Predicting Geographic Energy Production for Tandem PV Designs Using a Compact Set of Spectra Correlated by Irradiance

Emily C. Warmann and Harry A. Atwater, Member, IEEE

Abstract—We provide 20 direct spectra that capture the variation of the solar spectrum composition with intensity. We correlate the value of the air mass, aerosol optical depth at 500 nm, precipitable water, and ozone with the cumulative irradiance for direct sunlight with the use of National Solar Radiation Database (NSRDB). We use the values of these atmospheric parameters to generate spectra that represent their corresponding cumulative irradiance levels with the use of SMARTS multiple scattering and transmission model. By simulation of the performance of a solar cell design under these 20 spectra and combination of the intensity-specific performance with the relative frequency of each irradiance level at a particular location from the NSRDB, we can predict tandem cell energy production across the United States. Through comparison of the energy production of ideal tandem cells with two to ten subcells as predicted by our model to energy production integrated over one year’s worth of simulated spectra at nine locations across the United States as well as measured spectral irradiance from NREL’s solar observatory, we find the error in our energy production estimate to be under 5% for ten subcells and under 3% for up to five subcells. We demonstrate the utility of the approach with a selection of prospective and ideal multijunction bandgap combinations.

Index Terms—Photovoltaic systems, solar energy, solar power generation.

I. INTRODUCTION

As the photovoltaics field expands beyond single junction cells, it needs new methods to compare efficiency and predict energy production. In the field, the spectrum often deviates from the ASTM G-173 standard, and the efficiency of spectrally sensitive photovoltaics (PV) such as series-connected tandem or multijunction solar cells (MJSCs) suffer as current mismatch develops between the subcells [1]. Peak energy production may not correspond with peak efficiency under ASTM G-173 [2]. Conversely, comparison of designs on the basis of efficiency under the standard may not accurately predict their relative energy production [3]. Current solar resource datasets typically report only cumulative irradiance levels, that lack the variation in spectrum needed to predict tandem performance. Here we present 20 direct spectra that can estimate the spectral variation for any location where cumulative irradiance data are available. These 20 spectra allow evaluation of PV designs on a geographic basis.

A variety of methods can predict performance for spectrally sensitive PV. Data on atmospheric conditions such as pressure, aerosol optical depth (AOD), precipitable water (PW), and ozone and other pollutant levels input to a multiple scattering and absorption model such as SMARTS simulate atmospheric transmission [4]–[6]. This requires many spectra to capture variation over time, and the results will be location-specific.

An alternative to simulation of a full year’s worth of spectra is to identify a small number of spectra that represent the typical conditions at a location, such as the 50% air mass approach [7]. Yandt et al. generate generating spectra with AM values that range from 1 to 5, all other inputs held constant [8]. The relative prevalence annual of each AM value at a location determines the spectrum that best represents the typical spectral composition. The latitude determines the dominant air mass and make this method somewhat extensible, although it cannot analyze other geospatial factors without further tailoring.

Other approaches use a spectral data set and reduce it to a small number of representative spectra by averaging spectra grouped based on the wavelength that divides the incident spectrum into two equal power density bands, or the average photon energy (APE) [9], [10], based on the equivalent photocurrent ratio [11], or through a machine learning approach based on spectral features [12]. These methods can be accurate, but they depend on an existing spectral database and do not extend on a geographic basis.

While tools to optimize a PV design for energy production at specific locations are valuable, they are based on a vision of tailored cell design that may not be compatible with current cost trajectories for PV seeking efficiencies of scale. Rather, designs that can maximize energy production over a broad geographic extent will have higher economic potential. Similarly, while computational gains have eased the burden of simulation of a large number of spectra, finding model inputs that accurately represent a wide range of locations can be challenging and lies outside the main expertise of cell designers.
We propose a new tool to estimate energy production of spectrally sensitive PV designs. We have used the National Solar Radiation Database (NSRDB) to determine the correlation of the AM, AOD, PW and ozone with cumulative irradiance. With these inputs, we have used the SMARTS model to generate 20 spectra that represent average spectral composition correlated with cumulative irradiance that range from 0 to 1100 W/m². By simulation of the performance of PV designs under these spectra, we can analyze the designs’ spectral sensitivity and extend that analysis to a geographic basis by use of the location-specific frequency of those 20 cumulative irradiance levels as contained in the NSRDB. With this tool, we can predict the energy production of a tandem cell design across the United States after simulating performance under only 20 individual spectra.

