How Misuse of Solar Resource Datasets is Reducing Solar Industry Profits

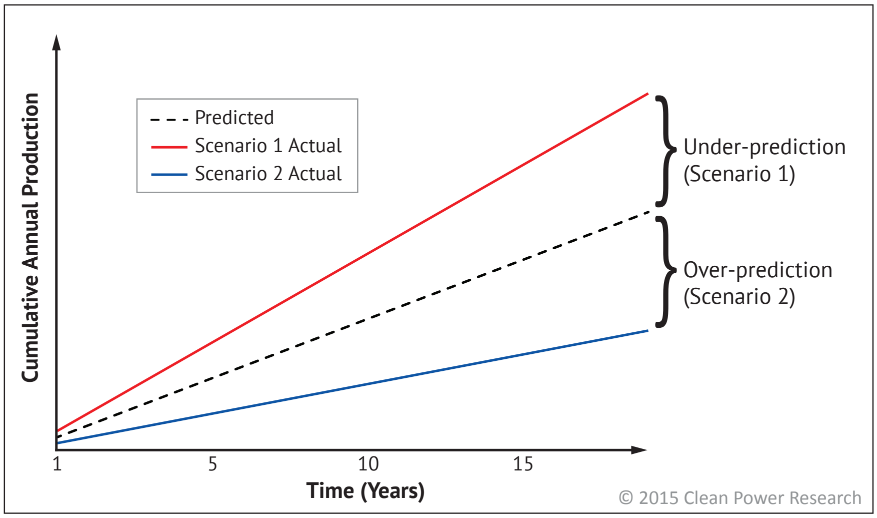
Choose the right solar dataset to maximize your returns

Adam Kankiewicz, Research Scientist
Gavin Novotny, Technical Product Manager
Clean Power Research

Clean Power Research®

Executive Summary

Accurately predicting solar photovoltaic (PV) system performance is critical for all solar industry participants. Third-party owners, system installers and independent engineers rely on accurate performance evaluations for quoting, performance guarantees and to assess financing potential. For these applications, small margins of error in performance modeling can have significant financial impacts, as suggested in Figure 1.



Scenario 2 represents a project in which projected return on investment is not achieved over the life of the project. **Scenario 1** represents a system producing more than projected. While at first glance under-prediction may seem positive, it can affect upfront decisions to proceed with a project even though it would be profitable in Scenario 1. A customer may decide against installing solar, or an investor may avoid involvement based on lower-than-actual performance estimates.

Figure 1: Cumulative Margin of Error vs. Solar Project Lifetime

Accurate performance modeling hinges on accurate solar resource data. NREL's Typical Meteorological Year data (previously TMY2, now TMY3) has long been considered the industry standard for solar resource data. While TMY3 data may be adequate for lead generation and initial estimates, it falls short of the precision required for lease and PPA quoting, performance guarantees and site evaluations.

This whitepaper demonstrates how the use of TMY3 data for applications requiring precise PV performance estimates creates avoidable performance risk, and jeopardizes project and portfolio profitability. Case study examples illustrate how alternatives such as SolarAnywhere® Typical Global Horizontal Irradiance [GHI] Year (TGY) satellite-based irradiance data can provide the precision and consistency required to support profitable, low-risk projects.

Table of Contents

An Introduction to Solar Resource Datasets 1

Dataset Fundamentals 3

The Path to Solar Energy: A History of TMY3. 4

Measurements with Satellite Data: A History of SolarAnywhere 5

Choosing a Solar Resource Dataset: Key Considerations 8

Consistent Data 8

Current Data 14

High Spatial Granularity and Continuity. 15

Maximize Returns by Choosing the Right Dataset 19

How to Get Started Using SolarAnywhere TGY 20

Highlights

Steps for Constructing Representative Data 4

Evolution of TMY Observations 6

The Relationship between TMY3 and the National Solar Radiation Database 7

Distinction Between TMY2 and TMY3 Datasets 8

The Nearest Neighbor Approach 12

Which Dataset Should I Use? 20

An Introduction to Solar Resource Datasets

The solar industry—including distributed (residential and commercial) and utility-scale developers—relies on a number of estimating tools to make decisions about investing in solar at specific sites. These tools are vitally important to the industry, and the output is used to determine everything from the design and cost of a rooftop system, to the financing terms for utility-scale systems. Accurate prediction of a PV system's performance over its lifetime can mean the difference between a profitable project and losses ranging from thousands (for small- to medium-scale distributed systems) to millions of dollars (for utility-scale systems).

The most popular solar estimation tools, including PVWatts®, System Advisor Model (SAM), PVsyst, Clean Power Estimator® and PowerBill® will return inaccurate results if the solar resource (i.e., irradiance) dataset used does not accurately represent the solar resource for the specific site being analyzed.

Figure 2 illustrates the inputs that go into a typical PV system performance model. Most inputs are site-specific and easily measurable, such as PV system specifications and array orientation. However, obtaining reliable solar resource information for a specific site can

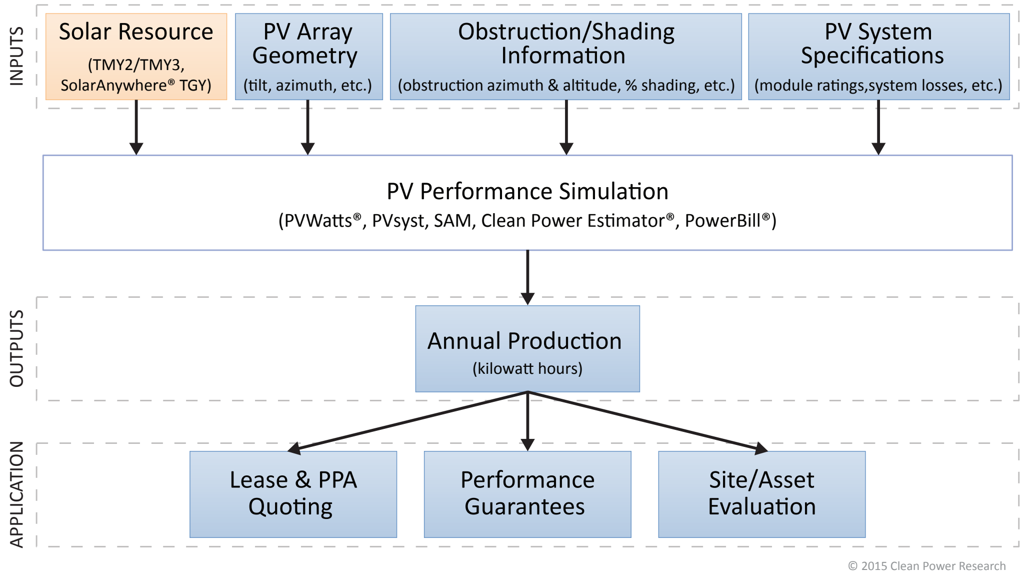


Figure 2: Modeling PV System Performance

be problematic. Accurate modeling requires a long-term history of a site's solar resource (ideally greater than 15 years). This makes ground-based measurement using on-site equipment time- and cost-prohibitive.

