

Modelling of Random Textured Tandem Silicon Solar Cells Characteristics: Decision Tree Approach

R.S. Kamath¹, R.K. Kamat²

¹ Department of Computer Studies, Chhatrapati Shahu Institute of Business Education and Research, University Road, 416004 Kolhapur

² Department of Electronics, Shivaji University, 416004 Kolhapur

(Received 04 July 2016; revised manuscript received 22 November 2016; published online 29 November 2016)

We report decision tree (DT) modeling of randomly textured tandem silicon solar cells characteristics. The photovoltaic modules of silicon-based solar cells are extremely popular due to their high efficiency and longer lifetime. Decision tree model is one of the most common data mining models can be used for predictive analytics. The reported investigation depicts optimum decision tree architecture achieved by tuning parameters such as Min split, Min bucket, Max depth and Complexity. DT model, thus derived is easy to understand and entails recursive partitioning approach implemented in the “rpart” package. Moreover the performance of the model is evaluated with reference Mean Square Error (MSE) estimate of error rate. The modeling of the random textured silicon solar cells reveals strong correlation of efficiency with “Fill factor” and “thickness of a-Si layer”.

Keywords: Silicon Solar Cell, Decision Tree, Efficiency, Rattle.

DOI: [10.21272/jnep.8\(4\(1\)\).04021](https://doi.org/10.21272/jnep.8(4(1)).04021)

PACS numbers: 85.30. – z, 88.40.fc

1. INTRODUCTION

The solar cells are turning into an undeniably well-known type of option vitality around the globe, utilizing photovoltaic and silicon solar cells to change over radiation into power. Solar power has had a long history of fizzled begins and restricted appropriation. The PV modules of silicon-based solar cells are extremely popular in the market due to their high efficiency and longer lifetime. In recent years, many research investigations are focused on the improvement of solar cell efficiency [1-3]. Scholarly literature in this regards reveals that the expense of thin film solar cells utilizing cost effective inorganic semiconductor materials, called "generation two" photovoltaics, is on a very basic level constrained by embodiment and parity of-framework expenses. To accomplish the expense and execution levels expected to contend in the wholesale vitality market, "third generation" photovoltaic technology requires not just a noteworthy change in productivity without adding apparently to substantial scale fabricating costs, additionally that it utilizes profuse, innocuous, stable, and robust materials. In this regards, recently our group reported investigations on the effect of device dimension on the silicon solar cell, by using the PC1D numerical simulation environment, wherein strong correlation of efficiency of the silicon solar cell with its size has been exemplified [4]. Furthermore among the conceivable methodologies that could meet these criteria, tandem or stack-cells of various bandgaps have been recognized as having high potential for productivity change, with efficiency confining in the range of 45 % and 50.5 % for two- and three-cell stacks individually [5].

In the present investigation, we demonstrate the modelling of tandem silicon solar cells characteristics using decision tree approach [11]. Decision trees are the traditional building blocks of data mining and the classic machine learning algorithm [12]. Because of its simple representation it is popular. The algorithm uses a recursive partitioning approach, implemented in the

“rpart” package. It is comparable to CART and ID3/C4 [16]. The reported decision tree consists of three rules with five nodes for modelling of tandem solar cell characteristics. The reported experiment is simulated in R and Rattle [15]. R is an open source tool for statistical data processing data mining. Rattle is a graphical data mining package offers GUI for R.

The rest of the paper is structured as follows; after a brief introduction, the second section deals with the theory of randomly textured tandem silicon solar cells. The third section portrays theoretical aspects of proposed decision tree model. The computational details, results and discussion are reported in the fourth section. The conclusion at the end divulges aptness of the decision tree for predicting the efficiency of tandem solar cell.

2. RANDOMLY TEXTURED TANDEM SILICON SOLAR CELLS: THEORETICAL CONSIDERATION

The tandem solar cell is more efficient than single junction solar cell owing to absorption of different band of electromagnetic spectrum by different layers of solar cell [6-8]. In the present investigation, we have considered randomly textured tandem silicon solar cells in the form of SiO₂/ITO/a-Si/c-Si/ZnO/Ag which is shown in Fig. 1.

