Datamining and analysis of the key parameters in organic solar cells

Moritz K. Riede[a,b], Andreas W. Liehr[b], Markus Glatthaar[a,b], Michael Niggemann[a,b], Birger Zimmermann[b], Tobias Ziegler[a], Andreas Gombert[a], Gerhard Willeke[a]

[a]Fraunhofer Institute for Solar Energy Systems (ISE), Heidenhofstr. 2, 79110 Freiburg, Germany;
[b]Material Research Center Freiburg (FMF), Stefan-Meier-Str. 21, 79104 Freiburg, Germany

ABSTRACT

The production process of organic solar cells (OSCs) is investigated and the effects of parameter variations on experimental results are analysed with the Principal Component Analysis (PCA). This statistical method is applied to an exemplar data set, in which the materials' concentration in the absorber solution and the spincoating speed of the absorber solution were varied intentionally. In addition to the remaining production parameters, the time intervals between the steps were included in the analysis. A large part of the variance in the experimental results can be explained with the evaporation conditions, the spincoating speed and the concentrations in the absorber solution. The PCA also confirms that the OSC is a complex and interdependent system, where one has to analyse the influence of several parameters at the same time in order to understand their effects on the OSC properties. The PCA results will be used to focus further experiments on the identified key parameters.

Keywords: Organic Solar Cells, Datamining, Principal Component Analysis, Electronic Laboratory Notebook

1. INTRODUCTION

In the recent years good progress has been made in the field of organic solar cells (OSC). The power conversion efficiencies rose from 2.5% to about 5% within the past four years[1] and many discoveries aiding the qualitative and quantitative understanding of OSC have been made. However, there are still many open questions regarding the fundamental physical understanding of the processes in an OSC and the relation between the production parameters and their effects on the OSC characteristics. This is caused both by the special properties of the charge carrier transport in organic semiconductors and by difficulties in experimental work towards solving these open questions. The investigated bulk hetero-junction OSCs show a complex behaviour, which is very sensible to parameter variations during production. Generally, an intentional variation of one single parameter during production is difficult. Further simultaneous variations of other parameters often affect the experimental results in the same order of magnitude as the intended variation. Although there are few steps necessary for manufacturing an OSC, the production steps have many degrees of freedom which often exhibit a complex interdependence. This complicates both the optimisation of the production steps and the development of a statistically firm understanding of the processes in an OSC.

An adequate sample of documented OSC preparation and validation records is the prerequisite to permit a sound statistical and physical analysis. Only an automation of the main OSC characterisation experiments and a continuous recording of all production parameters and experimental results for a large number of OSC can provide the data necessary for such an analysis. Such representative data sets are currently generated and all important parameters are stored in an electronic laboratory notebook. These data form the basis for a systematic statistical analysis. However, due to the above described complexity of the OSC, it is essential to investigate the effect of several parameter variations on the experimental results at once. The goal is to identify the dominating correlations between production parameters and OSC characteristics.

Further author information: (Send correspondence to M.K.R.)
M.K.R.: E-mail: moritz.riede@ise.fraunhofer.de, Telephone: 49 761 203 4795

Photonics for Solar Energy Systems, edited by Andreas Gombert, Proc. of SPIE Vol. 6197,
61970H, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.662974

The selection of suitable statistical methods is critical for being able to make physically meaningful deductions. There are several methods for analysing the dependencies in such data sets of which the Principal Component Analysis (PCA) was chosen.

2. PRINCIPAL COMPONENT ANALYSIS

All OSC manufacturing steps can be characterised by a set of parameters whose values are recorded for every OSC. Similarly, the measurement results' values, e.g. for the open circuit voltage V_{oc}, are collected. Both production parameters and measurement results can be treated as random variables. The term "random variable" is not a variable in the classical sense, but a function. The realisation of such a random variable describes here the mapping of an OSC production parameter or a measurement result to a scalar value for each solar cell.

One suitable method to analyse the effects of not one single, but of several random variables and their dependencies is the Principal Component Analysis[2, 3] (PCA). The PCA is a multivariate statistical method for numerically examining the correlation in a set of data. The goal is to identify latent, i.e. hidden factors, in the data which are determining the data's properties. It uses the fact that the variance in data is a measure of its information content.

