



**HAL**  
open science

## Leveraging Artificial Intelligence to Improve the Integration of Photovoltaic Energy into the Grid

Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, Philippe Blanc

► **To cite this version:**

Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, Philippe Blanc. Leveraging Artificial Intelligence to Improve the Integration of Photovoltaic Energy into the Grid. 2024. hal-04846823

**HAL Id: hal-04846823**

**<https://hal.science/hal-04846823v1>**

Preprint submitted on 18 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Transition Institute 1.5

L'ambition d'une véritable transition

## WORKING PAPER

#3 - novembre 2024

### Leveraging Artificial Intelligence to Improve the Integration of Photovoltaic Energy into the Grid<sup>1</sup>

GABRIEL KASMI<sup>a,b</sup>, LAURENT DUBUS<sup>b,c</sup>,

YVES-MARIE SAINT-DRENAN<sup>a</sup>, PHILIPPE BLANC<sup>a</sup>

<sup>a</sup>Mines Paris, Université PSL, Centre Observation Impacts Energie

<sup>b</sup>RTE France

<sup>c</sup>WEMC (World Energy & Meteorology Council, UK)

<sup>1</sup> This working paper summarizes the motivation, methodology, and results of G. Kasmi's PhD thesis, supervised by P. Blanc, Y.-M. Saint-Drenan and L. Dubus, and defended in April 2024. RTE France, the French Transmission Operator, funded this thesis and benefited from a subsidy from the ANRT (CIFRE thesis funding).



DOI : 10.23646/9nev-py65

## **Abstract**

In December 2023, the French photovoltaic (PV) installed capacity stood at 19 GWp. The French electricity transmission system operator (TSO) lacks power measurements for 20% of the fleet, which mostly correspond to small-scale (rooftop) systems. In the context of the rapid decarbonization of the electric mix, the PV installed capacity will continue to experience sustained growth in the coming years, and the so-called problem of poor PV observability threatens its long-term integration into the grid due to the uncertainties it creates. A better knowledge of the rooftop PV fleet, embodied in a nationwide technical registry recording the localization and characteristics of the PV installations, is necessary to improve PV observability. This working paper discusses how artificial intelligence (AI) can be reliably used to construct such a registry to improve the integration of rooftop PV into the grid.

## **Keywords**

Deep learning, Interpretability, Robustness, Reliability, Photovoltaic Energy, Observability

# 1. Introduction

## 1.1. Context: the lack of precise information on rooftop photovoltaic systems

In the current organization of the electric system, the transmission system operator (TSO) is responsible for the balance between the load and supply of electricity. It requires access to real-time measurements of electric production. While such measurements are available for thermal, wind and hydropower plants; photovoltaic (PV) energy is characterized by a great diversity of systems, ranging from large plants metered in real-time to small-scale, distributed systems. These systems, generally located on rooftops and whose installed capacity is lower than 36 kW<sub>p</sub>, amount to 22% of the PV installed capacity (in 2023), and their electric production is not metered by the TSO. The lack of measurements of the rooftop PV power production is referred to as the *lack of observability of the rooftop PV power production*.

To meet the decarbonization goals, the share of electricity in the energy supply needs to increase *and* to resort to low-carbon sources massively. According to the Intergovernmental Panel on Climate Change (IPCC), these energies, especially wind and solar energy, offer the highest potential for reducing CO<sub>2</sub> emissions by 2030 [1]. Therefore, renewable energy sources, in general, and PV energy, in particular, are currently experiencing sustained growth. In France, for example, the overall PV installed capacity could reach up to 200 GW<sub>p</sub> in 2050 [2], compared to "only" 18.6 GW<sub>p</sub> as of November 2023. Figure 1 summarizes the current PV installed capacity, its share in the electric mix, and the goals and forecasts according to the public authorities and RTE, the French TSO.

However, if the observability of rooftop PV remains the same, such projected installed capacities could have dire consequences for the electric grid: inaccurate measurements of the PV power production could lead to issues in grid management such as overgeneration or imbalances [4].

The lack of PV observability generally comes with poor knowledge regarding the rooftop PV fleet due to the lack of systematic registration of the individual PV systems' size and technical characteristics in many countries [5,6]. In the recent years, numerous works leveraged artificial intelligence and orthoimagery to construct registries of small-scale PV systems [5,7,8]. Among them, the DeepSolar project [9] was an important milestone as it was among the first work to construct a large-scale database of PV systems recording the surface and the localization of the systems.

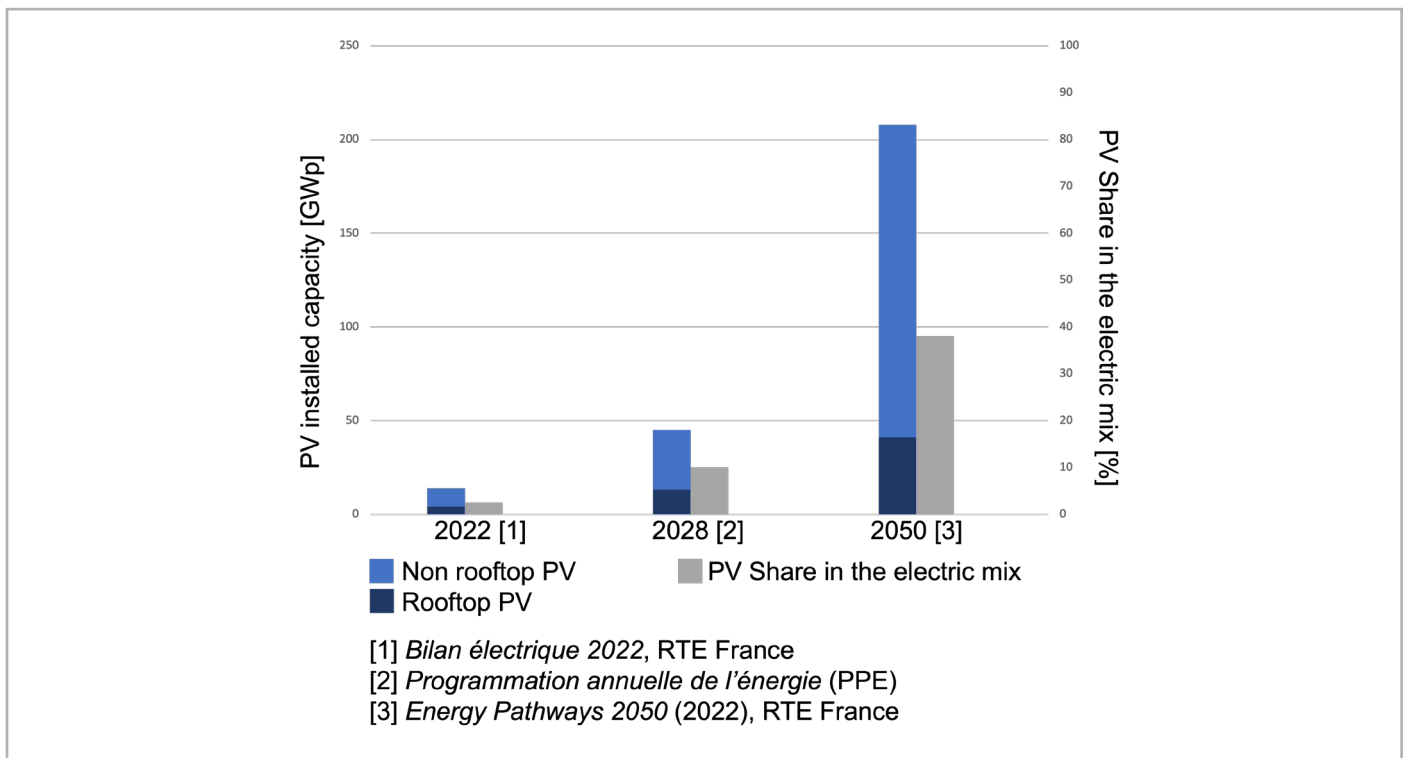


Figure 1: Expected PV share growth according to the Programmation Pluriannuelle de l'Énergie (PPE [3] and RTE's Energy Pathways 2050 [2]).