II. SPECTRAL MODEL

The NSRDB supplies simulated hourly cumulative irradiance estimates for 242 class 1, 618 class 2, and 594 class 3 sites across the United States [13]. The data are derived from historic hourly measurements of local temperature, pressure, and humidity; regional measurements of AOD, PW, and ozone from the AERONET sensor network; and a mixture of ground and satellite derived cloud observations, and comprise estimates for global horizontal, direct normal and diffuse irradiance with and without clouds. The database also includes the estimates for the AOD, PW, and ozone values used to generate the irradiance estimates.

For this analysis, we used the NSRDB data set from 1990 to 2010, restricted to the class 1 and class 2 sites. For each location we averaged the direct clear sky (without clouds) irradiance values for each hour over the 20 years of data to create an average for each hour of the year, and averaged the AOD, PW, and ozone values for each hour as well. To determine the correlation between these parameters with cumulative direct normal irradiance (DNI), we binned the DNI values into 20 ranges equally spaced between 0 and 1100 W/m². The air mass, AOD, PW, and ozone values corresponding to the DNI values in each range were then averaged across all locations and times. Fig. 1(a)–(c) shows the correlation of air mass, AOD, and PW with irradiance level. At low irradiance, change in the AM dominates. The AOD and PW correlations are strongest in the 800–1000 W/m² range, corresponding to the irradiance levels responsible for the bulk of the cumulative incident energy. Ozone (not shown) exhibits little correlation, consistent with prior results [4], [5].

We used these average values as inputs to SMARTS with the remaining inputs (pressure, atmosphere model, pollutant loading, etc.) taken from the input file that generates the ASTM G-173 standard spectrum [14]. The resulting spectra will represent the average incident clear sky spectrum for each of the 20 cumulative irradiance ranges.

Because the performance of MJSCs under the different spectra is determined by the current-matching among the subcells, we compared the model spectra by reweighting them with the same total flux as the AM1.5D reference. Fig. 2 shows the ratio of five reweighted model spectra to AM1.5D. Regions where the curves are greater than one represent energy ranges where subcells that absorb those photons will be over illuminated relative to AM1.5D, while subcells absorbing in regions where the curve is less than one will be current-starved and limit a series-connected ensemble. The irradiance levels that label the curves correspond to the upper limit of the range represented. As expected, the higher irradiance spectra get a larger percentage of the total flux from high energy photons, while the lower irradiance spectra are dominated by lower energy photons. Interestingly, the model
spectrum that represent irradiances of 1045–1100 W/m² also shows a higher contribution from red photons near the edges of the absorption bands, because of a narrowing of the water absorption bands in the extremely clear sky conditions required for irradiance values in this range.

III. MODEL VALIDATION

We used a two-pronged approach to validate our model. First we compared the model performance to a set of full year-simulated spectra for nine cities across the United States. This allowed comparison of model performance over a full year and at locations that vary in climate, elevation, and latitude. Our second validation exercise compared model performance against observed spectra as measured at the NREL Solar Radiation Research Laboratory. These spectra cover only a limited subset of the year and a single location; however, they are measured data. In both validation exercises we compared the model performance with the data set on the basis of APE, because this parameter has been shown to have a correlation with irradiance [15] and gives insight into spectral composition, and we compared the energy production of a set of test MJSC designs as predicted by the model with a full calculation under all spectra in the validation set.

We generated simulated spectra for one year at nine locations: Phoenix, AZ, Dallas, TX, Atlanta, GA, Denver, CO, Indianapolis, IN, Philadelphia, PA, Sacramento, CA, Seattle, WA, and Minneapolis, MN. These cities form a set of transects spanning three latitudes (~33°, ~38°, and 44°–47°) and multiple climate characteristics, specifically Mediterranean (CA and WA sites), arid/semiarid (AZ and CO), and humid subtropical/continental (TX, GA, IN, MN, PA sites). For each city we simulated a direct spectrum hourly for each day of the year with the use of the location-specific average AM, AOD, PW, and ozone value for that hour as reported in the NSRDB. The other inputs to the SMARTS simulation were taken from the ASTM G-173 standard spectrum file, corrected for the location’s elevation [14].

