Rte Le réseau de transport d'électricité

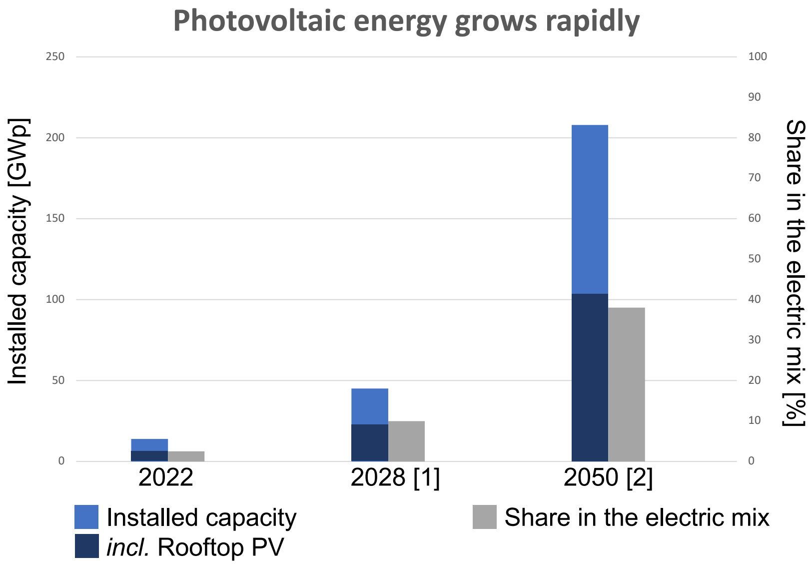
MINES PARIS | PSL

Enhancing regional PV power estimation using physics-based models, solar irradiance data and deep learning

Gabriel Kasmi, Augustin Touron, Philippe Blanc, Yves-Marie Saint-Drenan, Maxime Fortin, Laurent Dubus

1 Background and motivation

Rooftop PV lacks observability

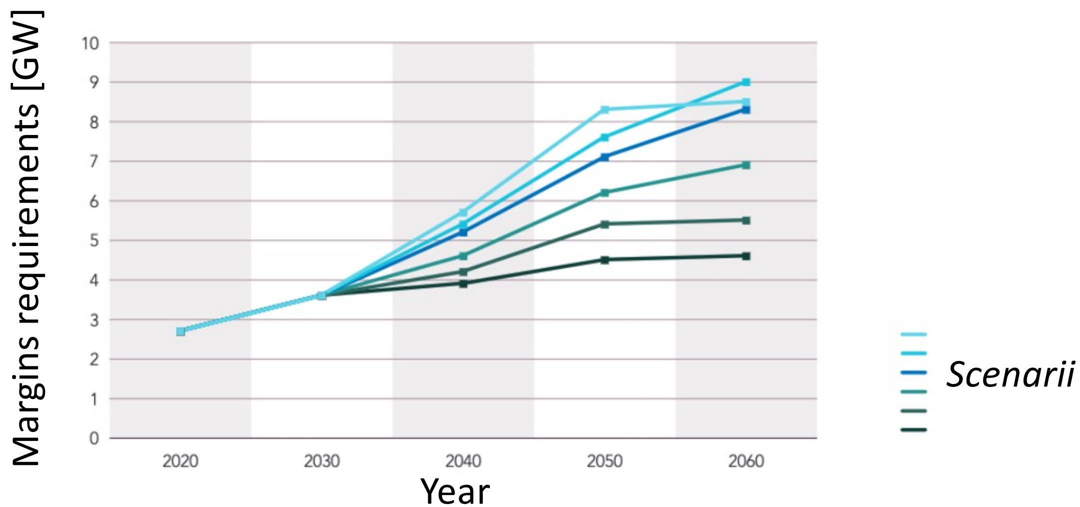


[1] *Programmation pluriannuelle de l'énergie (PPE)*
[2] Energy Pathways to 2050 (2022), RTE France

Rooftop PV: « **invisible » PV** (Shaker et al 2015).

➔ Increased **congestions**, **overgeneration**, **imbalances** (Pierro et al 2022).

Foreseen increase in margins requirements mostly due to the lack of observability of rooftop PV



How can we improve rooftop PV observability?

The two limitations to improving observability

1. TSOs need precise measurements or estimates of the production of the rooftop PV fleet

➔ *Correct parameters for rooftop PV*

2. TSOs need data to evaluate the accuracy of the the estimations

➔ *Ground truth production curves of rooftop PV*

How to address these limitations

➔ Three steps:

