Solar Power And Chemical Energy Systems
A Technology Collaboration Programme of the
International Energy Agency

Road Map for Creation of Advanced Meteorological Data Sets for CSP Performance Simulations

Lourdes Ramírez[a], Kristian Pagh Nielsen[b], Frank Vignola[c], Manuel Blanco[d], Philippe Blance[e], Richard Meyer[f]

[a] **CIEMAT, E-28040, Madrid, Spain.**

[b] **Danish Meteorological Institute, DK-2100 Copenhagen, Denmark.**

[c] **The Solar Radiation Monitoring Lab, University of Oregon, Eugene, OR 97403-1274, USA.**

[d] **The Cyprus Institute, P.O. Box 27456. CY-1645, Nicosia, Cyprus.**

[e] **ARMINES, F-75272 Paris Cedex 06, France.**

[f] **Suntrace GmbH, D-22757 Hamburg, Germany.**

Executive summary

The demands on the quality and contents of solar and meteorological data sets increase rapidly with the growth in Concentrating Solar Power (CSP) and other solar energy systems. This document is the final deliverable of the SolarPACES Project: “Beyond TMY,” and it is produced as a final and summarized information presented in the report *Discussion of currently used practices for: “Creation of Meteorological Data Sets for CSP/STE Performance Simulations”*.

High quality ground-based measurements of global, direct and diffuse solar radiation are of fundamental importance. This includes proper quality control and gap-filling procedures. Satellite-derived data site-adapted with the ground-based measurements provide global coverage. Uncertainty assessments of both these data sources are essential. Furthermore, long-term solar resource scenarios and auxiliary meteorological data are important. Reanalysis data from weather models and the prospects of realistic multi-year generation are all main topics that we summarize in this road map.

This roadmap aims to help non-experts in solar radiation or meteorological information in the understanding of the current state of the art and the needed steps for the Creation of Advanced Meteorological Data Sets for CSP Performance Simulations.

1. Introduction

A Concentrating Solar Power (CSP)/Solar Thermal Electric (STE) power plant is a substantial long-term investment. To evaluate the opportunities and risks associated with such a long-term investment requires careful technical and economic analysis. Usually, the results of such analysis are presented in what are known as feasibility studies.

Traditional CSP/STE feasibility studies start by defining an economic model to estimate the economic metrics that characterize the quality and attractiveness of the investment project associated with the construction and operation of the CSP/STE power plant. Typical economic metrics are the Levelized Cost of Energy (LCOE), the Internal Rate of Return (IRR), the Net Present Value (NPV), and the Debt Coverage Ratio (DCR).

Once the model is defined, the main challenge is to accurately estimate the technical and economic variables and parameters that are inserted into the model, such as the project’s Total Investment, the Annual O&M costs, the Annual Electricity Generation, the Discount Rates, the Equity-to-Debt ratio, etc.

The Annual Electricity Generation is the parameter that is characterized by the quality of the solar resource at the CSP/STE plant site and the technical performance of the CSP/STE technology selected for the solar power plant. To estimate it, one should first develop or acquire a year of relevant solar radiation and other meteorological data that is representative of the long-term meteorology at the solar plant site. When the solar resource is combined with the technical parameters that define the plan technology, configuration and operation strategy, a technical model of the plant is created and an estimate the Annual Electricity Generation is produced.

Often, only one yearly data set is used that is representative of the average meteorological year to be expected at the site in the long-term. Sometimes this is supplemented by estimates of the production in a bad year that will be exceeded with a certain probability.

While the above approach, combined with a sensitivity analysis of the economic variables and parameters of the economic model is useful to banks and other potential investors in the decision making process related to the decision of carrying out the investment, there are other more sophisticated approaches that can be pursued.

The one we think is worth exploring is a full stochastic approach, in which the following aspects are explicitly modeled and taking into account:

- The inter-annual variability of the solar resources and other meteorological variables.
- The uncertainty of these meteorological variables – mainly the DNI.
- The uncertainty of the technical input parameters.
- The uncertainty of the technical model used to determine the annual electricity yield.
- The uncertainty associated with all the different component costs, including commodity-like cost components, such as molten salt and other costs that determine the aggregate values of the plan investment and the annual O&M cost.