For this reason, the solar industry has turned to the use of modeled solar resource datasets such as TMY3, produced by the National Renewable Energy Laboratory (NREL), and SolarAnywhere Typical GHI Year (TGY), produced by Clean Power Research.

Not all solar resource datasets are created equal, however. When precise, high-accuracy performance modeling is needed to support quoting, performance guarantees and to assess financing potential, third-party owners, system installers and independent engineers require the highest-quality solar resource data available to accurately predict system performance and reduce project risk.

High-accuracy solar irradiance datasets have the following characteristics (these will be discussed in detail in the following section):

1. **Consistent data** - Deriving a solar resource dataset based purely on solar irradiance (GHI), and on a consistent, single-source of modeled irradiance data ensures that the most representative monthly irradiance data is chosen to construct a typical year dataset.

2. **Current data** - Long-term datasets that include solar resource data from the most recent year reflect regional weather trends such as air quality and drought.

3. **High spatial resolution and continuity** - To account for regional and micro-climate variability, high spatial resolution is critical for project sites that do not coincide with the location of the data site.

This paper explains the differences between NREL TMY3 and SolarAnywhere TGY datasets, and highlights the positive impact that using the correct solar resource data can have on quoting, performance guarantees and site evaluation—all applications that require precise performance modeling.

Dataset Fundamentals

There are many factors involved in the creation of typical year datasets, and understanding how each impacts the resulting typical year file is important to determining which dataset is appropriate for a particular application. This section explains how typical year files are constructed, and provides a brief history of the two typical year datasets discussed in this paper: NREL TMY3 and SolarAnywhere TGY.

Construction of a Typical Year File

Typical year solar resource datasets like NREL TMY3 and SolarAnywhere TGY provide one year of hourly irradiance data that can be used to simulate anticipated annual performance over the life of a PV system. Typical year datasets are constructed with hourly data that is most representative of the solar resource at a given location over a historic time period that spans 10 years or more.

To create a typical year dataset, monthly blocks of sequential, hourly data are pieced together to form an '8760' data file (8760 refers to the number of hours in a 365 day year). The historic month that most closely matches the long-term average for that month is selected for the dataset. In this way, the most representative data for each month is included in the typical year file. This approach also retains short-term variable weather patterns that would otherwise be muted if a simple average were used for each hour of the year.

For example, Figure 3 illustrates a typical year dataset construction. In this case, the dataset includes data from January 2000, February 2008, March 2006, and so on.

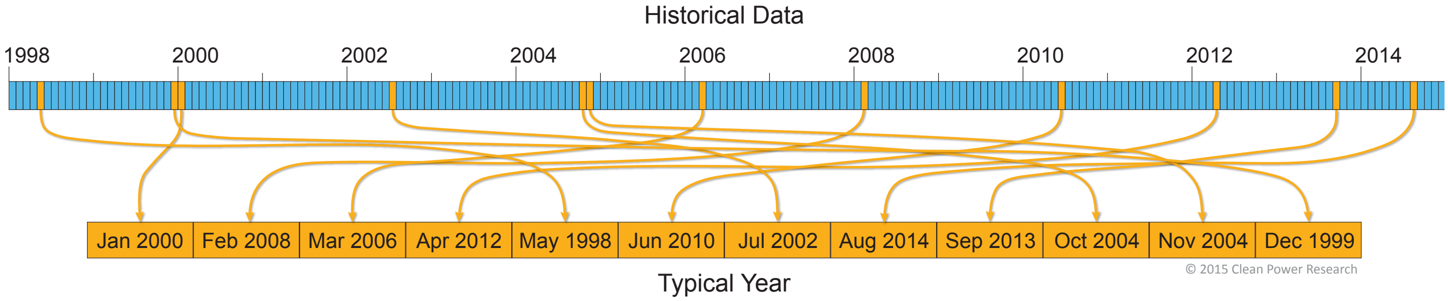


Figure 3: Example of Generic Typical Year Dataset Construction

Steps for Constructing Representative Solar Resource Files

1) For each month in the year (i.e., Jan., Feb., March, etc.), irradiance data is averaged across all the years included in the dataset (i.e., in the case of TMY3, 1991 through 2005 for most sites).

2) The historic month that most closely matches the long-term average for that month is selected for the dataset.

3) Hourly data from the most representative twelve selected months is combined to form a representative "8760" typical year file.

The Path to Solar Energy: A History of TMY3

The solar industry began using Typical Meteorological Year (TMY) datasets produced by the National Renewable Energy Lab (NREL) in the early 1980's. Originally developed to facilitate solar heating and cooling system simulations for buildings, NREL TMY data files (TMY, TMY2 and TMY3) are based on weather traits that influence overall building heating and cooling efficiency, including wind, temperature, humidity and solar resource.[1]

NREL TMY solar irradiance datasets were originally derived solely from NREL's Meteorological-Statistical (METSTAT) model.[2] This model incorporates both human and ground-based sensor observations, and estimates of cloud cover at single point locations.

Since the first NREL TMY dataset was released in 1977, NREL has released two additional versions: TMY2 in 1994, and TMY3 in 2007. As shown in Figure 4, data files for TMY3—the most common version used today—exist for 1,020 data sites in the continental United States, Alaska and Hawaii, and are derived from data through 2005.

[1] T. Freeman. 1979. "Evaluation of the "Typical Meteorological Years" for Solar Heating and Cooling System Studies Final Report." Altas Corporation.

[2] Maxwell, E.L., 1998. "METSTAT-The Solar Radiation Model Used in the Production of the National Solar Radiation Database (NSRDB)." Solar Energy, 62(4), 263-279

All TMY3 sites are not equal

Each TMY3 data site is classified based on the quantity and quality of historical data, as shown in Table 1. Class I sites offer the lowest uncertainty, being derived from 24 years of data with very little data missing. Class II sites are derived from just 12 years of data and have more interpolated data than Class I sites. Class III sites have the highest uncertainty due in large part to having large time periods with missing data. Bottom line: 78 percent of the TMY3 dataset is from high-uncertainty Class II and Class III sites sampled from just 12 years of historical data.

In addition to the uncertainty inherent in the three TMY3 classes, TMY3 datasets contain a mix of METSTAT and satellite-modeled irradiance data. This inconsistency in data sources, which will be discussed in more detail in the Consistent Data section below, introduces additional biases that contribute to uncertainty.