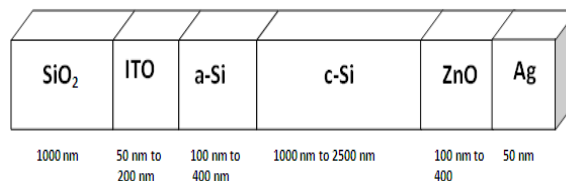


Fig. 1 – Schematic representation of random textured SiO₂/ITO/a-Si/c-Si/ZnO/Ag solar cell

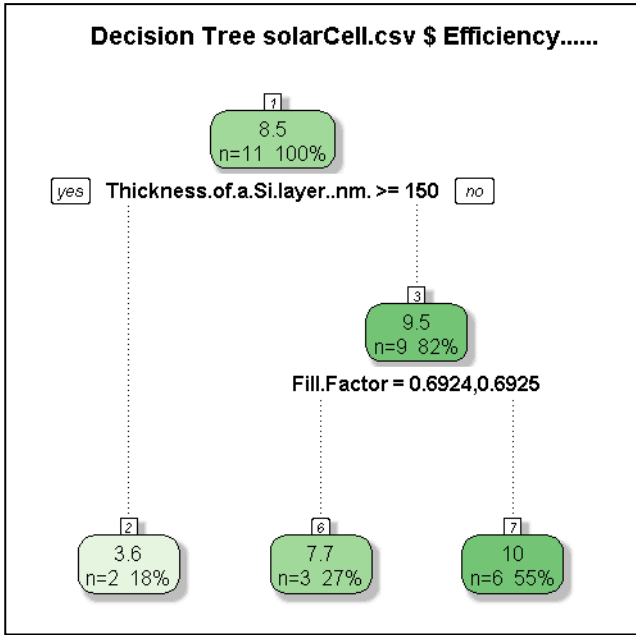


Fig. 2 – Decision tree for tandem solar cell efficiency

The data for decision tree modelling is acquired from time domain simulation of randomly textured tandem silicon solar cells using quadratic complex rational function approach [9]. Efficiency of solar cell is obtained by varying thickness of ITO layer, a-Si layer, c-Si layer and ZnO layer [9]:

- Efficiency of the tandem silicon solar cell increases with increase of thickness of ITO layer. The higher efficiency is obtained at 150 nm ITO thickness and efficiency tends to decrease further.
- Efficiency of silicon solar cell tends to decrease as thickness of the a-Si layer increases.
- Since the efficiency of the solar cell can be increased by absorbing a large portion of the infrared region, the efficiency tends to increase with increase of the thickness of the c-Si layer
- Variation in the ZnO layer thickness effects less on efficiency of solar cell

3. DECISION TREES: APTNESS FOR MODEL-LING

The decision tree is a very convenient and efficient representation of knowledge [16]. To model the characteristics of randomly textured tandem silicon solar cells, we have employed decision tree approach. It starts with a single root node that splits into multiple branches, leading to further nodes, each of which may further split or else terminate as a leaf node. Associated with each nonleaf node will be a test or question that determines which branch to follow [15]. The leaf nodes contain the decisions.

Fig. 2 shows decision tree derived in the present investigation represents efficiency of solar cell for different input combinations. That the root node of decision tree tests *thickness of a-Si*, has a value ≥ 150 continues down to the left side of the tree, otherwise right side of the tree. The next test down this right side of the tree is *fill factor* value. Thus it proceeds and will be able retrieve efficiency value of tandem solar cell [13].

Following are the set of tuning parameters varied in Rattle to obtain optimized decision tree model for tandem solar cell efficiency computation:

- Min split – Minimum number of observations that must exist in a node resulting from a split before a split will be performed 6
- Min Bucket – This is the minimum number of observations allowed in any leaf node of the decision tree 2
- Max Depth – This is the maximum depth of any node of the final tree 10
- Complexity – This parameter is used to control the size of the decision tree and to select optimal tree size.

