57º SBPO
SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
Gramado • RS
DE 5 A 9 DE OUTUBRO DE 2025

Leveraging Renewable Energy for Sustainable Development in the Amazon: A Data-Driven Approach

Renan Silva Santos
Pontifical Catholic University of Rio de Janeiro, Department of Industrial Engineering
Rio de Janeiro, Brazil
renansilvasantos@esp.puc-rio.br

Fernanda Araujo Baião
Pontifical Catholic University of Rio de Janeiro, Department of Industrial Engineering
Rio de Janeiro, Brazil
fbaiao@puc-rio.br

ABSTRACT

The Amazon rainforest, a region of immense biodiversity and ecological importance, faces numerous developmental challenges, including energy poverty and environmental degradation. Renewable energy, particularly solar energy, holds significant potential as a development driver for the region, offering sustainable solutions to these challenges. Data-driven approaches play a crucial role in enabling policymakers to harness this potential, providing insights for informed decision-making, optimizing resource allocation, and ensuring equitable and sustainable development outcomes. This expanded analysis explores the multifaceted role of renewable energy, focusing on solar power, as a catalyst for sustainable development in the Amazon, highlighting the importance of data-driven approaches in navigating the complexities of this endeavor.

KEYWORDS. Renewable energy; Sustainable development; Amazon rainforest; Data-driven approaches; Solar energy

Paper topics: AG&MA; AS&DS

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

1. Introduction

The Amazon rainforest, covering vast areas of South America, is renowned for its immense biodiversity and crucial role in global ecological balance. Despite its environmental significance, the region faces numerous developmental challenges, including energy poverty, limited infrastructure, and environmental degradation [Jean et al., 2021]. Many communities within the Amazon lack access to reliable and affordable energy sources, hindering economic development and quality of life [Ferreira et al., 2024].

Renewable energy, particularly solar power, holds significant potential as a catalyst for sustainable development in the Amazon. The region's high solar irradiance makes it an ideal candidate for solar photovoltaic (PV) technologies, which can provide clean, decentralized energy solutions to remote and isolated communities[Linguet et al., 2009]. Advancements in PV technology, decreasing costs, and the adaptability of both grid-connected and off-grid systems enhance the feasibility of solar energy deployment in the region [Alegre-Bravo and Lindsay Anderson, 2023].

However, harnessing this potential requires careful consideration of socio-environmental impacts, infrastructural limitations, and the diverse needs of indigenous and local communities [Jean et al., 2021]. Challenges such as deforestation, biodiversity loss, and community displacement must be addressed to ensure that renewable energy projects contribute positively to the region's development without compromising its ecological integrity [Velastegui-Montoya et al., 2022].

Data-driven approaches play a crucial role in enabling policymakers and stakeholders to navigate these complexities. By providing insights for informed decision-making, optimizing resource allocation, and ensuring equitable and sustainable development outcomes, data analytics empower the design and implementation of effective renewable energy strategies [Anderson et al., 2021]. Analyzing data on energy consumption patterns, resource availability, infrastructure needs, and community demographics helps tailor policies to specific contexts, maximizing the impact of renewable energy initiatives [Benes et al., 2024].

This analysis explores the multifaceted role of renewable energy as a driver for sustainable development in the Amazon. We highlight the importance of data-driven methodologies in assessing the potential of various renewable energy sources, evaluating socio-environmental impacts, and identifying the challenges and opportunities inherent in renewable energy development. By examining these factors, we aim to provide a comprehensive understanding of how renewable energy can contribute to addressing the Amazon's developmental challenges and promoting energy justice and inequality reduction in the region. The results are intended for policymakers and other stakeholders from the renewables sector, in order to help assess and prioritize investments

2. Literature Review

2.1. Renewable Energy Potential in the Amazon

The Amazon region possesses significant potential for various renewable energy sources, each offering unique advantages and facing specific challenges. Solar energy stands out due to the region's high solar irradiance, making it a viable solution for remote communities lacking access to conventional energy grids. Advancements in photovoltaic (PV) technologies, decreasing costs, and the flexibility of both grid-connected and off-grid systems enhance the feasibility of solar energy deployment. Successful implementations, such as solar projects in the Amanã Sustainable Development Reserve, demonstrate the effectiveness of solar energy in improving energy access and reducing inequality [Penteado et al., 2018].