While accuracy in prediction of energy production is the primary measure of model accuracy, we also examined the similarity of the model spectra to the larger observed and simulated spectral data sets. For each location, we sorted the simulated spectra into bins corresponding to the model irradiance ranges, and averaged the APE values of the spectra in each bin. Fig. 3(a) compares the APE of the model spectra and the bin average APE values for the nine cities. The black circles show the model APE values. The light gray circles, slightly offset to reduce overlap, correspond to the bin mean APE for each of the nine cities. The red X markers show the mean of the nine city bin means, and the error bars show the standard deviation in the mean among the nine cities.

In all cases, there is a trend of increase of APE with increase of irradiance until very high irradiance levels. The peak and subsequent decrease in APE at high irradiance is explained by the phenomenon shown in Fig. 2, and supports the conclusion that APE is not a unique spectral parameter [16]. Fig. 3 shows close agreement for irradiance levels higher than 200 W/m². The standard deviation in mean APE among the cities is roughly 0.006–0.015 eV for most bins, while the standard deviation in APE within a bin is roughly 0.2–0.4 eV, which indicate that once spectra are sorted into bins the differences between cities are small on average compared with the variation in the spectra themselves. Except for the two lowest bins, the model APE value for each bin falls within one standard deviation of the mean APE across cities, which suggests that the model is a good representation of the spectra in this data set. Generating model spectra on single-year data inputs rather than the full 20 year dataset produced slightly different APE values. The standard deviation in APE over the 20 subset models ranged from 0.001 to 0.006 eV in all but the lowest bin (0.01 eV). This suggests small interannual variation when averaged over all locations.

While this set of cities was chosen to span a wide range of latitudes, climate conditions, and elevations, the model spectra were generated with the use of input parameter values averaged across points that span the contiguous 48 states. Consequently the agreement between model and simulated spectra supports the model’s application to a wide range of geographic locations.

For further model validation we compared with a set of measured direct normal spectra from the NREL Solar Radiation Research Laboratory, recorded by the MS-7XX instrument in October 2015. To compare with our clear sky model, we selected days that show no cloud obstruction, and compensated for the limited range of the measured spectra, which extend only to 1650 nm (versus 4000 nm for the model spectra), by truncation of the model spectra and recenter of the bins around the reduced irradiance. The maximum irradiance considered remained 1100 W/m² and the minimum was 20 W/m², that produce 3873
measured spectra. As with the simulated spectra, we binned the measured spectra and averaged the APE values in each bin.

Fig. 3(b) shows the mean APE value and standard deviation for each bin along with the model APE values. The model and mean measured APE values agree at irradiance levels above 500 W/m², with the model value that fall within one standard deviation of the mean APE for 14 of the 18 irradiance ranges where comparison is possible. The measured lower irradiance spectra have higher APE values than the model or simulated spectra would predict, on average, perhaps because of remaining cloud effects. However the spectra in these low ranges constitute only 1.8% of the 3873 spectra and less than 0.01% of the cumulative irradiance in the data set. Error in this irradiance range will therefore have a small impact on the overall accuracy.

The second component of our validation exercise compares energy production predicted by the model method and a full calculation under all the spectra in each reference data set, with the use of a set of series-connected MJSCs with two to ten bandgaps optimized for the AM 1.5D standard spectrum [17]. We calculated the efficiency of these MJSCs with the use of an ideal detailed balance (Shockley–Queisser) calculation at 25 °C. In all cases we computed the power production of the test MJSCs under each validation spectrum (simulated or measured) and integrated over 1 h to calculate energy production, summing the energy production over all the spectra for each city (or for the whole observed data set) to determine the total energy production for the reference cases. We also calculated the power for each test MJSC under our 20 model spectra and then multiplied those power levels by the number of spectra in the corresponding irradiance bin for each data set, again integrating over 1 h to determine the model prediction. Table I shows the prediction error for the nine cities.

