Solar Energy Expansion Monitoring from Satellite Imagery

Emrecan Ulu, Jingyao Shi, Xiaoxuan Zhang

Project Description

This project applies Convolutional Neural Networks (CNNs) to detect solar panel installations from satellite imagery, with the goal of tracking renewable energy adoption patterns. By developing a model to identify utility-scale solar deployment, this study provides insights into how regional policies and socioeconomic factors influence the geographic distribution of solar energy infrastructure. The findings are intended to inform energy policy and infrastructure planning.

1. Data Collection and Preliminary Analysis

1.1 Methodology

To construct a dataset for training a CNN, geospatial data from OpenStreetMap (OSM) was integrated with satellite imagery from the Google Earth Engine (GEE) platform. The methodology was first developed and validated using data from Germany due to its high-quality and extensive OSM coverage for solar installations.

Solar Installation Data: Solar farm coordinates were obtained from OSM using Overpass API queries. The raw dataset was filtered to retain only utility-scale ground-mounted systems, excluding rooftop panels. After cleaning, this resulted in a primary dataset of 276,220 unique utility-scale points across Germany.

Satellite Imagery: Image patches were extracted from the Sentinel-2 Surface Reflectance collection on GEE. A time window of April-October 2022 was used to create a median composite of visible bands (B4, B3, B2), effectively masking clouds and reducing seasonal variability. Each patch was centered on a data point, covering 640 x 640 meters and exported at a resolution of 256 x 256 pixels.

Dataset Creation: To create a balanced dataset for binary classification, 10,000 image patches of utility-scale solar farms (positive samples) and 10,000 negative patches were collected. Negative samples were sourced from random locations, ensuring a buffer zone of 120 meters from any known solar installation to prevent data contamination.

1.2 Preliminary Data Analysis and Visualization (Germany)

Analysis of the spatial distribution of solar installations in Germany reveals distinct geographic patterns. The hexbin visualization below shows that installations are widespread, with particularly high densities in the southern states of Bavaria and Baden-Württemberg and additional clusters in eastern regions like Brandenburg.

The visualized patterns are a direct reflection of Germany's long-standing energy policy, the Erneuerbare-Energien-Gesetz (EEG). Early feed-in tariffs encouraged decentralized, widespread adoption by farmers, communities, and small businesses, leading to the broad distribution seen in Figure 1. The high concentration in southern states like Bavaria is not only due to better solar irradiation but also because these are economically powerful regions with available agricultural land suitable for such projects.

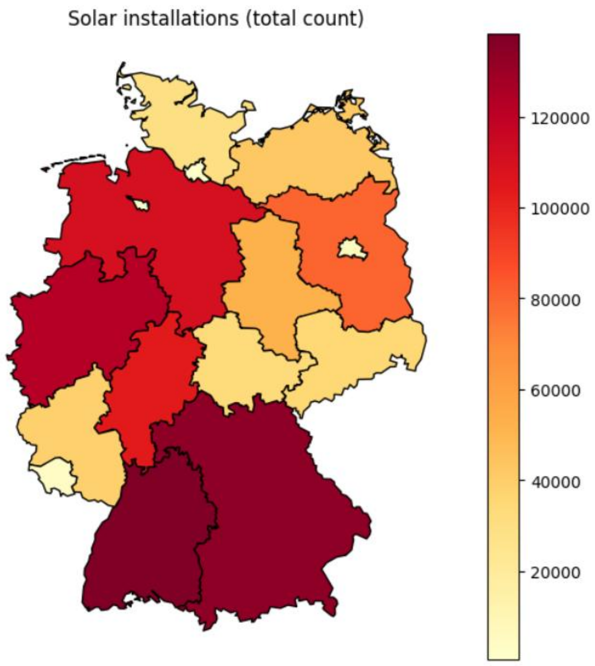

Figure 1: Total distribution of solar installations across German states

When normalized by land area, the pattern shifts. While southern states still show high density, smaller states and city-states like Berlin also rank highly due to installations being concentrated in a smaller area. This analysis indicates that solar deployment is not uniform and is influenced by factors such as solar irradiation, land availability, economic capacity, and, crucially, regional policy frameworks..