Wind energy potential, while influenced by the dense vegetation affecting wind speeds, presents opportunities along the Amazonian coast through offshore wind projects. Small-scale wind turbines could be suitable for isolated communities where wind conditions are favorable, providing an additional renewable energy option [Jannuzzi et al., 1999].

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

Biomass energy generation leverages the Amazon's vast agricultural resources, utilizing agricultural residues like sugarcane bagasse and corn stover. This approach can contribute to power generation and biofuel production. However, sustainable management practices are crucial to prevent deforestation and biodiversity loss associated with biomass exploitation [Miranda et al., 2024]. Small hydropower plants (SHPs) offer another avenue for renewable energy, especially suitable for riverside communities. These plants require careful consideration of socio-environmental impacts to minimize effects on aquatic ecosystems. Exploring low-impact alternatives like free-flow turbines and hybrid systems combining SHPs with other renewable sources can contribute to sustainable solutions [Jannuzzi et al., 1999].

2.2. Socio-Environmental Impacts and Challenges

The development of renewable energy projects in the Amazon must consider various socio-environmental factors and address significant challenges to ensure sustainable outcomes.

Deforestation and biodiversity loss can result from the construction of energy infrastructure, particularly large-scale hydropower plants, leading to habitat fragmentation and adverse effects on flora and fauna. The loss of ecosystem services and the impact on carbon sequestration capacities exacerbate environmental concerns. Implementing sustainable practices, such as careful site selection, minimizing land use changes, and creating ecological corridors, are essential mitigation strategies. Additionally, the relationship between deforestation, biomass production, and the need for sustainable forest management practices is vital to balance energy development with conservation efforts [Velastegui-Montoya et al., 2022].

Community displacement and cultural impacts are significant concerns associated with large infrastructure projects. The relocation of indigenous and riverside communities can lead to loss of cultural identity, disruption of traditional livelihoods, and social upheaval. Ensuring free, prior, and informed consent, along with fair compensation and the implementation of social and economic development programs, is crucial to preserve cultural heritage and promote equitable outcomes. Inclusive decision-making processes and respect for indigenous rights are fundamental to ethical project implementation [Velastegui-Montoya et al., 2022].

Greenhouse gas emissions associated with renewable energy projects, particularly methane emissions from hydropower reservoirs due to decomposing organic matter, pose environmental challenges. These emissions can offset some of the climate benefits of renewable energy. Alterations in hydrological regimes resulting from dam construction can impact aquatic life, fishing activities, and agriculture, along with increasing risks of erosion and siltation. Strategies to minimize emissions, such as designing reservoirs to reduce biomass flooding and implementing technologies to capture methane, along with maintaining ecological flow to preserve aquatic ecosystems, are necessary [Arantes et al., 2022].

Infrastructure limitations and accessibility issues present logistical challenges for transporting equipment and materials in the Amazon. The region's vast and difficult terrain, combined with a lack of developed transportation networks, increases the complexity and cost of project implementation. High investment costs necessitate public-private partnerships and integrated planning to optimize resource utilization and ensure the financial viability of renewable energy projects. The reliance on fossil fuels for power generation in isolated communities highlights the urgency of transitioning to renewable sources. Overcoming data scarcity on energy potential and socio-environmental impacts requires investing in scientific research, developing efficient monitoring systems, and strengthening the technical capacity of government agencies and local communities [Miranda et al., 2024; Jean et al., 2021].

