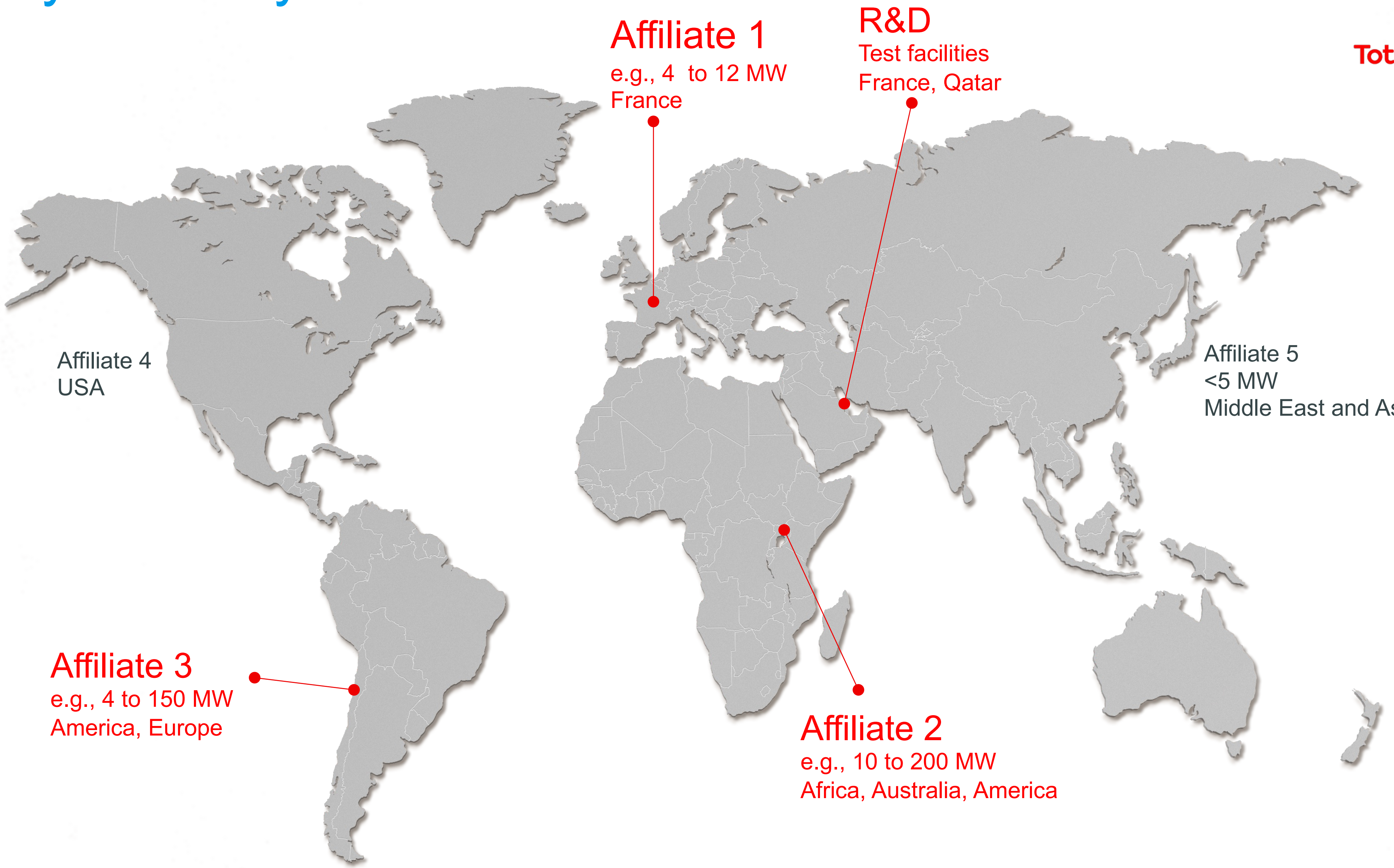
Towards a Digital Twin: strategy and research for the PV industry

EU PVPMC

Lluvia OCHOA, Nikola HRELJA, Julien CHAPON
9/11/2023

TotalEnergies

Diversity of PV systems



PV system
Location: France
Size: > 4 MW
Data quality: all

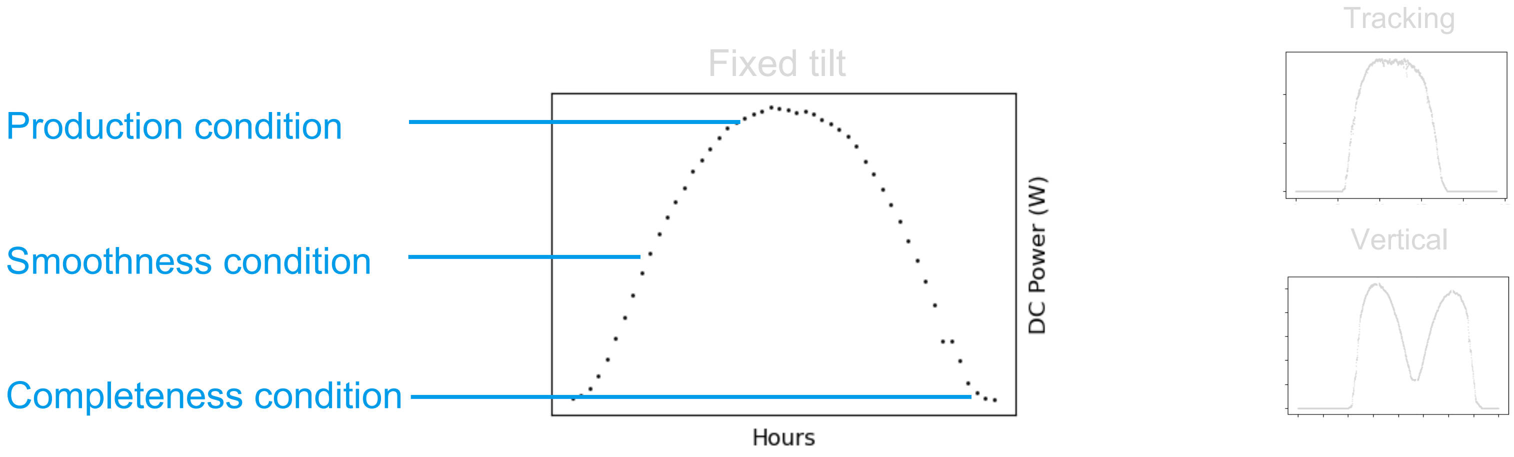
Method
Estimate production loss over the years including only **CSDs**. CSDs selected by the **shape of the daily power curves**. Estimation YoY per time segments (e.g., month, weeks)