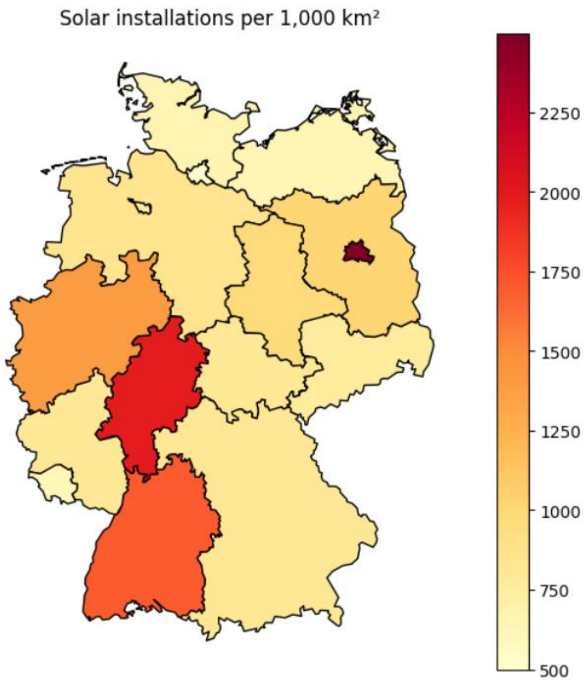

Figure 2: Normalized distribution of solar installations across German states

The two maps together tell a nuanced story. The absolute count map shows where the most infrastructure investment has occurred, highlighting economic hubs. The density map reveals where solar energy is most concentrated relative to land use, pointing to areas where policy and land availability have aligned most effectively. This preliminary analysis is critical because it confirms that the training data will carry a distinct "German signature"—characterized by medium-scale farms, often situated on or near agricultural land. This inherent bias is the key to understanding the model's performance in subsequent tests on different regions.

2. Deep Learning Model

2.1 Build CNN Model

A lightweight CNN was developed in PyTorch for the binary classification of image patches into "utility" (containing a solar farm) and "negative" (no solar farm).

Architecture: The model features stacked 3x3 convolutional layers with ReLU activations for feature extraction, followed by max-pooling and dropout layers to reduce dimensionality and mitigate overfitting. An adaptive average pooling layer ensures a fixed-size feature vector, which is then passed to a fully connected layer for classification.

Training: The model was trained using the Adam optimizer and cross-entropy loss. The training data was augmented with random cropping, flipping, and color jittering to improve robustness. Early stopping was employed to prevent overfitting, saving the model with the best validation loss.

2.2 Model Training and Validating with German Data

The model was trained on a 60/20/20 split of the 20,000-image German dataset. The training process showed smooth convergence, with training and validation losses decreasing together, indicating no significant overfitting.

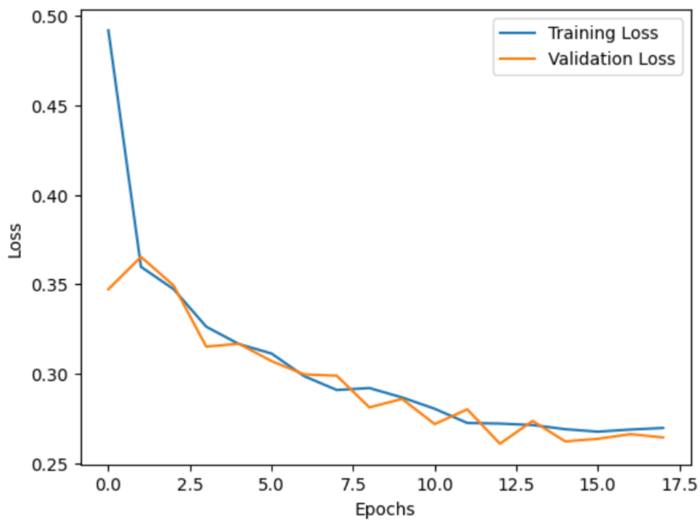

Figure 3: Training and validation loss curves for german data

On the held-out German test set, the model achieved an overall accuracy of 90.5%. Performance was well-balanced, with a high recall of 93.7% for the "utility" class, demonstrating its effectiveness in correctly identifying solar farms within its training domain.

class	precision	recall	f1-score	support
negative	0.945	0.8737	0.9079	2810
utility	0.8819	0.9488	0.9141	2793
accuracy			0.9111	5603
macro Avg	0.9134	0.9112	0.911	5603
weighted Avg	0.9135	0.9111	0.911	5603

Table 1: Classification Report for Solar Panel Detection with German Data

2.3 Model Deployment and Cross-Regional Testing (EU and US Datasets)

To assess the model's ability to generalize, it was tested on new, unseen datasets from the broader European Union (EU) and the United States (US) without retraining. EU dataset includes 500 positive and 500 negative images from Spain, Switzerland and Italy each. US dataset includes 500 positive and 500 negative images from California, Texas and Arizona each.