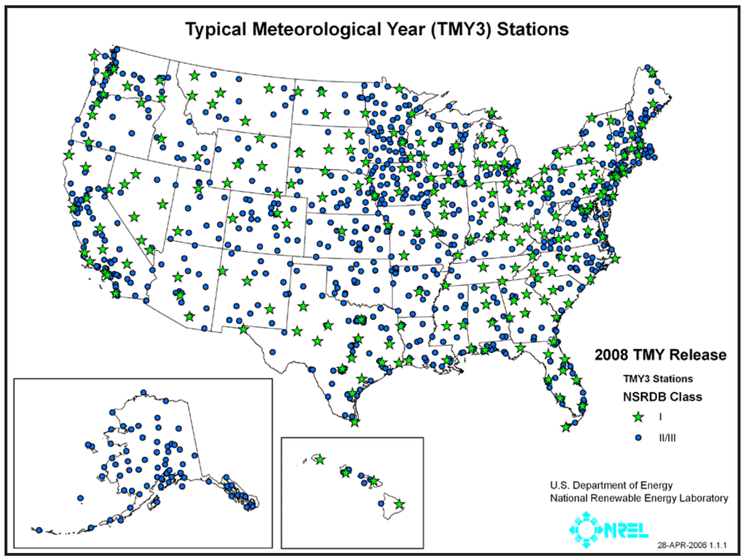


Figure 4: Geographic Dispersion of TMY3 Sites

	% of Sites	Uncertainty	# of Years Sampled
Class I	22%	Lowest	24
Class II	62%	Higher	12
Class III	16%	Highest	12

Table 1: Not All TMY3 Data Is Created Equal

Measurements with Satellite Data: A History of SolarAnywhere

With the availability of reliable geostationary satellite imagery along with improved computing resources in the early 2000's, a new method of modeling solar irradiance was developed by Dr. Richard Perez of the University at Albany (State University of New York).

Evolution of TMY Observations

Historically, humans reported hourly cloud cover observations at TMY3 sites. Through the 1990's and early 2000's, these human observations were phased out in favor of National Weather Service Automated Surface Observing System (ASOS) observations of cloud cover. With this transition, there was a notable decline in the accuracy of METSTAT data.

Figure 5, which shows METSTAT and Perez satellite model data for Fresno, Calif., illustrates the drastic fall in METSTAT GHI values in the early 2000's–even below volcano-influenced periods that resulted from eruptions of El Chichón (1982-1984) and Mount Pinatubo (1991-1994) that led to significant reductions of surface irradiance.[3] This is evidence of the negative effect that switching to ASOS equipment for cloud measurements had on the accuracy of METSTAT data. This decline in METSTAT model accuracy was a prime motivator for incorporating Perez satellite model data (represented by the orange line in the figure below) from 1998-2005 into the TMY3 dataset.[4]

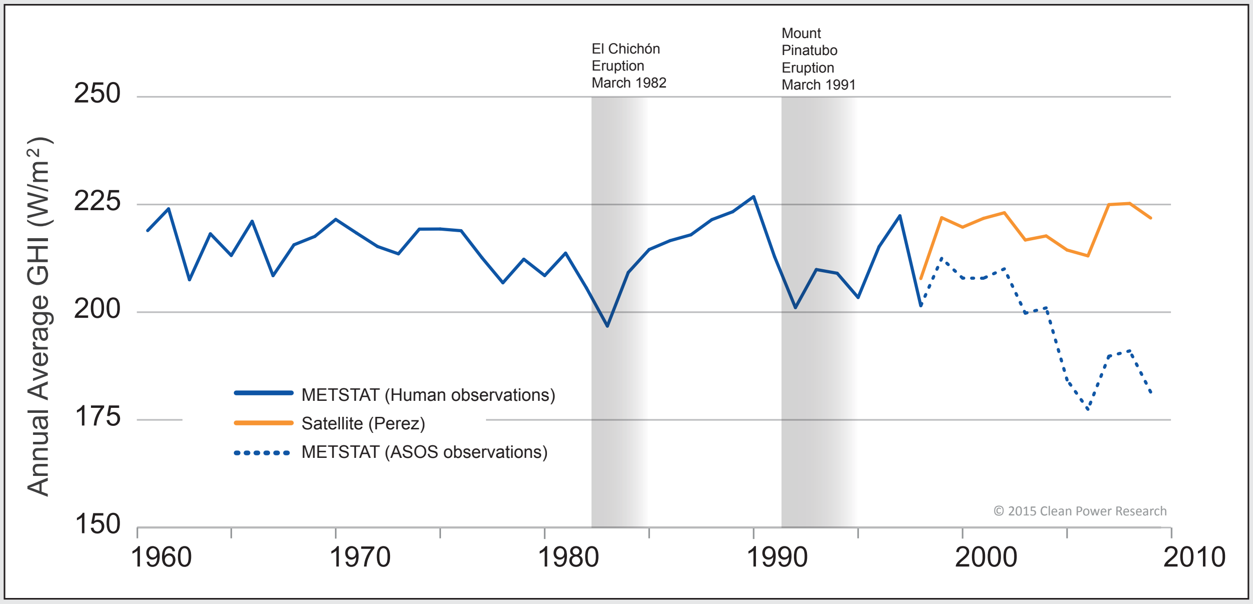


Figure 5: METSTAT and Perez satellite model annual average GHI for the Fresno Yosemite International Airport TMY3 location

[3] S. Wilcox and W. Marion. 2008. "Users Manual for TMY3 Data Sets." Page 18. National Renewable Energy Laboratory.

[4] S. Wilcox, et. al. 2007. "Completing Production of the Updated National Solar Radiation Database for the United States." National Renewable Energy Laboratory, etc.

Dubbed the “Perez satellite model,” this solar resource model offers full coverage across the United States, and produces temporally- and spatially-consistent data.

Today, SolarAnywhere time-series irradiance data based on the most current Perez satellite models are available from Clean Power Research under an exclusive partnership with the University at Albany. To generate the irradiance measurements that today comprise the product known as SolarAnywhere Data, hourly satellite images are processed to provide commercially available solar resource data that is highly accurate according to independent studies from organizations such as the University of California, San Diego.[5]

The Relationship between TMY3 and the National Solar Radiation Database

TMY3 files were distilled from meteorological data contained within the NREL National Solar Radiation Database (NSRDB). Contrary to common perception, less than two percent of the underlying solar irradiance data in the NSRDB comes from ground-based measurements.

The NSRDB is primarily composed of solar irradiance data derived from the METSTAT model. Satellite-based irradiance measurements from the Perez satellite model were later used to replace METSTAT irradiance data in TMY3 files between 1998 and 2005.

SolarAnywhere TGY is constructed from more than 16 years of up-to-date SolarAnywhere Data that includes data as current as the most recent complete year. SolarAnywhere TGY datasets are unique in that they have been designed specifically with solar PV modeling in mind. They employ a TGY methodology that ensures that consideration of other meteorological factors such as wind and temperature do not result in the selection of months with atypical irradiance patterns. Consistency in SolarAnywhere datasets is further preserved by the use of a single model (i.e., Perez) in the modeling of solar irradiance data.

Since its inception, SolarAnywhere coverage has expanded to include the continental United States, Hawaii, Mexico and parts of Canada. Today, SolarAnywhere provides historical data as well as typical year data.

With spatial resolutions as high as 1 km, SolarAnywhere is not site restricted, which enables an accurate representation of available irradiance at specific project sites that may not coincide with TMY3 stations. This is particularly important for coastal, mountainous and island locations that often exhibit sharp spatial gradients in solar resource.