Proof of concept
Inputs: power
Outputs: PLR
Aggregation: time-segment, then overall
Time range: more than 5 years
Filters: CSD filter

Scale-up potential
Needs weather data: no
Method complexity: easy
Time to validation: fast
Real-time: not applicable

PLR estimation without weather data

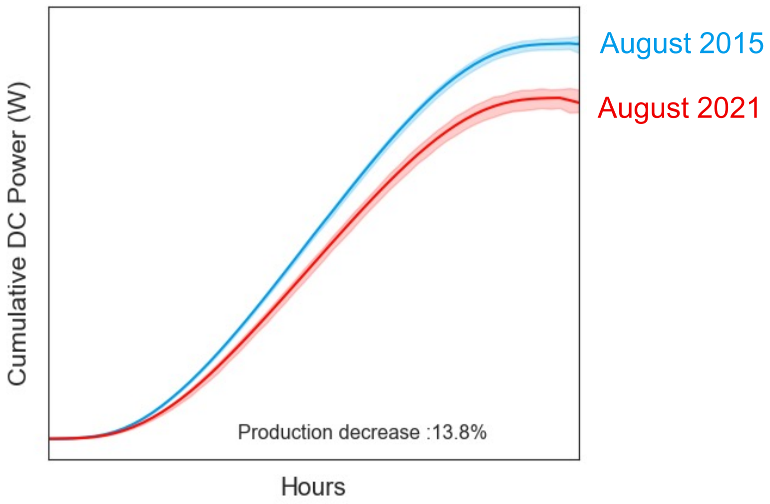
CSD filter



PLR estimation


Daily cumulative


Time-segment aggregation



CSD: clear sky day; PRL: Performance Loss Rate; YoY: year-on-year.
By L. Guillemot

PLR estimation without weather data

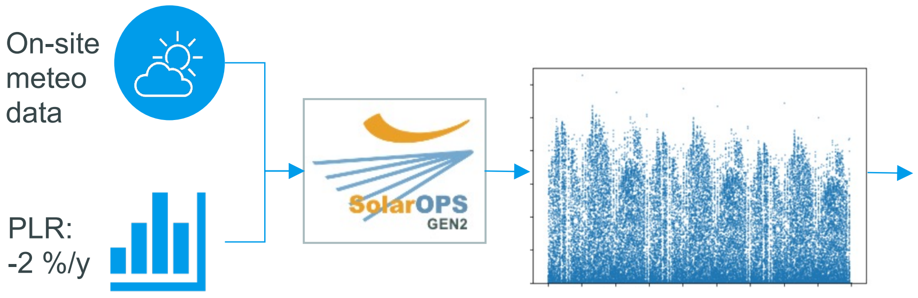
PV system
Location: France
Size: > 4 MW
Data quality: all

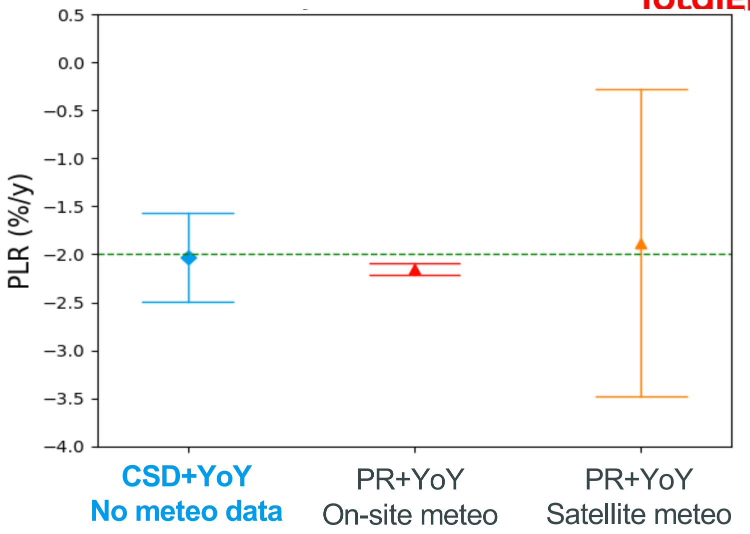
Method
Estimate production loss over the years including only **CSDs**.
CSDs selected by the **shape of the daily power curves**.
Estimation YoY per time segments (e.g., month, weeks)

Proof of concept
Inputs: power
Outputs: PLR
Aggregation: time-segment, then overall
Time range: more than 5 years
Filters: CSD filter

