Comparison of Typical Meteorological Year and On-site Measurements for Solar PV Site Selection

Diane Palmer* and Tom Betts
Centre for Renewable Energy Systems Technology, Loughborough University, LE11 3TU, UK
* Corresponding Author d.palmer@lboro.ac.uk

Abstract The standard approach when simulating photovoltaic output is to employ Typical Meteorological Year (TMY) data sets. This paper uses statistical analyses to investigate several TMYs. The objectives are: (1) to ascertain which TMY generation method most accurately reproduces solar conditions in the UK; (2) to discover the minimum time resolution at which TMYs delver reliable results; and (3) to determine if there are any circumstances in which TMYs can replace long-term time series.

Introduction For solar project planning, developers and financiers usually require long-term solar radiation data of high accuracy. This is often considered to be the most important site selection criterion. Investors also view possible errors in solar resource data as one of the likeliest causes of losses relative to expected return on investment. The solar data upon which project plans are based should comprise a time series of at least ten years and have known uncertainty in order to be judged "bankable" [1]. To secure preferable financing, analysis of on-site solar insolation needs to go beyond studying the long-term average irradiance. Daily, monthly, annual and project lifetime variability is required to calculate risk. Business banks usually specify quarterly operating reports. Production agreements might include time-of-day or seasonal price differences. Financial viability is also influenced by historical extremes, as well as mean, minimum and maximum seasonal and annual projected generation. Knowledge of all these statistics additionally facilitates plant design and operational planning.

TMY data sets are the photovoltaic (PV) industry standard for the solar resource. A TMY is a set of meteorological parameter(s) with representative values for every hour in a single year for a given geographical place. The one year of data values is selected as characteristic of the location from a long-term time series of at least 10 years. There are several ways of making this selection, reviewed by [2], and thus a choice of TMYs available for use.

The advantages of TMYs include reduced data storage requirements, speed of data access and lower cost than purchasing several years of raw measurement data. TMYs created from interpolated or satellite data are useful where there are no ground measurements available at the exact site of interest. They are compatible with industry-standard energy simulation software such as PVSyst [3] and can be used to produce electrical generation results quickly.

Typical Meteorological Year data is normally used for initial evaluations only. It is not judged adequate for further financial assessment of large-scale solar installations. The conventional wisdom is that TMY datasets do not sufficiently detail year-by-year weather variations which may cause a solar plant to fall short of long-term yield expectations. However, TMY data is much less expensive than purchasing long-term time series data from meteorological organisations or commercial satellite-data suppliers. Taking the UK as an example, this research compares global horizontal irradiance data from a free simple TMY (detailed below), a widely used commercial TMY (Meteonorm), UK Met. Office weather stations and satellite sources. The aim is to discover what influence the choice of dataset has and whether the difference in accuracy is sufficient to justify the extra cost of long-term datasets.

Statistical Analysis Several different types of TMY are compared to long-term global horizontal irradiation (GHI) measurements (10 years 2008-2017) from Sutton Bonington weather station, obtained from UK Met. Office MIDAS data [4]. This central location exhibits normal weather patterns for the UK. The yearly, monthly and daily sums of GHI in TMYs are compared with respective statistics in the long-term GHI data from the ground measurements. Hourly means and standard deviation are also calculated. Root Mean Square Error (RMSE) and normalised RMSE (nRMSE% normalised by mean) are used as a measure of how close the GHI in individual TMY data sets are to the long-term average measured data. Frequency distribution histograms for each specific data set are drawn up. Lastly, irradiation data from the separate sources is used to calculate hourly, daily, monthly and annual PV system output for a sample small commercial rooftop installation.

PVSyst is utilised to generate electrical yield for a hypothetical barn roof installation in Sutton Bonington, area 125 m^2, tilt 25° and azimuth 20° west, 3 × 4.2 kW inverters, 12 strings of 11 × 110W modules in series.

