The Journal of Open Source Software

SolarSpatialTools: A Python package for spatial solar energy analyses

Joseph Ranalli [1¶] and **William Hobbs** [2]

1 Penn State Hazleton, Hazleton, PA, USA **2** Southern Company, Birmingham, AL, USA ¶ Corresponding author

DOI: 10.21105/joss.06984

Software

- Review
- Repository
- Archive

Editor: Adam R. Jensen

Reviewers:

- @cwhanse
- @kbonney

Submitted: 16 May 2024
Published: 16 September 2024

Summary

Solar energy is a form of renewable energy whose resource (i.e., sunlight) is available on the earth's surface with a relatively low energy density. This type of resource inherently requires spatial distribution of collection infrastructure in order to achieve increased generation scale. This is true both in the case of distributed (e.g., rooftop solar) and centralized generation. As international responses to climate change promote growing interest in solar energy, there is a corresponding growth of interest in tools for working with distributed solar energy data that possesses these characteristics. This package, `SolarSpatialTools`, aims to contribute to that need by providing research codes for spatial analyses of solar energy data and resources.

Statement of need

As mature packages already exist for supporting general analysis and modeling of solar energy systems, such as `pvlib-python` (Anderson et al., 2023) and `pvanalytics` (Perry et al., 2022), this package is not intended to serve as a replacement, a competitor, or to fragment those communities. Rather, `SolarSpatialTools` serves to collect codes for several tasks that are out-of-scope for `pvlib-python` and `pvanalytics`, but are still of general interest to the research community. Where appropriate, capabilities of `SolarSpatialTools` are contributed to `pvlib-python` or `pvanalytics`. For example, a Python language port of the Wavelet Variability Model (Lave & Kleissl, 2013) contained in the MATLAB `pvlib` package (Andrews et al., 2014) was first developed within `SolarSpatialTools` but was contributed to `pvlib-python` in 2019. `SolarSpatialTools` primarily grew out of personal research codes developed by the lead author under the name `solartoolbox`, but as tools have reached a level of maturity that attracted interest of a broader audience, it has been prepared as a package for more general public use.

To be more specific, a variety of analytical techniques related to solar energy are documented in literature, but are not already implemented by existing packages in part due to their relatively high complexity relative to those packages' intended scope. For example, techniques for processing cloud motion vectors (CMVs) from spatially distributed data sets are documented in the literature, such as the method by Jamaly & Kleissl (2018) and that by Gagné et al. (2018). Implementation of these techniques is laborious, requiring calculation of mutual correlation between all possible sensor pairs within a distributed data set. This fundamentally leads to a need to handle data types (i.e., simultaneous time series for each sensor) that are not aligned with the primary focus of the existing packages. Further, the number of calculation steps that are specialized for these CMV calculations makes them unattractive for inclusion in existing solar energy packages, without leading to an extreme broadening of scope to adapt to this singular use case. At the same time, the level of detail in those calculation steps makes them potentially difficult for other investigators to individually implement on a consistent and optimized basis. As they serve a common need within solar energy research, they are implemented in a well documented way by SolarSpatialTools to help alleviate this challenge.

Features

There are three capabilities of the SolarSpatialTools package that are most likely to be of interest for a general audience. These main capabilities are contained in the following modules:

- signalproc: tools for performing signal processing analyses across multi-sensor networks of solar energy data
- cmv: tools for computing the cloud motion vector from spatially distributed sensor networks
- field: tools for analyzing the relative positions of spatially distributed measurement units via cloud motion

These three main capabilities are also supported by extended documentation and tutorials in an additional directory of the package:

- demos: demonstration codes and sample data to help users get started with the package

Signal Processing

The signalproc module was developed as part of efforts to analyze aggregation of irradiance by spatially distributed plants, but may also be applicable to other signal processing tasks. This approach is used by the Wavelet Variability Model (Lave & Kleissl, 2013), the model of Marcos et al. (2011) and the Cloud Advection Model (Ranalli & Peerlings, 2021), which was developed by the lead author based on the physical intuition of Hoff & Perez (2010). The module contains codes for implementing these types of models using a transfer function paradigm. Some wrappers are provided for scipy (Virtanen et al., 2020) signal processing functions to simplify their application on the data type conventions used by this package. A demonstration of the signal processing capability as it pertains to comparing the different spatial aggregation models is provided in the demos directory of the package (signalproc_demo.py).