In such a model, all the input variables and parameters are considered probability distributions. The challenge is to determine these distributions. How to determine the probability distribution of the Annual Electricity yield of the CSP/STE plant is the overarching theme of this document. Obviously, it starts with how to model the probability distribution of the solar radiation and other relevant meteorological variables. In this report we discuss the factors affecting this distribution.

Economic support from the IEA Technology Collaboration Programme for Solar Power and Chemical Energy Systems (SolarPACES) for meeting and travel expenses connected to making this report is highly appreciated. Additional funding from the Energy Development and Demonstration Program (EUDP) of the Danish Energy Agency, The Energy Trust of Oregon, The National Renewable Energy Laboratory, Bonneville Power Administration, the German Ministry of Economic Affairs and Energy under contract 0325643C, Suntrace GmbH, the Danish Meteorological Institute, and CIEMAT is also acknowledged.

2. Ground-based measurements

Background

A typical-sized Concentrating Solar Power (CSP)/Solar Thermal Electric (STE) power plant with a size in the order of 100 MW may require an investment volume of about 400 million USD. An in-situ measurement campaign at the site is highly recommended to achieve financial close. Nowadays satellite-derived global irradiance data are available for all potential sites. However, satellite-derived direct normal irradiance has much larger uncertainties. Therefor a local measurement campaign of at least one year of duration has several relevant advantages:

- The information is directly related to the location that has to be characterized. No consideration on spatial relation has to be made.

- It is not possible to derive meteorological information in very high frequency (minutes or seconds) directly from satellite images or numerical weather prediction (NWP) models. Such high frequency can be essential for the plant simulation.
- A measurement campaign of solar radiation during a year aims to check or to adjust locally products and estimations for long term as those coming from satellite images of reanalysis with NWP models.
- The meteorological information like temperature, humidity and wind speed can be assessed and compared with the output from NWPM.
- The investment in the instrumentation could be used latter in the power plant operation if it is built in the selected location or at a future location for new characterization.

On the other hand, some important points have to be considered in relation with a solar radiation measurement campaign:

- Solar radiation instruments have to be maintained at least once a week. The instruments have to be cleaned and performance has to be monitored to ensure proper operation. Otherwise problems can occur that typically result in a negative bias in the estimated solar resource.
- Some instruments are more robust for isolated locations as it is the case of rotating shadow band irradiometers, but this device has its own limitations that result from the spectral response characteristics of the pyranometer used or uncertainties in the correction algorithms used to remove systematic deviation. Direct normal measurements with pyrheliometers are more accurate but require even more frequent maintenance – preferably daily. Decision of whether to use one type of solar radiation instrument or another has to be assessed and could be different at different stages of the project.

In addition to the measurement campaign previous to the power plant erection, local measures are also needed for the solar field characterization and the assessment of the real plant performance.

Future steps:

1. Recommendations on meteorological stations including devices selection for each stage of the project. Each stage has different characteristics and need different instrumentation, time frequency and accuracy. Measurements of the three solar variables (global, diffuse and direct normal) have to be done in the final stage of the feasibility assessment. For the commissioning of large CSP plants it is recommended to use at least one better several pyrheliometers.
2. The WMO network of meteorological stations also should be upgraded to accommodate the need of the solar energy industry; instruments for wide-spread global monitoring of direct and diffuse irradiances are needed.
3. Recommendations should be made on the number of stations needed to ensure proper characterization of the solar resource for performance estimations. Recommendations for adequate maintenance of ground-based solar radiation measurements in the WMO framework should be made. Today several of the highest level solar radiation measurement stations promoted by renewable energy programmes such as World Bank's ESMAP Tier 1 station almost fulfill BSRN standard. Only few additions such as infrared measurements using a pyrgeometer would be needed to be of additional value also for climate science.

3. Quality control and gap-filling

Background

Quality control is an important point in the solar radiation data treatment. The common goal for any type of solar radiation assessment is to calculate the sum of the solar radiation during a period. This can be expressed as daily, monthly or annually irradiation. This point differs from the case of wind energy assessment. For wind energy assessments the reference values have been the average wind speed and a well-defined distribution function. Thus, wind measurements are conducted to characterize the local behavior of wind distributions.