(i) Reliable *characterization* of the PV installed capacity

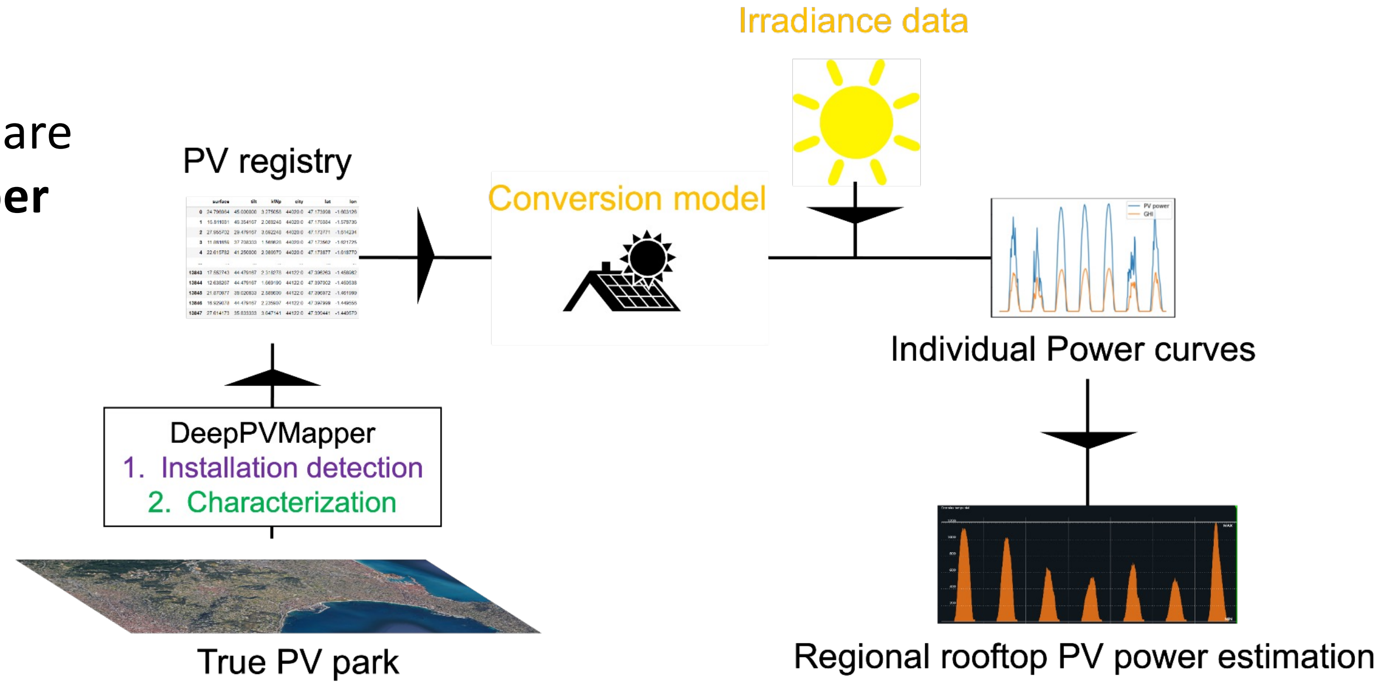
(ii) Integrate this characterization for improving PV power estimation models

(iii) Evaluate the accuracy of the estimation method on ground truth power generation measurements

2 Estimating rooftop PV power production

A new model for estimating rooftop PV power production

- Rooftop PV characteristics are collected by **DeepPVMapper** (Kasmi et al, 2022)
- Four sources of errors:
  1. Detection
  2. Characterization
  3. Conversion model
  4. Irradiance data



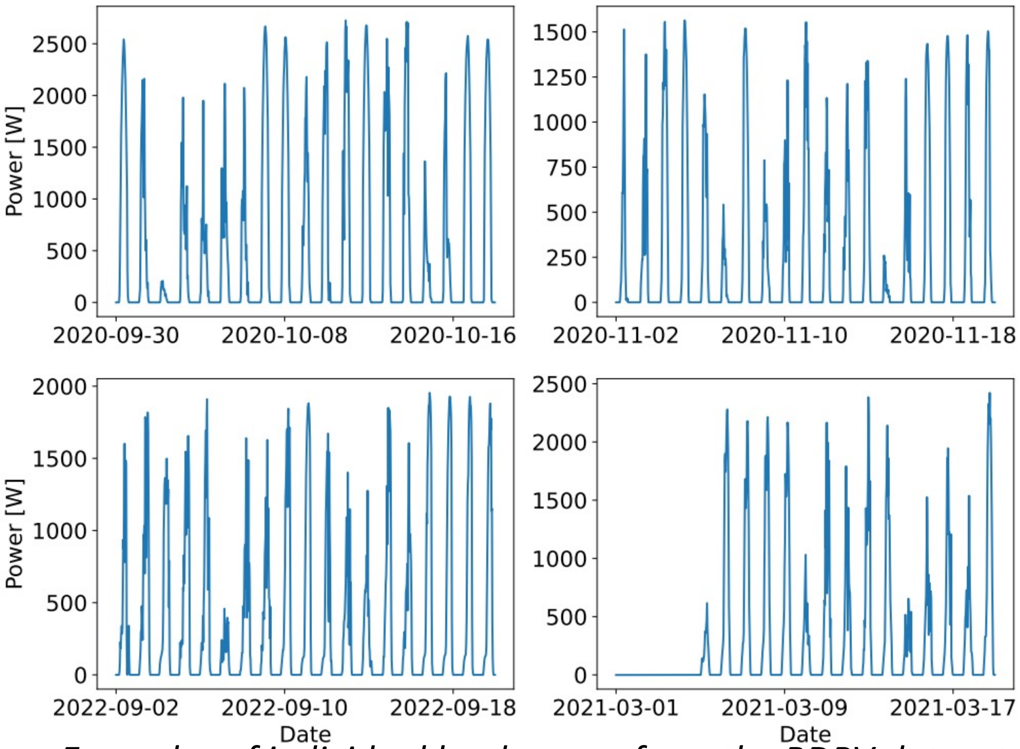
Flowchart of the proposed approach

3 Assessing the validity of our approach

Validation data: the BDPV dataset

- We need of reference records (i.e., **individual production curves and not total energy records**) of the rooftop PV production.