[5] Jamaly, Mohammad, and Kleissl, Jan. 2012. “Validation of SolarAnywhere Enhanced Resolution Irradiation in California.” Department of Mechanical and Aerospace Engineering, University of California, San Diego.

Choosing a Solar Resource Dataset: Key Considerations

When selecting a dataset for use in modeling PV output, there are three key considerations. Data should:

1. **Be consistent**
2. **Be current**
3. **Have high spatial resolution and continuity**

In the following sections, we'll examine how use of solar irradiance datasets with these characteristics leads to lower uncertainty in estimating future output of solar PV systems.

Distinction Between TMY2 and TMY3 Datasets

While TMY3 is the most current dataset of its kind from NREL, its predecessor, TMY2, is still used by some solar industry participants. The content of this paper focuses on TMY3; however, the concepts addressed in "Choosing a Solar Resource Dataset: Key Considerations" also apply to TMY2 data with the following distinctions:

- **Consistent data** - TMY2 files were distilled solely from METSTAT modeled data spanning the period from 1961–1990. Perez satellite data are only available after 1998.
- **Current data** - Because data from TMY2 is 25 to 55 years old, near term weather trends such as air quality and drought are not accounted for.
- **High spatial resolution and continuity** - The TMY2 dataset contains data from only 239 sites, which drastically reduces spatial resolution.

Consistent Data

Reliable typical year data is dependent on the selection of the most representative solar resource month, irrespective of other environmental factors. Ambient temperature and wind speed are factored in independently by the solar estimation tools highlighted in Figure 2 on page 1. NREL TMY data files were originally developed for analyzing the impacts of solar

[6] T. Freeman. 1979. "Evaluation of the "Typical Meteorological Years" for Solar Heating and Cooling System Studies Final Report." Altas Corporation.

heating and cooling on a building's performance,[6] and therefore include the influence of multiple weather traits such as wind, temperature, humidity and solar resource, as shown in Table 2. Of the ten indices used to generate TMY3 files, only two are components of solar irradiance, and these are given only a 50 percent weighting factor. As a result, weather factors such as humidity, temperature and wind can directly influence the typical month selection process, leading to atypical solar months being selected for TMY3 files.

Index	TMY Weighting	TGY Weighting
Max Dry Bulb Temp	5%	
Min Dry Bulb Temp	5%	
Mean Dry Bulb Temp	10%	
Max Dew Point Temp	5%	
Min Dew Point Temp	5%	
Mean Dew Point Temp	10%	
Max Wind Velocity	5%	
Mean Wind Velocity	5%	
Mean Global Horizontal Irradiance (GHI)	25%	100%
Mean Direct Normal Irradiance (DNI)	25%	

Table 2: Weighting of Meteorological Parameters Used in Typical Year Processing

Similarly, the use of more than one source of modeled irradiance data can lead to increased uncertainty in "8760" files. TMY3 sites contain irradiance data modeled by both the METSTAT and the Perez satellite models.

Recognizing the accuracy and quality offered by Perez satellite modeled irradiance data, NREL incorporated satellite data to enhance the overall quality of the TMY3 dataset. To generate TMY3 datasets, NREL first selected the typical months based entirely on METSTAT irradiance data. Once the TMY3 distillation process was completed, NREL replaced METSTAT-based irradiance data with Perez satellite-based irradiance data for all TMY3 typical months chosen between 1998 and 2005. As a result, nearly all of the TMY3 sites contain a mix of both METSTAT and satellite-based irradiance data, as illustrated by the examples in Figure 6.

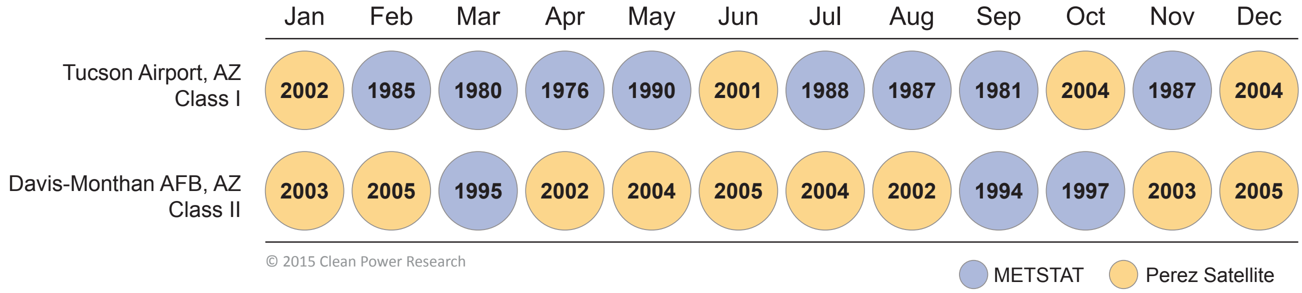


Figure 6: Example TMY3 Irradiance Data Sources (METSTAT and Perez Satellite)

After this replacement, however, TMY3 data was not re-processed to ensure that the most representative months of data were included in the dataset. The intent of incorporating data from the Perez satellite model was to improve the accuracy of TMY3 files by replacing METSTAT data that had demonstrated declining accuracy (see "Evolution of TMY3 Observations" on page 6). However, supplanting METSTAT data in this manner without re-executing the typical year selection process leads to inclusion of irradiance data that is not representative of the typical month. This is true regardless of the class.

In contrast to TMY3 data, SolarAnywhere TGY datasets are based solely on SolarAnywhere satellite-derived data. Exclusive use of the most current version of the Perez satellite model irradiance data is central to the accuracy of SolarAnywhere TGY irradiance data. This consistency ensures that the most representative monthly irradiance data are chosen to construct a typical year dataset, thereby increasing accuracy in PV system performance modeling.

The following case study from the Minneapolis / St. Paul region of Minnesota demonstrates how the construction of TMY3 files can lead to the inclusion of atypical irradiance data, and an inaccurate PV production estimate as a result.

Minneapolis/St. Paul Case Study: An example of poor TMY3 monthly choices

This case study highlights how the high variance of TMY3 data within the Minneapolis/St. Paul region can lead to significant variations in estimated production. Figure 7 and Figure 8 provide a comparison of the annual average GHI predicted by TMY3 and SolarAnywhere TGY. In Figure 8, SolarAnywhere TGY 10 km gridded data form the background of the image (the orange to red tiles), and TMY3 data from the five regional TMY3 data sites are superimposed for comparison (filled circles). The color within each of the circles corresponds to annual average GHI as shown in the scale below the image.

