

Time Dilated Bundt Cake Analysis of PV Output

Giray Ogut Bennet Meyers Stephen Boyd

Stanford University

IEEE PVSC, 6/10/2024

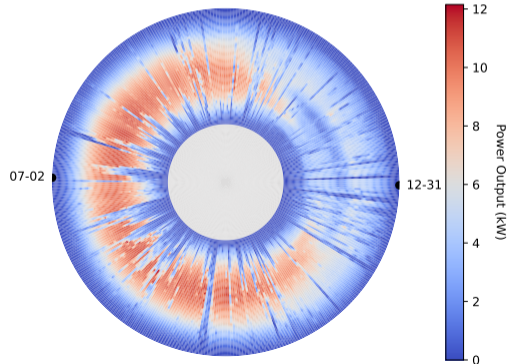
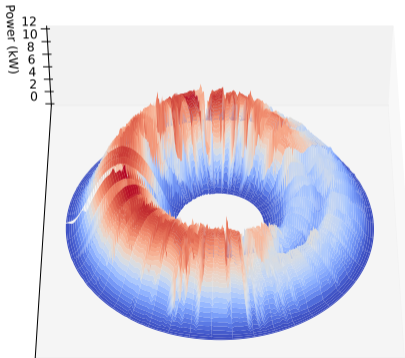
Overview

- ▶ **objective:** Model power generation of a PV system with seasonal variation
- ▶ **application:** Estimate marginal quantiles and detect for clear sky intervals
- ▶ **approach:** White-box ML model using convex optimization
 - low computational requirements (e.g. single laptop)
 - works with missing data (even after removing % 70)
 - interpretable and auditable results

- ▶ **prior work:** Physical models using geographic coordinates, atmospheric parameters etc. [Kambezidis1997, Stein2012, Antonanzas2019] or black-box models such as neural networks [Zhandire2017, Mosaico2019, Masucco2019]

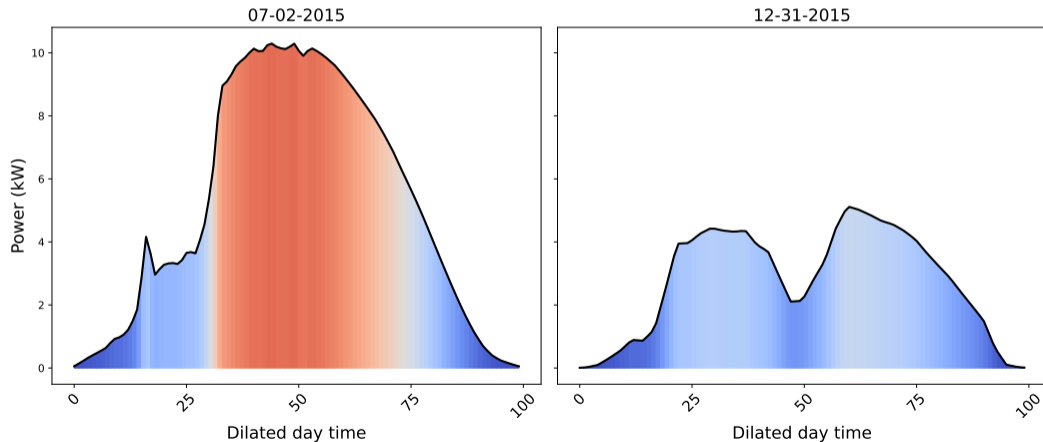
Bundt cake representation

- ▶ express daytime, time dilated data $X \in (\mathbf{R}_+ \cup \{?\})^{D \times M}$ as a "Bundt" cake
- ▶ D is the number of days and M is the number of intervals in a day
- ▶ allow for missing data $X_{d,m} = ?$



Bundt cake representation (continued)

- ▶ each "slice" of the cake corresponds to a day
- ▶ slices start and end near zero (*i.e.* standing waves)
- ▶ allows comparison of different days on equal footing



Method

- ▶ to get from raw data to Bundt cake representation, we do
 - PV sunrise/sunset estimation
 - time dilation

- ▶ using Bundt cake representation, we can
 - estimate marginal quantiles
 - detect clear sky

Fitting smooth multiperiodic basis functions

- ▶ define a set of basis functions, $\phi_i \quad i = 1, \dots, N$ such as

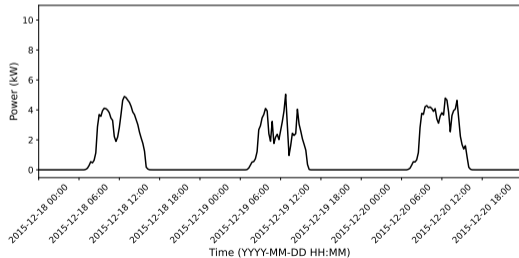
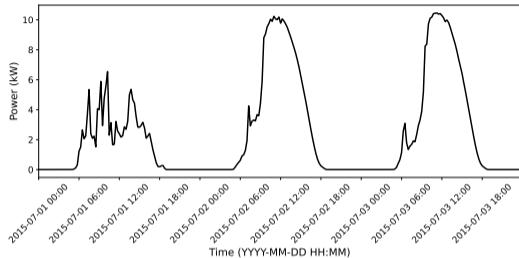
$$\cos(2\pi 3t/P_1), \quad \sin(\pi t/P_2), \quad \cos(2\pi 3t/P_1) \sin(\pi t/P_2)$$