- We gathered data **from 1000 installations spread over France** to construct our reference. This data comes from the non profit ''asso BDPV''

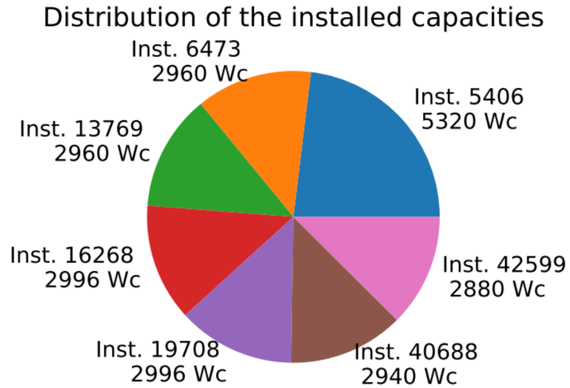


Examples of individual load curves from the BDPV dataset

Representative cells

- We extract representative cells (convex hulls) from the BDPV dataset.
- These cells are representative of the industrial cases of RTE.

Example of a representative cell (9,087 km²)

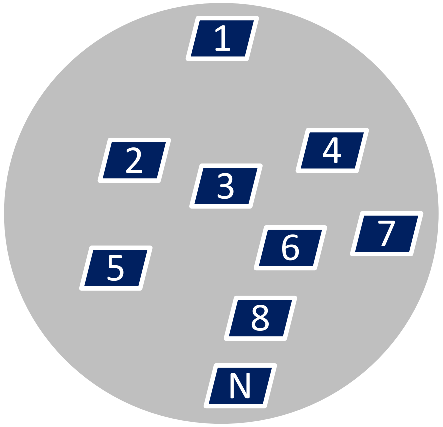


Distribution of the installed capacities

4 Analysis of the detection error

Methodology

What happens if we detect too few or too many installations?



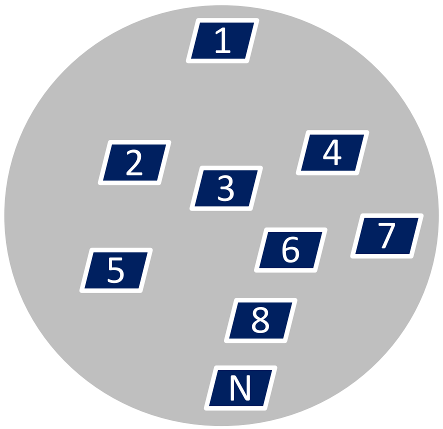
We focus on the power curves only

Each cell contains N installations

We define **n_ref < N** the number of installations that will compose the **reference** fleet (which will be approximated)

Methodology

What happens if we detect too few or too many installations?



1. Define the **reference** power curve and the approximation

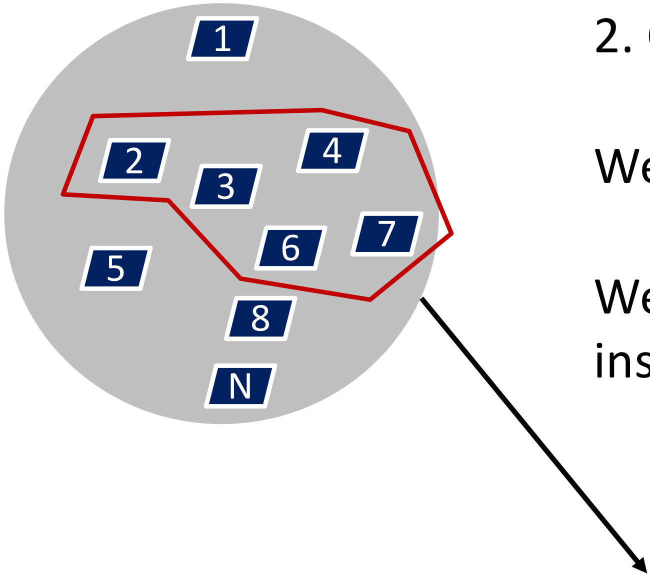
We compute the powerset of N:

$$P(N) = \{\{1\}, \{1,2\}, \{1,2,3\} \dots \}\backslash\{\emptyset\}$$

The references are all the subsets that contain **n_ref** installations

Methodology

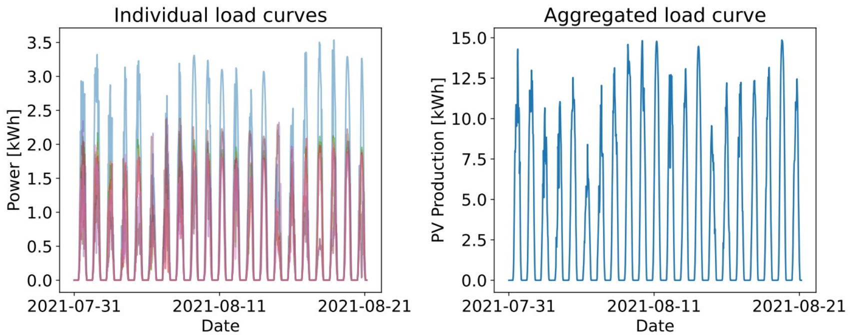
What happens if we detect too few or too many installations?