As can be observed in Figure 7, the regional GHI variance among the five TMY3 stations is approximately 23 W/m^2. This is substantially larger than the regional GHI variance of

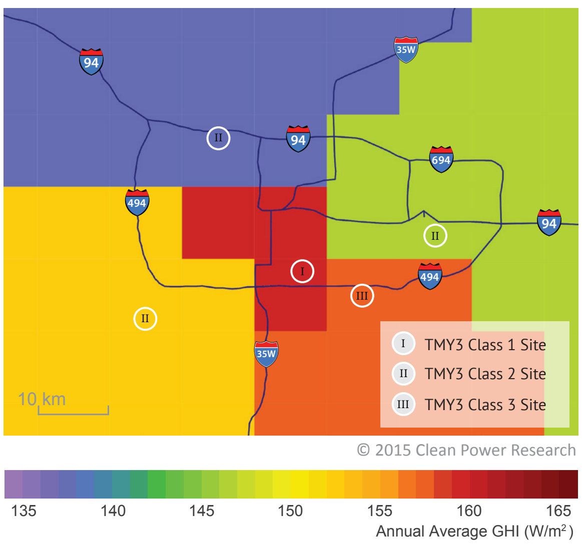


Figure 7: TMY3 Nearest Neighbor Gridded 10 km Annual Average GHI in the **Minneapolis/St. Paul Region**

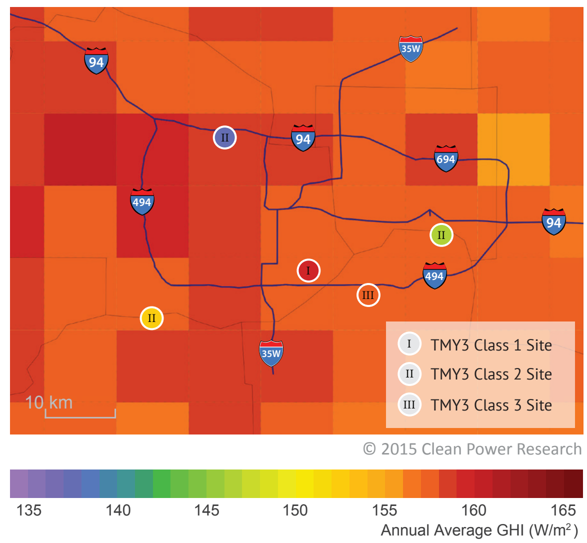


Figure 8: SolarAnywhere TGY Gridded 10 km Annual Average GHI in the **Minneapolis/St. Paul Region**

The Nearest Neighbor Approach

When using TMY3 data, the nearest neighbor method of selecting which data site to use for a particular project is typically applied. This methodology simply extrapolates TMY3 data from the nearest data site to the PV project site, as shown in Figure 7. Figures 8 and 9 illustrate how TMY3 annual average GHI compares to SA TGY annual average GHI.

NREL advises that a distance of 25 km (which can decrease according to local topographical conditions) is the maximum distance for which a TMY3 station should be used to evaluate a project site.[7] When using the nearest neighbor approach, the TMY3 station may be outside of the 25 km recommended maximum.

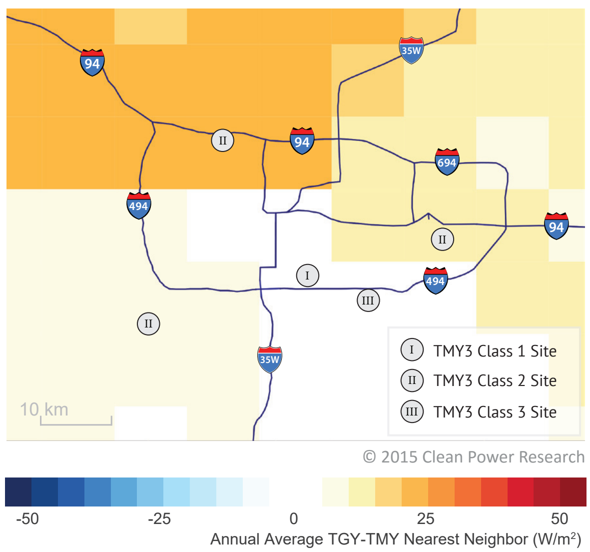
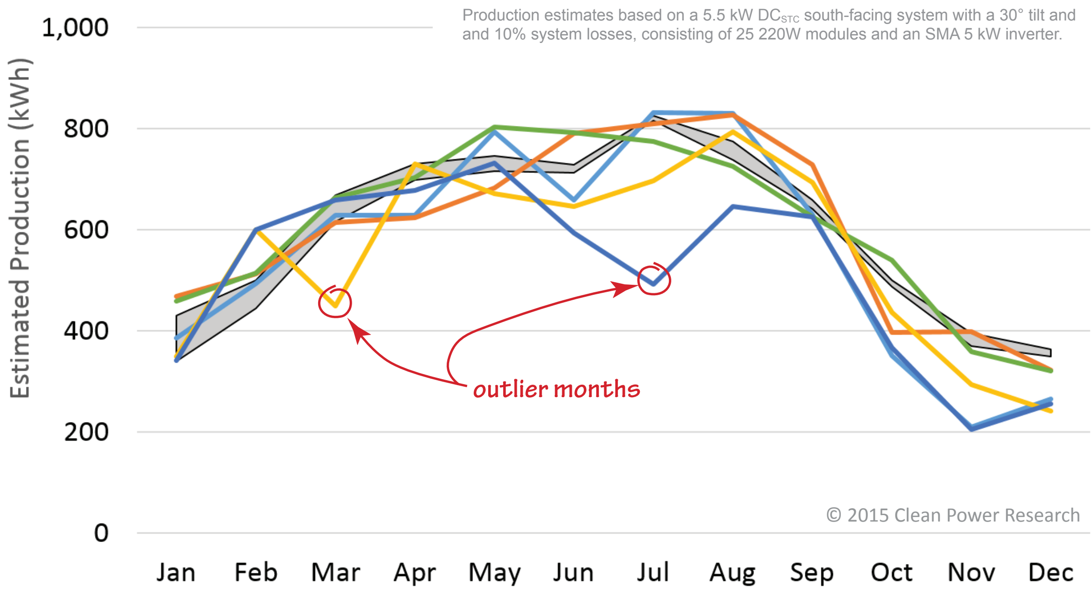


Figure 9: Average Annual Difference between SolarAnywhere TGY and TMY Nearest Neighbor in the **Minneapolis/St. Paul Region**

approximately 5 W/m² seen in the SolarAnywhere TGY gridded dataset shown in Figure 8. Figure 9 combines the data displayed in Figure 7 and Figure 8 to illustrate the absolute difference between SolarAnywhere TGY and TMY3 for each 10 km grid, with light areas representing little to no difference and darker areas representing higher difference.

[7] J. Dean, A. Kandt, K. Burman, L. Lisell and C. Helm. 2009. "Analysis of Web-Based Solar Photovoltaic Mapping Tools." National Renewable Energy Laboratory.

Figure 10 illustrates how the variance between the TMY3 datasets impacts production estimates. In this example, TMY3 production estimates vary by as much as 15 percent from the baseline TMY3 Class I site at the Minneapolis/St. Paul International Airport.