As Table I shows, the error increases with the number of subcells, from 0.6% for two subcells (averaged over locations) to 3.1% for ten subcells, and the standard deviation in error increases from 0.4 to 1.2 percentage points. The error for five subcells is 3.1% or less and the error for ten subcells is less than 5% for all locations. The validation with the use of observed spectra was only possible for the MJSCs with 2–4 subcells, because the spectra are truncated at an energy higher than the lowest bandgap for the MJSCs with more subcells. The model prediction for the two, three, and four subcell MJSCs had errors of 0.2%, 1.3%, and 1.5%, respectively, when compared with the full 3873 spectra calculation. These values are consistent with the average errors for the nine cities despite coming from a distinct data set. Further, when comparing the energy production error for sets of 20 4, 6, and 8 J cells (efficiency ~2 percentage points below ideal), the standard deviation of the error for designs compared across a single location (0.5, 0.6, and 1 percentage points, respectively) was less than half the standard deviation in error for a single design compared between locations (1.2, 1.8, 2.3 percentage points). This suggests the model will be more accurate at identification of performance differences between designs than geographic performance trends in a single design because of geographically correlated errors.

The spectra generated by this method assume “clear sky” conditions, or the absence of clouds. The error estimate detailed above also uses the clear sky estimates of cumulative irradiance from the NSRDB to serve as a fair comparison; however, this will overstake the actual cumulative irradiance for a location by 10–50%. The NSRDB includes estimates of irradiance with both clear sky and cloudy conditions, which together provide an hourly estimate of cloud transmission. Once this cloud correction is applied, the distribution changes; however, the error in the corrected estimate is within 0.05 percentage points of the error for the clear sky estimates and the geographic pattern is unchanged.

### IV. Application of Model

Once the efficiency versus irradiance curve is established for an individual design under the 20 spectra, the performance of that design over the entire United States can be determined based on the average frequency of each irradiance level at different

<table>
<thead>
<tr>
<th>Number of subcells</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix</td>
<td>0.5%</td>
<td>1.6%</td>
<td>2.1%</td>
<td>1.9%</td>
<td>2.4%</td>
<td>2.9%</td>
<td>2.8%</td>
<td>3.1%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Dallas</td>
<td>0.6%</td>
<td>1.4%</td>
<td>2.1%</td>
<td>1.3%</td>
<td>1.8%</td>
<td>2.7%</td>
<td>2.4%</td>
<td>2.6%</td>
<td>3.2%</td>
</tr>
<tr>
<td>Atlanta</td>
<td>0.5%</td>
<td>1.0%</td>
<td>1.4%</td>
<td>0.9%</td>
<td>1.1%</td>
<td>1.8%</td>
<td>1.5%</td>
<td>1.7%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Sacramento</td>
<td>0.7%</td>
<td>1.4%</td>
<td>1.7%</td>
<td>1.4%</td>
<td>1.7%</td>
<td>2.2%</td>
<td>2.0%</td>
<td>2.2%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Denver</td>
<td>0.1%</td>
<td>0.4%</td>
<td>0.5%</td>
<td>0.5%</td>
<td>0.7%</td>
<td>0.8%</td>
<td>0.7%</td>
<td>0.8%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Indianapolis</td>
<td>0.7%</td>
<td>2.3%</td>
<td>2.8%</td>
<td>3.0%</td>
<td>3.5%</td>
<td>4.0%</td>
<td>4.0%</td>
<td>4.4%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Seattle</td>
<td>0.1%</td>
<td>1.2%</td>
<td>1.6%</td>
<td>1.6%</td>
<td>2.0%</td>
<td>2.3%</td>
<td>2.3%</td>
<td>2.5%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>1.3%</td>
<td>2.7%</td>
<td>3.3%</td>
<td>3.1%</td>
<td>3.6%</td>
<td>4.2%</td>
<td>4.1%</td>
<td>4.4%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>1.1%</td>
<td>1.9%</td>
<td>2.3%</td>
<td>2.0%</td>
<td>2.3%</td>
<td>2.9%</td>
<td>2.7%</td>
<td>2.9%</td>
<td>3.3%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>0.6%</td>
<td>1.5%</td>
<td>1.9%</td>
<td>1.7%</td>
<td>2.1%</td>
<td>2.6%</td>
<td>2.5%</td>
<td>2.7%</td>
<td>3.1%</td>
</tr>
</tbody>
</table>
locations from the NSRDB multiplied by the power produced under the corresponding spectra and integrated over 1 h.