2.3.1 EU Test Results:

class	precision	recall	f1-score	support
negative	0.6502	0.622	0.6358	1500
utility	0.6375	0.6651	0.651	1499
accuracy			0.6435	2999
macro Avg	0.6438	0.6436	0.6434	2999
weighted Avg	0.6438	0.6435	0.6434	2999

Table 2: Classification Report for Solar Panel Detection with EU Data

The model's performance dropped significantly on the broader EU dataset, achieving an accuracy of 64.4%. While it still performed better than random chance, the confusion matrix shows a high number of errors, particularly 567 false negatives (utility farms misclassified as negative).

The model's struggle on the broader EU dataset is a clear indicator of regional diversity in solar farm design. The 567 false negatives (a recall of only 66.5% for "utility") suggest that nearly one-third of solar farms in the EU test set did not match the visual signature learned from German examples. These could be Italian agrivoltaics projects with crops growing between panels, Spanish farms in more arid settings, or installations with different mounting technologies, all of which present a different texture and context than the model was trained on. Furthermore, the 502 false positives indicate that the model was frequently confused by other features in the European landscape, such as large-scale greenhouses or specific industrial complexes, misidentifying them as solar farms. This balanced error profile shows the model is not simply failing, but is actively confused by the heterogeneity of European energy and land-use policies

.

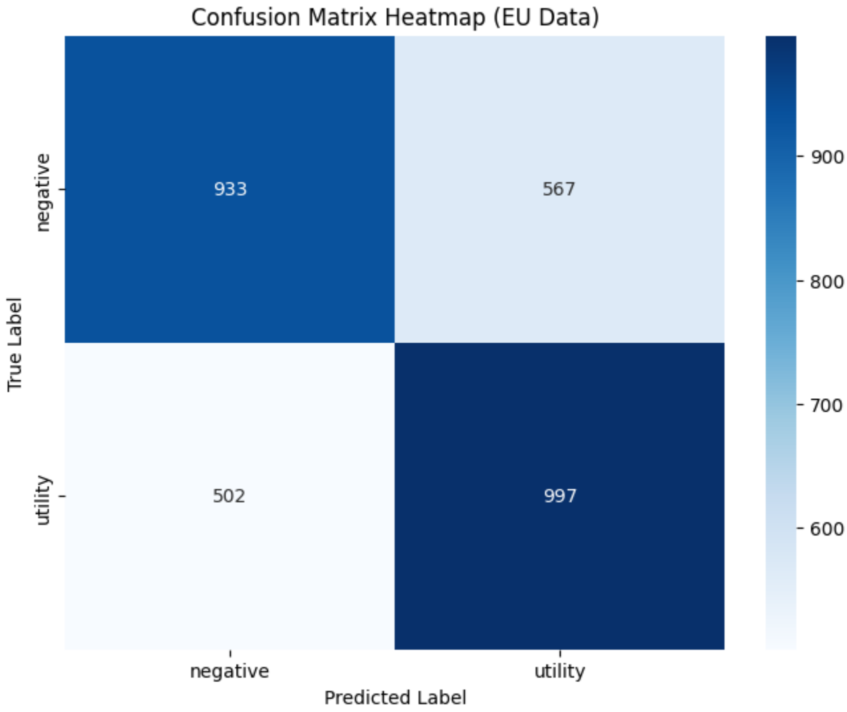

Figure 4: Confusion Matrix Heatmap for EU data

2.3.2 US Test Results:

class	precision	recall	f1-score	support
negative	0.5364	0.8605	0.6609	1498
utility	0.6481	0.2568	0.3679	1499
accuracy			0.5586	2997
macro Avg	0.5923	0.5587	0.5144	2997
weighted Avg	0.5923	0.5586	0.5143	2997

Table 3: Classification Report for Solar Panel Detection with US Data

The performance degradation was even more severe on the US dataset, with accuracy falling to 55.9%. The most notable failure was the extremely low recall for the "utility" class at only 25.7%. This means the model failed to identify nearly three-quarters of the actual utility-scale solar farms in the US test set.

The poor generalization performance highlights a classic machine learning challenge: covariate shift. The visual characteristics (e.g., size, layout, surrounding landscape) of solar farms in the US and other EU countries are significantly different from those in Germany on which the model was trained.