	Class	Annual Production	Variance
TMY3 Minneapolis / St. Paul Int'l Airport	Class I	7,282 kWh	Base
TMY3 South St. Paul Muni	Class III	7,180 kWh	-1.4%
TMY3 Flying Cloud	Class II	7,044 kWh	-3.3%
TMY3 St. Paul Downtown Airport	Class II	6,606 kWh	-9.3%
TMY3 Minneapolis / Crystal	Class II	6,198 kWh	-14.9%
SolarAnywhere TGY Range for the 5 TMY3 locations — Minimum		7,044 kWh	Base
SolarAnywhere TGY Range for the 5 TMY3 locations — Maximum		7,235 kWh	2.71%

Figure 10: Monthly Average GHI of TMY3 and SolarAnywhere TGY Gridded 10 km Data in the **Minneapolis / St. Paul Region**

As can be seen in the chart, production estimates using TMY3 data from three of the TMY3 stations align reasonably well with SolarAnywhere TGY production estimates throughout the year. However, outlier months can clearly be seen for the St. Paul Downtown Airport

(March) and the Minneapolis/Crystal (July) TMY3 stations. This is illustrative of the effect of basing Class II and III sites on 12 years of data, which leads to a greater likelihood that the TMY3 weighting process will result in the selection of outlier months. This effect is less pronounced for Class I stations that have 24 years of data.

Current Data

Both TMY3 and SolarAnywhere TGY are derived from long-term datasets containing 10-plus years of data. As shown in Figure 11, TMY3 files for Class I sites are derived from 24 years of historical data modeled between 1976 and 2005, and most TMY3 files for Class II/III TMY3 sites contain 12 years of data modeled between 1991 and 2005 (excluding atypical months resulting from the volcanic eruptions of El Chichón in 1982-1984, and Mount Pinatubo in 1991-1994). SolarAnywhere TGY data files are based on 16 years of irradiance data from 1998 through the most recent complete calendar year (e.g., 2014).

A key difference between TMY3 and SolarAnywhere TGY is the consideration of near-term data. While the latest irradiance data included in TMY3 files were modeled through 2005, SolarAnywhere TGY is based on data as current as the most recent complete year to account for near-term climate variations.

Year-over-year, the advantage of including near-term data becomes more pronounced, as SolarAnywhere is able to adapt to regional weather trends and rapidly changing environmental conditions such as air quality.

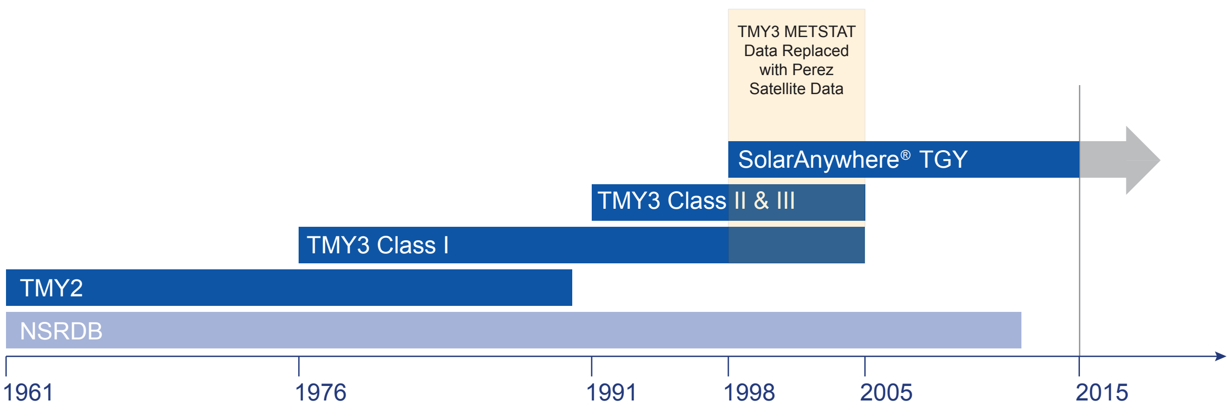


Figure 11: Typical Year Dataset Timeline

High Spatial Granularity and Continuity

Accounting for regional climate variability is critical to accurately predicting PV system performance. This is true not only for islands and coastal regions, but for any inland region that exhibits similar microclimatic characteristics, such as within the mountainous California Coast Ranges and the Front Range of Colorado. In order to capture micro-climate characteristics, data with high-spatial resolution is required.

The site-restricted nature of TMY3 data is a severe limitation. As noted by NREL, TMY3 irradiance data should not be used for project sites more than 25 km removed from the nearest TMY3 station.[8] Adhering to this recommendation, TMY3 irradiance data is only valid for approximately 24 percent of the United States (excluding Alaska) when taking into account all TMY3 sites. When considering only the most accurate TMY3 data at Class I sites, only 6 percent of the United States is covered. Figures 12 and 13 illustrate SolarAnywhere TGY 10 km coverage as compared to TMY3 coverage.

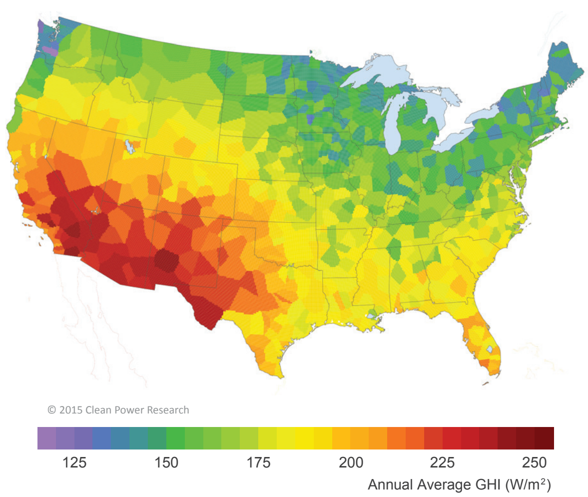


Figure 12: TMY3 Annual Average GHI

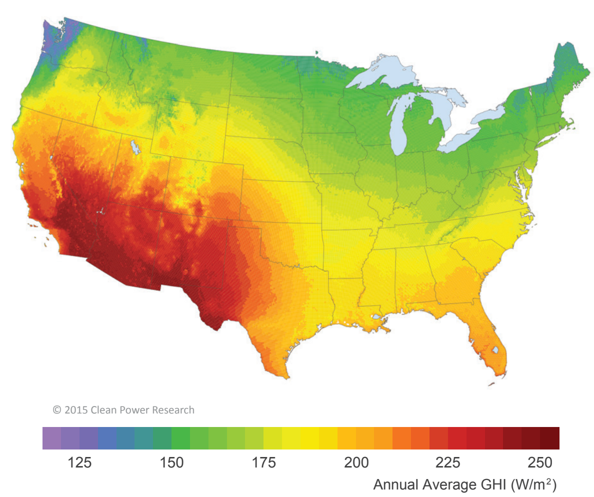


Figure 13: SolarAnywhere TGY Annual Average GHI

With SolarAnywhere TGY gridded data available in resolutions as high as 1 km, locational deficiencies induced by low density TMY3 data are eliminated. The case studies that follow for San Diego, Calif., and Oahu, Hawaii, will demonstrate how large inaccuracies can result from extrapolating irradiance data to nearby project locations.