Cloud Motion Vector Calculation

The cmv module contains tools for calculating the cloud motion vector from a spatially distributed data set. Two methods from the literature are implemented, that of Jamaly & Kleissl (2018) and that of Gagné et al. (2018). These methods are both based upon computation of the relative time delay between individual sensors but utilize different techniques to process those into a global cloud motion vector. This module depends upon signalproc for some of its computations. A demonstration of the cloud motion vector calculation capability is provided in the demos directory of the package (cmv_demo.py) along with a Jupyter notebook with detailed explanations (cmv_demo.ipynb).

Field Analysis

The field module contains an implementation of the method developed by the authors (Ranalli & Hobbs, 2024a, 2024b) for comparison of a plant's layout from its design plan with that inferred from relative cloud motion across the plant. The method produces a prediction of a single reference sensor's apparent position on the basis of the relative delay between it and other nearby sensors. The application relies on the availability of two distinct cloud motion vectors, which allow triangulation of the sensor's planar position. The implementation depends on both signalproc and cmv. It is demonstrated in several of the codes in the demos directory including field_demo.ipynb, and field_demo_detailed.ipynb. Aspects of automating the process (Ranalli & Hobbs, 2024b) are demonstrated by automate_cmv_demo, field_reassignment_demo and field_demo_full_process. The last of these demonstrations also exemplifies parallelization of the implementation to speed up the processing for an entire plant.

Demos

The demos directory includes a variety of demonstration codes and explanatory Jupyter notebooks for the tools in the package, as described in the preceding sections. These demonstrations make use of a few sample datasets that are included in h5 files. Two samples are subsets of distributed irradiance network timeseries taken as a subset of the HOPE Melpitz campaign (Macke et al., 2017). One hour of sample data is available with the dataset's native sample rate of 1 s, while a longer four-day subset is available with 10 s resolution. Two additional sample data sets consist of combiner-level data from operational photovoltaic generation plants. Each is taken from a different plant and consists of five, distinct one hour periods of 10 s resolution time series of combiner current. These periods are chosen as those known to experience a high degree of variability due to cloud motion, making them suitable for use with the CMV and signal processing analyses. Data from these plants are anonymized to prevent identification of proprietary data; combiner locations are only given in relative east and north spatial coordinates and their generation magnitudes are scaled to an arbitrary value. As the analytial techniques contained in this package are primarily based on the variability of the signals, the anonymization process does not affect the utility of the data for the purposes of the demonstrations, and in particular, the plant data are used to demonstrate the field module.

Additional Modules

The remaining modules in SolarSpatialTools are somewhat less likely to be of general interest, but serve either a specialized or supporting purpose to the primary functionality:

- dataio: prewritten functions for downloading and preprocessing distributed solar irradiance data specifically from the HOPE (Macke et al., 2017) and NRCAN (Pelland et al., 2021) measurement campaigns.
- irradiance: a wrapper for pvlib-python.clearsky_index for easier processing of multiple simultaneous timeseries.
- spatial: tools for performing vector and geographic projection operations necessary for other modules.
- stats: calculations for some simple metrics used in solar energy. The variability metrics variability_index (Stein et al., 2012) and variability_score (Lave et al., 2015) may not presently be implemented by other packages and might be of some interest to other users.

Acknowledgements

Work on SolarSpatialTools was funded by Penn State Hazleton and Penn State School of Engineering Design and Innovation.