The performance of the model is calculated by using Mean Square Error between expected output and estimated output, given in equation (1). The Y_i represents the observed value of the i^{th} observation, where, $i = 1, 2, \dots, n$ and \hat{Y}_i denote the predicted value of the i^{th} observation [13]. The difference $(Y_i - \hat{Y}_i)$ is termed as an error.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

4. COMPUTATIONAL DETAILS, RESULTS AND DISCUSSION

This section explores details of experiment conducted for the modelling of tandem silicon solar cells characteristics using decision tree approach. R and Rattle are used to analyze model structure, configuring tuning parameters to obtain optimized DT model [18]. The model is conceived as Multi-Input Single Output. It works basically with six inputs viz. thickness of SiO_2 , ITO, a-Si, c-Si, ZnO and silver layer. Efficiency of solar cell is considered as output variable [17]. The dataset for decision tree modelling is acquired from time domain simulation of randomly textured tandem silicon solar cells using quadratic complex rational function approach [9]. Dataset reveals that higher efficiency is evident at 150 nm ITO layer thickness, 100 nm a-Si layer thickness, and 2500 nm c-Si layer thickness.

Fig. 3 summarizes the decision tree for tandem solar cell characteristics modelling. It is the text view of resultant decision tree and also highlights the key interface widgets that need to deal with to build a tree. The tree has built to predict the value of the variable “Efficiency” based on the remainder of the variables in the dataset supplied [14]. Variables actually used in tree construction are “Fill factor” and “thickness of a-Si layer”. Fig. 4 shows the rule set which is a textual representation of tandem solar cell decision tree model. The decision tree shown in fig 2 translates to the rules, where each rule corresponds to one pathway through the decision tree, starting at the root node and terminating at a leaf node.

Thus derived optimized decision tree entails values for tuning parameters such as Min split, Min bucket, Max depth and complexity are 6, 2, 10 and 0.01 respectively. Performance evaluation of the model is summarized in Table 1. This complexity table explains iterations and associated change in the accuracy of the model as new levels are added to the tree. We are most likely interested in the cross-validated error, which is the *xerror* column of the table. The CP (complexity

parameter) value reveals that as the tree splits into more nodes, the complexity parameter is reduced. But we also note that the cross validation error starts to increase as we further split the decision tree. This tells the algorithm to stop partitioning, as the error rate is not improving.

Table 1 – Complexity Table for DT model

level	CP	nsplit	rel error	xerror	xstd
1	0.76616	0	1.000000	1.20111	0.31371
2	0.19570	1	0.233841	0.54733	0.23574
3	0.01000	2	0.038144	0.31371	0.25321

5. CONCLUSION

In the present paper, we have reported decision tree

(DT) modelling of randomly textured tandem silicon solar cells characteristics. A decision tree model is one of the most common data mining models. It is popular because the resulting model is easy to understand. The reported investigation depicts optimum decision tree architecture achieved by tuning parameters such as Min split, Min bucket, Max depth and complexity. DT model, thus derived is easy to understand and entails recursive partitioning approach implemented in the rpart package. Result concludes that DT prediction is a suitable approach since the resulting analysis is much more accurate and precise. Consistent with our earlier reported investigations [4], the modelling of the randomly textured tandem silicon solar cells demonstrates strong correlation of efficiency with “Fill factor” and “thickness of a-Si layer”.

```
n= 11
node), split, n, deviance, yval
  * denotes terminal node

1) root 11 75.20826000  8.463545
 2) Thickness.of.a.Si.layer..nm.>=150 2  2.47620300  3.608400 *
 3) Thickness.of.a.Si.layer..nm.< 150 9 15.11054000  9.542467
 6) Fill.Factor=0.6924,0.6925 3  0.01448451  7.733967 *
 7) Fill.Factor=0.6928,0.6929 6  0.37803510 10.446720 *

Regression tree:
rpart(formula = Efficiency..... ~ ., data = crs$dataset[crs$train,
  c(crs$input, crs$target)], method = "anova", parms = list(split = "information"),
  control = rpart.control(minsplit = 6, minbucket = 2, maxdepth = 10,
    cp = 0.0102, usesurrogate = 0, maxsurrogate = 0))

Variables actually used in tree construction:
[1] Fill.Factor          Thickness.of.a.Si.layer..nm.

Root node error: 75.208/11 = 6.8371

n= 11
      CP nsplit rel error  xerror  xstd
1 0.76616     0  1.000000  1.20111  0.52302
2 0.19570     1  0.233841  0.54733  0.23574
3 0.01020     2  0.038144  0.31371  0.25321
```