[8] J. Dean, A. Kandt, K. Burman, L. Lisell and C. Helm. 2009. "Analysis of Web-Based Solar Photovoltaic Mapping Tools." National Renewable Energy Laboratory.

San Diego Case Study: Western U.S. Coastal Gradients

In the San Diego region, the solar resource is lowest near the coast and improves toward inland regions. Figures 14 and 15 show that six TMY3 stations are congregated near the coast, with the next available station located approximately 90 miles inland in the Imperial Valley region of California.

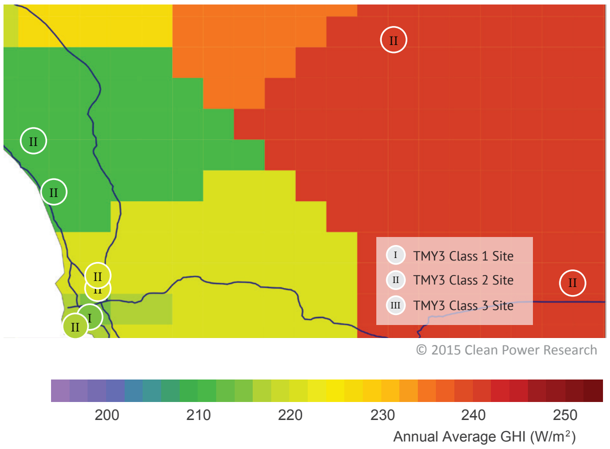


Figure 14: TMY3 Nearest Neighbor Gridded 10 km Annual Average GHI in **San Diego, Calif.**

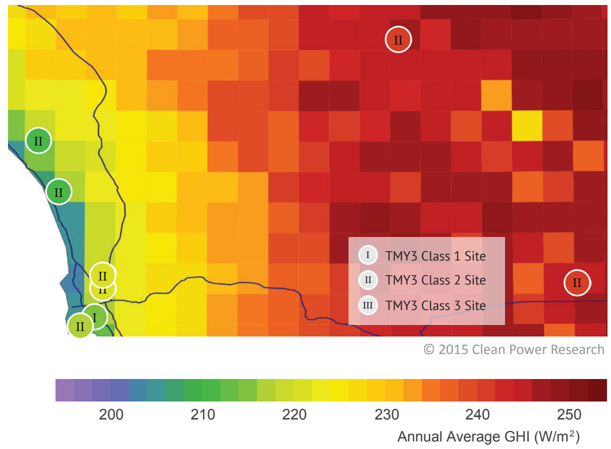


Figure 15: SolarAnywhere TGY Gridded 10 km Annual Average GHI in **San Diego, Calif.**

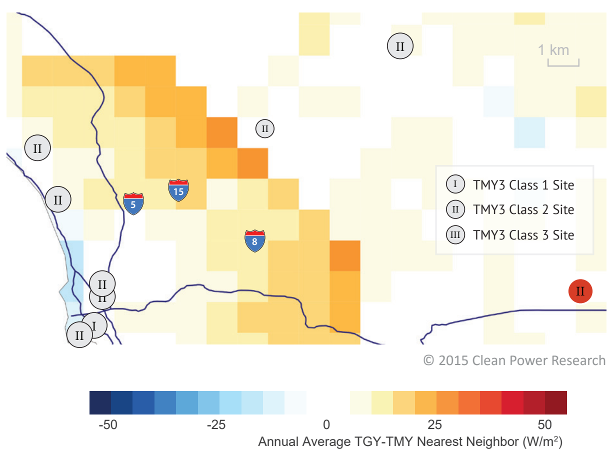


Figure 16: Average Annual Difference between SolarAnywhere TGY and TMY Nearest Neighbor in **San Diego, Calif.**

With this uneven TMY3 coverage distribution, it's very difficult to accurately evaluate the solar resource for sites located between the coastal and inland region where a sharp solar resource gradient exists. The inability of TMY3 data to capture these gradients is clearly demonstrated in Figure 16, which displays the difference between between SolarAnywhere in Figure 15 and TMY3 in Figure 14.

Both 10 km and 1 km SolarAnywhere TGY data sources provide accurate coverage between the lower coastal solar resource and the much higher inland solar resource. Due to the sharp gradients in coastal California topography, use of 1 km SolarAnywhere TGY data shown in Figure 17 provides the most reliable solar resource assessments.