2.3. Opportunities for Sustainable Development and Energy Justice

Despite the challenges, renewable energy development in the Amazon presents significant opportunities for promoting sustainable development and energy justice.

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

Renewable energy projects can contribute to reducing inequality by providing equitable access to clean and affordable energy. This access can drive economic development, improve quality of life, and enhance access to essential services such as education and healthcare. Generating employment opportunities in the renewable energy sector can stimulate local economies and provide skills development for community members. Focusing on inclusive development ensures that the benefits of renewable energy reach all communities, regardless of location, income, or ethnicity, thus addressing issues of energy poverty and social inequity [Velastegui-Montoya et al., 2022].

Data-driven approaches are indispensable for informed decision-making in leveraging renewable energy effectively. They empower policymakers with evidence-based insights for optimizing resource allocation, assessing environmental and social impacts, and fostering transparency and accountability. Analyzing data on energy consumption patterns, resource availability, infrastructure needs, and community demographics helps tailor policies and projects to specific contexts, maximizing the impact of renewable energy initiatives. Data analytics also enables technological innovation and optimization of renewable energy systems, enhancing efficiency, reducing costs, and maximizing the reliability of infrastructure [Benes et al., 2024].

Community engagement is enhanced through data transparency and accessibility. Utilizing data visualization tools and participatory mapping platforms allows communities to understand project plans, voice concerns, and contribute to design and implementation. This participatory approach empowers local populations, promotes ownership, and ensures that renewable energy projects align with community values and aspirations. Incorporating case studies and concrete examples, such as citizen science initiatives where local communities participate in monitoring and data collection, demonstrates the practical application of data-driven approaches in promoting renewable energy and energy justice. These initiatives showcase how renewable energy can serve as a catalyst for sustainable development in the Amazon, addressing both environmental sustainability and social equity [Facchinelli et al., 2022].

3. Methodology

The methodology for analyzing potential solar investments in the Amazon region comprised several stages, beginning with data acquisition and preprocessing, followed by feature extraction, clustering, and the calculation and analysis of priority scores. To ensure robust results, a variety of methods and experimental plans were tested to ascertain the most effective approach for assessing solar energy potential.

3.1. Data Acquisition and Preprocessing

Data was sourced from multiple publicly available databases, encompassing socio-economic, geographical, and energy-related attributes of municipalities in the Amazon region. The acquisition process involved collecting information from government and energy regulatory agencies to ensure comprehensive coverage. Key data sources included municipality coordinates obtained via the Brazilian Institute of Geography and Statistics (IBGE) API [IBGE], historical solar radiation data retrieved from the Photovoltaic Geographical Information System (PVGIS) API provided by the European Commission's Joint Research Centre (JRC) [PVGIS], and socio-economic indicators such as Human Development Index (HDI), illiteracy rate, and electricity access percentages acquired from the Institute of Applied Economic Research (IPEA) API [IPEA]. The time-series data used ranged from 2008 to 2023.

The preprocessing phase involved several critical steps [Kirchner et al., 2016]. First, data cleaning addressed missing values and inconsistencies; numerical features with missing values were imputed using the median to enhance robustness against outliers, and municipalities lacking essential data were removed. Second, data integration involved merging the geographical coordinates, solar radiation data, and socio-economic indicators into a unified dataset based on municipality

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

identifiers. Third, feature engineering was applied to transform the time series solar radiation data into a set of descriptive features, as detailed in the next section. Finally, data normalization was performed using Min-Max scaling, ensuring all features ranged between 0 and 1 and contributed equally during the clustering phase. This preprocessing allowed to obtain a consistent and unified dataset with information from all obtained sources. Specifically, the feature engineering step made that instead of a set of at least 65 features for the time series data [Barandas et al., 2020], a total of 10 features was used, making the problem computationally less demanding and retaining only essential information from the series.