Data The TMYs employed in this analysis are as follows:

a) A data set constituted by taking the average of every hour for each of the 10 years 2008-2017 in the Sutton Bonington GHI weather station data. This approach, from now on termed “TenYrHrAvg”, is not strictly a TMY as such, and is included for purposes of comparison only.

b) A simple TMY, subsequently called “SimpleTMY” calculated from the 10-year Sutton Bonington GHI measurements thus:

i. For each month in the year (i.e., January, February, March, etc.), the ten-year monthly mean of irradiance data is obtained for the period 2008-2017.

ii. The historic month that most closely matches the ten-year mean for that month is identified i.e. the most “typical” month.

iii. The hourly data values from the twelve typical months are combined to create the typical year file. This comprises January 2011, February 2015, March 2010, April 2014, May 2015, June 2014, July 2015, August 2009, September 2013, October 2010, November 2012 and December 2010.

c) Another simple TMY calculated as described in (b) above but taking MIDAS data interpolated to a point 50 km from Sutton Bonington as input, rather than the weather station data itself. This is called “KrigeTMY”.

d) A Meteonorm [5] file for Sutton Bonington. Meteonorm is based on spatial interpolation of monthly weather station averages of GHI, supplemented by Meteosat (2–3 km resolution) images where ground measurements are sparse. (34% satellite data was used in the case of this site.) Next, hourly values are stochastically generated from the monthly averages. Although again, not a “true” TMY, a single year of data is produced from the values of several.

e) A NASA-SSE [7] TMY for Sutton Bonington. NASA-SSE (Surface Meteorology and Solar Energy programme) are worldwide monthly data, average of 1983-2005 EUMETSAT satellite measurements, ≈ 111 km resolution. The TMY was generated using Meteonorm.

f) A PVGIS v.5 [6] TMY for Sutton Bonington. This takes as its inputs data from satellite-based solar radiation data (CM-SAF ≈ 5 km resolution) and reanalysis climate data (air temperature, relative humidity and wind speed). The ISO 15927-4 procedure is used to construct the TMY. That is, typical months are selected by comparing the distribution of each meteorological parameter in each month with the long-term distribution of that parameter and month (Finkelstein-Schafer statistic, FS). The FS method is judged better than the mean because it chooses months with less extreme daily values which are closer to the long term daily mean. The ISO-15927-4 method assigns an equal weight to air temperature, solar radiation, and relative humidity.

Results and Interpretation When compared to the long-term 2008-2017 GHI time series for Sutton Bonington (SB 2008-2017), the average hourly means, annual sums, average daily sums and average monthly sums of TenYrHrAvg, SimpleTMY and Meteonorm differ by at most 0.1%. This suggests that these 3 TMYs are suitable for pre-feasibility studies and for estimating overall results for the complete 25-year lifetime of a solar project. On the other hand, the remaining TMYs (KrigeTMY, NASA-SSE-TMY and PVGIS-TMY) differ from SB 2008-2017 by 6%, 3.5% and 3.5% respectively. In the cases of KrigeTMY and NASA-SSE-TMY, it is the low spatial resolution of the GHI data which is causing the inaccuracy, rather than the method of TMY generation. The two TMY generation methods have previously just given accurate results when used in SimpleTMY and Meteonorm. But the spatial resolution of both KrigeTMY and NASA-SSE-TMY is poor: KrigeTMY is based on data interpolated to a distance of 50 km and NASA-SSE-TMY GHI inputs are grid squares of over 100 × 100 km. For PVGIS-TMY, it is unlikely to be the underlying satellite-derived GHI data which is engendering the differences because this is similar to that used in Meteonorm. Therefore, the problem must lie in either the choice of data additional to GHI used in the TMY generation (air temperature, relative humidity and wind speed), or in the weighting of the meteorological parameters in the TMY creation.

When annual PV system output is calculated using PVSyst for the various TMYs, the results are analogous to those for annual GHI. Meteonorm annual yield for Sutton Bonington differs from that calculated for 2013 (an average year in terms of standard deviation and sum of GHI) by only 0.1%. Annual yields obtained from NASA-SSE-TMY and PVGIS-TMY differ by 1% and 4%.

In temperate climates such as the UK, weather conditions vary considerably from year to year.

In the ten-year period under investigation (2008-2017), the annual sum of GHI ranged from 896 MWh/m² in 2012 to 1046 MWh/m² in 2009 – a difference of 149 MWh/m² or 14%. Naturally, TMYs do not capture this variability. The annual sum of solar radiation in TMYs remains the same. The difference in total annual GHI between a TMY and the actual measured value at Sutton Bonington may be 8-18%, depending on which TMY and which year. This is similar to the natural year-on-year variation of ground-based measurements.

Turning to analysis on a monthly basis, most of the TMYs investigated follow the sinusoidal curve of SB 2008-2017, with monthly sums of GHI rising from January to June, and then falling to December (Fig. 1). The exceptions are NASA-SSE-TMY which underestimates in June and PVGIS-TMY which overestimates in March and April, underestimates in May, and overestimates again in June and July.

