Towards a Validation Protocol for Satellite-Based Solar Resource Data

Ian Grant[1], Manajit Sengupta[2], Fernando Camacho[3], Dominique Carrer[4] and Zhuosen Wang[5]

[1] Bureau of Meteorology, Melbourne (Australia)

[2] National Renewable Energy Laboratory, Golden CO (USA)

[3] University of Valencia, Valencia (Spain)

[4] Météo-France, Toulouse (France)

[5] University of Maryland, College Park MD (USA)

Abstract

Satellite-based solar resource data are valued for their spatial extensiveness and long temporal record. Validation of such data, including accurate assessment of their uncertainties against reference datasets, is essential to give users confidence in and maximum benefit from the data. More broadly, the international Earth observation from space community estimates many geophysical parameters from satellite data, such as, to name just a few, precipitation, sea surface temperature and atmospheric aerosol amount, and validates them. In particular, the Committee on Earth Observation Satellites (CEOS), an international body that coordinates the use of civilian non-meteorological Earth observing satellites, has made considerable progress in developing standard validation protocols for numerous satellite-derived parameters. These protocols recommend good practices to support the benchmarking and validation of satellite datasets. This paper describes the benefits and development approach of a validation protocol for satellite solar resource data that is being developed under the auspices of the CEOS Working Group on Calibration and Validation (WGCV) Land Product Validation (LPV) subgroup. The protocol development has the support of International Energy Agency (IEA) / Photovoltaic Power Systems (PVPS) / Task 16 *Solar resource for high penetration and large scale applications* and will build on the work of Task 16 and previous IEA Tasks for solar resource assessment.

Keywords: solar resource, satellite, validation, protocol

1. Introduction

Satellite-derived solar resource data are recognised as valuable for being spatially extensive and for providing long historical records, often extending back decades. Acceptance and optimum use of solar resource data depend on having a thorough and reliable characterisation of their uncertainties. Furthermore, users often have access to data from multiple data providers and need to understand the differences between datasets in terms of their respective uncertainty magnitudes and their relative strengths and weaknesses. Thus, benchmarking of satellite-derived solar data—validating their data values and quantifying their uncertainties—is key to the confident and informed use of the data.

For satellite-derived data, uncertainties are usually based on comparisons with quality ground measurements. Different satellite datasets can also be intercompared. While there is some commonality to how such comparisons are reported—mean bias error (MBE) and root mean squared error (RMSE) are almost invariably included in the reported metrics—the choice of metrics reported and the manner of their generation can vary between various dataset assessments. There is currently no standardised method for accuracy assessment of satellite datasets (Sengupta et al., 2015). A suite of metrics that assesses a range of data characteristics of interest to users in the solar energy sector, including, for instance, dataset completeness and higher order statistical properties, can provide a comprehensive picture of dataset quality. Benchmarking can be improved by careful approaches to using ground-based reference data that consider the selection of sites spanning a range of conditions, the instrumentation and its maintenance, and quality control of the data. Furthermore, benchmarking reports will be intercomparable if they follow a standardised approach. While some work has been done to establish a standardised benchmarking approach for solar resource data, there is scope to consolidate this work in a single recommended validation protocol.

doi:10.18086/swc.2019.45.05 Available at http://proceedings.ises.org

The Committee on Earth Observation Satellites (CEOS) is an international body that coordinates the operation and utilisation of civilian non-meteorological Earth observing satellites. CEOS, through its Land Product Validation (LPV) Subgroup, has well developed processes for developing good practice recommendations for the validation of a large range of satellite-derived biogeophysical data products. The CEOS processes thus offer a path to formalising an internationally standard recommended validation protocol for satellite solar resource data.

A brief description of the CEOS structure and activities is given here. CEOS Members are agencies that operate Earth observing satellites, while CEOS Associate Members are other agencies that utilize data from such satellites. CEOS conducts activities to further its aims through its hierarchical structure, which resembles the structure by which the International Energy Agency (IEA) organizes its work through its Technology Collaboration Programmes (TCPs). For instance, the Photovoltaic Power Systems (PVPS) TCP oversees thematic Tasks such as Task 16 *Solar resource for high penetration and large scale applications*, which in turn supervises Subtasks that are groups of Activities. CEOS's component groups develop work packages in response to gaps and needs identified by their technical experts, or to requests flowing down the CEOS hierarchy, sometimes from outside such as from the United Nations.

CEOS's Working Group on Calibration and Validation (WGCV) coordinates CEOS's work to ensure long-term confidence in the accuracy and quality of satellite-based Earth observation data and products. WGCV's Land Product Validation (LPV) Subgroup is dedicated to fostering the quantitative validation of global land products derived from remote sensing data and to relay results so they are relevant to users. A significant LPV activity to further these aims is the development and publication of a set of validation protocols that recommend approaches for the validation and characterization of uncertainty of satellite-derived data for specific geophysical variables relevant to the land surface. These protocols are intended to recommend "good" practices rather than necessarily "best" practices, since a WGCV priority is to encourage consistency in validation practices across the community of data producers and users. The protocols can be adopted by data producers, be they CEOS agencies or others, or anyone else who validates satellite data products. Recently published LPV validation protocols include land surface albedo (Wang et al., 2019) and land surface temperature (Guillevic et al., 2018).