## 1.2. The lack of reliability of existing rooftop PV mapping techniques

Unfortunately, we cannot directly apply DeepSolar to France as we need more information on PV systems than the sole surface area. As our goal is to estimate the PV power production, we need the installed capacity, tilt, and azimuth angles of the PV systems, which are the minimal necessary information to estimate the PV power production [10].

The fundamental limit to deep learning-based methods such as DeepSolar is their sensitivity to *distribution shifts* [11], which leads to an unpredictable loss in accuracy if we were to apply the algorithm straightforwardly to France. This sensitivity to distribution shifts results in a lack of reliability of deep learning methods.

This working paper presents the results of Kasmi (2024) [12] on how we can improve the reliability of artificial intelligence (AI) methods, namely deep learning algorithms, to map rooftop PV systems. We present a framework intended to improve the reliability of such algorithms and highlight the mild conditions that need to be satisfied to make deep learning-based remote sensing on orthoimagery a suitable method for constructing a nationwide registry of rooftop photovoltaic (PV) installations intended to improve the observability of PV power production in France.

The remainder is organized as follows: section 2 presents the data used to construct the PV registry and the existing data sources on PV systems in France, section 3 presents and discusses the three pillars to enforce a robust AI system, section 4 assesses the relevance of the PV registry for improving rooftop PV observability and section 5 concludes.

# 2. Data: available sources and needs

## 2.1. Geographical information system (GIS) data

The main data source used in this work is the database BDORTHO<sup>2</sup> of the IGN [13], the French National Institute of Geographical and Forest Information (IGN). The BDORTHO contains aerial orthorectified imagery, covering the whole French territory. Images are updated every three years on a rolling basis (each year, a third of the territory is updated), and the images have a ground sampling distance (GSD) of 20cm/pixel.

We leverage the BDTOPO of the IGN, in addition to the BD ORTHO. It records the location of the buildings, roads, and other infrastructure as geolocalized polygons. This database is referred to as a shoebox model as buildings are registered as 2D geolocalized polygons and is useful to merge PV polygons located on the same rooftop. Finally, we had access to numerical surface models. These models indicate the elevation of the ground and can therefore enable us to compute the tilt angle of rooftops. The IGN gradually releases the database LiDAR HD<sup>3</sup> [14], which is accurate enough to compute the tilt and azimuth angles of individual houses. Unfortunately, the LiDAR HD was not available for the whole of France by the time the PhD thesis was completed.

## 2.2. Existing PV data

**The Registre national d'installations (RNI).** The National registry of installations (RNI) is the reference database regarding electric production facilities. This data is aggregated by RTE and is accessible on the *Open Data Réseau Energies* portal<sup>4</sup> [15]. Due to privacy reasons, only city-wise aggregations of PV systems with an installed capacity below 36 kWp is publicly accessible. Therefore, the RNI records, for each city, the total number of systems and their cumulative installed capacity. It is updated every three months.

**RTE's internal data.** RTE has access to the disaggregated data regarding PV systems below 36 kWp. However, this database only features the city and the installed capacity of the PV system and does not contain any information regarding the technical characteristics of the systems.

<sup>2</sup> BD ORTHO data is accessible at the following URL: <https://geoservices.ign.fr/bdortho>.

<sup>3</sup> LiDAR HD data is accessible at the following URL: <https://geoservices.ign.fr/lidarhd>.

<sup>4</sup> This portal is accessible here: <https://opendata.reseaux-energies/>.

**Crowdsourcing data.** The thesis relied on data provided by the non-profit association *Base de données photovoltaïque* (BDPV)<sup>5</sup>. This association proposes to PV system owners to input the characteristics of their installation to see whether it works normally or not. BDPV eventually gathered a detailed database of 28,000 PV systems with a comprehensive set of technical characteristics and their precise location. We also had access to PV power time series for 1,700 PV systems. Finally, one can mention the data coming from OpenStreetMap (OSM). OSM is a collaborative project aiming at constituting an open-source geographical database. Most PV power plants, and some rooftop PV systems are registered in OSM as geolocalized polygons.

### 2.3. Training dataset: the BDAPPV database

**Table 1: Overview of the data records of the training dataset BDAPPV**

Provider	Total number of samples	Positive samples (Share [%])	Number of installations linked with installations characteristics (Share [%])
Google	28,807	13,303 (46.18)	8,019 (27.84)
IGN	17,325	7,686 (44.36)	3,658 (21.11)

We needed training data before deploying a model over France to map PV systems. To this end, we leveraged the database of BDPV and set up crowdsourcing campaigns to annotate images of PV systems. This work led to the training database *Base de données d'apprentissage profond photovoltaïque* (BDAPPV, [16]).

This dataset contains annotated images of 28,000 PV panels in France and neighboring countries. This dataset also proposes annotations of images that depict the same PV panel but from two different image providers: images coming from the Google Earth Engine [17] and from the IGN [13]. We have double annotations for more than 8,000 PV systems. The dataset also contains the technical characteristics of the PV systems. Table 1 summarizes the number of samples and the share of positives images featured in BDAPPV.

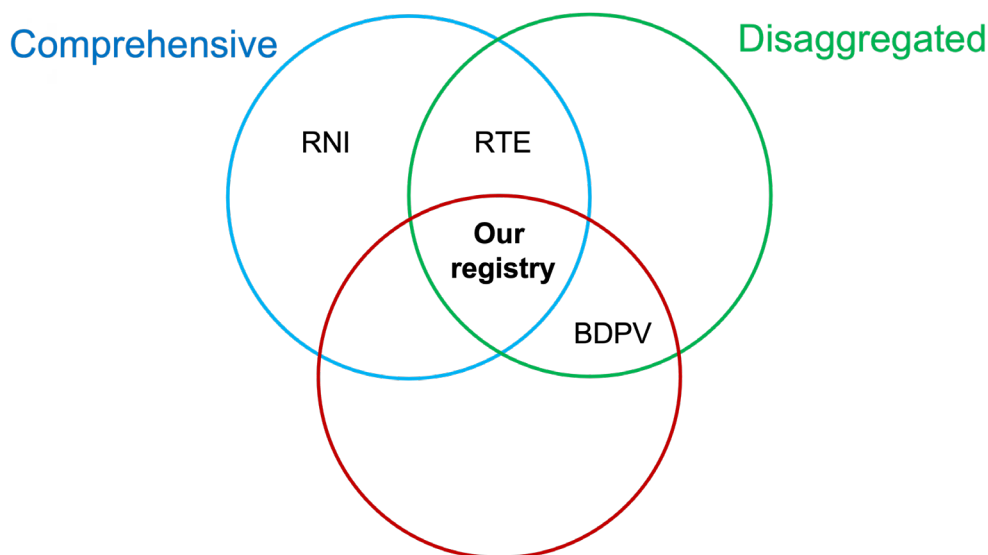
<sup>5</sup> Website: <https://asso.bdpv.fr>.

## 2.4. Data needs for the PV registry

The main end goal of our PV registry is to estimate the PV power production of small-scale systems. This estimation is based on a physical model of the PV system and requires a limited set of characteristics. Saint-Drenan *et al.* [10] showed that the tilt and azimuth angles, the installed capacity and the localization were sufficient for a satisfactory estimation of the PV power production using solar irradiance data and a physical model.

The aim of the registry is to gather this data at the scale of France. Therefore, it should satisfy three main criteria: it should be as comprehensive as possible, disaggregated and record the technical characteristics of the PV systems.

None of the data sources mentioned in section 2.2. simultaneously satisfies these three criteria. Figure 2 illustrates our requirements and the extent to which existing data sources satisfy them. Our registry must satisfy the three criteria *simultaneously*.



**Technical characteristics: tilt, azimuth and installed capacity**

Figure 2: Venn Diagram summarizing the data requirements of the PV registry.