- ▶ more sophisticated bases such as splines can also be used
- ▶ express any model parameter $\theta \in \mathbf{R}$ as

$$\theta = \sum_{i=1}^N \alpha_i \phi_i(t).$$

- ▶ fitting the coefficients α_i is formulated as a convex optimization problem
- ▶ see our paper & open source package
<https://stanford.edu/~boyd/papers/bundt.html> & `pip install spcqe`

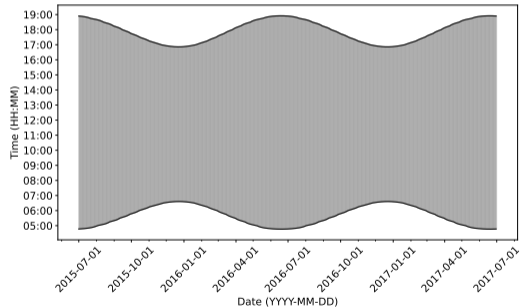
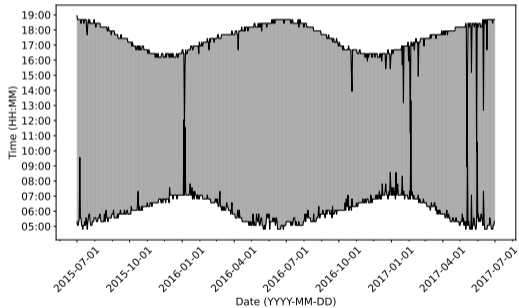
Numerical example



- ▶ residential PV system in Southern California
- ▶ 15-minute average power from 07/01/2015 to 06/30/2016 (no missing data)
- ▶ clear seasonal and daily patterns
- ▶ shade in winter, fog early in the day during summer

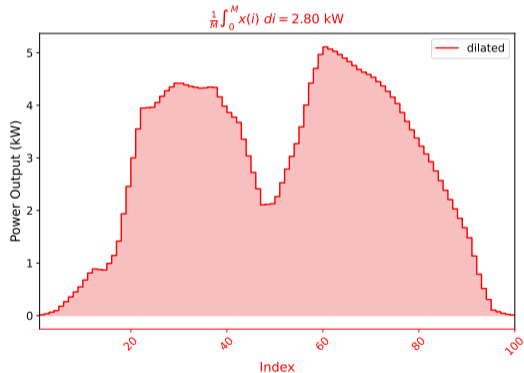
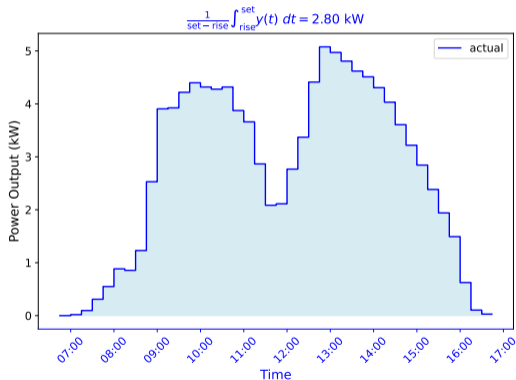
Smooth PV sunrise/sunset estimation

- ▶ PV sunrise/sunset is different from astronomical sunrise/sunset
- ▶ basic thresholding results in noisy estimates and multiple transitions in a day
- ▶ we come up with smooth estimates

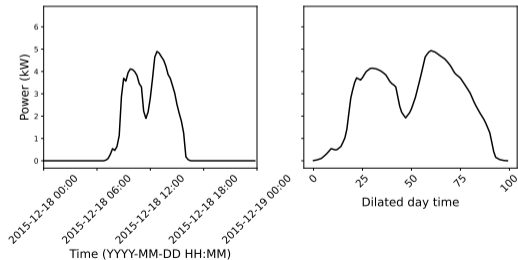
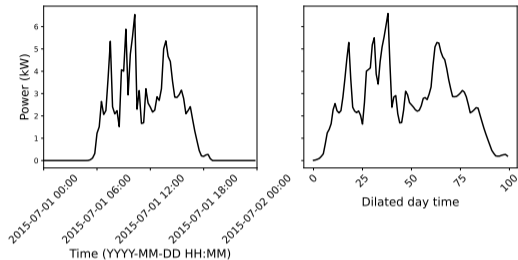