In the case of solar radiation the characterization of the distribution function it is not agreed as for the wind energy case, and long series are needed for the long-term characterization of the local solar resource. When energy values have to be addressed quality control and gap filling methodologies become essential. A compromise of the level of restrictive filters minimizing the gaps has to be achieved, due to gaps have to be fulfilled before any daily, monthly or annual energy calculation. Filters have to be carefully applied mainly when the data are suspicious as a result of minimal maintenance. A station logbook or other ways of documentation should be kept, so that cleaning events can be identified and corrections applied.

When there are measurements for the three solar radiation variables (global, diffuse and direct normal) filters can be applied following well known recommendations including geometrical relationships between them. But when any of the three variables are not recorded, the closing equation cannot be performed, which is very powerful way for identifying errors. Issues with too high irradiance values can be detected relatively well, while too low values may remain undetected. Thus, both phenomena: removal of the highest values and missed detection of wrong low values contribute to a systematic underestimation of the expected direct normal irradiance. Even when all three variables are recorded, failures in the tracker have to be specifically detected.

Future steps:

1. Determine a standard method for adequate direct normal irradiance filters, including failure detection on tracker systems.
2. Determine a standard method for adequate and harmonized gap filling methodology. Options based in different interpolation models or historical patterns have to be compared and agreed.
3. Determine a standard method for the application of soiling correction.

4. Satellite-derived data

Background

The main advantage of using satellite-derived solar radiation is the provision of long-term solar radiation and meteorological data in places without ground-based measurements. Setting up new ground stations is necessary for CSP project planning, but it is impossible to wait for many years until enough data for a reliable yield analysis are available. Due to the inter-annual variability of DNI, at least 20 years of data are recommended for CSP resource assessments. The minimum of 10 years as mentioned by the SolarPACES guiSmo Report

(2017) may be misleading as the past 10 to 15 years have been extraordinary stable and sunny in many regions compared to earlier years.

The accuracy of satellite-derived data is generally lower than that of well-maintained and appropriate ground-based measurements. Depending on the local conditions, systematic errors may still remain and an adaptation to take account of local effects should be performed. Where local measurements are available and after using such adaptations, the accuracy of the long-term data sets is suitable for CSP resource assessment.

Satellite-derived solar radiation data sets are based mainly on the monitoring of clouds by geostationary satellites. No information for other meteorological variables like temperature, humidity, pressure or wind speed can be derived reliably from these.

Satellite-derived solar radiation databases are often available as commercial products; some of them provide useful meta-data and validation reports. Few satellite-derived solar radiation data sets are free and available to public access. For the local adaptation and validation of all satellite-derived data bases there are available only a limited number of solar radiation measurements stations around the world. Thus data bases behavior have to be carefully tested mainly in regions without well-known measurement stations.

The gridded satellite data have a limited spatial resolution and consists of image-like data taken at discrete time steps. Currently, the highest spatial resolution of the relevant geostationary weather satellites is in the range of one kilometer and the temporal frequency is down to five minutes, but typically rather 15 to 30 minutes.

Future steps:

1. Improvements in the temporal and spatial resolutions. Main issue to improve DNI from satellite is to solve the parallax problem by considering 3-dimensional radiative transfer. These would require to consider also cloud bottom and cloud top heights, which is a major challenge to retrieve from geostationary meteorological satellites.
2. Determine a standard method for adequate, harmonized blind testing to assess the differences of the satellite-derived solar radiation from ground-based measurements. As most satellite solar radiation retrievals are tuned by ground-based measurements it is very difficult to derive their actual uncertainty not influenced by data incest issues. Therefore uncertainty for new sites tends to be underestimated. For avoiding data incest it would be necessary to use ‘secret’ stations, which have not been made available to the model developers earlier.

5. Uncertainty

Background:

All measurements have **uncertainties**[1]. These uncertainties can be random or vary depending on other variables such as temperature or the intensity of incident radiation. To a certain degree, **random effects** average out as longer time intervals are involved. **Systematic effects,** that depend upon the environment or configuration under which the measurements are conducted, average to specific values with finite differences from the true value. While different systematic effects can cancel with other systematic effects they can also be

---

[1] Words printed in bold type are defined in the glossary for the uncertainty discussion.

additive. With careful studies, the systematic effects can be measured and modeled. With a sufficient knowledge of these systematic effects and measurement of the environmental conditions that produce these systematic effects, the contribution to the difference between the measurement and the true value uncertainty can be significantly reduced.