Fig. 4 shows an example analysis based on the model: the performance gain relative to the ideal 3 J tandem for ideal 4, 6, 8, and 10 J cells (Fig. 4(a)–(d)) mapped over the contiguous United States. In Fig. 4(e), a table shows the efficiency of these ideal designs relative to the ideal 3 J tandem. The map in Fig. 4(a) for the 4 J tandem has a very narrow range on the color map and shows a performance improvement of approximately 6% relative to the 3 J across the entire area. This 6% improvement is predicted by the AM1.5D efficiency of the 4 J relative to the 3 J, as shown in Fig. 4(e). In Fig. 4(b), the map shows a 10–13% efficiency gain for the 6 J relative to the 3 J. The geospatial pattern is the same as that of the 4 J; however, the range of improvement is broader and everywhere lower than the 15% improvement predicted by the AM1.5D comparison, which suggests a lower energy return on the design challenge of addition of subcells. Fig. 4(c) and (d) show a diminishing benefit to additional subcells past six, with the 8 J showing only a slight improvement over the 6 J relative performance map. The 10 J map in Fig. 4(d) shows the first change in the geospatial pattern. In addition to having a narrow range of relative performance, the areas that show the highest increase (roughly 14%) in performance relative to the 3 J are in the eastern portions of the country. The southwest region has relatively lower performance increase of around 13%. In all areas, the improvement is far lower than that predicted by the 10 J efficiency under AM1.5D, which is 21% higher than the 3 J reference efficiency. The map for the 12 J design (not shown) shows a performance improvement relative to 3 J of 9%–10% with a pattern similar to the 10 J map, which indicates that the peak annualized efficiency and energy production will be achieved with a design of 8–10 subcells depending on the particular location. While the performance range on the maps is close to the average error in Table I, this general efficiency trend is consistent with prior work on the energy production of multijunction PV [17], [18].

As a second test case, consider the case of two MJSCs, each with four subcells and an efficiency of 52% under the AM1.5D spectrum (versus the optimal efficiency of 54%). Fig. 5 shows the subcell bandgaps of these two candidate MJSCs as well as the optimal 4 J bandgaps tabulated in Fig. 5(a), while the maps of the one sun annualized efficiency—that is the total annual energy production divided by the total annual irradiance—for the two candidates are presented in Fig. 5(b) and (c). Interestingly, although both designs have equal efficiency under the standard AM1.5D spectrum, their expected energy production varies on a geospatial basis. Design A is both more efficient everywhere on the map and more consistent in performance compared with design B. These differences are greater than the expected error in energy production for tandems with four subcells, and the full-year calculation for the nine cities displays the same performance pattern.

Here designs A and B have substantial differences in the bandgaps of three out of four subcells yet also have near identical efficiency under AM1.5D. The striking performance difference presented in Fig. 5 is not intuitively obvious based only on the bandgaps. While the model application here uses only an ideal Shockley–Queisser cell model, a more detailed analysis which incorporate device physics and optical system performance will allow the designer to make detailed comparison of prospective designs with the use of a broader set of relevant parameters, and
the result will give performance predictions on the geospatial basis relevant for large-scale deployment.

V. CONCLUSION

These 20 direct spectra that represent the spectral composition of sunlight in different cumulative irradiance ranges allow the prediction of energy production for prospective PV designs by combination of irradiance-dependent performance with existing location-specific cumulative irradiance data. Unlike previous efforts to capture real-world spectral variation, this method requires no spectrally resolved irradiance measurements, neither atmospheric data, nor additional spectrum simulation.

We have used these 20 direct spectra to predict the annualized efficiency of prospective series-connected MJSC designs optimized for the AM1.5D spectrum. Mapping these annualized efficiency predictions across the contiguous United States shows the utility of the 20 spectra for analysis of designs for energy production potential on a broad geographic basis. With this tool we can predict performance differences in designs that have identical efficiency under the AM1.5D spectrum and obtain more realistic estimates of the benefit of increasing the number of subcells in a design, both useful for assessment of the economic potential of prospective designs.

ACKNOWLEDGMENT

E. C. Warmann would like to thank Carissa Eisler and Pilar Espinet-Gonzales for many useful discussions.

REFERENCES