The random variables in this case are given by the OSCs' recorded production parameters and experimental results. Assuming that there are p of those random variables, plotting their values results in a scatterplot in a p-dimensional Cartesian coordinate system spanned by the p axes of the random variables. The PCA's approach is to find a new set of axes for the data such that the first new axis is lying in the scatterplot in a way that its variance, as measure of information content, is maximised along this axis. The second axis is chosen to capture the second highest variance while being orthogonal to the first axis and so on. All new axes are a linear combination of the axes given by the original random variables.

Mathematically, this can be achieved by choosing a new coordinate system in which the correlation of the random variables is minimized. The transformation to the new coordinate axes is given by the eigenvectors of the covariance matrix. Let the data be described by a p-dimensional random vector $\mathbf{X}$ composed of the scalar-valued random variables X_i $(i = 1, \ldots, p)$, e.g. concentration in the blend solution, spincoating speed, etc.. After subtracting the mean from each X_i, the $p \times p$-dimensional covariance matrix is given by $\mathbf{\Sigma}$ with elements

$$\Sigma_{i,j} =< X_i X_j >= \mathrm{Cov}(X_i, X_j). \tag{1}$$

Often the X_is are normalised quantities. $\mathbf{\Sigma}$ is always symmetric and positive definite. Hence it is possible to diagonalise it through an orthogonal transformation such that

$$\mathbf{\Lambda} = \mathbf{\Gamma}^T \mathbf{\Sigma} \mathbf{\Gamma}, \tag{2}$$

with $\mathbf{\Lambda}$ being diagonal and having the eigenvalues λ_α $(\alpha = 1, \ldots, p)$ of $\mathbf{\Sigma}$ chosen in decreasing order as diagonal entries. Since $\mathbf{\Sigma}$ is positive definite, its eigenvalues are non negative. The matrix $\mathbf{\Gamma}$ for the transformation has the corresponding normalized eigenvectors $\mathbf{e}_\alpha$ as columns. The new random vector with the coordinate axes is defined as

$$\mathbf{Y} = \mathbf{\Gamma}^T \mathbf{X}. \tag{3}$$

The Y_α are called principal components (PCs) of $\mathbf{X}$ and are uncorrelated random variables. Their variance $\mathrm{Var}(Y_\alpha)$ is given by λ_α. Y_α's elements $Y_{i,\alpha}$ can be considered the X_i's contribution to λ_α, the so-called loading. Given that the variance measures the information content and the principal components Y_α are sorted by descending λ_α often the first few principal components can adequately describe the variations in the data, because they contain most of the data's variance.

However, the PCA does not differentiate between production parameters and experimental results. It treats all equal and shows only their correlation. Still it is possible from its results to identify parameters in the production process of organic solar cells (OSCs), which have a significant influence on the device properties and performance. Once those parameters are found, their physical origin on their influence on the OSC performance has to be further investigated.

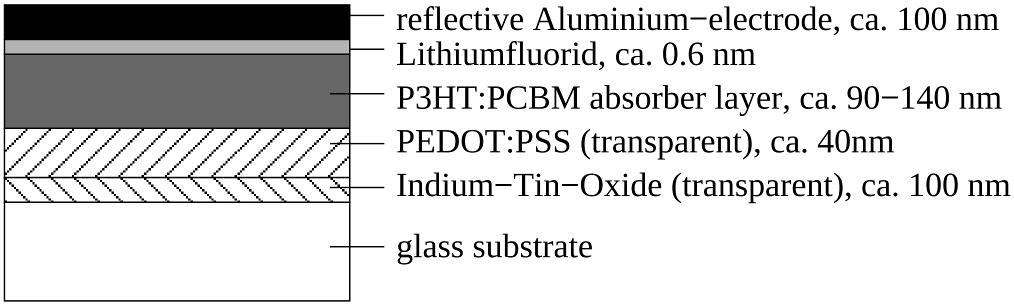

Figure 1. The layer structure glass|ITO|PEDOT:PSS|P3HT:PCBM|LiF|Aluminium of the OSCs investigated. The ITO, which is coated with a thin PEDOT:PSS layer, acts as anode for the P3HT:PCBM absorber layer. The cathode is realised through an evaporated Lithiumfluorid|Aluminium layer. Both PEDOT:PSS and absorber layer are spincoated from solution.