# 3. Three pillars for a reliable AI

The framework for improving the reliability of deep learning-based algorithms is based on the following three pillars:

- Monitoring: As deep learning models can fail unpredictably, it is necessary to monitor their accuracy over the operational design domain (ODD, [18]), i.e., the data on which they are used in production. This evaluation is at most indirect, as by definition, there is no labeled data on this domain.
- Auditing: We argue that it is necessary to gain intuition on how models work to be able to understand when they might fail and what the practitioner can expect from them. To this end, we leverage explainable AI (XAI) techniques. These methods enable us to understand how the model works, e.g., by shedding light on the important areas of an image for making a prediction.
- Robustness: When deployed in production, the input data will be subject to some alterations. While it seems complicated to improve the robustness in general, we can at least identify the primary source of variability in the input data and improve the model's invariance to such alterations.

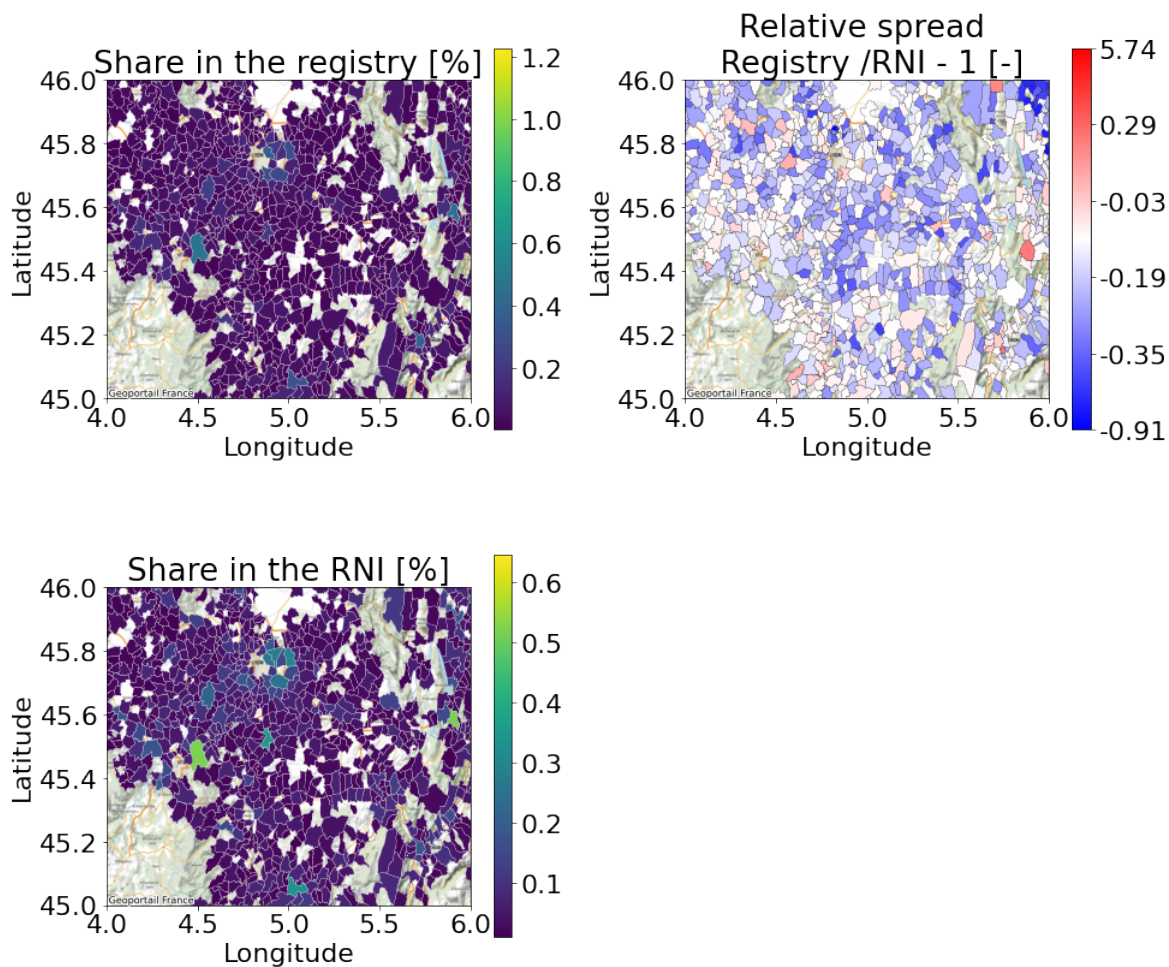
## 3.1. First pillar: the monitoring of the quality of the model predictions

With the monitoring, we aim to track the model's performance over its ODD. In our case, we trained a model on a training dataset and deployed it over IGN orthoimages covering all of France. Therefore, our ODD, referred to as the mapping area, is France.

The standard way of evaluating a model's accuracy is to compare its prediction against ground-truth labels. We do not have such ground-truth labels in France, so we have to rely on the closest available data, the *Registre National d'Installations* (RNI). The RNI indicates the total number of installations and the cumulated installed capacity for each city.

To monitor the accuracy of the registry with the RNI, we rely on unsupervised model evaluation methods [19]. These approaches automatically compare the model outputs with an external data source. In our case, we aggregate the detections of our model to estimate the installed capacity and the number of installations. We then compare

these estimations with the RNI records. We called this method the downstream task accuracy (DTA) as it enables us to evaluate a model's accuracy according to metrics that are relevant for the downstream or final users: for instance, we can compute the average error in the estimation of the aggregated installed capacity and derive metrics indicating whether the model locally overestimates or underestimates the aggregated installed capacity.

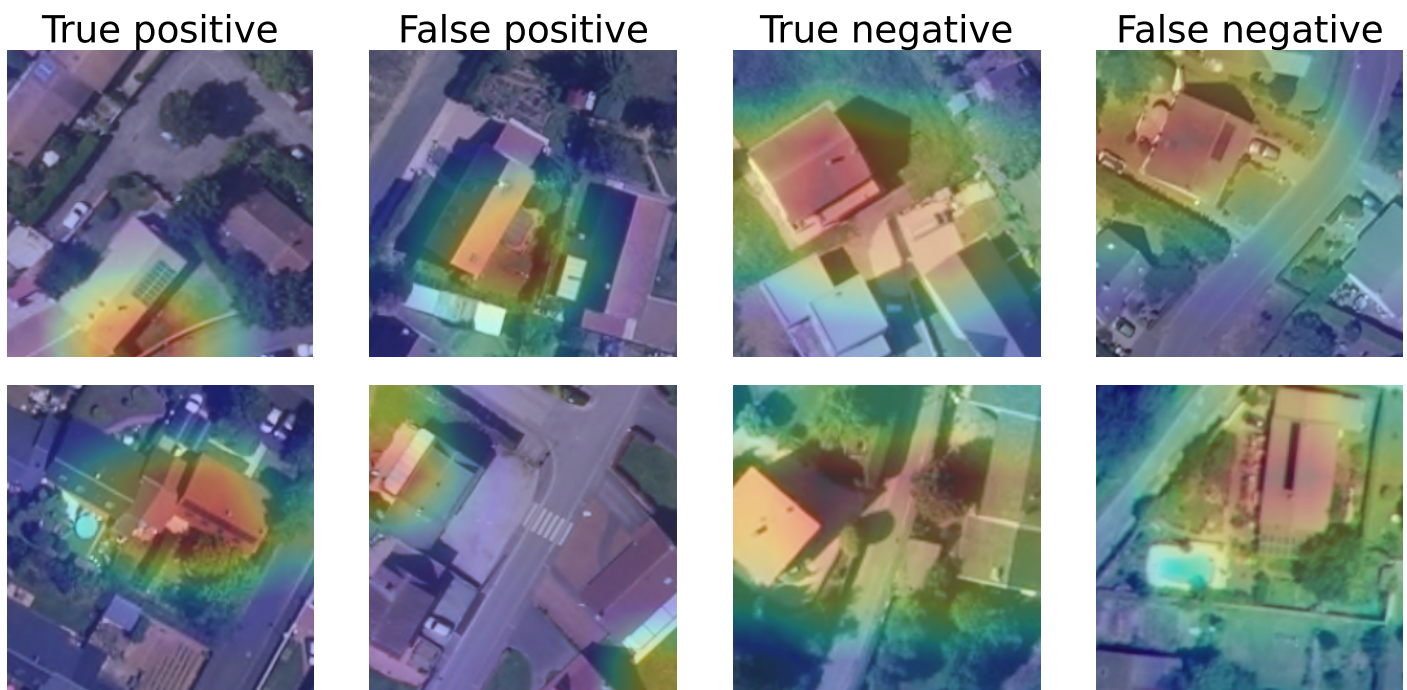


**Figure 3: Comparison between the distribution of the installed capacity estimated and reported in the RNI. The leftmost plot displays the relative spread between the two.**

As seen in Figure 3, comparing the model's aggregations with the RNI enables quickly identifying cases where the model made wrong estimations. In Kasmi *et al.* (2022) [6], we leveraged the DTA to show that the model's performance varies significantly during deployment. We also quantified the performance drop encountered by rooftop PV mapping algorithms, which was highlighted by earlier works [20,21]. Deploying a model based on 3D-PV-Locator (Mayer *et al.*, 2022) [8] and fine-tuned over France using the BDAPPV dataset [16], we documented a 30-percentage point accuracy drop when shifting from the training dataset to the mapping area. The main question is to understand why such a performance drop occurs.