3.2. Feature Extraction from Time Series Data

A crucial aspect of the project involved extracting meaningful features from the solar radiation time series for each municipality, rather than using the raw data directly. This process aimed to capture essential characteristics of solar radiation patterns. Feature extraction was primarily conducted using two libraries: TSFEL (Time Series Feature Extraction Library), which extracted a broad range of statistical, temporal, and fractal features like mean, standard deviation, entropy, autocorrelation, and fractal dimensions [Barandas et al., 2020]; and catch22, which provided 22 canonical time series features designed for diversity and complementarity [Lubba et al., 2019].

Following extraction, feature selection methods were applied to reduce dimensionality and retain the most relevant features [Venkatesh and Anuradha, 2019]. These methods included removing low-variance features using a Variance Threshold, selecting features highly correlated with preliminary clustering labels via the ANOVA F-test, and identifying features with the highest Mutual Information relative to clustering labels. The resulting set of extracted and selected features, combined with socio-economic indicators, formed the final dataset for clustering and analysis.

3.3. Experimental Plan and Method Variations

To identify the optimal strategy for clustering and prioritizing municipalities, multiple experimental setups were designed and tested. These experiments varied the clustering algorithms, feature selection methods, and preprocessing techniques employed.

3.3.1. Clustering Methods

Several clustering algorithms were evaluated for their suitability. K-means clustering was used as a primary method due to its simplicity and efficiency, with tests run for various numbers of clusters, k. Hierarchical clustering, specifically agglomerative (bottom-up) approaches with different linkage methods (ward, complete, average), was tested to explore potential hierarchical relationships. Additionally, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) was employed to identify density-based clusters of arbitrary shapes and detect outliers [Kirchner et al., 2016].

The impact of different feature selection and preprocessing methods on clustering performance was assessed. Experiments included using all extracted features as a baseline, applying a Variance Threshold (removing features with variance below 0.05, resulting in 10 features), selecting the top 10 features based on ANOVA F-scores, choosing features with the highest Mutual Information relative to cluster labels, and using only the 22 canonical features extracted by catch22. The catch22 set is noted for strong performance in time-series classification tasks and minimal redundancy [Lubba et al., 2019].

The quality of the resulting clusters was quantitatively assessed using established metrics [Kapil and Chawla, 2016]. The **Silhouette Score** (S) measured how similar an object is to its own cluster compared to others, calculated as:

$$S = \frac{b - a}{\max(a, b)} \tag{1}$$

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

where a is the average intra-cluster distance and b is the average nearest-cluster distance.

The **Davies-Bouldin Index (DB)** evaluated the average similarity between clusters, with lower values indicating better separation:

$$DB = \frac{1}{k}\sum_{i=1}^{k}\max_{j\neq i}\left(\frac{\sigma_i + \sigma_j}{d_{ij}}\right) \tag{2}$$

where σ_i is the average distance between points in cluster i and its centroid, and d_{ij} is the distance between the centroids of clusters i and j.

The **Calinski-Harabasz Index (CH)** measured the ratio of between-cluster dispersion to within-cluster dispersion, where higher values suggest better-defined clusters:

$$CH = \frac{\mathrm{Tr}(B_k)}{\mathrm{Tr}(W_k)} \times \frac{n-k}{k-1} \tag{3}$$

Here, $\mathrm{Tr}(B_k)$ is the trace of the between-cluster dispersion matrix, $\mathrm{Tr}(W_k)$ is the trace of the within-cluster dispersion matrix, n is the total number of data points, and k is the number of clusters.

To aggregate these three metrics into a single performance measure, a **Figure of Merit (FoM)** was defined for each experimental configuration i:

$$FoM_i = \frac{S_i - S_{min}}{S_{max} - S_{min}} + \frac{DB_{max} - DB_i}{DB_{max} - DB_{min}} + \frac{CH_i - CH_{min}}{CH_{max} - CH_{min}} \tag{4}$$

This FoM normalizes each metric score between 0 and 1 (with DB inverted as lower is better) and sums them, providing a composite score where higher values indicate better overall clustering performance across the metrics.