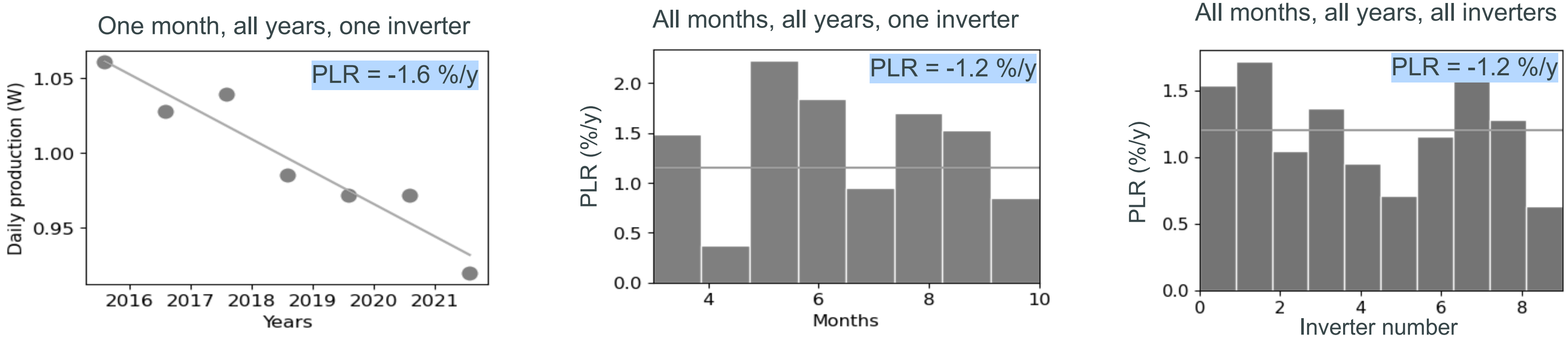
Scale-up potential
Needs weather data: no
Method complexity: easy
Time to validation: fast
Real-time: not applicable

Results: synthetic data





Results: Site #1



CSD: clear sky day; PRL: Performance Loss Rate; YoY: year-on-year.
By L. Guillemot

PV system

Location: Uganda
Size: 10 MW
Data quality: good, then bad

Method

Classify inverter malfunction events using **simple tests on the electrical data at inverter level**. Discriminate from events at transformer- and plant-levels to identify individual inverter faults.
Must be **fast and explainable**.

Proof of concept

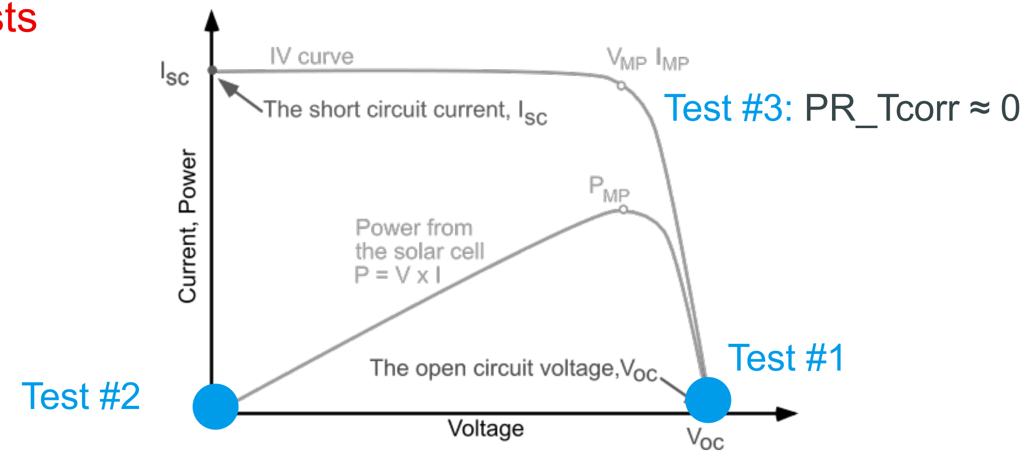
Inputs: P, I, U. GPOA (optional)
Outputs: Events of inverter faults
Aggregation: no
Time frequency: 1-5 min
Filters: outliers, missing values, day/night

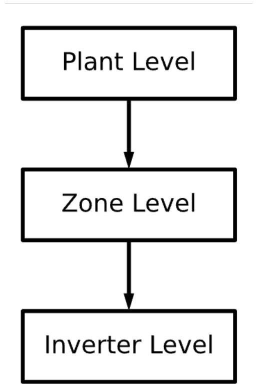
Scale-up potential

Needs weather data: optional
Method complexity: easy
Time to validation: fast
Real-time: yes