When financing a solar electric system, the probability that the system will exceed a certain level of performance (**probability of exceedance**) – the P-level – is important for risk evaluation in the financial model. Usually the most probable value P50, relating to the 50% level of exceedance, which should be exceeded by half of the cased, is taken as the base case. To approach the distribution of the expected CSP yield most financial models also assume a normal distribution. Under this assumption for quantifying this distribution curve only one additional point is required. Many banks simply take the P90 level, which should be exceeded in 90% of the cases.

But also this P90 value has its uncertainty due to uncertainties in the underlying measurements and model data and above mentioned other sources of uncertainty. Financers often take a conservative approach and subtract some of the uncertainty in the probability of exceedance (give it a **haircut**) from the system production estimates. For example, if the P90 probability of exceedance has a 3% uncertainty, then a financial analysis may give the yield gained from a P90 data set an additional 3% reduction, often called a 'haircut'.

Current practice used the determine uncertainty estimates are difficult to defend with a high degree of confidence. The values often assume that many random effects cancel for long time periods and that extremes caused by systematic effects don't contribute significantly to the overall uncertainty. For example the DNI measurements have larger uncertainties when the sun is very low in the sky, but the uncertainty in the DNI measurement more towards the median solar elevation is used in the analysis.

The **G.U.M.** methodology of uncertainty analysis is thestandard for quantifying uncertainties in science. For such relatively complex outdoor measurands such as solar radiation measurements it is still a major effort to follow the GUM approach. Long-term goal should be to identify individual uncertainty levels for each and every measurement, that depending on instrumentation, maintenance, calibration date and time of day may be different for each time-step. The GUM methodology can also be applied to the analysis of system performance and this will help provide a solid basis for determining the uncertainty in the performance levels.

The future goal of uncertainty analysis is to more reliably determine the uncertainty in the measurements. It should be determined, if it is possible to reduce these uncertainties with modeled adjustments that account for systematic effects that increase uncertainty. Additionallythe uncertainty of the modeled or combined values needs to be given.

Future steps:

1. Determine a standard method to measure the uncertainties of all relevant measurements that are used in STE system performance estimates. It is recommended that these uncertainties are determined in the field under settings in which the instrument is used and not under laboratory conditions where some systematic uncertainty effects can be masked.
2. Characterize and model systematic effects and determine the degree to which these models reduce the uncertainty in the data values and system performance estimates.
3. Develop a check list with the uncertainties associated with different sources of input data and the likely effect of these uncertainties on system performance estimates. Create a document that illustrates how to use these uncertainties to determine the overall uncertainty in system performance.

4. Evaluate and characterize uncertainties used in modeling and/or modifying irradiance data used in system analysis.
    a. For example, methods that modify satellite-derived irradiance values using ground-based measurements.
    b. Models that simulate short interval irradiance from hourly data.
5. Characterize uncertainties into random effects that will be reduced with a long-term analysis and uncertainties introduced by systematic effects that can be accounted for by modeling.

Glossary of uncertainty terminology

G.U.M.: The Guide to the Expression of Uncertainty in Measurement (GUM) is a product of the Joint Committee for Guides in Metrology (JCGM). JCGM is producing a series of documents to establish general rules for evaluating and expressing uncertainty in measurement. The GUM methodology is becoming a standard in the analysis of irradiance and other meteorological measurements.

http://www.bipm.org/en/publications/guides/gum.html
http://www.iso.org/sites/JCGM/GUM-introduction.htm

Uncertainty: Uncertainty is an expression of doubt about how well the result of the measurement represents the value of the quantity being measured. (G.U.M.)

Error: Error is the difference between the measured or modeled value and the true value. While error and uncertainty are often used interchangeably, they are distinct concepts based on the same measurements.

> **Random effect – Random error**: Random error is presumed to arise from stochastic, temporal, and spatial variations in the measurements. Random error usually can be reduced by increasing the number of measurements. When the GUM methodology is used, these are characterized as random effects.
>
> **Standard Deviation**: Standard deviation is a measure of the uncertainty of the mean due to random effects.
>
> **Systematic effect – Systematic error**: A systematic error arises from a recognized quantity on the measurement. This is now called a systematic effect. This effect can be quantified and if large enough, a correction factor can be applied to compensate for the effect. As with the random effect, the uncertainty caused by a systematic effect can be reduced but it cannot be eliminated.