An experiment plan systematically varied the clustering algorithms, number of clusters, and feature selection methods. For each combination, clustering was performed, evaluation metrics were calculated, and the results were aggregated to identify the best-performing configurations based on the FoM and individual metric scores.

3.4. Clustering Analysis

Clustering algorithms were applied as outlined in the experimental plan to identify groups of municipalities with similar potential for solar investments. The optimal number of clusters was determined by running the experiments across various cluster numbers and parameter combinations, evaluating the results using the metrics described previously. Based on these evaluations, the optimal clustering configuration was identified. For instance, initial results suggested that k-means clustering with $k = 4$ and features selected via mutual information maximization yielded strong performance.

3.5. Priority Scoring

A **Priority Score (PS)** was developed and calculated for each municipality to assess its suitability for solar investment. This score represented a weighted combination of selected features, inspired by similar works in Multi Criteria Decision Making [Williams and Cai, 2025; Şahin, 2021], designed to emphasize areas with both high solar potential and significant socio-economic need. The formula used was:

$$PS = w_1 \times R_{\text{max}} + w_2 \times (1 - E) + w_3 \times (1 - \text{HDI}) \tag{5}$$

where R_{max} represents the normalized maximum solar radiation value for the municipality, E is the normalized proportion of households with electricity access, and HDI is the normalized Human Development Index. The weights w_1, w_2, w_3 reflect the relative importance assigned to each component; in this study, they were set to $w_1 = 0.3$, $w_2 = 0.5$, and $w_3 = 0.2$, prioritizing electricity access needs while also considering solar resources and overall development level. This scoring mechanism aimed to identify municipalities where solar investments could have the greatest impact.

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

3.6. Analysis and Interpretation

Following the clustering and scoring, municipalities were ranked based on their Priority Scores, allowing for detailed analysis of the top-ranked candidates. The clustering results were utilized to identify prevailing patterns and group municipalities exhibiting similar characteristics. The analysis focused on identifying clusters characterized by high average maximum solar radiation, indicating strong solar energy generation potential, and examining the socio-economic profiles within each cluster to pinpoint regions with greater need due to factors like low electricity access and lower HDI.

This comprehensive methodology provided a systematic, data-driven approach for assessing solar investment potential in the Amazon region. By integrating clustering techniques, time series feature extraction, and a priority scoring system sensitive to socio-economic context, the analysis identified municipalities poised for impactful solar energy projects, offering a scalable framework adaptable to other regions for renewable energy planning.

4. Results

4.1. Feature Extraction and Selection

As detailed in the Methodology, time series feature extraction techniques were employed using TSFEL and catch22 libraries to capture the temporal dynamics of solar radiation data. Subsequently, feature selection methods, including Variance Threshold, Mutual Information, and ANOVA F-Test, were applied to reduce dimensionality and identify the most informative features for clustering. The effectiveness of these different feature sets was evaluated during the clustering experiments. The retained features used in the clustering analysis are shown in Table 1.

Table 1: Description of Features and Indicators. *Italic* indicates features extracted from the time series of solar irradiation

Feature / Indicator	**Data Type**	**Description**
Illiteracy rate (%)	Float	Illiteracy rate for people over 15 years old
GDP per capita (R$)	Float	Municipality GDP normalized by population
Rural Population	Integer	Number of people located in the municipality's rural area
BPC Recipients	Integer	Number of BPC recipients in the municipality
% households with electricity	Float	% of households with access to electricity in the municipality
ECDF Percentile	*Float*	*Proportion of data points less than or equal to a given value*
Interquartile Range	*Float*	*Spread of the middle 50% of the data*
Kurtosis	*Float*	*Measure of data distribution tailedness and outlier presence*
Max	*Float*	*Highest value in the time series*
Mean Absolute Deviation	*Float*	*Average absolute difference from the mean*
Median Absolute Deviation	*Float*	*Robust measure of variability from the median*
Peak-to-Peak Distance	*Float*	*Difference between maximum and minimum values*
Standard Deviation	*Float*	*Measure of data point spread from the mean*
Variance	*Float*	*Square of standard deviation quantifying data variability*
Maximum Fractal Length	*Float*	*Complexity or self-similarity of the time series*