The websites for CEOS and its sub-bodies just described are:

- CEOS: http://ceos.org/
- WGCV: http://ceos.org/ourwork/workinggroups/wgcv/
- LPV: https://lpvs.gsfc.nasa.gov/
- CEOS cal/val portal: calvalportal.ceos.org/ .

LPV validation protocols typically devote a significant part of the document to recommendations on the collection of ground-based validation data, covering aspects such as the location and spatial homogeneity of field sites, instrumentation and quality control. Given that the Baseline Surface Radiation Network (BSRN, https://bsrn.awi.de/) already exists to provide high quality ground-based solar data, a radiation validation protocol does not need to cover that topic in depth.

2. Satellite solar data

Solar radiation is a key environmental variable for many natural and manmade systems. Besides being the fuel for solar energy generation, it also drives weather, climate, evapotranspiration, plant growth and building thermal behaviour. There are, therefore, many applications of solar data beyond the solar energy sector. The World Meteorological Organization's Global Climate Observing System (GCOS) classifies downwelling surface solar radiation as an Essential Climate Variable (ECV; https://gcos.wmo.int/en/essential-climate-variables) and includes it as a core variable to be measured to high quality in its proposed GCOS Surface Reference Network (GSRN; WMO, 2019). GCOS considers solar radiation from a climate perspective, and its requirements for solar radiation monitoring include coarser spatial and temporal resolution and more stringent stability requirements than do solar energy applications, and an interest in only global horizontal irradiance (GHI) as one component of the surface energy budget rather than its direct and diffuse components.

Several operational agencies produce publicly available solar datasets from satellite data. The spatial coverage is variously national, global or the region of the globe covered by one geostationary satellite. Some of these are limited to serving climate applications and typically comprise only GHI with a resolution of tens of km. Examples are datasets produced by:

- The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) through its Satellite Application Facilities on Climate Monitoring (CM-SAF; https://www.cmsaf.eu) and Land Surface Analysis (LSA-SAF; https://landsaf.ipma.pt)
- The US National Aeronautics and Space Administration (NASA) from its CERES instrument (https://eosweb.larc.nasa.gov/project/ceres/ceres_table)
- The Japan Aerospace Exploration Agency (JAXA; https://www.eorc.jaxa.jp/ptree/index.html).

Other datasets are intended to support solar energy applications and deliver datasets that are at finer spatial resolution and include direct and/or diffuse components. Examples include:

- The US National Renewable Energy Laboratory's (NREL) National Solar Radiation Database (NSRDB; https://nsrdb.nrel.gov/)
- NASA's Prediction Of Worldwide Energy Resources (POWER, formerly Surface meteorology and Solar Energy; https://power.larc.nasa.gov/) data
- The radiation service of the European Commission's Copernicus Atmospheric Monitoring Service (CAMS; https://atmosphere.copernicus.eu/)
- The Australian Bureau of Meteorology's satellite-based solar data (http://www.bom.gov.au/climate/data-services/solar-information.shtml).

3. Previous work

IEA coordinates expert groups working on energy research topics through its TCPs. IEA/PVPS Task16 *Solar resource for high penetration and large scale applications* (http://www.iea-pvps.org/index.php?id=389) brings together an international group of some one hundred self-funded technical experts from around twenty nations to advance solar resource assessment and forecasting for the solar energy sector. The goals of Task 16 include lowering the uncertainty of satellite-derived solar data and establishing or contributing to the international benchmarking of datasets. IEA/PVPS Task 16 builds on the work of a series of earlier IEA tasks of fixed duration on solar resource data and forecasting, most recently Task 36 and Task 46 of IEA's Solar Heating and Cooling (SHC) Programme.

Beyer et al. (2009), which was published under the auspices of IEA/SHC Task 36, consolidates much of the current practice in satellite solar data validation. This report made recommendations for, among other topics, reporting point comparisons at ground stations and map-based relative intercomparison of satellite datasets. Its recommended metrics included MBE, RMSE and measures of the difference between probability distributions.

4. Protocol structure

The radiation validation protocol can be modelled on previous LPV protocols such as the recently published protocol for land surface albedo (Wang et al., 2019), with adaptation to incorporate prior work on the benchmarking of solar data and to identify any necessary future work.

The scope of the albedo protocol includes:

- definitions;
- requirements for satellite-based data;
- spatial and geometrical aspects of satellite observations;
- sources of uncertainty;
- sources, characteristics and uncertainties of reference data;
- global and seasonal representativeness of validation;
- stability evaluation;
- status of current validation capacity and methods;
- validation metrics;
- the stratification of performance statistics.

All of these issues are appropriate for coverage in the solar radiation protocol.