3. EXPERIMENTS

To quantify the influence of parameters in the OSC production process on the device properties and performance, the PCA is as a first step applied to a small exemplar data set consisting of manufacturing parameters and measurement results of 62 organic solar cells on seven different substrates. Solar cells on those substrates, which had obvious defects, had been removed from the analysis before. The OSCs were built in the standard structure, as shown in figure 1*. Indium-Tin-Oxide (ITO)-coated glass was used as substrate. Poly-3,4-ethylenedioxythiophene doped with polystyrene sulfonic acid (PEDOT:PSS, formulation AI4083) was bought from HC-Stark and the absorber solution was made by blending poly(3-hexylthiophene) (P3HT, Fraunhofer Institute for Applied Polymer Research IAP) and 1-[3-(methoxycarbonyl)propyl]-1-phenyl-[6.6]C_{61} (PCBM, Solene B.V.) in chlorobenzene with relative ratio kept at 1:1 by weight. The following parameters were intentionally varied during production:

- Absolute P3HT and PCBM concentration in the absorber solution
- Spincoating speeds of the absorber solution

The PEDOT:PSS was spin-coated ($2000rpm$) under ambient conditions onto a cleaned ITO-substrate, which had been exposed to an UV-ozone treatment for $600s$ before. The PEDOT:PSS layer was cured at 200C for $600s$ directly after spincoating and the substrates were transferred into a glovebox with an inert gas atmosphere, in which all following steps are carried out. The photoactive absorber layer was spincoated from solutions with different absolute concentrations of 1:1 P3HT:PCBM. Lithiumfluorid (LiF) and Aluminium were thermally evaporated with layer thicknesses of $0.6nm$ and $\sim 100nm$ respectively. The time intervals between production steps and the conditions during evaporation were not intentionally set to specific values and hence differed between substrates. All production parameters were recorded electronically for each substrate in an electronic laboratory notebook, being realised as MySQL database and HDF5 file. The V_{oc}, the short circuit current density J_{sc} and the fill factor FF were taken from automated current-voltage measurements at simulated AM1.5g ($100mW/cm^2$) corrected for the spectral miss match between calibrated reference diode and OSC.

4. RESULTS AND DISCUSSION

The following five production parameters were chosen for the PCA and tested for their significance in the OSC production: the absolute concentration of 1:1 P3HT:PCBM in the absorber solution (Concentration); the absorber solution's spincoating speed (1st Abs-SC-Speed); the time between spincoating of the absorber solution and begin of evaporation (Δt Abs-SC to Evap); the start pressure of evaporation (Evaporation p_{start}) and the evaporation rate of aluminium. Table 1 contains the principal components (PCs) of the production parameters and V_{oc} as columns with decreasing component variance. The PCs' eigenvalues and the variance they contain is

---

[*] "|" denotes an interface, ":" is used for mixtures

shown in table 2. The same analysis is repeated for J_{sc} and FF, but only the first three PCs are shown in table 3, because they capture almost all of the variance in the data. The power conversion efficiency η is not included in the tables because it is at the constant illumination intensity the same linear combination of V_{oc}, J_{sc} and FF for all OSCs.

Table 1. The principal components (PCs) of the PCA as columns with decreasing component variance (see table 2). The PCs are linear combinations of the initial random variables. Their directions and weights are given by the loading factors. The three most influential production parameters of the first three PCs are written in bold.