### 3.2. Second pillar: auditing the decision process of the model

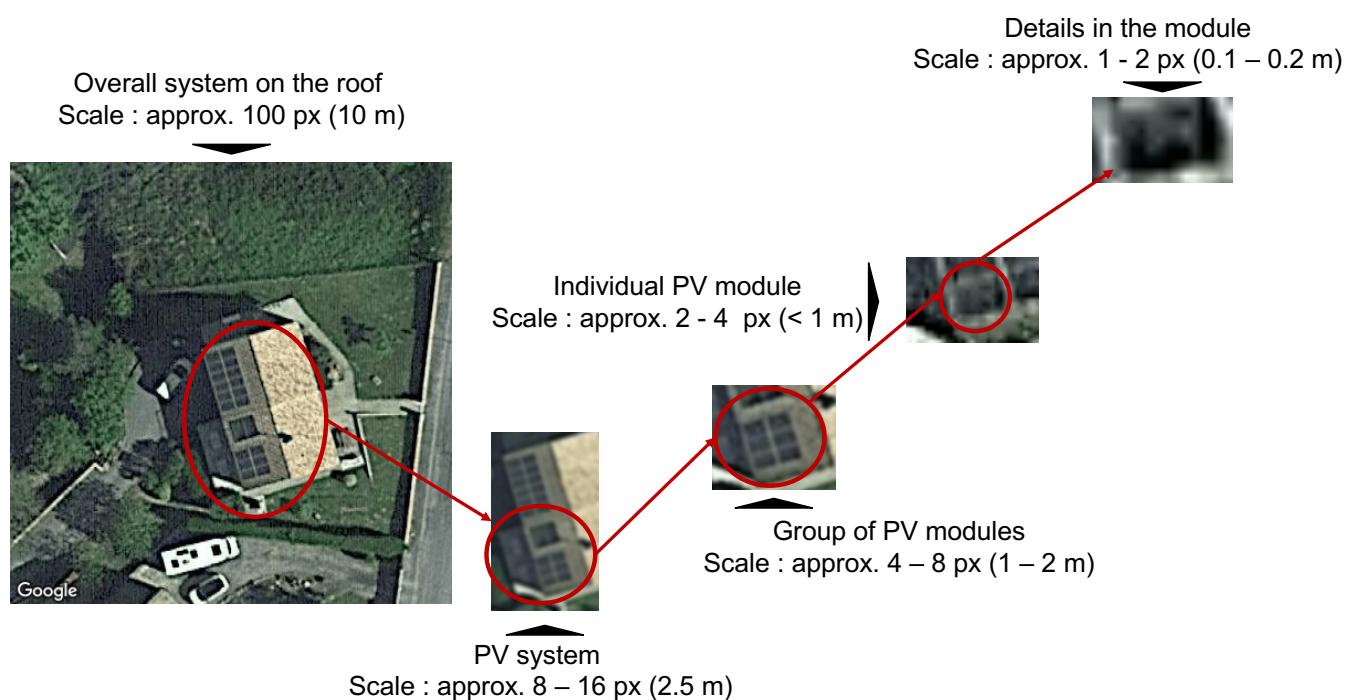
The auditing ensures that the classification model detects PV panels for the correct reasons. By correct, we mean that the model should detect PV panels by relying on relevant components of the image and not spurious factors (e.g., rooftops or nearby objects such as pools). To carry out this assessment, we initially relied on a class of feature attribution method called the Gradient Class Activation Map (GradCAM, [22]). This method highlights the areas on the input image that contributed the most to a model's decision. Figure 4 shows examples of model explanations computed with the GradCAM.



**Figure 4: Model explanations using the GradCAM [22] for some true positives, false positives, true negatives and false negatives. The redder, the higher the contribution of an image region to the predicted class.**

The GradCAM was sufficient to reveal that the model does not rely on spurious factors and focuses on where PV panels are located. However, it failed to explain false positives and false negatives. Indeed, we can see from Figure 4 that in the case of positives (true or false), the model focuses on a specific area of the image. When false detections arise, the model seems to confuse a given factor which *resembles* a PV panel with an actual PV panel. On the other hand, in the case of negative detections, the model does not seem to focus on any specific area. In the case of false negatives, the model does not “hook” to the PV panel despite it being there. More broadly, we can suppose that false detections are caused by the model relying on different scales and that false predictions can arise depending on whether patterns are present at different scales.

To better understand why false predictions arise, we need to assess what the model sees on the input image, not only where it looks, as traditional feature attribution methods do. Towards this end, we introduced the wavelet scale attribution method (WCAM, [23]). Disentangling the scales provides more information about *what* the model is seeing compared to traditional attribution methods. Indeed, as illustrated in Figure 5, a single spatial location corresponds to different scales, which do not precisely encompass the same semantical content. At the finest scales (a few pixels), a PV panel corresponds to individual modules. In contrast, at larger scales, the main feature that arises is the overall frame (i.e., a rectangle) or the gridded pattern.



**Figure 5: Decomposition of the scales of a PV panel. Source: Kasmi *et al.* (2023) [24].**

The WCAM builds on the sensitivity analysis method of Fel *et al.* (2021). This method consists in sampling random perturbation masks. In the original method, these masks are applied to the input image. The model's sensitivity to the masks' perturbation is then recorded through the variation in the predicted probability. Suppose the occluded area is not important for the prediction. In that case, the variation of the predicted probability is small, and if the variation is high, it means that the occluded area is important. Importance is evaluated using Sobol sensitivity analysis [26].

To highlight the important scales, we perturb the dyadic wavelet transform of the image instead of the image itself. The dyadic wavelet decomposition [27] consists in decomposing an image into different scales. As the wavelet transform is invertible, we

can reconstruct a perturbed image from its perturbed wavelet transform. The model is evaluated on the perturbed image, but this time, we can trace back the importance of the perturbations in the wavelet domain instead of the image domain.

The outcome of our approach is a heatmap in the wavelet domain. This representation highlights the important scales in the model's prediction. An interesting feature for the application to PV panel classification, the WCAM, revealed that for a single spatial location (highlighted in standard feature attribution methods), the model relies on different scales. These scales correspond to structural elements such as details within the PV modules or the gridded pattern. We refer the reader to [23] for further details on the WCAM.