Fig. 3 – Summary of the Decision Tree for tandem solar cell characteristics modelling

```
Tree as rules:

Rule number: 7 [Efficiency.....=10.44671666666667 cover=6 (55%)]
  Thickness.of.a.Si.layer..nm.< 150
  Fill.Factor=0.6928,0.6929

Rule number: 6 [Efficiency.....=7.73396666666667 cover=3 (27%)]
  Thickness.of.a.Si.layer..nm.< 150
  Fill.Factor=0.6924,0.6925

Rule number: 2 [Efficiency.....=3.6084 cover=2 (18%)]
  Thickness.of.a.Si.layer..nm.>=150

[1] 1 3 5 4 2
```

Fig. 4 – Textual representation of tandem solar cell decision tree model

REFERENCES

1. S.P. Bremner, M.Y. Levy, C.B. Honsberg, *Prog. Photovoltaic.: Res. Appl.* **16** No 3, 225 (2008).
2. A.M. Mozharov, A.D. Bolshakov, G.E. Cirilin, D.A. Kudryashov, A.S. Gudovskikh, I.S. Mukhin, Z.I. Alferov, *J. Phys.: Conf. Series* **690** No 1, 012041 (2016).
3. M. Boccard, M. Despeisse, J. Escarre, X. Niquille, G. Bugnon, S. Hanni, M. Bonnet-Eymard, F. Meillaud, and C. Ballif, *IEEE J. Photovoltaic.* **4** No 6, 1368 (2014).
4. A.B. Nawale, R.A. Kalal, A.R. Chavan, T.D. Dongale, R.K. Kamat, *J. Nano- Electron. Phys.* **8** No 2, 02002 (2016).
5. "GCEP Research » Blog Archive » Nanostructured Silicon-Based Tandem Solar Cells", Web.stanford.edu, 2016. [Online]. Available: <http://web.stanford.edu/group/gcep/cgi-bin/gcep-research/all/nanostructured-silicon-based-tandem-solar-cells/>.
6. J.A. Weicht, F.U. Hamelmann, G. Behrens, *J. Phys.: Conf. Series* **682** No 1 012017 2016.
7. T.F. Schulze, L. Korte, E. Conrad, M. Schmidt, B. Rech, *J. Appl. Phys.* **107** No 2, 023711 (2010).
8. S. Ji, C.S. Hwang, P. Jeong, S. Lee, K.S. Lee, *Thin Solid Films* **598**, 226 (2016).
9. T.D. Dongale, S.V. Katkar, K.V. Khot, K.V. More, S.D. Delekar, P.N. Bhosale, R.K. Kamat, *J. Nanoeng. Nanomanuf.* **6**, 103 (2016)
10. S.H. Park, I. Shin, K.H Kim, R.Street, A. Roy, and A.J. Heeger *Adv. Mater.* **27** No 2, 298 (2015).
11. R. Liu, A. Kumar, Z. Chen, A. Agrawal, V. Sundararaghavan, and A. Choudhary, *A Data Mining Approach in Structure-Property Optimization. 2014. TMS Annual Meeting & Exhibition, Symposium of Data Analytics for Materials Science and Manufacturing* (Feb. 16-20, San Diego, CA).
12. Ghanshyam Paliania, Chenchen Wang, Xun Jiang, Sanguthevar Rajasekaran, Ramamurthy Ramprasad, *Sci. Report.* **3** (2013).
13. C.C. Fischer, K.J. Tibbetts, D. Morgan, G. Ceder: *Nat. Mater.* **5** No 8, 641 (2006).
14. A. Agrawal, P.D. Deshpande, A. Cecen, G.P. Basavarsu, A.N. Choudhary, S.R. Kalidindi, *Integrating Mater. Manufacturing Innovation* **3** No 8, 1 (2014).
15. R. Kamath, R. Kamat, *Educational Data Mining with R and Rattle, River Publishers Series in Information Science and Technology* (River Publishers: Netherland: 2016).
16. W. Graham, *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*, Springer.
17. Katja Hansen, Grégoire Montavon, Franziska Biegler, Siamac Fazli, Matthias Rupp, Matthias Scheffler, O Anatole von Lilienfeld, Alexandre Tkatchenko, and Klaus-Robert Müller, *J. Chem. Theory Comput.* **9**, 3404 (2013).
18. Richard LeSar, *Statistical Analysis and Data Mining* **1**, 372 (2009).