Probability of exceedance: In this document, the probability of exceedance is determined from an analysis of system performance under meteorological situations that the system is likely to experience. For example P50 is estimated system performance that will be exceeded 50 percent of the time. P90 is the system performance that will be exceeded 90 percent of the time. The input parameters can be measured or modeled irradiance and other relevant meteorological parameters along with system performance characteristics.

6. Long-term data variability

Background

Even when some works show the statistical distribution of the solar radiation in different places around the world, these results are frequently obtained considering different time resolutions, and they are not definitive about the distribution function or about the way to characterize the distribution with short measurements periods. Long-term data solar radiation variability is limited to a few studies of the impact of major volcanic eruptions and a few climate modeling studies. Trends due to increasing or decreasing levels of aerosol pollution is an issue for using TMY data sets based on data from the recent decades for future resource assessments.

Future steps:

1. In the case of yearly direct normal irradiation, it should at least be tested whether the distribution is Gaussian. If not, it can be tested whether the Weibull, log-normal or Gumbel is the best choice.
2. Years with volcanic eruptions that are Plinian or ultra-Plinian eruptions should be analyzed separately, as these will often be outliers compared with the general statistical distribution.
3. The uncertainties of P90 values can be estimated with Monte-Carlo simulations of given statistical distributions. These depend on the length of the long-term data set used to derive them. A shortcoming of this method is that only should be used if the premises of the central-limit theorem are met.
4. Long term changes and climate scenarios have to be included in future profitability studies, due to their impact in the lifetime of the power plant. In particular, aerosol pollution scenarios are important.

7. Auxiliary meteorological data

Background

Long-term auxiliary meteorological data come from the station network operated by the members of the World Meteorological Organization (WMO). These data include wind, wind gusts, temperature, humidity and visibility. Additional meteorological data needed for CSP/STE plant simulations include low level extinction and upper air winds for solar power tower plants.

The use of all-sky cameras to observe clouds and aerosols both for minute scale forecasting and for cloud cover characterization has trended strongly in recent years. For cloud cover characterization these data resolves features that the satellite-images cannot resolve. These cloud cover features are very important for high-frequency solar radiation variability.

For Central Receiver plants the sunlight reflected at the heliostats has to cross substantial distances through the atmosphere near the ground. As the optical losses caused by extinction along this path can be sufficient it is required to quantify these losses. Recently, methods have been suggested for estimating the level extinction from standard meteorological. This could be AOD, DNI or preferably visibility measurements.

With advent of the ‘internet of things’ large quantities of ‘big data’ of low quality are becoming available. A major issue is, if and how these can be quality assured. The quality of the many different types of Big Data for meteorology could be assessed. This is an overall issue in the field of meteorological science.

In particular, regular and surveillance camera data can potentially be used to provide high coverage cloud cover and visibility data that can complement high quality ground-based data and satellite-derived data.

Future steps:

1. The use of all-sky camera data for cloud and aerosol observations should be standardized.
2. Methods for estimating extinction near the surface should be tested further and standardized. More work is needed to standardize these methods and estimate their uncertainties.
3. The usefulness of consumer camera images and surveillance camera data (Big Data) for cloud and visibility observations should be investigated.

8. Reanalysis data

Background

Numerical weather prediction models are developed for weather forecasting purposes. In the re-analysis mode the NWP models are run using all available initialization input data. Available reanalysis products cover different periods in different temporal and spatial resolutions. Re-analysis products are operated by institutional bodies, but sometimes their results are not free.

Re-analysis datasets are made with respect to climatological purposes and therefore cover more than 30 years. For long-term purposes they need to be harmonized as the underlying observations have been made with a series of instruments, while the reanalysis is per definition use a single NWP physics and data assimilation scheme.

Recent comparisons of solar radiation data from reanalysis products have demonstrated that in general, its results are still far less accurate then satellite-derived solar radiation data. Even so, there are specific developments that show a relevant improvement when using post-processing treatments. These types of post-processing treatments are also commonly applied in the case of satellite-derived models.