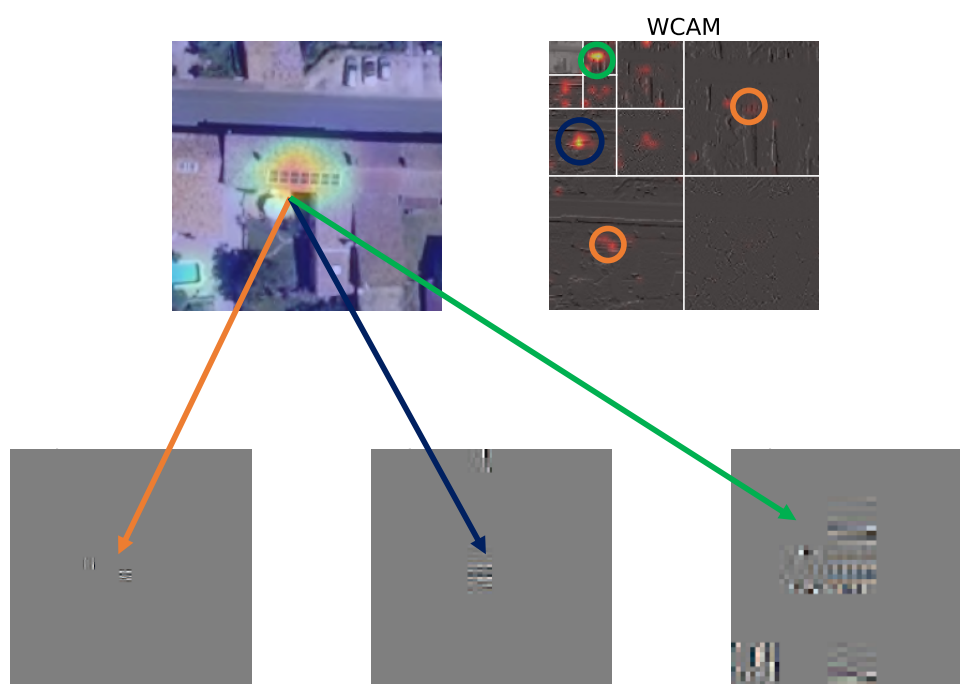


Figure 6: Decomposition of the scales of a PV panel. Source: Kasmi *et al.* (2023) [24].

We used the WCAM to carry out analyses to explain the false predictions. We highlighted that false positives arise when a gridded pattern appears on the input image. On the other hand, in [24,28] we studied how the acquisition conditions disrupted some frequency ranges (which corresponded to scales) and could lead to false negatives.

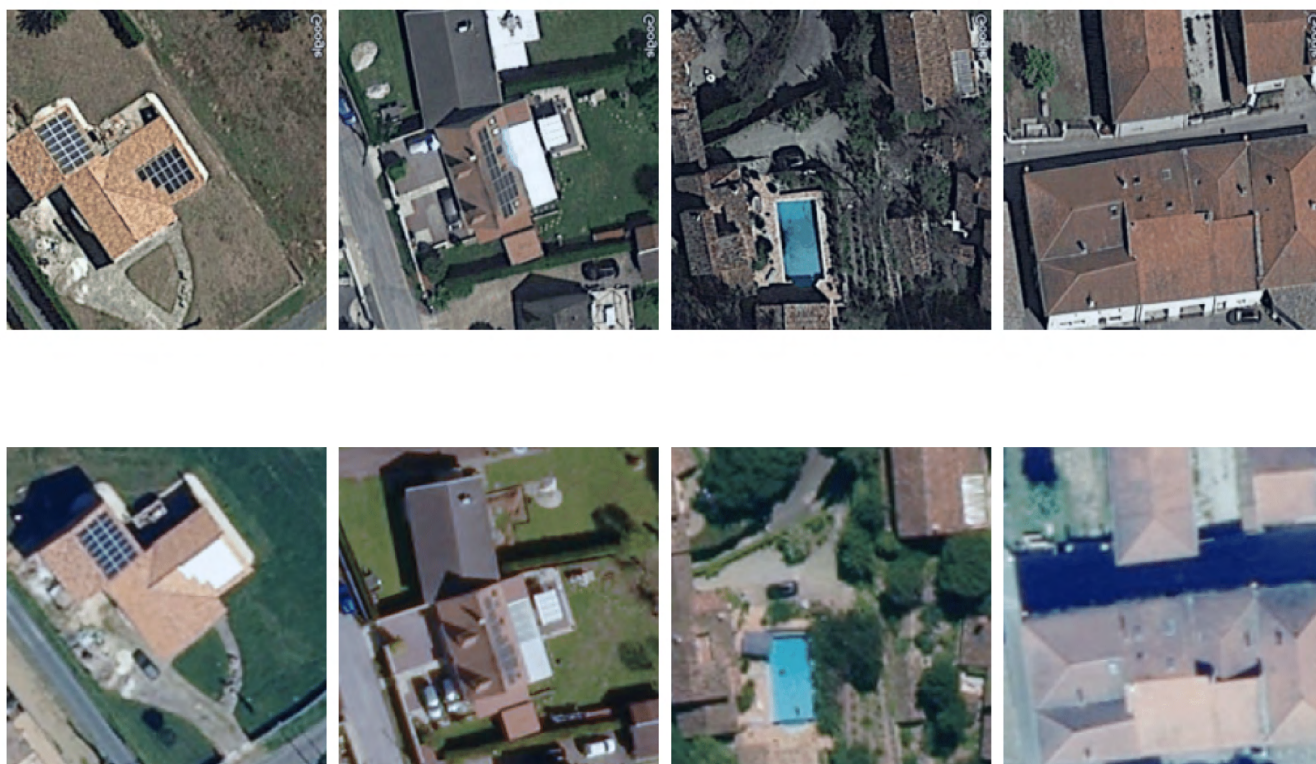
### 3.3. Third pillar: improving the robustness of the model to input perturbations

Analyzing the model's decision with the WCAM showed that the classification model relies on different scales. While dealing with the false positives is difficult, we can at least

focus on the false negatives. In [24], we showed that the varying acquisition conditions significantly altered the model's performance during its real-life deployment.

To improve the model's robustness to varying acquisition conditions, we introduced a new data augmentation technique: the wavelet perturbation. This method alters the image's wavelet transform to force the model to rely on various scales. We aim to ensure that if the acquisition conditions disrupt a scale, the model learns to rely on other, less perturbed scales to make its prediction.

We compared our wavelet perturbation with other popular data augmentation methods to improve the model's robustness to image corruptions. We evaluated our method against AugMix [29], RandAugment [30] and AutoAugment [31] and achieved state-of-the-art results. Our benchmark dataset is BDAPPV [16]: we train a ResNet [32] model on Google images and evaluate its accuracy on IGN images depicting the same PV panel. This benchmark is a natural case study for varying acquisition conditions.



**Figure 7: Examples of images from BDAPPV [16] depicting the same PV panel from two different providers, thus mimicking the shift in acquisition conditions. Source: Kasmi *et al.* (2023) [24].**

In addition to improving the robustness to varying acquisition conditions, we introduced PyPVRoof [33], a Python package for extracting the characteristics of rooftop PV systems. The motivation for introducing this package was that existing works [34,35,36] lacked

standardization. Each method had data requirements, and the various methods introduced could not be compared. We aimed to systematically compare existing approaches and design a set of methods that could be used in different cases regarding data availability. Practitioners can use PyPVRoof to extract the tilt and azimuth angles, surface, installed capacity, and localization of rooftop PV systems no matter the complementary data they have at their disposal (no data, 3D LiDAR data or a survey on PV systems).

The last step towards building DeepPVMapper was to provide a comprehensive benchmark of classification and segmentation models, evaluate the gains brought by the two-step approach introduced by DeepSolar, and bring several improvements to the pipeline to minimize false detections and speed up the computations. DeepPVMapper improves over existing works by being more robust to varying acquisition conditions, less prone to false detections, and more flexible than previous works to extract the characteristics of PV systems as it accommodates different cases of complementary data availability (e.g., 3D LiDAR data). Finally, we designed DeepPVMapper to easily support updates for the classification and segmentation model, thus facilitating its utilization in places other than France. The source code of DeepPVMapper is accessible on an online repository [37].

# 4. Assessment of the impact on rooftop PV observability

## 4.1. Mapping rooftop PV systems in France

We deployed DeepPVMapper over France to detect rooftop PV systems. As of April 2024, we deployed DeepPVMapper over 38 *départements* (covering nearly 175,000 km<sup>2</sup>) and recorded the tilt and azimuth angles of the systems and their installed capacity, surface, and pointwise localization. Figure 8 illustrates the resulting PV registry. The area covered by DeepPVMapper is currently the world's second largest after DeepSolar (who mapped rooftop PV systems over the continental United States) and the largest with this level of detail: Mayer *et al.* [8] covered Northrhine-Wesphalia (35,000 km<sup>2</sup>) with a similar level of details.