2. Compute the reference power curve

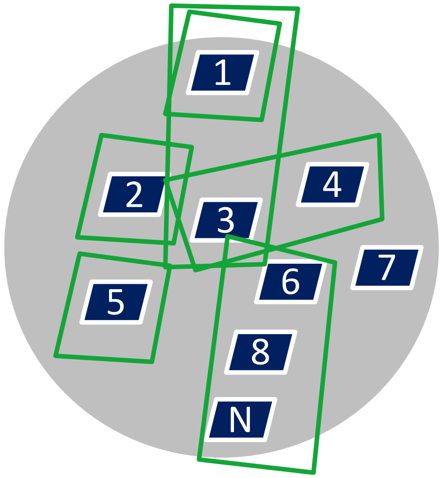
We compute the installed capacity

We aggregate the power curves of n_ref installations



Methodology

What happens if we detect too few or too many installations?



3. Approximate the reference power curve with an increasing number of installations

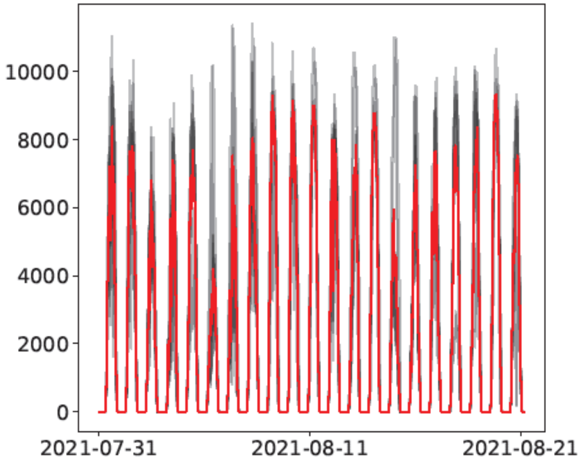
We approximate the reference power curve with **n_sample** installations, **n_sample** varies between 1 and N

We compute the aggregated power curve and rescale it

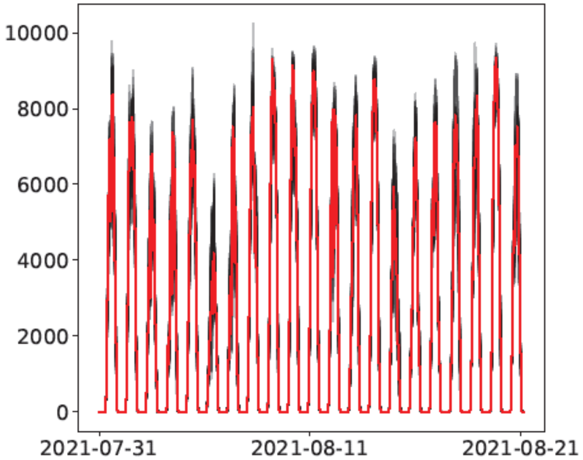
We evaluate the accuracy with the RMSE

Methodology

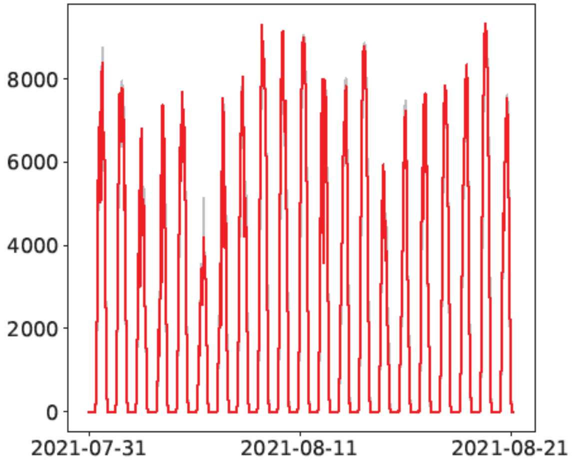
Increasing number of installations in the sample