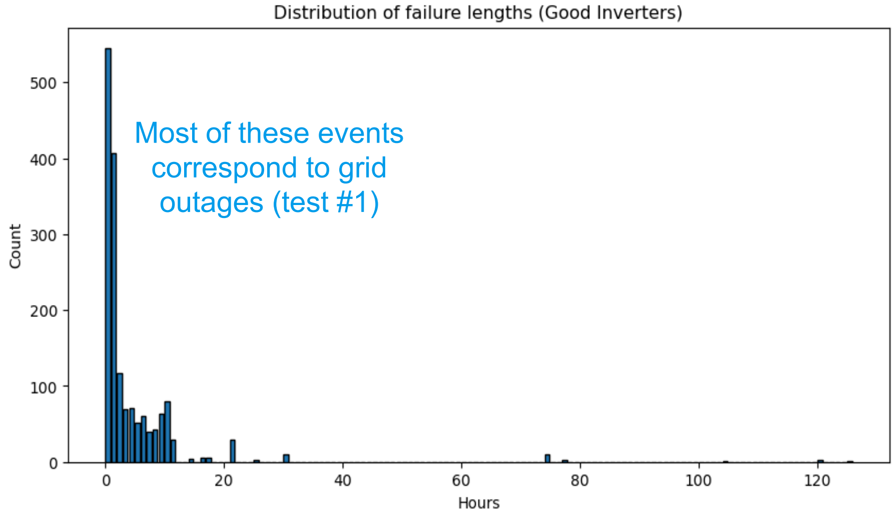
Inverter problems

Tests

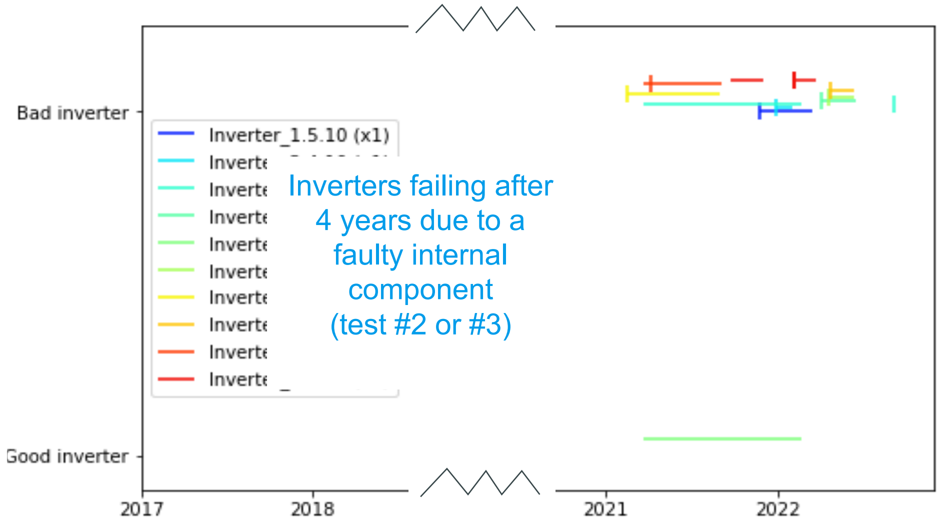




Results: Site #2



Detection of inverters' total breakdown



PR_Tcorr: temperature-corrected Performance Ratio
By: C. Becot, A. Zubair et al.

PV system
Location: France
Size: 4.5, 12 MW
Data quality: bad

Method
Inspection of PR_Tcorr.
Estimation of soiling rate using satellite data and **empirical dust deposition models**.

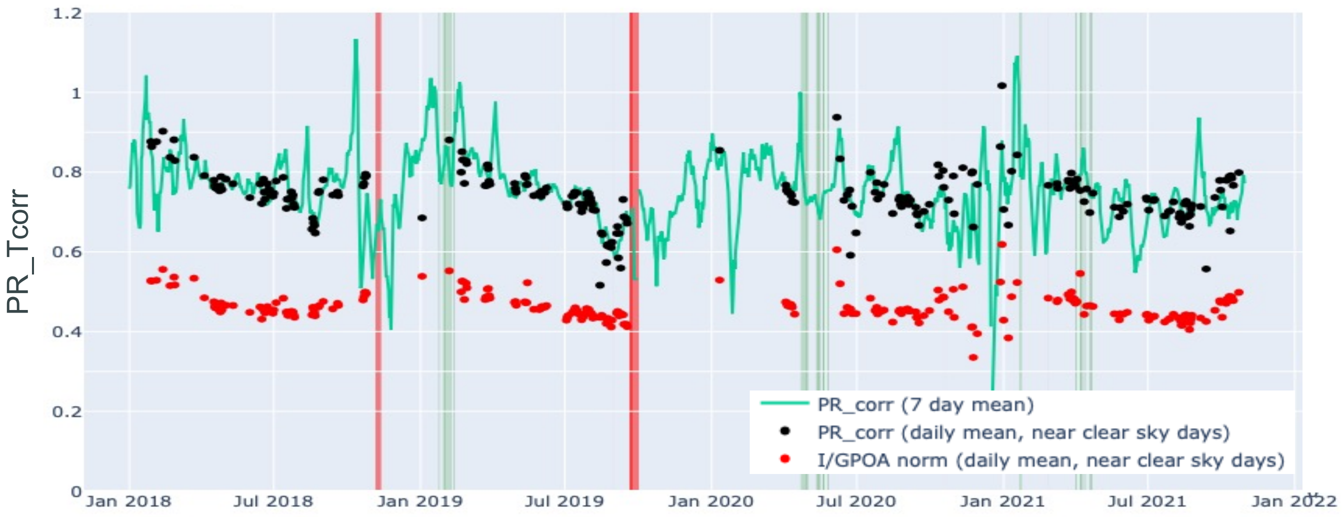
Proof of concept
Inputs: electrical data and atmospheric data
Outputs: PR_Tcorr, soiling rate
Aggregation: daily
Time range: at least one year
Filters: 10h-16h, irrad>400W

Scale-up potential
Needs weather data: yes
Method complexity: medium
Time to validation: long
Real-time: possible, probably not necessary