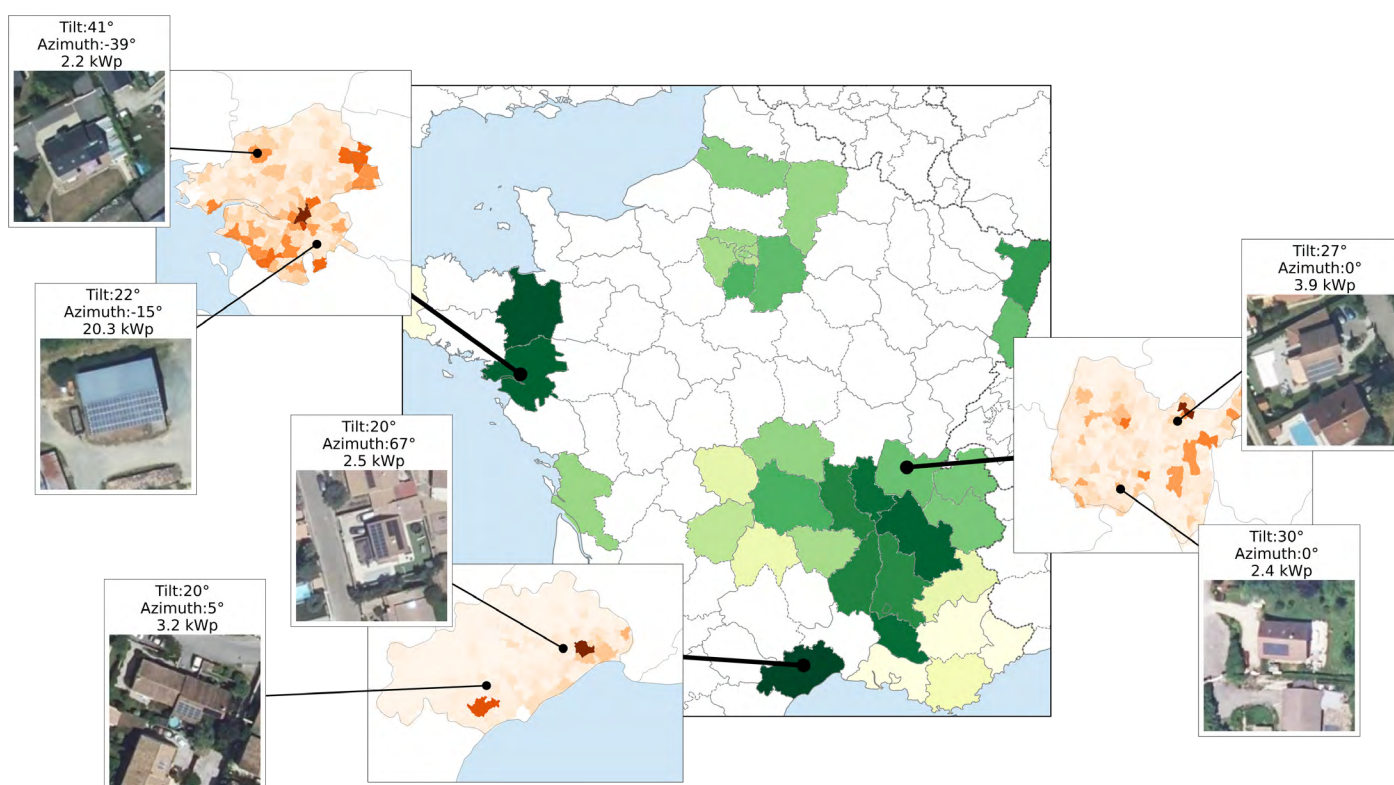


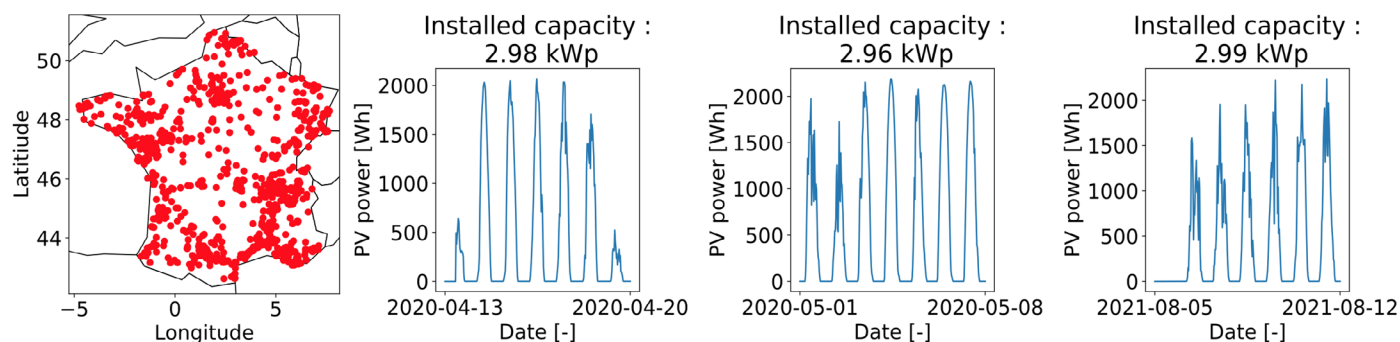
Figure 8: Overview of the French départements mapped with DeepPVMapper (as of April 2024). The greener, the higher the aggregated installed capacity. White départements have not been mapped yet.

## 4.2. Assessing the gain for PV power production estimation

The registry enables us to address the first prerequisite to improve rooftop PV observability: improving the knowledge regarding the PV fleet. The second prerequisite is to estimate



these systems' PV power production accurately. This is done in two steps. First, we use a conversion model (PVWatts, [38]) that takes as input the PV system's characteristics and solar irradiance data (coming from the Copernicus Atmospheric Monitoring Service (CAMS, [39])). The model outputs a load factor expressed in  $W_h/W_p$ . Second, we compare the estimation of the PV power production with ground truth PV yield time series. These time series come from BDPV. After data curation, we have about 900 records of PV power yields, covering all of France.



**Figure 9: Examples of PV power time series.**

We feed the conversion model with the PV systems' parameters estimated in DeepPVMapper's generated registry. We compare this parameterization with the best possible parameterization, i.e., with ground truth parameters on the PV systems coming from BDPV. We refer to this parameterization as the "Oracle". The estimation error is about 10% with a conversion model parameterized with DeepPVMapper and 8% with the Oracle. These results show that we need relatively little information on the rooftop PV systems to derive reasonably accurate estimations of their PV power production, and thus that our nationwide registry contains enough details for carrying out estimations of the rooftop PV power production. Using a conversion model feed with PV system's parameters obtained with deep learning-based methods is a suitable approach as the resulting estimation of the PV power production is competitive with the first best.

Our approach, introduced in Kasmi *et al.* (2024) [40], consists in estimating the rooftop PV power production for each system. In practice, we have about a million systems to consider. Therefore, we need to consider the scalability of our approach. In particular, we need to ensure that there are no systematic biases in estimating the system characteristics that could lead to an overall estimation error larger than the individual estimation error. We compared the behavior of the PV power estimation error of the oracle with DeepPVMapper. The error behavior is the same for both parameterizations, indicating no evidence of potential biases in estimating the system's parameters with DeepPVMapper.

# 5. Conclusion and perspectives

**Answer to the scientific question.** The scientific question was whether deep learning-based remote sensing on orthoimagery could be a suitable method for constructing a nationwide registry of rooftop photovoltaic (PV) installations intended to improve the observability of PV power production in France. More broadly, it raises the question of whether deep learning models are mature enough to be safely used in industrial pipelines. This work's central contribution is identifying quality and dependability standards and proposing a methodology to verify that the deep learning model and the generated data meet these standards. The necessary conditions are the ability to monitor the model's data and to audit its decision process. Therefore, it is necessary to have complementary data and to define relevant KPIs against which the data produced by the deep learning model will be monitored. Standard feature attribution techniques are insufficient for auditing the model's decision as they do not assess what models see. Our WCAM provides a first step towards addressing this issue. Finally, having a robust and accurate model is desirable but insufficient to achieve the required level of trust in the data and the decision process, as user's trust comes from his or her ability to monitor the data and audit the model. Therefore, deep learning and Earth observation data are suitable because one has enough additional data to monitor the model during its deployment.