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

4.2. Clustering Experiments

An extensive experimental plan, detailed in the Methodology, was executed to determine the most effective clustering approach. This involved testing various combinations of clustering algorithms (k-Means, Hierarchical, DBSCAN with varying parameters), feature sets (all features, variance threshold reduced, ANOVA selected, mutual information selected, catch22 only), and numbers of clusters. Full experiment plan results is available in this link

Cluster quality for each experiment was assessed using the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index, aggregated into the Figure of Merit (FoM). Table 2 summarizes the performance of the top configurations identified through this systematic evaluation.

Table 2: Summary of Top Performing Clustering Experiment Results

Algorithm	Feature Set	Parameters	Sil.	DBI	CHI	FoM
k-Means	Mutual Info Features	$k = 4$	0.454	0.767	**1280.3**	**2.886**
Hierarchical	Mutual Info Features	$k = 4$, linkage=complete	0.423	**0.744**	1156.0	2.882
DBSCAN	Variance Threshold	$\varepsilon = 0.5$, min_samples=5	**0.484**	0.774	616.1	2.155

Based on the aggregated FoM and individual metrics, the k-means algorithm using $k = 4$ clusters and features selected via mutual information maximization demonstrated the best overall performance among the tested configurations.

4.3. Clustering Results

Following the experimental evaluation, the k-means clustering result with $k = 4$ and mutual information-selected features was chosen for in-depth analysis.

4.3.1. Cluster Characterization

Analysis of the data within each of the four clusters revealed distinct profiles:

- **Cluster 0**: Characterized by municipalities with the highest maximum solar radiation values alongside moderate socio-economic indicators.
- **Cluster 1**: Comprised municipalities showing high socio-economic development (higher HDI and electricity access) but possessing lower solar radiation potential.
- **Cluster 2**: Grouped municipalities with generally lower socio-economic indicators and moderate solar radiation potential.
- **Cluster 3**: Contained municipalities exhibiting high solar radiation potential coupled with lower socio-economic development levels.

4.3.2. Visualization of Clusters

The geographical distribution of these four clusters across the Amazon region is visualized in the interactive map shown in Figure 1, which can be accessed from this link.

4.4. Priority Scoring and Identification of High-Potential Municipalities

The Priority Score (PS), defined in the Methodology section (Equation 5), was calculated for each municipality to rank them for solar energy investment suitability, using weights $w_1 = 0.3$, $w_2 = 0.5$, and $w_3 = 0.2$. The formula $PS = 0.3 \times R_{\text{max}} + 0.5 \times (1 - E) + 0.2 \times (1 - \text{HDI})$ prioritizes areas balancing high solar potential (R_{max}) with significant need, indicated by lower electricity access (E) and lower human development (HDI).