The model's performance on US data represents a catastrophic failure of generalization. The confusion matrix is dominated by 1,114 false negatives, meaning the model looked at a vast number of American solar farms and failed to recognize them.

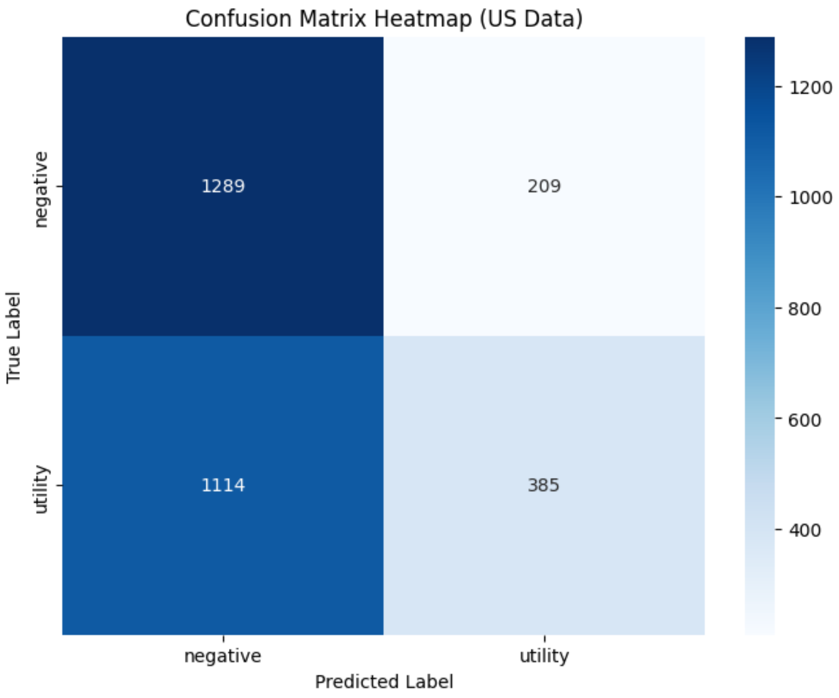

Figure 5: Confusion Matrix Heatmap for US data

This is a direct result of extreme covariate shift driven by three factors:

Scale: US federal policies and land availability in states like California and Arizona encourage the construction of massive, sprawling solar plants that are orders of magnitude larger than the average German farm. The model, trained on smaller plots, likely interprets these vast, uniform arrays as an unidentifiable ground texture rather than a collection of objects.

Environment: The model was trained on images from Germany's temperate climate, characterized by green croplands and forests. It has no context for the desert and arid scrubland environments common for US solar farms, where the color palette, lighting, and shadows are completely different.

Layout: The relatively low number of false positives (209) combined with the massive number of false negatives indicates the model was highly biased toward the "negative" class. It rarely encountered features that resembled a German solar farm, so it defaulted to classifying almost everything as "negative," leading to its failure to detect the vast majority of true installations.

3. Energy Policy Analysis

3.1 Analysis Using the Predictive Model

The stark difference in model performance between Germany, the broader EU, and the US serves as a powerful, data-driven indicator that solar deployment patterns are not globally uniform. These differences are heavily influenced by distinct national and regional energy policies, which shape the economic viability, permitting processes, and physical characteristics of solar projects.

Dimension	USA	EU
Main support mechanisms	Inflation Reduction Act (2022) provides technology-neutral tax credits: §48E ITC and §45Y PTC, 30% base with bonuses for domestic content and energy communities; credits are transferable and available as direct pay (U.S. Congress, 2022).	EU member states apply competitive renewable auctions under RED III and national schemes with contracts for difference or market premiums (EU, 2023). RED III establish accelerated permitting (1 year for repowering, 2 years for new projects) and “go-to areas” to speed up solar deployment (EU, 2023).
2030 targets	Target of 100% clean electricity by 2035, ~400 GW of solar expected by 2030 (U.S. Congress, 2022).	EU overall binding target: 42.5% renewables by 2030 under RED III (EU, 2023)

Table 4: Targets in EU and USA

Germany (High Accuracy): The model was trained here. German policy historically included feed-in tariffs and now uses auctions, with a strong role for municipal utilities (Stadtwerke) and citizen cooperatives. This may lead to more standardized, medium-scale projects in rural or repurposed industrial areas, creating a consistent visual signature that the model learned effectively.