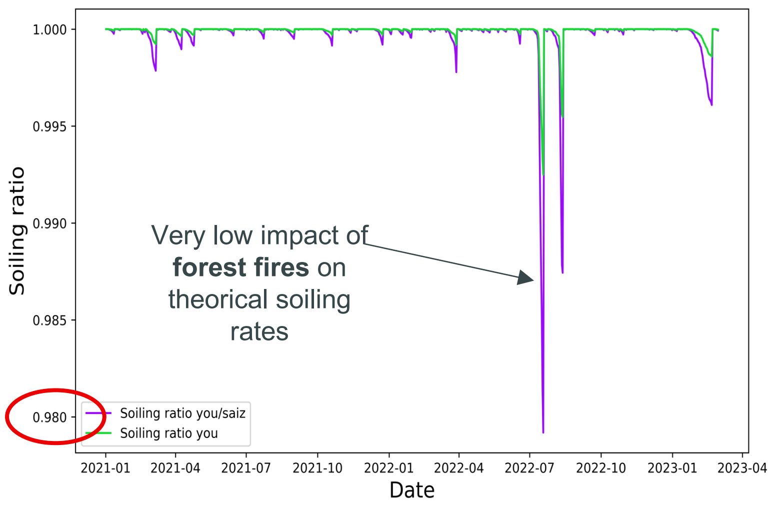
Soiling estimation

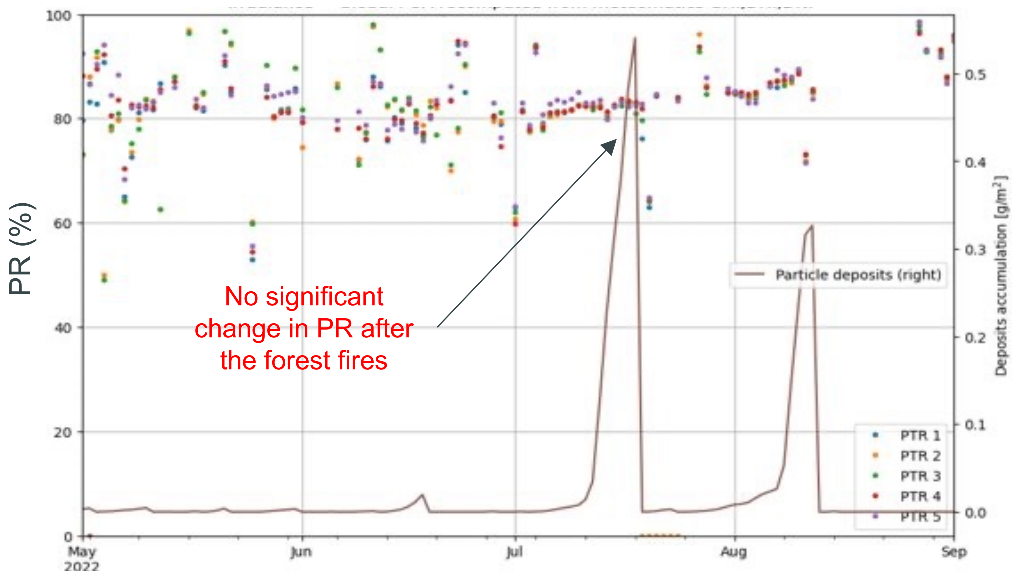
Site #1

Although a pattern can be observed, it cannot be corroborated due to poor data quality.



Site #3





PR_Tcorr: temperature-corrected Performance Ratio
By: C. Becot, F.Salmi, E. Le-Borgne, A. Toumiranta et al.

PV system

Location: Qatar
Size: test facility
Data quality: excellent

Method

Signal decomposition approach. Implementation of Python library solar-data-tools by B. Meyers

Proof of concept

Inputs: power
Outputs: soiling rate
Aggregation: no
Time frequency: 1-20 min
Filters: no

Scale-up potential

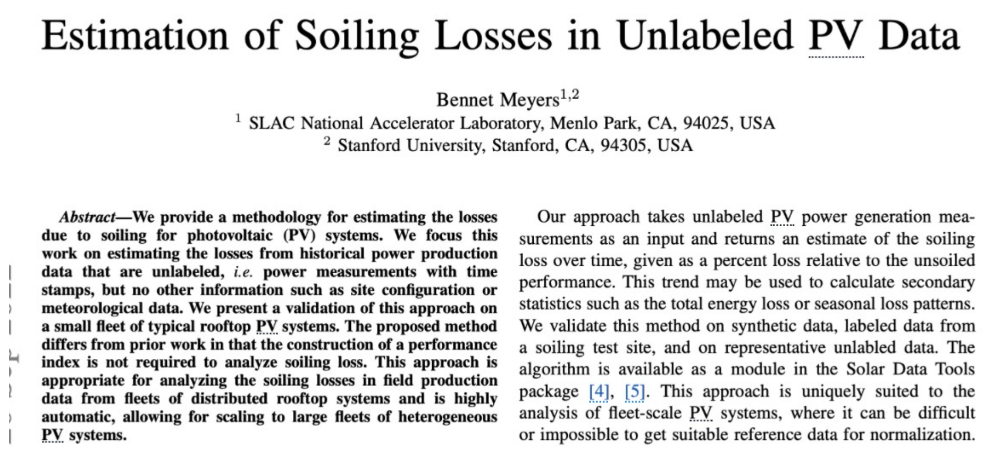
Needs weather data: no
Method complexity: difficult
Time to validation: long
Real-time: not necessary

Soiling estimation

TotalEnergies

Test facility #1

- String cleaned every 2 months
- Reference cell cleaned weekly
- Real soiling ratio estimated with string and reference cell
- Severe soiling
- PVInsight works very well



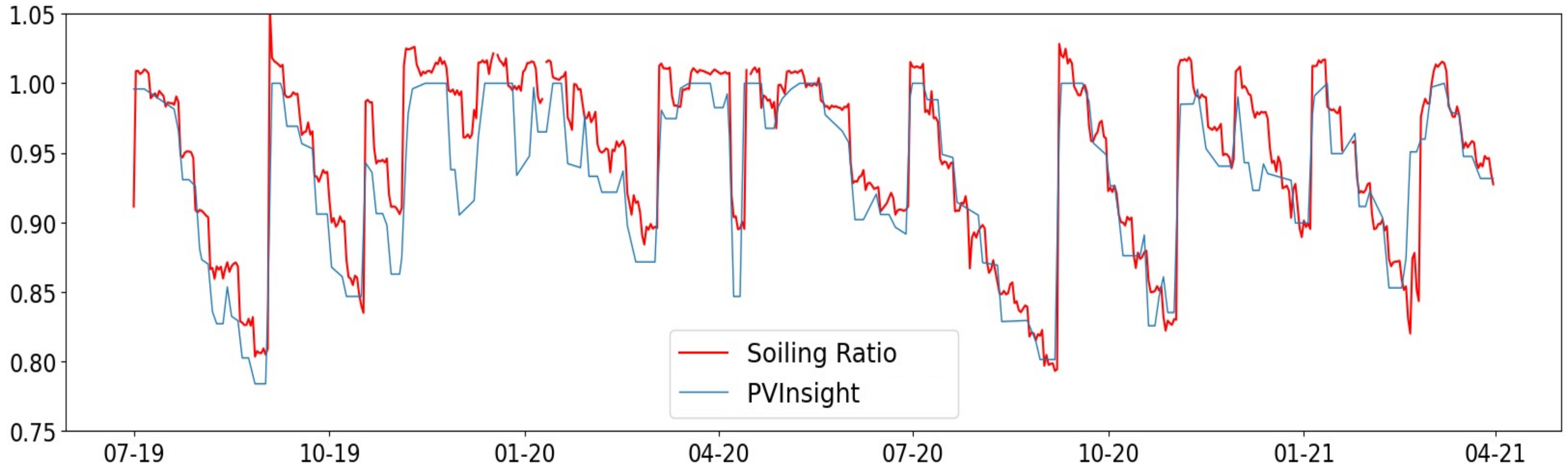
Estimation of Soiling Losses in Unlabeled PV Data