	$\mathbf{V_{oc}}$					
	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
Concentration	-0.20	**-0.63**	-0.34	-0.67	0.02	0.00
1st Abs-SC-Speed	-0.16	0.34	**-0.91**	0.18	-0.01	0.00
Δt Abs-SC to Evap	**-0.54**	0.12	0.14	-0.04	-0.82	-0.04
Evaporation p_{start}	**0.54**	-0.14	-0.14	0.03	-0.44	0.69
Evaporation Rate Aluminium	**0.54**	-0.14	-0.14	0.03	-0.37	-0.73
V_{oc}	-0.22	-0.66	-0.06	0.72	0.00	0.00

Table 2. The PCs eigenvalues and the amount of variance they capture (both as single and cumulated value) in the transformed coordinate system. Already the first three PCs explain more than 98% of the variance in the data.

	Eigenvalue	% of Variance	Cumulated %
PC 1	3.23	53.8	53.8
PC 2	1.86	31.1	84.8
PC 3	0.84	14.0	98.8
PC 4	0.07	1.2	100.0
PC 5	0.00	0.0	100.0
PC 6	0.00	0.0	100.0

Table 3. The first three PCs describing the correlation of the investigated production parameters and J_{sc} (left) and FF right. The percentage of the variance these PCs capture is shown in brackets. Together they contain more than 98% of the variance in both cases. Hence the remaining three PCs are omitted.

	$\mathbf{J_{sc}}$				**FF**		
	PC 1 (52.6%)	PC 2 (31.1%)	PC 3 (15.4%)		PC 1 (53.5%)	PC 2 (31.2%)	PC 3 (14.8%)
Concentration	-0.16	0.68	-0.12		-0.20	0.67	-0.07
1st Abs-SC-Speed	-0.21	-0.04	0.96		-0.19	-0.18	-0.96
Δt Abs-SC to Evap	-0.55	-0.10	-0.12		-0.54	-0.14	0.14
Evaporation p_{start}	0.55	0.12	0.11		0.54	0.15	-0.13
Evaporation Rate Aluminium	0.55	0.12	0.11		0.54	0.15	-0.13
J_{sc}	-0.12	0.70	0.12	FF	-0.20	0.67	-0.11

The PCs show in each case similar loadings of variables, which can be grouped into three categories. PC 1 is dominated by parameters belonging to the evaporation process. PC 2 has a high loading of the concentrations in the absorber solution, whereas the spincoating speed of the absorber solution has the most influence on PC 3. The other remaining PCs contribute only little information. The main loadings of the first three PCs will be discussed in the following paragraphs.

PC 1: The first PC is dominated by the time interval between spincoating of the absorber solution and cathode-evaporation and the evaporation step itself. The distribution of time intervals in the data does not suggest an artificial correlation due to the order of production. A longer time between absorber solution spincoating and Aluminium cathode evaporation seems to benefit V_{oc} and FF, while the correlation with J_{sc} is not that strong. One possible explanation could be residual solvent in the absorber layer when evaporating the Aluminium cathode, leading to a degraded interface, although the absorber layer on the substrates was dried between $25h$ and $30h$ in the glovebox. An evaporation too early after spincoating of the absorber solution might also explain the substantial reduction in the FF sometimes seen in experiment by introducing kinks into the IV-curve in the fourth quadrant. The PC 1 has similar loading of the time interval between those two steps and the evaporation parameters start pressure und rate of aluminium evaporation. The data is however not conclusive what caused this strong correlation and it can not be excluded that they might be artefacts of the preparation conditions. Their effects on the solar cell parameters will be investigated in further studies.

PC 2: The second PC is dominated by the concentration in the absorber solution and has a large contribution of the V_{oc}, J_{sc} and FF as well. The relative signs tell that reduced concentration in the absorber solution leads to a reduced V_{oc}, J_{sc} and FF. All these effects were observed in the experiments and attributed to shunts through gaps in the apparently non closed absorber layer. Additionally, PC 2 has in the case of V_{oc} and FF an additional contribution from the spincoating speed of the absorber, which is negligible in the case of J_{sc}. Something which otherwise would not have been noticed. The polymer concentration in the absorber solution has probably the main influence on the absorber solution's properties by changing the viscosity. Together with the spincoating speed, the concentration in the absorber solution has a large influence on the absorber layer's thickness and the absorber layer morphology.[4] The effects seen here are described in more detail in the next paragraph (PC 3) with the influence of the spincoating speed.