United States (Very Low Recall):

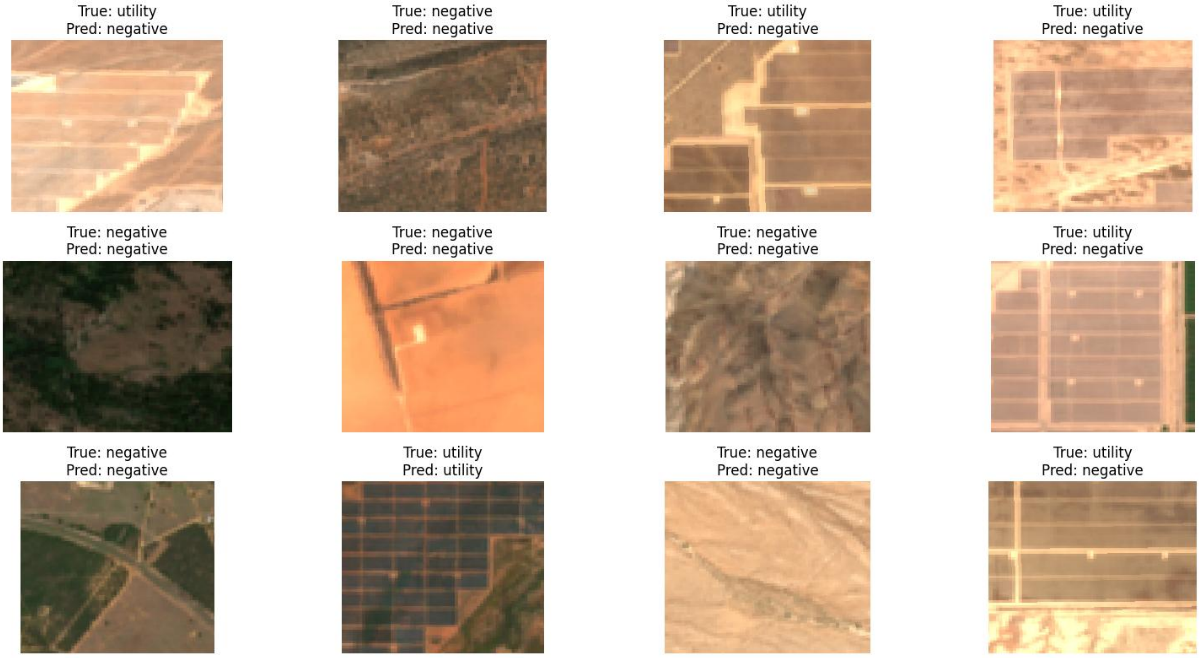

Figure 6: Examples in US test dataset

The model failed dramatically here. US policy is dominated by federal tax credits (ITC/PTC), but permitting and grid connection rules vary immensely by state. This fragmented approach, combined with vast, arid land in the Southwest, encourages the development of massive, utility-scale solar farms that look visually distinct from typical German installations. The model, trained on smaller German farms, likely failed to recognize these large-scale layouts and different environmental contexts (e.g., desert vs. cropland).

European Union (Moderate Accuracy):