Bennet Meyers[1,2]
[1] SLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA
[2] Stanford University, Stanford, CA, 94305, USA

***Abstract*—We provide a methodology for estimating the losses due to soiling for photovoltaic (PV) systems. We focus this work on estimating the losses from historical power production data that are unlabeled, *i.e.* power measurements with time stamps, but no other information such as site configuration or meteorological data. We present a validation of this approach on a small fleet of typical rooftop PV systems. The proposed method differs from prior work in that the construction of a performance index is not required to analyze soiling loss. This approach is appropriate for analyzing the soiling losses in field production data from fleets of distributed rooftop PV systems and is highly automatic, allowing for scaling to large fleets of heterogeneous PV systems.**

Our approach takes unlabeled PV power generation measurements as an input and returns an estimate of the soiling loss over time, given as a percent loss relative to the unsoiled performance. This trend may be used to calculate secondary statistics such as the total energy loss or seasonal loss patterns. We validate this method on synthetic data, labeled data from a soiling test site, and on representative unlabled data. The algorithm is available as a module in the Solar Data Tools package [4], [5]. This approach is uniquely suited to the analysis of fleet-scale PV systems, where it can be difficult or impossible to get suitable reference data for normalization.

Results



By M. Povéda et al.

PV system
Location: Uganda
Size: 10 MW
Data quality: good, then bad

Method
For each time interval (e.g., week, month) solve an **optimization** problem to find the parameters of the diode model that best fit the **MPP** or OC data. Using all the time segments, evaluate the **evolution** of these parameters over time.

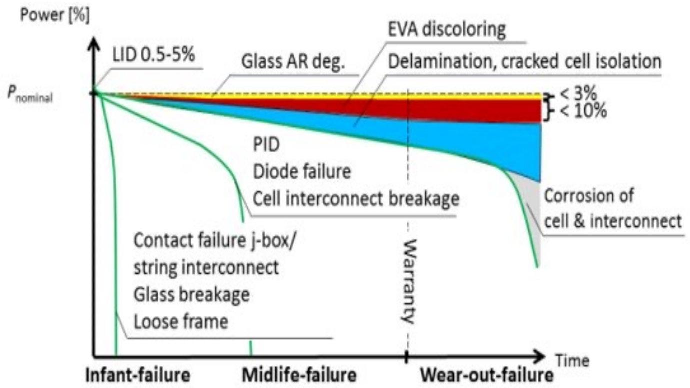
Proof of concept
Inputs: MPP data or OC data.
Outputs: Isc, Io, Rs, Rsh, etc
Aggregation: daily, then time segment
Time range: years
Filters: in progress

Scale-up potential
Needs weather data: yes
Method complexity: difficult
Time to validation: long
Real-time: not applicable

IV Curve parametrization

TotalEnergies

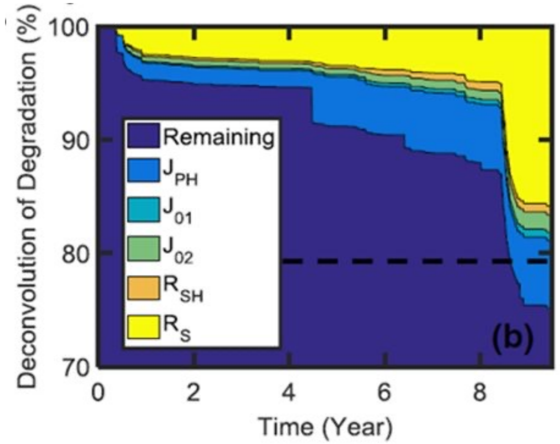
The Suns-Vmp method (Sun, X. et al., 2009)



MPP data

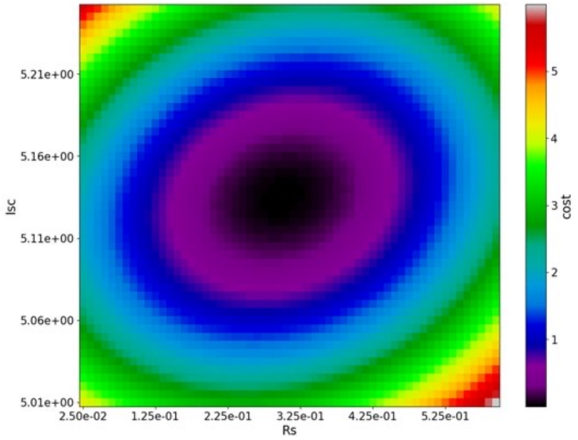
Optimization
For each time segment

Model parameters

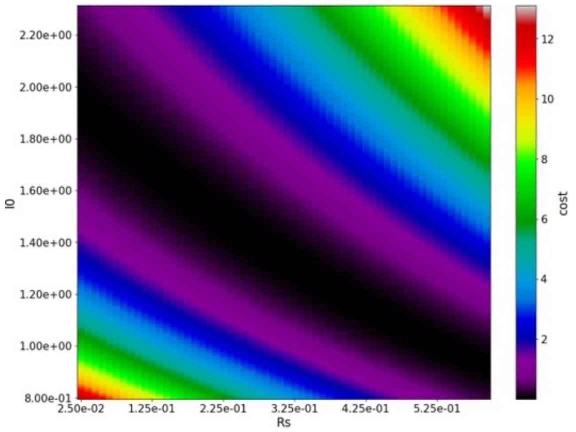


Preliminary results Site #2

Single solution
R_s with I_{sc} is identifiable



Multiple solutions
R_s with I_0 non-identifiable



By N. Hrelja, N. Harder, C. Becot

Sun, X., Chavali, R.V.K., Alam, M.A., 2019. Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power point—the Suns-Vmp method. Prog. Photovoltaics Res. Appl. 27, 55–66. https://doi.org/10.1002/pip.3043