**Contributing to the integration of PV into the grid.** This work contributes to improving PV observability to the extent that it provides valuable additional information regarding the geographical distribution of small-scale PV systems. It also contributes to the improvement of the estimation of the power production of small-scale PV systems. The thesis showed that our proposed approach, based on remote sensing and weather data enables an accurate estimation of the PV power production of rooftop PV systems, at the scale of the installation. Our results are promising, yet further work is needed to assess the superiority of the proposed approach compared to the TSO's existing methods.

**Broader impact.** The thesis work showed that deep learning tools can be relevant for acquiring detailed knowledge on rooftop PV systems. Using traditional methods such as surveys or self-reports can be time consuming, whereas the remote sensing of PV systems

using deep learning algorithms is a fast and relatively cheap solution. However, despite relatively accurate and cost efficient, deep learning is not necessarily very reliable. The thesis work discussed a series of conditions that needed to be satisfied so that the users can trust the data obtained with such methods. I believe that the main requirement lies in improving the practitioner's understanding regarding the functioning of the model in order to improve the critical thinking of the end users towards these tools. The framework introduced in this thesis is a first step towards this goal.

# References

- [1] IPCC. Summary for Policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2021. doi:10.1017/9781009157896.001. Type: Book Section.
- [2] RTE France. *Energy Pathways to 2050*. Technical report, RTE France, 2022.
- [3] République française. Décret n° 2020-456 du 21 avril 2020 relatif à la programmation pluriannuelle de l'énergie, 2020.
- [4] Marco Pierro, Fabio Romano Liolli, Damiano Gentili, Marcello Petitta, Richard Perez, David Moser, and Cristina Cornaro. Impact of PV/Wind Forecast Accuracy and National Transmission Grid Reinforcement on the Italian Electric System. *Energies*, 15(23):9086, November 2022. ISSN 1996-1073. doi:10.3390/en15239086.
- [5] Jordan M. Malof, Rui Hou, Leslie M. Collins, Kyle Bradbury, and Richard Newell. Automatic solar photovoltaic panel detection in satellite imagery. In 2015 *International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 1428–1431, Palermo, Italy, November 2015. IEEE. ISBN 978-1-4799-9982-8. doi:10.1109/ICRERA.2015.7418643.
- [6] Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, and Philippe Blanc. Towards Unsupervised Assessment with Open-Source Data of the Accuracy of Deep Learning-Based Distributed PV Mapping. In Thomas Corpetti, Dino Ienco, Roberto Interdonato, Minh-Tan Pham, and Sébastien Lefèvre, editors, *Proceedings of MACLEAN: MACHINE LEARNING FOR EARTH OBSERVATION Workshop co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2022)*, Grenoble, France, September 19-22, 2022, volume 3343 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2022.
- [7] Bala Bhavya Kausika, Dedip Niemijer, Iris Reinkem, Peter Blokker, and Gerard Hooy. GeoAI for detection of solar photovoltaic installations in the Netherlands. *Energy and AI*, 6:100111, December 2021. ISSN 2666-6548. doi:10.1016/j.egyai.2021.100111.
- [8] Kevin Mayer, Benjamin Rausch, Marie-Louise Arlt, Gunther Gust, Zhecheng Wang, Dirk Neumann, and Ram Rajagopal. 3D-PV-Locator: Large-scale detection of rooftop-mounted photovoltaic systems in 3D. *Applied Energy*, 310:118469, March 2022. ISSN 0306-2619. doi:10.1016/j.apenergy.2021.118469.
- [9] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. *Joule*, 2(12):2605–2617, December 2018. ISSN 2542-4351. doi:10.1016/j.joule.2018.11.021.

[10] Yves-Marie Saint-Drenan, S. Bofinger, R. Fritz, S. Vogt, G.H. Good, and J. Dobschinski. An empirical approach to parameterizing photovoltaic plants for power forecasting and simulation. *Solar Energy*, 120:479–493, October 2015. ISSN 0038-092X. doi:10.1016/j.solener.2015.07.024.

[11] Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Wei Hua, Michihiro Yasunaga, Richard Lanus Phillips, Irena Gao, Tony Lee, Etienne David, Ian Stavens, Wei Guo, Berton Earnshaw, Iman Hafeez, Sara M. Beery, Jure Leskovec, Anshul Kundaje, Emma Pierson, Sergey Levine, Chelsea Finn, and Percy Liang. WILDS: A Benchmark of in-the-Wild Distribution Shifts. In *Proceedings of the 38th International Conference on Machine Learning*, pages 5637–5664. PMLR, July 2021.

[12] Gabriel Kasmi. *Enhancing the Reliability of Deep Learning Algorithms to Improve the Observability of French Rooftop Photovoltaic Installations*. PhD thesis, Université Paris sciences et lettres, Paris, April 2024.

[13] IGN. BD ORTHO® | Géoservices, 2024. Online. Accessed on 01.10.2024. URL: <https://geoservices.ign.fr/bdortho>.

[14] IGN. LiDAR HD | Géoservices, 2024. Online. Accessed on 01.10.2024. URL: <https://geoservices.ign.fr/lidarhd>.

[15] ODRÉ. *Registre national des installations de production et de stockage d'électricité (au 30/04/2022)*, 2022.

[16] Gabriel Kasmi, Yves-Marie Saint-Drenan, David Trebosc, Raphaël Jolivet, Jonathan Leloux, Babacar Sarr, and Laurent Dubus. A crowdsourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata. *Scientific Data*, 10(1):59, January 2023. ISSN 2052-4463. doi:10.1038/s41597-023-01951-4.

[17] Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202:18–27, 2017. Publisher: Elsevier.

[18] Frédéric Boisnard, Ryma Boumazouza, Mélanie Ducoffre, Thomas Fel, Estèle Glize, Lucas Hervier, Vincent Musso, Agustin Martin Picard, Antonin Poché, and David Vigouroux. Guidelines to explain machine learning algorithms. Technical report, DEEL, 2024.

[19] Hui Zhang, Jason E. Fritts, and Sally A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260–280, May 2008. ISSN 1077-3142. doi:10.1016/j.cviu.2007.08.003.

[20] Tim De Jong, Stefano Bromuri, Xi Chang, Marc Debusschere, Natalie Rosenski, Clara Schartner, Katharina Strauch, Marion Boehmer, and Lyana Curier. Monitoring Spatial Sustainable Development: Semi-automated analysis of satellite and aerial images for energy transition and sustainability indicators, 2020.

[21] Jordan M Malof, Boning Li, Bohao Huang, Kyle Bradbury, and Artem Streltso. Mapping solar array location, size, and capacity using deep learning and overhead imagery, 2019.

[22] Ramprasath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, 128(2):336–359, February 2020. ISSN 1573-1405. doi:10.1007/s11263-019-01228-7.

[23] Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, and Philippe Blanc. Assessment of the Reliability of a Model's Decision by Generalizing Attribution to the Wavelet Domain. In *XAI in Action: Past, Present, and Future Applications workshop at NeurIPS 2023*. arXiv, September 2023. doi:10.48550/arXiv.2305.14979.

[24] Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, and Philippe Blanc. Can We Reliably Improve the Robustness to Image Acquisition of Remote Sensing of PV Systems? In *NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning*, 2023. doi:https://doi.org/10.48550/arXiv.2309.12214.

[25] Thomas Fel, Romain Cadene, Mathieu Chalvidal, Matthieu Cord, David Vigouroux, and Thomas Serre. Look at the Variance! Efficient Black-box Explanations with Sobol-based Sensitivity Analysis. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 26005–26014. Curran Associates, Inc., 2021.

[26] I. M Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112, January 1967. ISSN 0041-5553. doi:10.1016/0041-5553(67)90144-9.