n_ref = 4, n_sample = 1

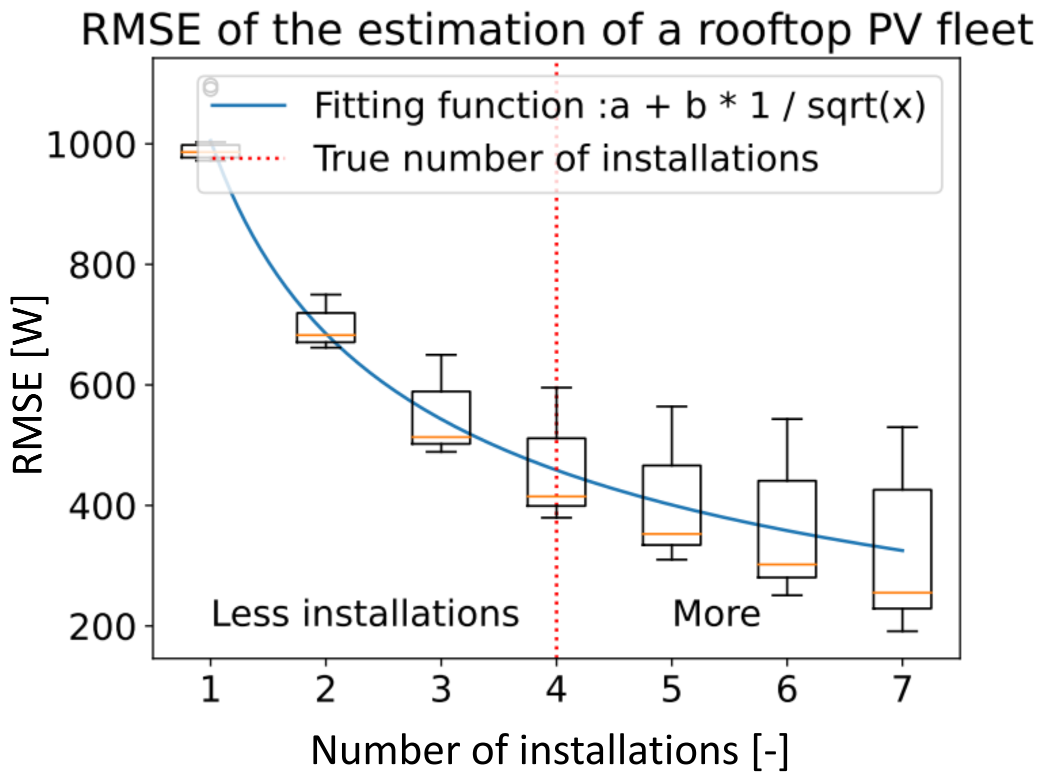


n_ref = 4, n_sample = 4



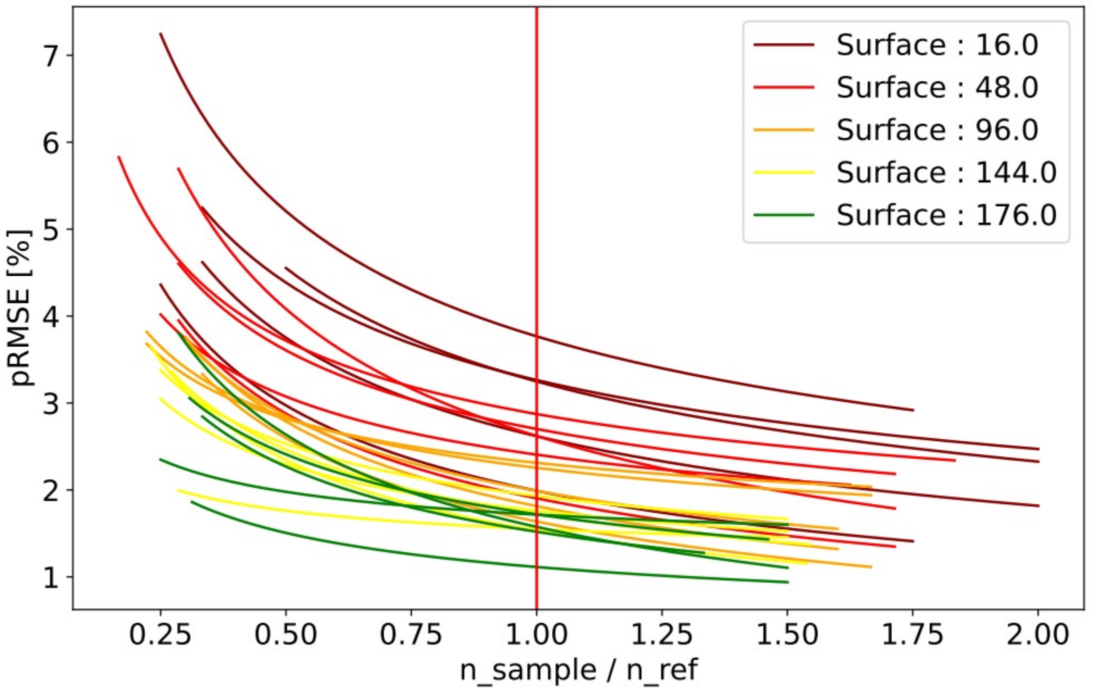
n_ref = 4, n_sample = 7

Results I



RMSE decreases with n_samples

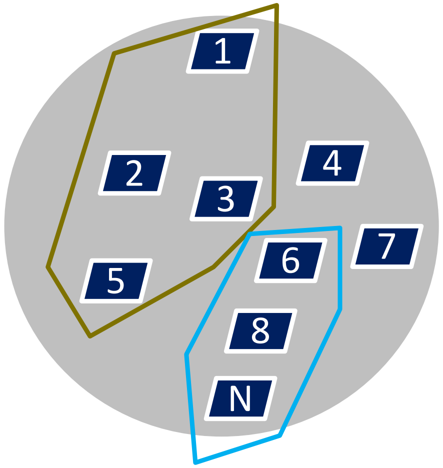
Results II



RMSE decreases with N

Analysis of the detection error

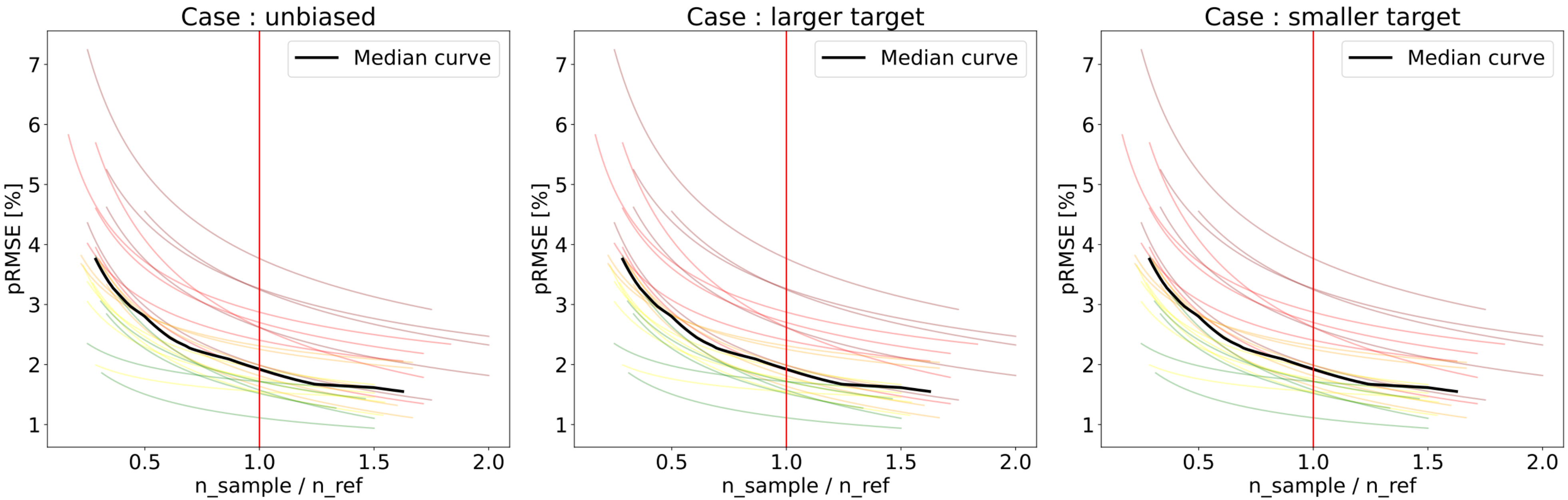
What happens if we detect too few or too many installations?



What if we slightly bias the reference ?

We define the reference as the largest or the smallest installations of the reference cell.

Results III



Our results hold even if there is a slight bias between the reference and the approximation

What's next ?

- We will use our representative cells to evaluate the characterization and modelling errors

- We can also deploy the current estimation method and evaluate the gain of our approach.

Check-out our work!

A crowdsourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata. Scientific Data, 10(1), 59.

Towards unsupervised assessment with open-source data of the accuracy of deep learning-based distributed PV mapping. In *Workshop on Machine Learning for Earth Observation (MACLEAN), in Conjunction with the ECML/PKDD 2022*.

Paper

Git

Thank you!

Questions ?

Not all PV are measured equally

Photovoltaic energy

Plants ($\geq 1\ MWp$)

Measurable : RTE has access to real time measurements of the production

Decentralized $< 1\ MWp$)

Small plants

Rooftop PV

Observable : RTE can reconstruct the PV production from *ex post* measurements

- **No measurement** of the production (either real time or *ex post*)
- **No comprehensive and exhaustive registry** (i.e. with sufficient information to estimate the production)

Proposed approach

- Cluster the installations detected by DeepPVMapper by Voronoi cell and compute the aggregated installed capacity.
- Compute an aggregated PV power curve using the characteristics, a conversion model and irradiance data
- Rescale the power curve using the true installed capacity computed from the RNI.
- Outcome: aggregated production curves for each substation.