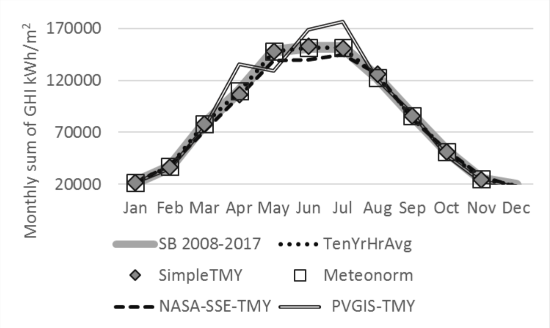

Figure 1. Monthly sum of GHI for the long-term time series and 5 TMYs

Monthly system outputs obtained from PVSyst exhibit a similar pattern. Hourly analysis reveals that PVGIS-TMY only demonstrates this two-peaked distribution, with monthly GHI rises in April and July, between 9 am and 6 pm (the most productive hours for PV). From 4 am to 8 am, it underestimates, and at 7 pm it overestimates. When the hourly sums are totalled for the year, PVGIS-TMY underestimates before noon and overestimates after noon (Fig. 2).

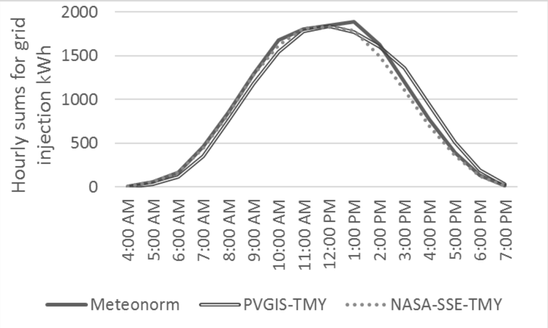

Figure 2: Hourly electrical yield for 3 TMYs

The emulation of long-term time series by TMYs was further studied by RMSE. All the TMYs were found to have an annual nRMSE of *c.* 1%, apart from KrigeTMY (2%). Given the differences between the TMYs, this statistic is not very informative, and further measures were investigated.

In terms of standard deviation, SimpleTMY most closely matches SB 2008-2017, differing from it by 1.5 kWh/m² GHI. This is followed by Meteonorm (Δ = 4.7 kWh/m²), KrigeTMY (Δ = 7.4 kWh/m²), PVGIS-TMY (Δ = -7.7 kWh/m²) and lastly TenYrHrAvg (Δ = 23.3 kWh/m²). It is well known that simple averaging of yearly data underestimates the amount of variability, which is why this technique is not generally adopted. Apart from TenYrHrAvg, none of the TMYs have greater differences to the long-term standard deviation than any of the ten years in the data from which it is constituted. Interestingly, the standard deviation of shorter term 5-year time series for Sutton Bonington for 2008-2012 and 2013-2018, differ from SB 2008-2017 by 1.8 kWh/m² GHI and -1.4 kWh/m² GHI only. This suggests that shorter term time series may be as good as TMYs.

To assess how well the hourly distribution of TMYs matches that of SB 2008-2017, irradiation frequency distribution charts were constructed (Fig.3).

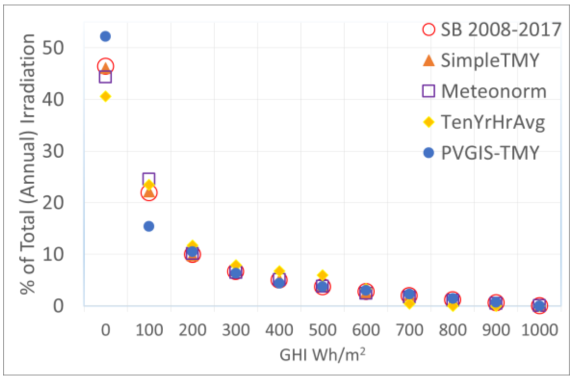

Figure 3: % of Year in irradiation bins (0-1000 Wh/m²) for each of 4 TMYs compared to long-term time series

It may be seen that SimpleTMY closely follows the distribution of SB 2008-2017. TenYrHrAvg displays an unrealistic clustered distribution, with too few high and low values, and too many mid-range values. This is characteristic over-smoothing resulting from the basic averaging employed by this dataset. Meteonorm slightly underestimates zeroes and creates too many 1-100 Wh/m² values of GHI. (From other graphs, not included here, it also underestimates over 600 Wh/m²). To a much greater extent, PVGIS-TMY overestimates zero values and underestimates in the 100 GHI

Wh/m^2 bin. (It also overestimates over 800 Wh/m^2). In the UK, a substantial proportion of of solar energy is produced under conditions of low irradiation [8], so large inaccuracies in simulating these values can impact largely on yield estimates.