Many PV sites, many solutions

1. Performance Loss Rate estimation
2. Identification of inverter problems
3. Soiling estimation
4. IV curve parametrization: source causes of degradation
5. Clipping and curtailment
6. Machine learning-based anomaly detection (supervised, unsupervised)

Integrated Power

Renewables	Flexible generation	Storage	Trading	Customers
First quartile target Strengthen renewables **industrialization**	CCGTs to **complement renewable** production	Manage **intermittency** to capture value: BESS, hydro	**Capture volatility** and maximize asset value	**Deliver clean firm power to large B2B** Supply B2C & EV charge

Integration to deliver clean firm power

> 100 TWh production by 2030

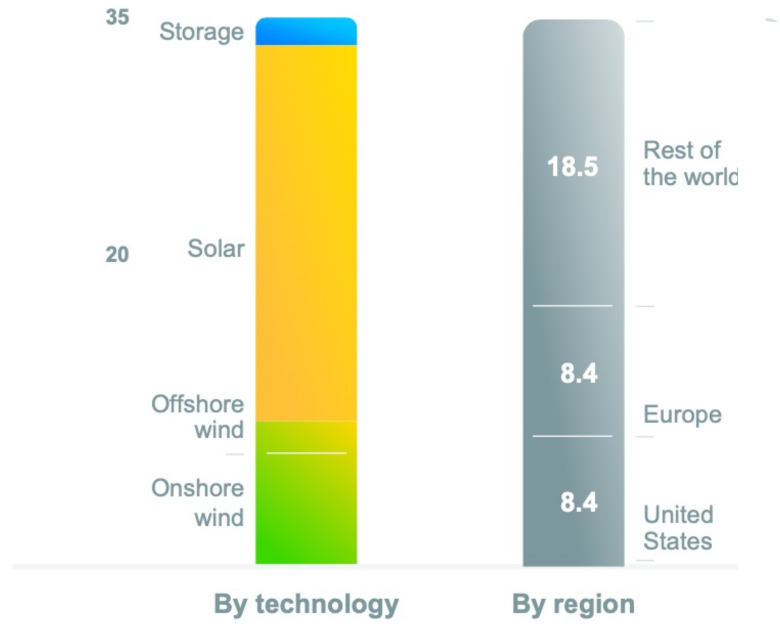
A secure portfolio of 35 GW by 2025

Gross installed capacity

GW

Breakdown

GW



Recommendations to Business Units

Benchmark of commercial products and open-source Python libraries

Benchmark #1:
Monitoring platform A

- Meant for production, operations and maintenance
- Useful KPIs, functionalities to manage PV solar fleet
- Inaccurate in terms of performance analytics

Benchmark #2:
Python Library B

- Accurate in terms of performance analytics
- No need of weather data
- Only usable by Python programmers
- Currently dependent of an external commercial solver

TotalEnergies

- Internal R&D tools
- Same level of accuracy than external options
- Next developments can be prioritized according to the needs of other Business Units
- On-going work to enhance data quality assessment

Benchmark #3:
Assessment of 9 commercial monitoring platforms (by N. Ngoun)

Benchmark #5:
Evaluation of Python Library C (in progress, by N. Hrelja)

Benchmark #4:
Assessment of 5 commercial monitoring platforms (by D. Roisse, outside R&D team)

Standardization

TotalEnergies

Data quality grading

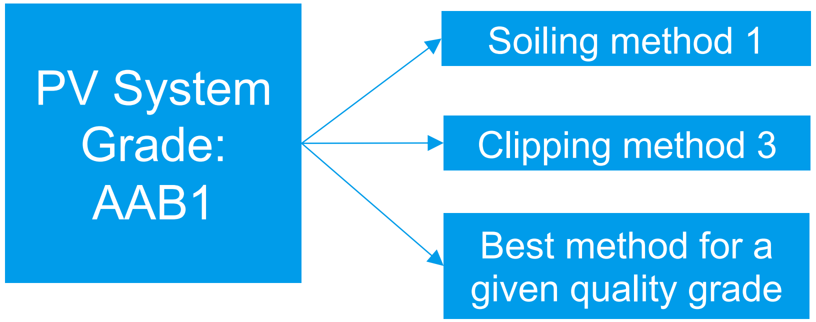
Temporal availability grading. Taken from Lindig et al. 2022

Letter Grade	Outliers (%)	Missing percentage (%)	Longest Gap (days)
A	Below 10	Below 10	Below 15
B	10–20	10–25	15–30
C	20–30	25–40	30–90
D	Above 30	Above 40	Above 90
P/*F*	Time series $\geqslant$ 24 months = > Pass		

Future work: Spatial availability grading

Number grade	Electrical data	Level	On-site meteo data
1	P, I, U	String	GHI, Tamb
2	P, I, U	String	No
3	P, I, U	Inverter	GHI, Tamb
N	P	Plant	No

Future work: selection of the best method depending on data quality



Ultimate goal = Decision-making tool for PV plant diagnosis

Lindig, S., Theristis, M., Moser, D., 2022. Best practices for photovoltaic performance loss rate calculations. Prog. Energy 4, 022003. https://doi.org/10.1088/2516-1083/ac655f