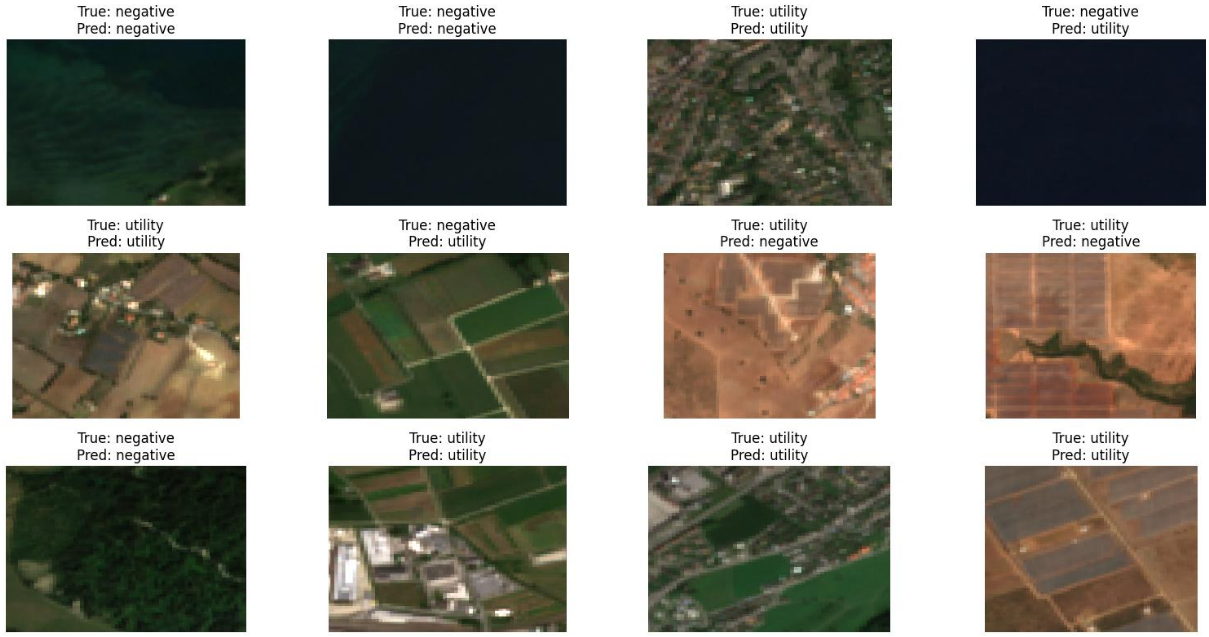

Figure 7: Examples in eu test dataset

The EU's Renewable Energy Directive (RED) provides a framework, but implementation varies. Italy's focus on agrivoltaics, Spain's use of auctions and "just transition nodes," and Switzerland's subsidies for unique alpine plants create diverse visual patterns. A model trained solely on German data would struggle to capture this heterogeneity, leading to the observed drop in accuracy. For instance, a solar farm integrated with crops (agrivoltaics) has a different texture and layout than one on a former industrial site.

3.2 Implications for Energy Transition Monitoring

This analysis demonstrates that a "one-size-fits-all" model for monitoring solar deployment is insufficient. The failure of the German-trained model to generalize is not a failure of the CNN architecture itself, but rather a reflection of real-world policy impacts on infrastructure development.

The model's predictive power is a proxy for the visual uniformity of solar installations. Therefore, monitoring where a specific model succeeds or fails can itself be a tool for identifying regions with unique deployment characteristics, prompting further policy-specific investigation.

4. Conclusion and Discussion

This project successfully demonstrated the development of a CNN capable of detecting utility-scale solar farms from satellite imagery with over 90% accuracy within its training region (Germany). However, testing revealed that the model's performance does not generalize well to other regions like the US and broader EU, a finding attributed to significant differences in solar deployment patterns driven by divergent energy policies.

The key takeaway is that energy policy directly shapes the physical and spatial characteristics of renewable energy infrastructure. This has critical implications for the use of machine learning in global monitoring: models must be trained on diverse, representative global datasets to be effective.

5. Directions for Future Research

Expanding the Training Dataset: Incorporate labeled data from a wide variety of countries and climates to build a more robust and generalizable model, following the approach of studies like Kruitwagen et al. (2021).

Adopting Segmentation Models: Transition from classification to semantic segmentation (e.g., using a U-Net architecture) to extract precise installation footprints, enabling more accurate capacity estimation.

Multi-spectral Imagery: Utilize the full range of spectral bands from Sentinel-2 (e.g., near-infrared) to help the model better distinguish solar panels from other features like water bodies or dark soils, reducing classification errors.

6. References

- Kruitwagen, L., Story, K., Friedrich, J., Byers, L., Skillman, S., & Hepburn, C. (2021). A global inventory of photovoltaic solar energy generating units. Nature, 598(7882), 604–610.

- Bundesministerium der Justiz. (2023). Erneuerbare-Energien-Gesetz 2023 (EEG 2023). BGBl. I Nr. 221.

- European Union. (2023). Directive (EU) 2023/2413 of the European Parliament and of the Council of 18 October 2023 on the promotion of energy from renewable sources (RED III).

- Gobierno de España. (2020). Real Decreto 1183/2020de 29 de diciembre, de acceso y conexión a las redes de transporte y distribución de energía eléctrica.

- Ministerio para la Transición Ecológica y el Reto Demográfico (MITECO). (2021). Plan Nacional Integrado de Energía y Clima 2021–2030 (PNIEC).

- Ministero della Transizione Ecologica. (2021). Piano Nazionale Integrato per l'Energia e il Clima 2021–2030.

- U.S. Congress. (2022). Inflation Reduction Act of 2022, Pub. L. No. 117–169, 136 Stat. 1818.