**, *10*, 395–401.

(20) Rech, J. J.; Bauer, N.; Dirkes, D.; Kaplan, J.; Peng, Z.; Zhang, H.; Ye, L.; Liu, S.; Gao, F.; Ade, H.; You, W. The crucial role of end group planarity for fused-ring electron acceptors in organic solar cells. *Materials Chemistry Frontiers* **2019**, *3*, 1642–1652.

(21) Ahmed, S.; Kalita, D. J. End-capped group manipulation of non-fullerene acceptors for efficient organic photovoltaic solar cells: a DFT study. *Physical Chemistry Chemical Physics* **2020**, *22*, 23586–23596.

(22) Sahu, H.; Ma, H. Unraveling Correlations between Molecular Properties and Device Parameters of Organic Solar Cells Using Machine Learning. *The Journal of Physical Chemistry Letters* **2019**, *10*, 7277–7284.

(23) de Wergifosse, M.; Grimme, S. Nonlinear-response properties in a simplified time-dependent density functional theory (sTD-DFT) framework: Evaluation of the first hyperpolarizability. *The Journal of Chemical Physics* **2018**, *149*, 024108.

(24) Martynov, A. G.; Mack, J.; May, A. K.; Nyokong, T.; Gorbunova, Y. G.; Tsivadze, A. Y. Methodological Survey of Simplified TD-DFT Methods for Fast and Accurate Interpretation of UV–Vis–NIR Spectra of Phthalocyanines. *ACS Omega* **2019**, *4*, 7265–7284.

(25) Duan, R.; Han, G.; Qu, L.; Yi, Y. Importance of molecular rigidity on reducing the energy losses in organic solar cells: Implication from geometric relaxations of A-D-A electron acceptors. *Materials Chemistry Frontiers* **2021**, *5*, 3903–3910.

(26) Abarbanel, O.; Hutchison, G. Machine Learning to Accelerate Screening for Marcus Reorganization Energies. *ChemRxiv* **2021**,

Graphical TOC Entry