[27] S.G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989. ISSN 1939-3539. doi:10.1109/34.192463.

[28] Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, and Philippe Blanc (2024). Space-scale Exploration of the Poor Reliability of Deep Learning Models: the Case of the Remote Sensing of Rooftop Photovoltaic Systems. *arXiv preprint arXiv:2408.07828*.

[29] Dan Hendrycks, Norman, Ekin Dogus Cubuk, Barrett Zoph, Justin Gilmer, and Balaji Lakshminarayanan. AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty. In *8th International Conference on Learning Representations, ICLR 2020*, Addis Ababa, Ethiopia, April 26-30, 2020. OpenReview.net, 2020.

[30] Ekin D. Cubuk, Barrett Zoph, Jonathan Shlens, and Quoc V. Le. RandAugment: Practical automated data augmentation with a reduced search space. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 3008–3017, June 2020. doi:10.1109/CVPRW50498.2020.00395. ISSN: 2160-7516.

[31] Ekin D. Cubuk, Barrett Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V. Le. AutoAugment: Learning Augmentation Policies from Data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[32] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[33] Yann Trémenbert, Gabriel Kasmi, Laurent Dubus, Yves-Marie Saint-Drenan, and Philippe Blanc. PyPVRoof: a Python package for extracting the characteristics of rooftop PV installations using remote sensing data, September 2023. <https://doi.org/10.48550/arXiv.2309.07143>

[34] Ayobami S. Edun, Kirsten Perry, Joel B. Harley, and Chris Deline. Unsupervised azimuth estimation of solar arrays in low-resolution satellite imagery through semantic segmentation and Hough transform. *Applied Energy*, 298:117273, September 2021. ISSN 03062619. doi:10.1016/j.apenergy.2021.117273.

[35] Brenda So, Cory Nezin, Vishnu Kaimal, Sam Keene, Leslie Collins, Kyle Bradbury, and Jordan M. Malof. Estimating the electricity generation capacity of solar photovoltaic arrays using only color aerial imagery. In 2017 IEEE *International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1603–1606, Fort Worth, TX, July 2017. IEEE. ISBN 978-1-5090-4951-6. doi:10.1109/IGARSS.2017.8127279.

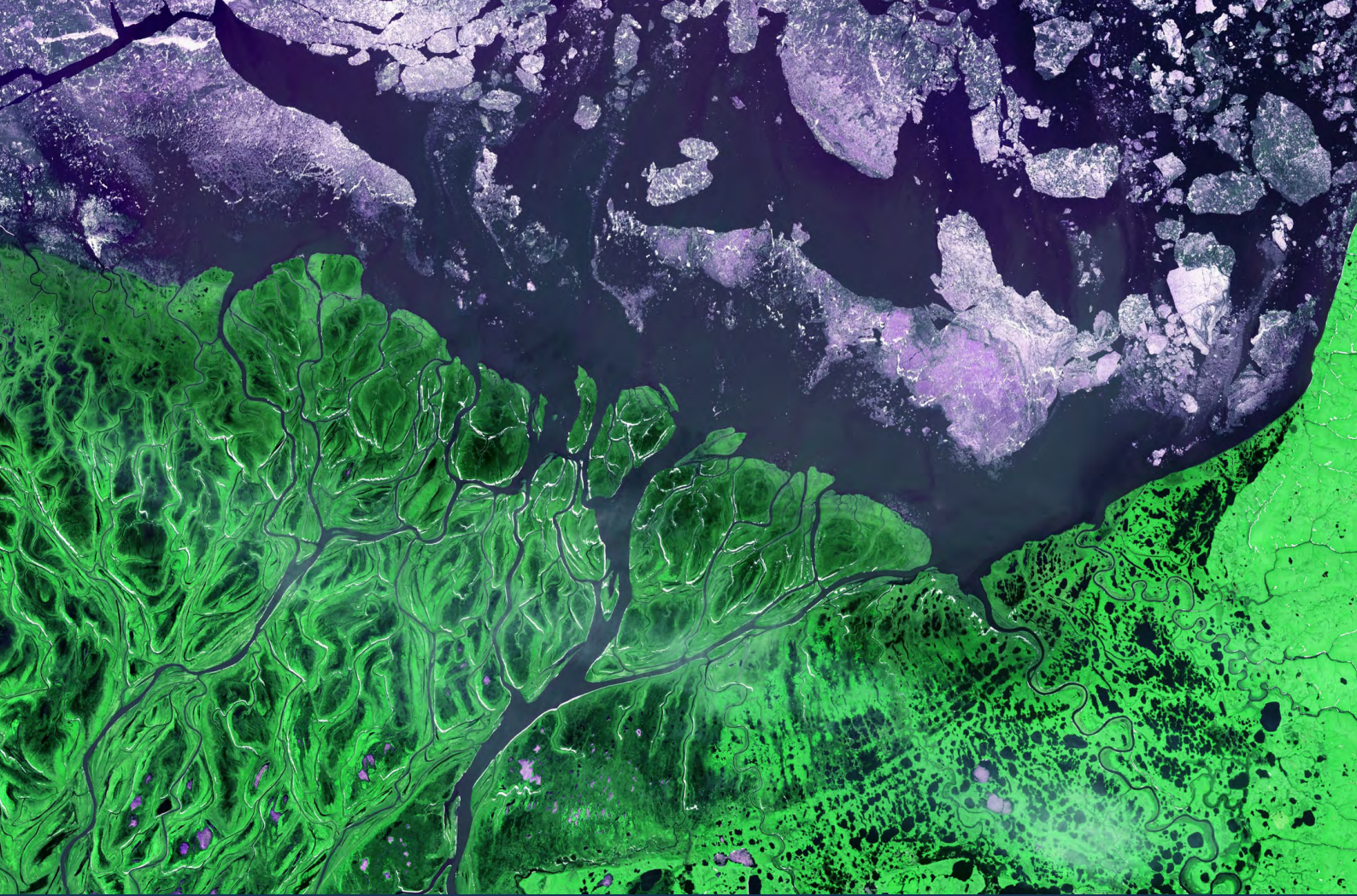
[36] Benjamin Rausch, Kevin Mayer, Marie-Louise Arlt, Gunther Gust, Philipp Staudt, Christof Weinhardt, Dirk Neumann, and Ram Rajagopal. An Enriched Automated PV Registry: Combining Image Recognition and 3D Building Data. In *NeurIPS 2020 Workshop on Tackling Climate Change with Machine Learning*, 2020.

[37] Gabriel Kasmi, Dubus Laurent, Blanc Philippe, and Yves-Marie Saint-Drenan. DeepPVMapper, September 2023d. URL:<https://doi.org/10.5281/zenodo.8380321>.

[38] A. Dobos. PVWatts Version 5 Manual. Technical Report NREL/TP-6A20-62641, 1158421, NREL, September 2014.

[39] Zhipeng Qu, Armel Oumbe, Philippe Blanc, Bella Espinar, Gerhard Gesell, Benoît Gschwind, Lars Klüser, Mireille Lefèvre, Laurent Saboret, Marion Schroedter-Homscheidt, and Lucien Wald. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, 26(1):33–57, February 2017. ISSN 0941-2948. doi:10.1127/metz/2016/0781.

[40] Gabriel Kasmi, Augustin Touron, Philippe Blanc, Yves-Marie Saint-Drenan, Maxime Fortin, and Laurent Dubus. Remote-Sensing-Based Estimation of Rooftop Photovoltaic Power Production Using Physical Conversion Models and Weather Data. *Energies*, 17(17):4353, August 2024b. ISSN 1996-1073. doi:10.3390/en17174353.



*Ce travail a été réalisé dans le cadre du parcours doctoral de  
The Transition Institute 1.5 (TTI.5)*

*TTI.5 est un institut de recherche de Mines Paris - PSL dédié à  
la transition bas carbone*

## **CONTACT**

 [the-transition-institute.minesparis.psl.eu](https://the-transition-institute.minesparis.psl.eu)

 [tti.5@minesparis.psl.eu](mailto:tti.5@minesparis.psl.eu)