Solar radiation is different from many other geophysical variables in its fine scale spatial and temporal variability.

Solar radiation has the spatiotemporal variability of drifting cloud fields, which motivates ground observations to be made at 1 minute or even 1 second sampling, and satellite solar data products to be produced with temporal steps as small as the 10 or 15 minutes imaging cycle of the most recent generation of geostationary satellites. This variability introduces unique challenges in comparing satellite and ground data, such as significant spatial variation within a satellite imager's pixel or a satellite product grid cell, the greater temporal variability of point ground observations than that of spatially averaged satellite observations, and the dependence of the errors of dataset intercomparison on spatial and temporal resolution.

The radiation protocol's recommended metrics will include those already in common use for solar data. However, the existing LPV protocols include metrics characterising dataset properties that have not generally been reported for solar data but could be considered as worthwhile additions to solar benchmarking. Table 1 presents the metrics recommended by the albedo protocol for assessment of satellite data against ground data. The completeness metrics in this table are, for instance, possibly new and useful for solar radiation. Some metrics are not appropriate for solar radiation, such as the temporal "smoothness" metric that is suitable for land surface variables that vary on the scale of weeks but not for the far more variable solar radiation. Conversely, solar radiation metrics should include the measures of PDF difference developed in IEA/PVPS Task 36, as well as possibly differences in the representation of temporal variability or ramps in datasets.

Tab. 1: Common practice and recommended good practice for validation, from the CEOS surface albedo validation protocol

Quantity	Current practice	Good practice, add:
Accuracy	Bias; absolute bias	Median error Median and percentiles of residuals Box-plots of residuals vs. Albedo
Precision	Standard deviation	Median absolute deviation Median 3 point difference
Uncertainty	Root mean square error	Scatter plot of match-ups Median and percentiles of absolute residuals, RMSE Box-plots of absolute residuals vs. Albedo
Completeness		Gap size distribution Gap length
Stability		Time series average, standard deviation, and regression slope Mean error per decade

The validation protocol needs to serve all applications of solar radiation data, and the selection of metrics might need to be adapted to specific data products and their application. For instance, metrics of fine scale temporal variability might be applicable to 15-minute solar time series but not daily totals.

The protocol document should identify knowledge gaps where future research is needed. For instance, to what degree does the land surface in the vicinity of a ground site need to be homogeneous? What is an adequate global distribution of ground sites? What quality control should be applied to ground data? What approach should be taken to deal with gaps in ground or satellite data to be compared as temporal aggregates such as daily totals or monthly averages? Is there a need to define separate validation requirements for the global, direct and diffuse components?

5. Conclusion

The widespread adoption of a standard validation protocol for satellite-based solar resource data would improve the confidence in and use of such data and improve the intercomparability of dataset assessments. Such a protocol is being developed through CEOS's Land Product Validation Subgroup with the broad support of IEA/PVPS Task 16.

6. References

Beyer, H. G., Martinez, J. P., Suri, M., Torres, J. L., Lorenz, E., Müller, S. C., Hoyer-Klick, C., Ineichen, P., 2009.

Management and Exploitation of Solar Resource Knowledge: Ca–Contract No. 038665— D 1.1.3. Report on Benchmarking of Radiation Products.
http://www.mesor.org/docs/MESoR_Benchmarking_of_radiation_products.pdf

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J. A., Remedios, J., Román, M., Camacho, F., 2018. Land Surface Temperature Product Validation Best Practice Protocol. Version 1.0, in: Guillevic, P., Göttsche, F., Nickeson, J., Román, M. (Eds.), Best Practice for Satellite-Derived Land Product Validation. Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/lst.001.
https://lpvs.gsfc.nasa.gov/PDF/CEOS_LST_PROTOCOL_Feb2018_v1.1.0_light.pdf

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., et al., 2015. Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications. Technical Report NREL/TP-5D00-63112, 236 pp., National Renewable Energy Laboratory, USA.
https://www.nrel.gov/docs/fy15osti/63112.pdf

Wang, Z., Schaaf, C., Lattanzio, A., Carrer, D., Grant, I., Román, M., Camacho, F., Yu, Y., Sánchez-Zapero, J., Nickeson, J., 2019. Global Surface Albedo Product Validation Best Practices Protocol. Version 1.0, in: Wang, Z., Nickeson, J., Román, M. (Eds.), Best Practice for Satellite Derived Land Product Validation. Land Product Validation Subgroup (WGCV/CEOS), doi: 10.5067/DOC/CEOSWGCV/LPV/ALBEDO.001.
https://lpvs.gsfc.nasa.gov/PDF/CEOS_ALBEDO_Protocol_20190307_v1.pdf

WMO, 2019. GCOS Surface Reference Network (GSRN): Justification, requirements, siting and instrumentation options. GCOS-226, World Meteorological Organization, Geneva, Switzerland.
https://library.wmo.int/index.php?lvl=notice_display&id=21404#.XT-g4RqP6JA