References

Anderson, K. S., Hansen, C. W., Holmgren, W. F., Jensen, A. R., Mikofski, M. A., & Driesse, A. (2023). Pvlib python: 2023 project update. *Journal of Open Source Software*, *8*(92), 5994. https://doi.org/10.21105/joss.05994

Andrews, R. W., Stein, J. S., Hansen, C., & Riley, D. (2014). Introduction to the open source PV LIB for python Photovoltaic system modelling package. *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, 0170–0174. https://doi.org/10.1109/PVSC.2014.6925501

Gagné, A., Ninad, N., Adeyemo, J., Turcotte, D., & Wong, S. (2018). Directional Solar Variability Analysis. *2018 IEEE Electrical Power and Energy Conference (EPEC)*, 1–6. https://doi.org/10.1109/EPEC.2018.8598442

Hoff, T. E., & Perez, R. (2010). Quantifying PV power Output Variability. *Solar Energy*, *84*(10), 1782–1793. https://doi.org/10.1016/j.solener.2010.07.003

Jamaly, M., & Kleissl, J. (2018). Robust cloud motion estimation by spatio-temporal correlation analysis of irradiance data. *Solar Energy*, *159*, 306–317. https://doi.org/10.1016/j.solener.2017.10.075

Lave, M., & Kleissl, J. (2013). Cloud speed impact on solar variability scaling – Application to the wavelet variability model. *Solar Energy*, *91*, 11–21. https://doi.org/10.1016/j.solener.2013.01.023

Lave, M., Reno, M. J., & Broderick, R. J. (2015). Characterizing local high-frequency solar variability and its impact to distribution studies. *Solar Energy*, *118*, 327–337. https://doi.org/10.1016/j.solener.2015.05.028

Macke, A., Seifert, P., Baars, H., Barthlott, C., Beekmans, C., Behrendt, A., Bohn, B., Brueck, M., Bühl, J., Crewell, S., Damian, T., Deneke, H., Düsing, S., Foth, A., Girolamo, P. D., Hammann, E., Heinze, R., Hirsikko, A., Kalisch, J., … Xie, X. (2017). The HD(CP)^2 Observational Prototype Experiment (HOPE) – an overview. *Atmospheric Chemistry and Physics*, *17*(7), 4887–4914. https://doi.org/10.5194/acp-17-4887-2017

Marcos, J., Marroyo, L., Lorenzo, E., Alvira, D., & Izco, E. (2011). Power output fluctuations in large scale pv plants: One year observations with one second resolution and a derived analytic model. *Progress in Photovoltaics: Research and Applications*, *19*(2), 218–227. https://doi.org/10.1002/pip.1016

Pelland, S., Gagné, A., Allam, M. A., Turcotte, D., & Ninad, N. (2021). Spatiotemporal Interpolation of High Frequency Irradiance Data for Inverter Testing. *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 0211–0218. https://doi.org/10.1109/PVSC43889.2021.9518827

Perry, K., Vining, W., Anderson, K., Muller, M., & Hansen, C. (2022). *PVAnalytics: A Python Package for Automated Processing of Solar Time Series Data* (NREL/PR-5K00-83824). National Renewable Energy Lab. (NREL), Golden, CO (United States). https://www.osti.gov/biblio/1887283

Ranalli, J., & Hobbs, W. B. (2024a). PV Plant Equipment Labels and Layouts Can Be Validated by Analyzing Cloud Motion in Existing Plant Measurements. *IEEE Journal of Photovoltaics*. https://doi.org/10.1109/JPHOTOV.2024.3366666

Ranalli, J., & Hobbs, W. B. (2024b, June). Automating Methods for Validating PV Plant Equipment Labels. *52nd IEEE PV Specialists Conference*.

Ranalli, J., & Peerlings, E. E. M. (2021). Cloud advection model of solar irradiance smoothing by spatial aggregation. *Journal of Renewable and Sustainable Energy*, *13*(3), 033704. https://doi.org/10.1063/5.0050428

Stein, J. S., Hansen, C. W., & Reno, M. J. (2012). The Variability Index: A New and Novel Metric for Quantifying Irradiance and PV Output Variability. *Proceedings of the World Renewable Energy Forum*, 13–17.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Walt, S. J. van der, Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2