PC 3: The third PC is dominated by the spincoating speed of the absorber solution; however the correlation with the solar cell parameters is lower than in the first two PCs. For a given concentration in the absorber solution, the spincoating speed first affects the thickness of the wet film on the substrate while spincoating. The thickness of the wet film in turn affects the film's drying conditions and hence probably the morphology formation.[4] Determination of morphology effects is currently beyond this analysis presented here. However, the primary effects of the spincoating conditions can be seen in the analysis. The real changes in absorber thickness, an important parameter for OSCs, can be estimated from absorption measurements of the complete OSC (see figure 2) by comparing local minima in the spectra with results from optical absorption simulations of the OSC layer system.[5] The measured absorption is normalised, because primarily the shape is of interest here. A PCA with the concentration in the absorber solution, the spincoating speed and the position of the local minimum as variables (indicated in figure 2 (right)) confirms the strong correlation (not shown). The approximate layer thickness as deduced from the indicated minima is shown in the figure as well. It is interesting to note that substrates, whose absorber solutions have been spincoated at $1000rpm$, appear to have nearly the same layer thickness despite a different concentration in the absorber solution. Still, they differ much in their solar cell parameters, which is picked up in all PC 2s.

Comparing the loading of the PC 3 for V_{oc}, J_{sc} and FF reveals some differences. For V_{oc} the concentration in the absorber solution has a higher loading, albeit PC 3 having little contribution from V_{oc}. For J_{sc} and FF, the concentration's loading on PC 3 is lower, but the correlation with the measurement result is large enough to be significant. Another notable difference is the change in the concentration's relative sign in the loading from J_{sc} with respect to V_{oc} and FF. Generally J_{sc} has a non-linear behaviour with respect to the layer thickness, as is shown in optical simulations.[6] The non-linearity is caused by the optical interference pattern formed by reflections off the Aluminium cathode. The simulated behaviour of the J_{sc} shows a shallow local minimum in the region of the estimated absorber layer thicknesses with little changes.

The original analysis included the time intervals between PEDOT:PSS-spincoating onto the ITO substrate and the absorber solution spincoating onto the dried PEDOT:PSS layer. Various treatments of the ITO electrode have been investigated and significant changes in its electronic properties, depending on the method, have been reported.[7] The effect of some treatments is described to be only temporary,[8] leading to a short interval in which the ITO surface has the desired properties. However, comparing this directly with the used dataset, it was not possible to confirm these effects here. Rather the substrates' order of preparation led to the correlation

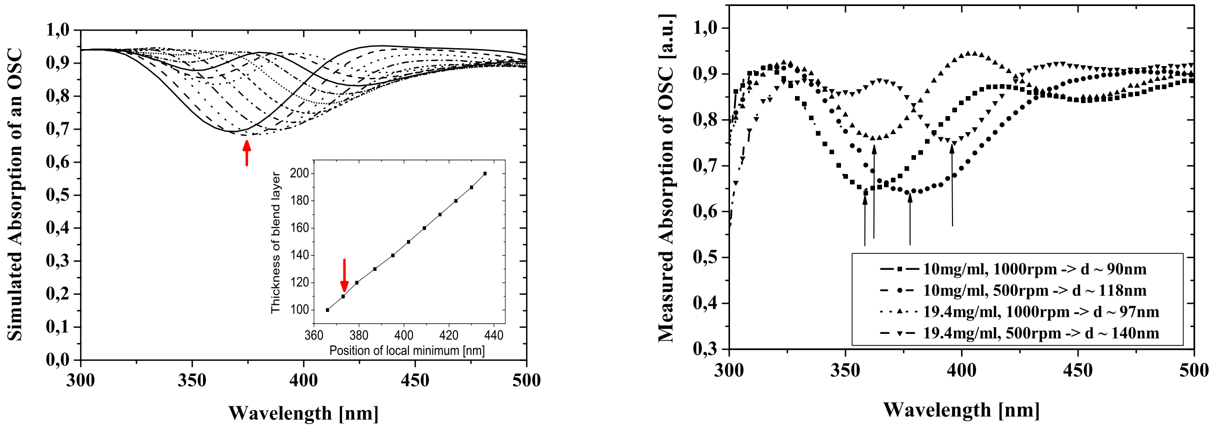

Figure 2. Simulated absorption features of an organic solar cell with different absorber layer thicknesses.[5] The inset shows the absorber layer thickness as function of the position of the indicated local minimum. On the right measured absorption of OSCs with different spincoating speeds and concentrations in the absorber solution are shown. The position of the local minimum in indicated and the deduced estimates for the absorber layer thickness shown.