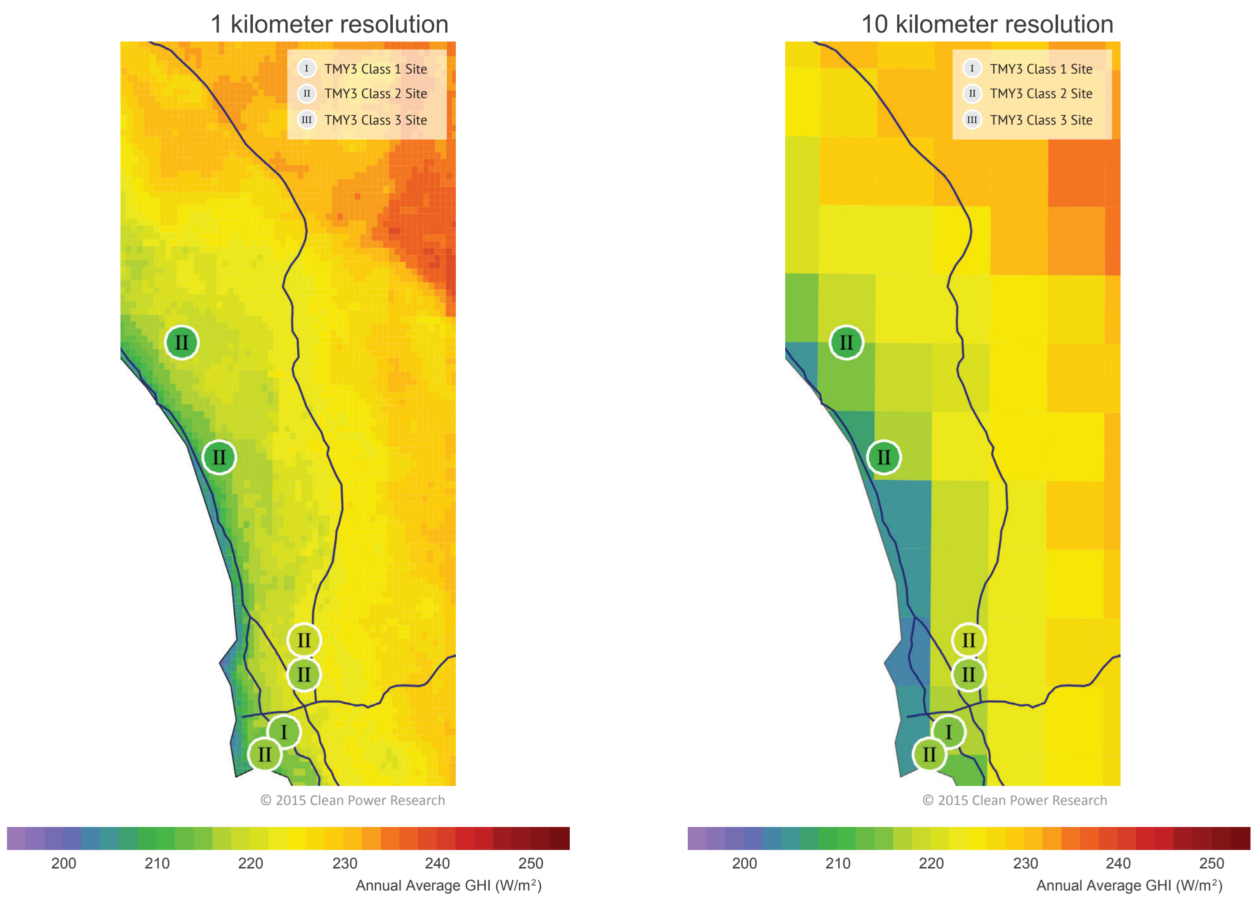


Figure 17: Comparison of Gridded 1 & 10 km SolarAnywhere TGY and TMY3 Annual Average GHI in **San Diego, Calif.**

Oahu, Hawaii Case Study: Island Climates

Similar to San Diego, the lack of TMY3 stations throughout the Hawaiian Islands injects high uncertainty when it comes to evaluating solar energy potential. The 1 km SolarAnywhere TGY data shown in Figure 19 provides evidence of the impact of topography on the solar resource of Oahu. The TMY3 representation of Oahu's solar resource in Figure 18 provides a stark contrast to the variations illustrated by SolarAnywhere TGY data.

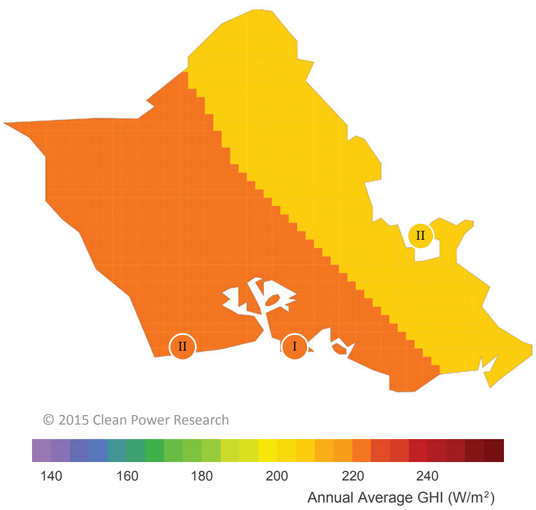
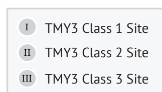


Figure 18: TMY3 Nearest Neighbor Gridded 1 km Annual Average GHI in **Oahu, Hawaii**



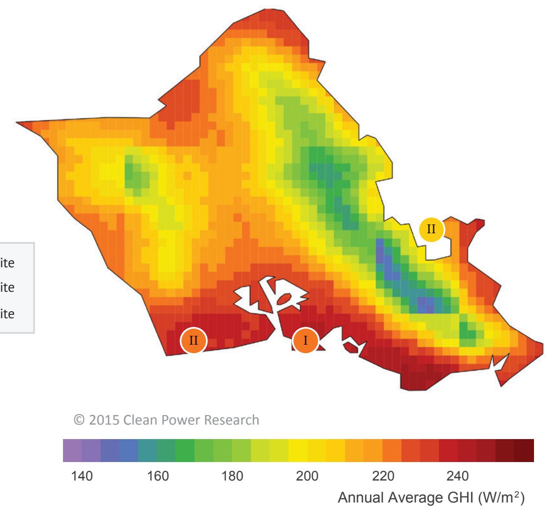


Figure 19: SolarAnywhere TGY Gridded 1 km Annual Average GHI in **Oahu, Hawaii**

Not only is TMY3 data unable to capture the solar resource gradients observed throughout the island's interior, coastal gradients in areas immediately adjacent to the TMY3 stations are misrepresented as well. Figure 20 shows that using TMY3 data on Oahu can drastically misrepresent solar resource potential, even for project sites very near to TMY3 stations.

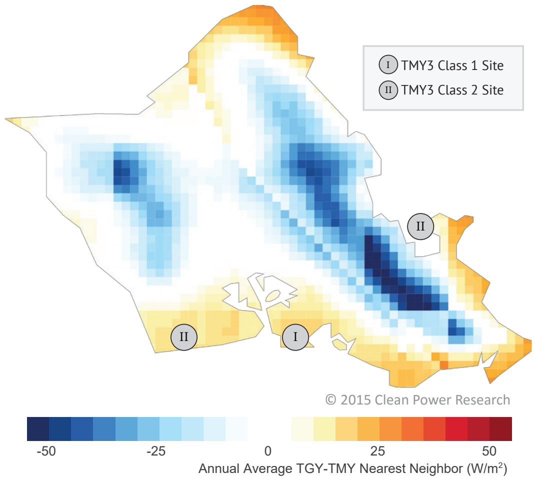


Figure 20: Average Annual Difference between SolarAnywhere TGY and TMY Nearest Neighbor in **Oahu, Hawaii**

Maximize Returns by Choosing the Right Dataset

TMY3 data has been invaluable in supporting the development of solar throughout the United States, and it will continue to be a valuable low-cost tool for approximating solar resource potential. However, as the case studies in this paper illustrate, TMY3 presents potentially severe limitations around accuracy and geographic specificity for applications that require precise solar resource estimation. These limitations can propagate substantial negative effects to a project's risk assessment and return on investment.

For those banking their bottom line profits on actual solar production matching expected production, consistent, current solar resource data is critical. Satellite-derived solar resource data offers these characteristics, along with the high spatial granularity needed to estimate production in microclimates.

Which Dataset Should I Use?

Use Cases	TMY3	SolarAnywhere TGY
Lead Generation / Initial Estimates	✓	✓
Lease and PPA Quoting		✓
Setting Performance Guarantees		✓
Utility-scale Site/Asset Evaluation		✓

How to Get Started Using SolarAnywhere TGY

Clean Power Research's APIs offer exclusive programmatic access to PV production simulations using SolarAnywhere TGY datasets. This scalable, enterprise-grade software-as-a-service (SaaS) also includes other calculations critical to solar quoting, including electric bill and incentive savings along with financial modeling. For more information on the Clean Power Research APIs, please visit http://www.cleanpower.com/solutions/software/ or contact info@cleanpower.com to get started.

Utility-scale developers and independent engineers looking for site-by-site SolarAnywhere TGY data downloads, please visit www.solaranywhere.com for more information on this service or contact info@cleanpower.com to get started with a subscription.

Clean Power Research®
Powering Intelligent Energy Decisions

10604 NE 38th Place Ste 100 | Kirkland, WA 98033
+1 425.242.4174 | www.cleanpower.com