Voronoi cells around substations

Definition of the sources of errors

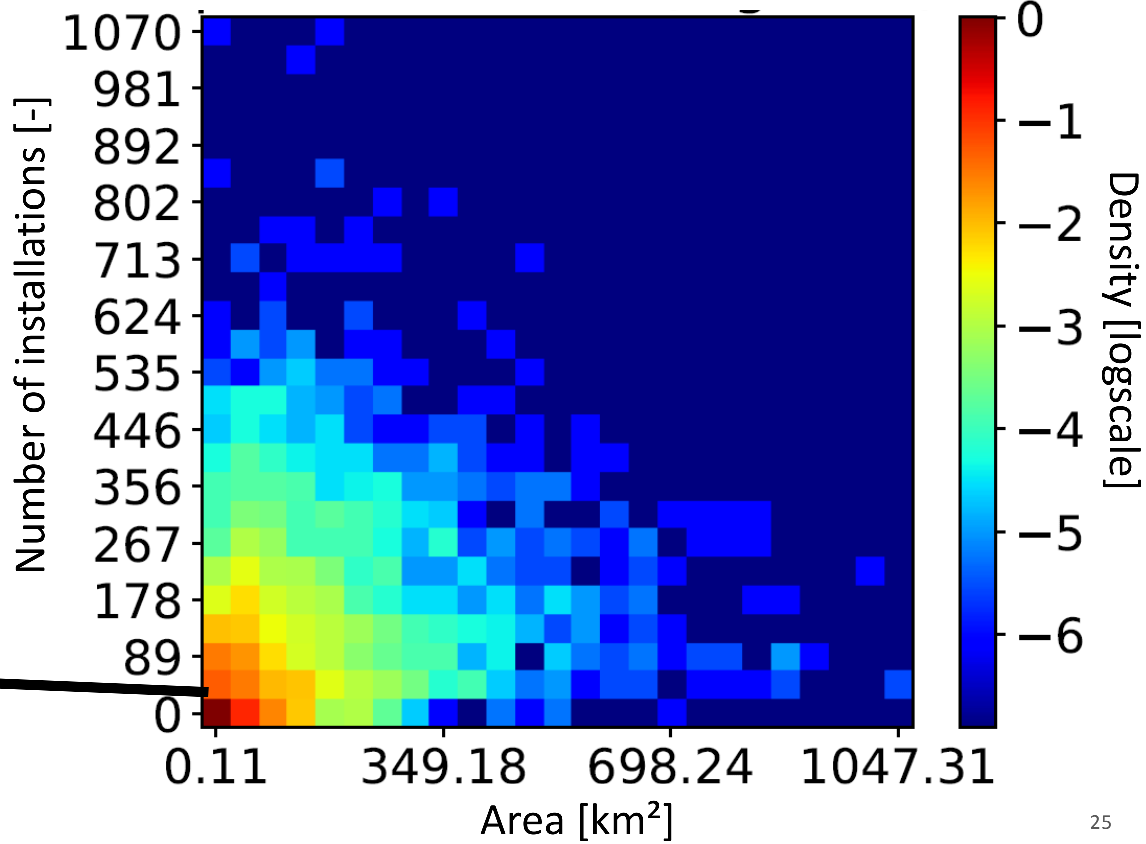
- We decompose the error as follows:
  - ✓ **Detection** error: what happens if we underestimate or overestimate the number of installations in the Voronoi cell?
  - ✓ **Characterization** error: what happens if the characteristics are wrong?
  - ✓ **Modelling** error: residual error (conversion model and irradiance data), estimated by evaluating the mismatch between a simulation and the true load curve

Deriving representative cells

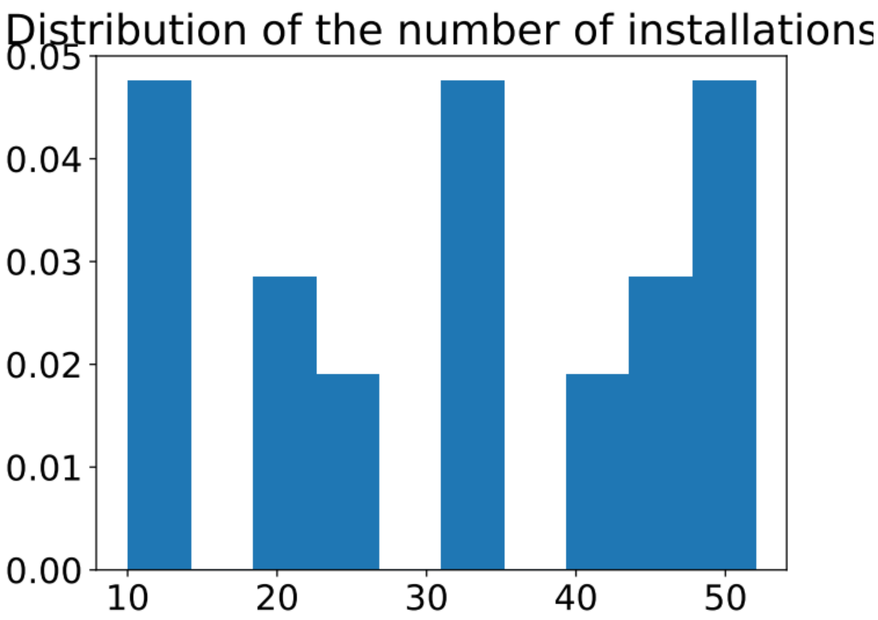
- We estimate the density of installations in the Voronoi cells using the *Registre national d'installations*
- We look for subsets of the BDPV dataset that has similar characteristics: it corresponds to <u>representative cells</u>