with the time intervals. The substrates from an absorber solution of low concentration were all accidentally prepared with shorter time intervals before the spincoating steps than substrates from absorber solutions with higher concentrations. When the low concentration lead to low V_{oc}, J_{sc} and FF, this undesired correlation was picked up by the PCA. Hence these time intervals were omitted in this analysis, but will be addressed with a larger dataset, where such undesired correlation from the production order will be avoided.

5. CONCLUSION

We are investigating the correlation between production parameters and measurement results with a new approach. All relevant experimental data is systematically recorded and the PCA used as statistical analysis method. As a first step, this method was applied to an exemplar data set of 62 organic solar cells, for which several parameter variations at once were considered. Their influence was quantified and possible explanations given. The main influence on the measurement results was due to evaporation of the cathode layer, followed by the concentration of P3HT and PCBM in the absorber solution and the spincoating speed of the absorber solution. The PCA has shown a high interdependence between the parameters. Their physical origin will be investigated in further experiments and the PCA will be used for the larger available data of both production parameters and experimental results. With this analysis, it should be possible to identify the dominating correlations between production parameters and OSC characteristics. The results will then be used to direct the experiments at the influential key parameters to determine the factors governing the OSC behaviour. However, great care is necessary to avoid artificial correlations resulting from the production process.

ACKNOWLEDGMENTS

The authors thank Nicholas Keegan for making the solar cells, Christian Wawrzinek and Sébastien Braun for their work on the webinterface and the Fraunhofer Institute for Applied Polymer Research IAP for the P3HT donation.

REFERENCES

1. K. Lee, J. Y. Kim, W. Ma, and A. Heeger, "New architecturs for thermally stable high efficient polymer solar cells," in *Organic Photovoltaics VI*, Z. H. Kafafi and P. A. Lane, eds., *Proceedings of SPIE* **5938**, pp. 59380B–1–12, 2005.
2. K. Person, "On lines and planes of closest fit to system of points in space," *Philosophical Magazine* **2**(6), pp. 559–572, 1901.
3. J. Honerkamp, *Stochastic Dynamical Systems*, John Wiley & Sons, New York, 1993.
4. U. Zhokhavets, G. Gobsch, H. Hoppe, and N. S. Sariciftci, "Anisotropic optical properties of thin poly(3-octylthiophene)- films as a function of preparation conditions," *Synthetic Metals* **143**(1), pp. 113–117, 2004.
5. T. Ziegler, *Optische Spektroskopie an Organischen Solarzellen.* Diploma thesis, Universitaet Freiburg, Fakultaet fuer Mathematik und Physik, 3 2006.
6. M. Niggemann, T. Ziegler, M. Glatthaar, M. Riede, B. Zimmermann, and A. Gombert, "Optical near field phenomena in planar and structured organic solar cells," in *Photonics for Solar Energy Systems*, *Proceedings of SPIE* **6197**, 2006.
7. J. S. Kim, M. Granstrm, R. H. Friend, N. Johanson, W. R. Salaneck, and F. Cacialli, "Indium-tin oxide treatments for single- and double-layer polymeric light-emitting diodes," *J. Appl. Phys.* **84**, pp. 6859–6870, 1998.
8. A. Berntsen, Y. Croonen, R. Cuijpers, B. Habets, C. Liedenbaum, H. Schoo, R. J. Visser, J. Vleggaar, and P. Weijer, "Polymer light-emitting diodes: from materials to devices," in *Organic Light-Emitting Materials and Devices*, Z. H. Kafafi, ed., *Proceedings of SPIE* **3148**, pp. 264–271, 1997.