Standardization

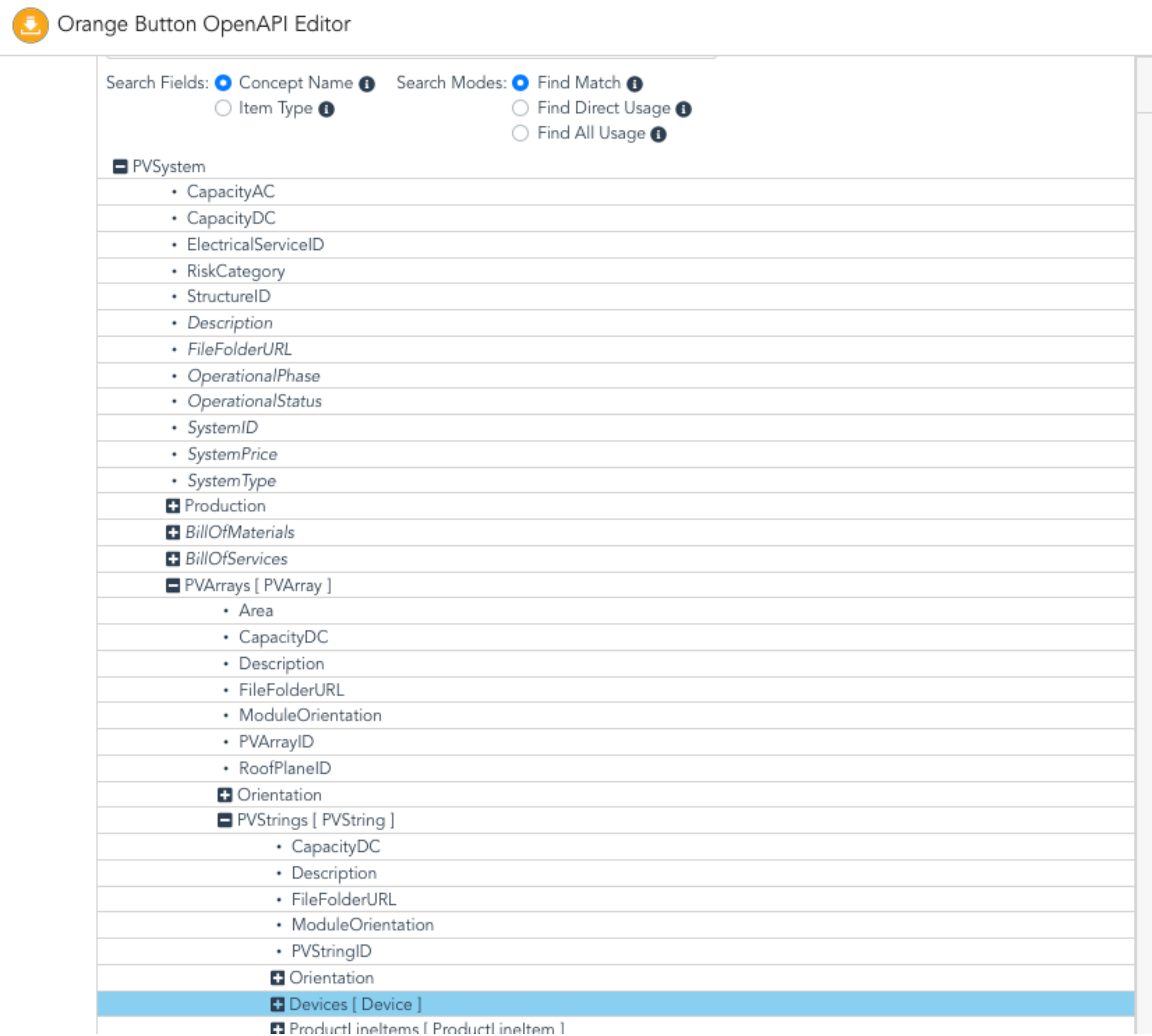
Data standards

- MINES Paris https://libinsitu.readthedocs.io/en/latest/

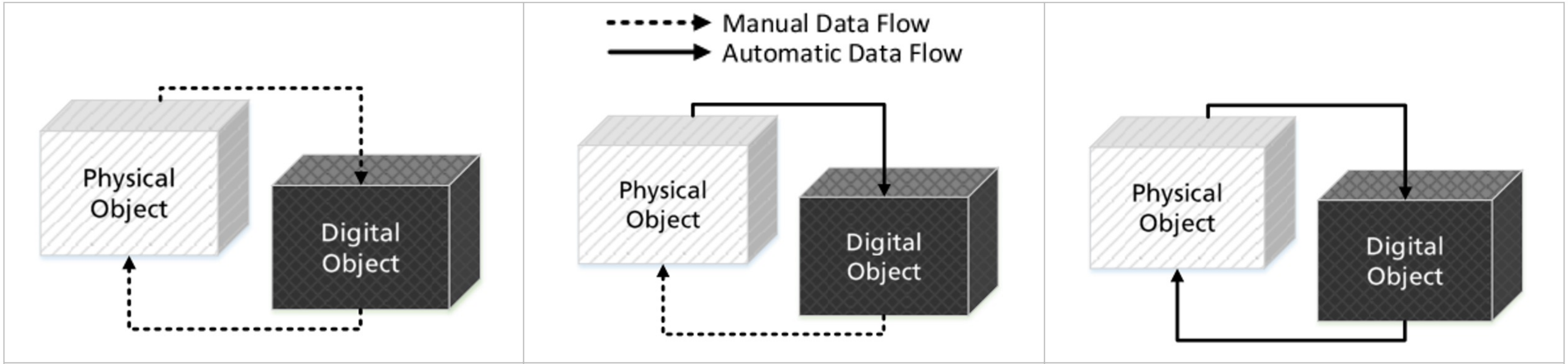
- Orange Button https://myorangebutton.com

- FAIR

- Massel, L., Shchukin, N., Cybikov, A., 2021. Digital twin development of a solar power plant. E3S Web Conf. 289, 03002. https://doi.org/10.1051/e3sconf/202128903002



Ambition: a true Digital Twin



a) Digital Model | b) Digital Shadow | c) Digital Twin

We are here

”As complex as needed, as simple as possible”

Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification, in: IFAC-PapersOnLine. pp. 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474

Conclusion

TotalEnergies

- Algorithms
- Recommendations
- Standardization
- Digital Twin
- Sharing knowledge

Federated learning?
Open sourcing?

Thank you