While satellite-derived models have been developed focused on solar radiation variables, NWP models were focused on meteorological variables such as temperature, precipitation and humidity, while the irradiance at the ground has not been a main focus. Assimilating cloud data with the methods of the satellite-irradiance models have prospects for significantly improving reanalysis data sets. As in the case of satellite-derived data, NWP model data are gridded area averages with a limited spatial resolution with accumulated irradiation as output. For the upcoming ERA 5 global reanalysis data the spatial resolution will be 31 km and the output frequency 1 hour. Regional reanalysis data sets can have spatial resolutions down to the kilometer scale.

Reanalysis data sets are typically the only option to obtain reliable long-term information about some other meteorological parameters needed for CSP simulations such as wind, temperature, humidity and pressure, when ground-based measurement are not available.

Future steps:

1. The cloud assimilation methods of reanalysis models need to be improved.
2. Improvements in the temporal and spatial resolution are needed.
3. Dedicated kilometer scale regional reanalysis models should be run for regions with high solar resources.

4. The methods for adequate, harmonized blind testing to assess the differences of the NWP modeled solar radiation and ground-based measurements should be improved with a focus on solar resource assessment.

9. Realistic multi-year generation

Background

In carrying out feasibility or other types of economic assessment studies of CSP/STE systems, it is of great interest to model the uncertainty and variability associated to different inputs of the economic model used in the study. A sound and elegant approach to carry out this modeling is to associate each input variable with a probability distribution. A key input in any economic model of a CST system is the annual energy yield of the system. The traditional approach to generate an estimate of the annual yield is to feed the Typical Meteorological Year (TMY) at the location of the CST system to an energy model of the system. But this approach does not produce a probability distribution of the energy yield. To produce a probability distribution the energy model of the CST system has to be fed with a large series of realistic meteorological years which are consistent with the TMY and with the long term expectations of the different meteorological variables within the meteorological year (DNI, ambient temperatures, relative humidity, wind speed and direction, etc.)

An approach to generate these large series of realistic meteorological years it is to automatize the treatment of available meteorological measurements in a given site and use them to generate full annual series of realistic meteorological years with high resolution solar radiation data. The most common procedure assumes a distribution function for the annual and monthly values, and sampling these distributions values of infinite months sequences can be obtained.

Future steps

1. Determine a standard methodology for adequate and harmonized assessment of realistic multi-year generation.
2. Include the new methodology in the most common feasibility assessment tools.

10. Summary of recommendations for future works

Ground-based measurements

1. Recommendations on meteorological stations for selection of measurement devices should be made and related to development stages of solar projects and its size in terms of MW.
2. Recommendations should be made to improve the solar resource measurements in the WMO network of meteorological stations. As countertrade to climate science solar energy may consider to equip high level solar radiation measurement stations with additional instruments primarily of value for climatology such as pyrgeometers to log the down-welling flux in the infrared.

Quality control and gap-filling

3. Determine a standard method for detection of erroneous direct normal irradiance values.
4. Determine a standard method for adequate and harmonized gap filling of missing data.
5. Determine a standard method for the application of soiling correction.

Satellite-derived data

6. Determine a standard method for adequate, harmonized blind testing to assess the differences of the satellite-derived solar radiation vs. the ground-based measurements.

Uncertainty

7. Determine a standard method to measure the uncertainties of all relevant measurements that are used in STE system performance estimates.
8. Develop a check-list with the uncertainties associated with different sources of input data and the likely effect of these uncertainties on system performance estimates.
9. Characterize uncertainties into random effects that will be reduced with a long-term analysis and uncertainties introduced by systematic effects that can be accounted for by modeling.

Long-term data variability

10. In the case of yearly direct normal irradiation, it should at least be tested whether the distribution is Gaussian. If not, it can be tested whether the Weibull, log-normal or Gumbel is the best choice.
11. Long term changes and climate scenarios have to be included in future profitability studies, due to their impact in the lifetime of the power plant. In particular, aerosol scenarios are important.

Auxiliary meteorological data

12. The use of all-sky camera data for cloud and aerosol observations should be further established.
13. Methods for estimating atmospheric extinction near the ground should be analyzed on accuracy and standardized.

Reanalysis data

14. Dedicated kilometer scale regional reanalysis models should be run for regions with high solar resources.

Realistic multi-year generation

15. Determine a standard methodology for adequate and harmonized assessment of realistic multi-year generation.
16. Include the new methodology in the most common feasibility assessment tools.