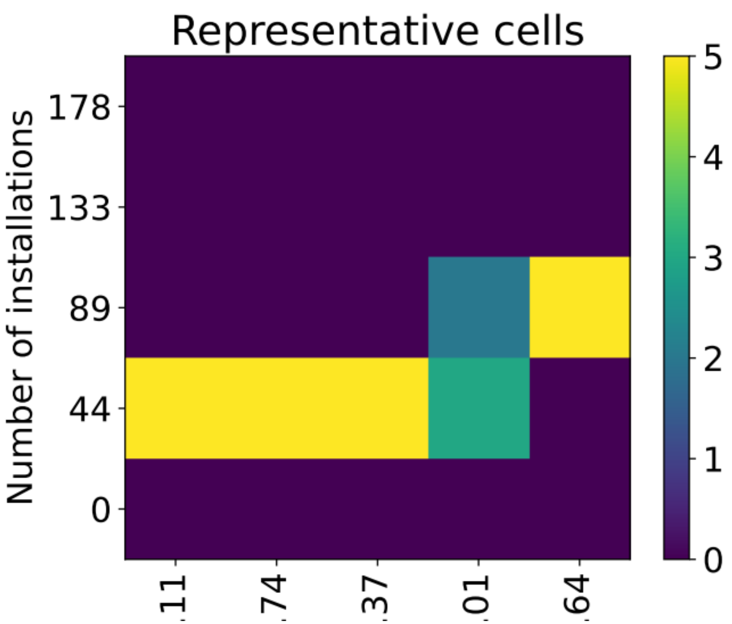
Installations of a representative cell (9km²)

Density of installations per voronoi cell (logscale)



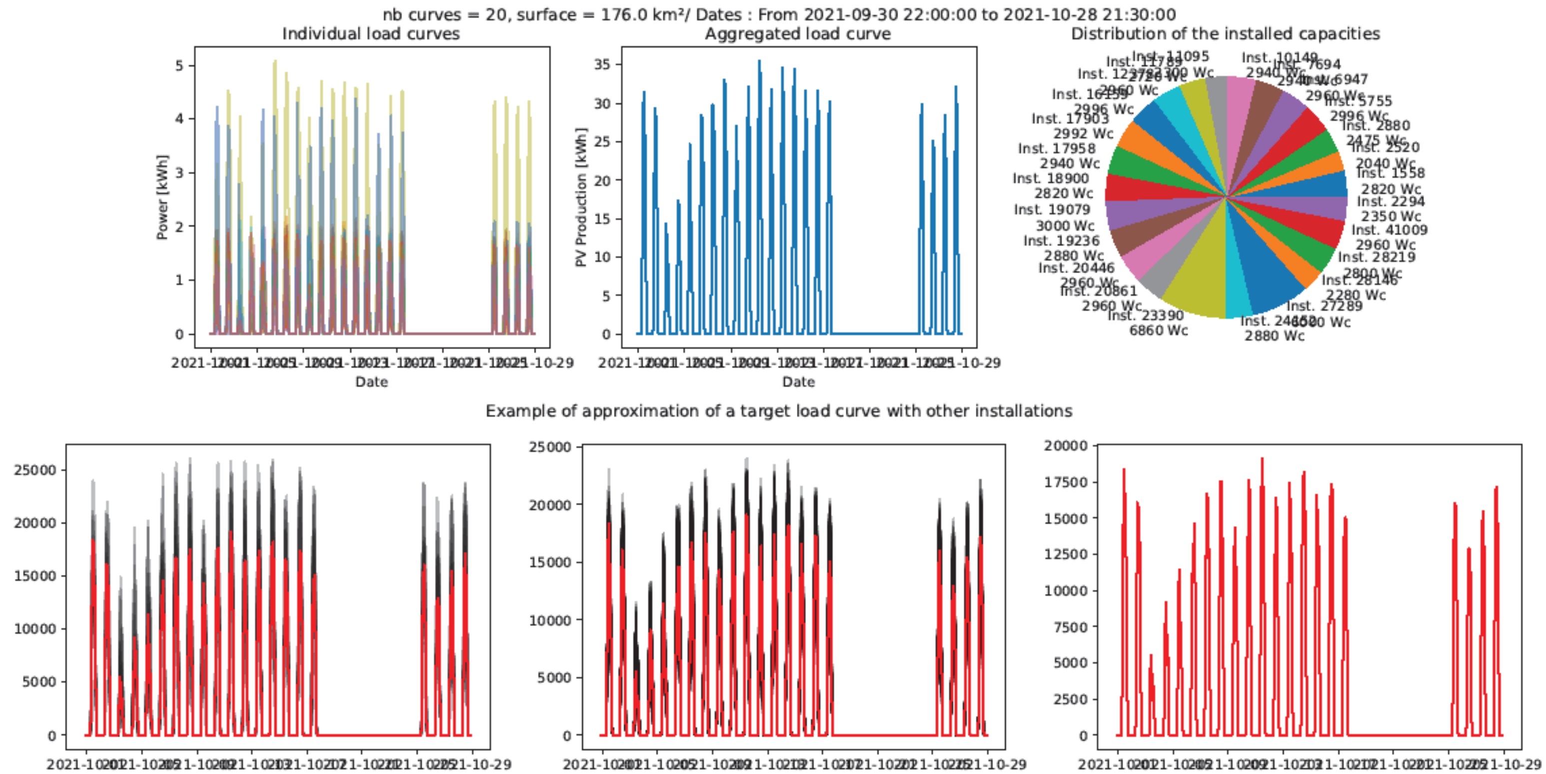
Representative cells





Analysis of the detection error: oversampling is less harmful than undersampling

Large fleet of installations



Analysis of the detection error: oversampling is less harmful than undersampling

We can see that the **fit is better with more installations**, even if we oversample the PV fleet