In distribution (although not in quantity), the GHI values calculated by Meteonorm somewhat resemble those of a wet year like 2012. By contrast, PVGIS-TMY mimics a hot year like 2018. For both, the resemblance is greater in higher irradiation bins. Thus, these so-called typical years are not really typical. This problem may also be found in the seemingly well performing SimpleTMY. The process underlying this dataset has selected December 2010 values for the TMY. This was the coldest December for 100 years. Probably the average irradiation resulted from lack of pyranometer readings due to snow cover being compensated by snow reflectance, rather than being a true average.

Finally, daily performance of the TMYs was examined by looking at daily system output from PVSyst. All TMYs were discovered to have a unique pattern of daily grid injection. None of the patterns coincided with that of any actual year.

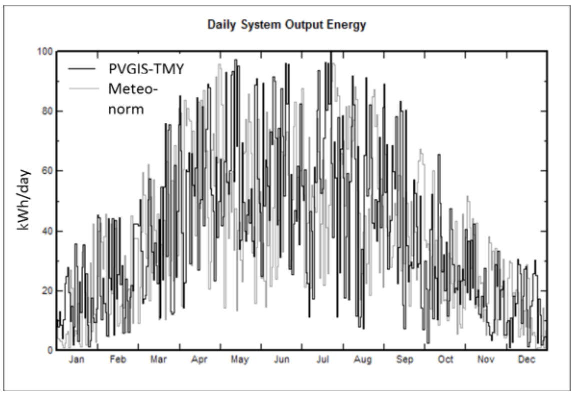

Figure 4: Daily output for a sample system calculated from Meteonorm and PVGIS-TMY

Discussion Taking an overview of the statistical measures used here, SimpleTMY delivers results closest to the long-term time series, followed by Meteonorm. Both employ quality ground measurements. Including extra parameters to GHI e.g. temperature can avoid the selection of atypical months, but these must be weighted specifically for PV. (PVGIS-TMY functions poorly for solar, due to being developed for building energy performance.)

Some TMYs function much better than others. Quality data inputs are as important for realistic results as the TMY model and the individual effect of each is difficult to separate. The more effective TMYs only can deliver project lifetime and annual sums of GHI. These TMYs may also provide average monthly sum, average seasonal sum, average daily sum and average hourly sum. These more accurate TMYs can also provide specific monthly sums e.g. January, June etc (Fig. 1.) and therefore seasonal figures. However, no TMY is suitable for daily and hourly analyses (Figs 3 and 4), nor is variability is covered.

Conclusion TMYs are actually not that typical of the weather phenomena they represent. They may produce the “correct” annual total of GHI, but not accurately distributed over 8760 hours or 365 days. There is room for improvement in their development. A shorter term time series, e.g. 5 years, or even a single non-extreme year (e.g. 2013, in this case) of high quality ground measurements can produce more practical results than a TMY.

References

[1] T. Cebecauer and M. Suri,”Typical Meteorological Year data: SolarGIS approach”, Energy Procedia 69 pp. 1958 – 1969, 2015

[2] E.F.M. Abreu, P. Canhoto, V. Prior and R. Melicio, “Solar resource assessment through long-term statistical analysis and typical data generation with different time resolutions using GHI measurements”, Renewable Energy, 127 pp 398 – 411, 2018

[3] A. Mermoud and B. Wittmer, PVSyst v. 6.7.8, University of Geneva, Switzerland http://www.pvsyst.com.

[4] UK Met Office, “MIDAS: UK Hourly Weather Observation Data. NCAS British Atmospheric Data Centre,” 2006. [Online]. http://catalogue.ceda.ac.uk/uuid/916ac4bbc46f7685ae9a5e10451bae7c.

[5] J. Remund and S. Kunz, Meteonorm v. 7.2, Meteotest, Switzerland, www.meteonorm.com

[6] T. Huld, PVGIS v.5, JRC, Ispra, Italy, re.jrc.ec.europa.eu/pvg_tools/en/tools.html

[7] NASA Langley Research Center (LaRC) Hampton, VA (USA), NASA-SSE, https://power.larc.nasa.gov/

[8] C. N.Jardine and K. Lane, Photovoltaics in the UK: An introductory guide for new consumers, ECI Research Report 27, 2003. https://pork.ahdb.org.uk/media/73759/photovoltaics-in-the-uk-an-introductory-guide-for-new-consumers.pdf