Time dilation

- ▶ identify daytime between estimated sunrise and sunset
- ▶ map each daytime interval to a fixed period M
- ▶ time dilation preserves the energy of the signal (*i.e.* same area under the curve)

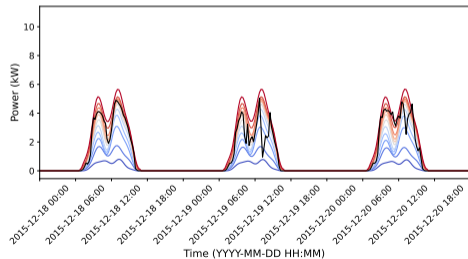
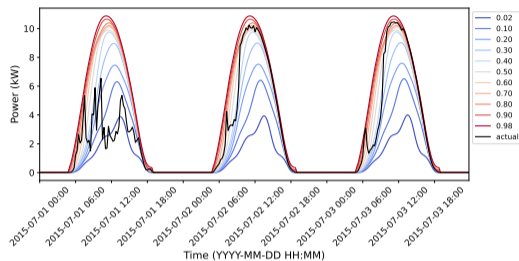


Time dilation (continued)



- ▶ shorter days are dilated more
- ▶ we can compare different days on equal footing

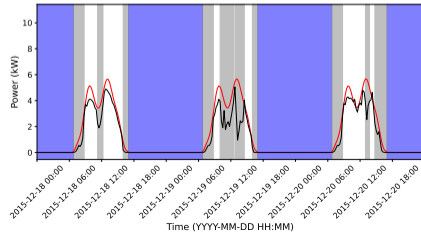
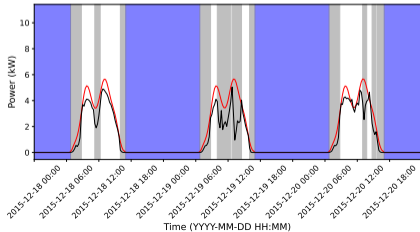
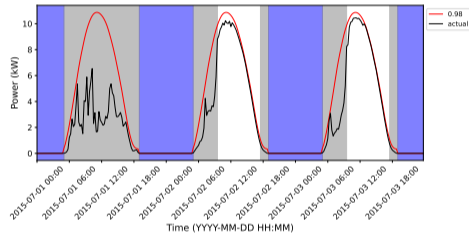
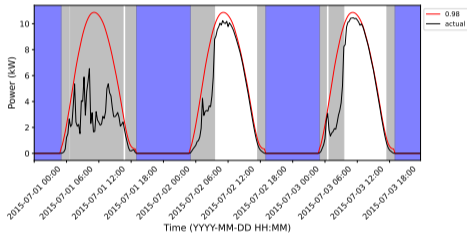
Smooth marginal quantiles



- ▶ estimating quantiles means estimating CDF at each time
- ▶ very powerful: estimate whole distribution from a single point!
- ▶ quantiles adapt to the seasonal variation
- ▶ quantiles can be used for various applications such as anomaly detection, missing data infill, etc.

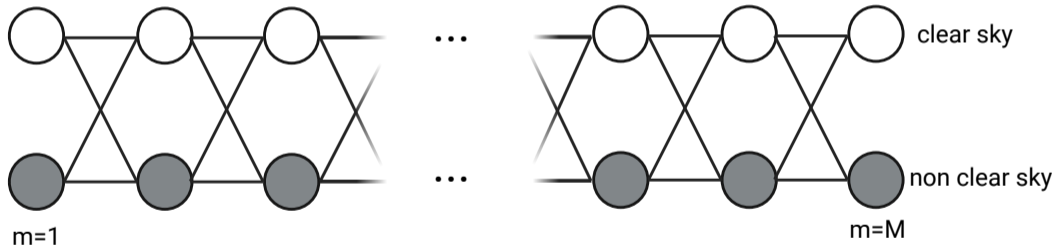
Clear sky detection

- ▶ naïve method: thresholding on clear sky index
- ▶ leads to many transitions between clear and non clear sky intervals
- ▶ introduce a post-processing step that smooths out clear sky estimate



Clear sky detection (continued)

- ▶ form graph: 2 nodes/interval (clear, non-clear)
- ▶ detected clear sky is the shortest path which minimizes transitions and mismatches
- ▶ solve for each day separately using dynamic programming ($\mathcal{O}(M)$)



Conclusion & next steps

- ▶ introduced a white-box ML model for power output of a *single* PV system
- ▶ showed how to estimate quantiles and detect clear sky intervals

- ▶ will start fleet scale analysis with *multiple* PV systems
- ▶ do joint Gaussianization and copulas
- ▶ include heterogeneous data (temperature, wind, etc.) in analysis

This material is based on work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number 38529.