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

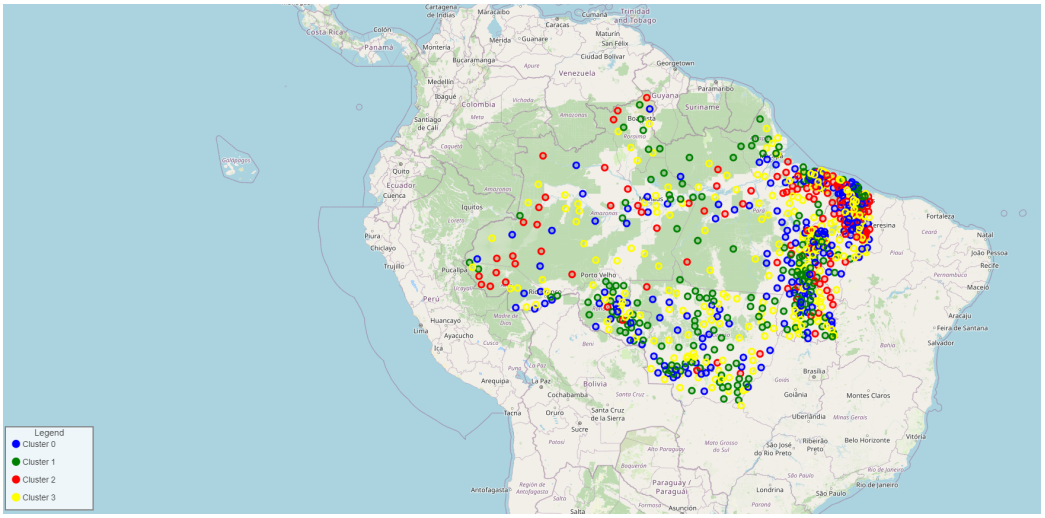


Figure 1: Geographical Distribution of Clusters in the Amazon Region

Table 3: Top 10 Municipalities for Solar Investment Based on Priority Score

Municipality	Priority Score	Cluster
Uiramutã, RR	0.869	2
Santa Isabel do Rio Negro, AM	0.759	0
Jordão, AC	0.727	2
Campinápolis, MT	0.705	0
Amajari, RR	0.691	2
Marajá do Sena, MA	0.679	2
Cachoeira do Piriá, PA	0.675	0
Melgaço, PA	0.664	0
Marechal Thaumaturgo, AC	0.658	2
Guajará, AM	0.645	0

4.4.1. Top Municipalities Identified

Table 3 presents the top 10 municipalities identified based on the highest Priority Scores, indicating their strong candidacy for solar investment.

These results highlight municipalities that not only possess favorable solar resources but also exhibit socio-economic conditions suggesting that solar investments could yield significant developmental benefits. Notably, many top-ranked municipalities fall into Cluster 2 (moderate solar, lower socio-economic) or Cluster 0 (high solar, moderate socio-economic), aligning with the priority score's aim.

4.5. Analysis and Interpretation

The combined analysis of clustering results and priority scores offers valuable insights for targeted solar investment. Clusters 0 and 3 contain municipalities with high solar radiation potential, but those in Cluster 2 (often ranking high in PS) represent areas where moderate solar potential coincides with lower socio-economic development, indicating a potentially higher impact from investment. The prioritization mechanism effectively highlighted municipalities where lower HDI and electricity access amplify the need for intervention. Consequently, investments in these high-priority areas, often located in more remote or underserved regions as seen in Figure 1, could

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

significantly improve living conditions and foster sustainable development. This analysis confirms the utility of integrating environmental (solar radiation) data with socio-economic indicators to guide effective and equitable energy investments.

5. Conclusions

This study successfully demonstrated a robust, data-driven methodology for identifying and prioritizing municipalities within the Amazon region suitable for solar energy investment. By integrating solar radiation data, socio-economic indicators, advanced time series feature extraction, and clustering techniques, the analysis provided a nuanced understanding of regional potential and needs.

The key conclusions drawn from this work are fourfold. First, the use of time series feature extraction methods proved effective in capturing essential patterns within solar radiation data, which enhanced the subsequent clustering process. Second, systematic experimentation identified an optimal clustering configuration – specifically, k-means with $k = 4$ using features selected via mutual information – that effectively grouped municipalities based on relevant characteristics. Third, the developed Priority Score served as a valuable tool, successfully combining solar potential with socio-economic need to pinpoint municipalities where investments are likely to have the most significant positive impact. Finally, the analysis revealed that high-priority municipalities are often situated in remote areas characterized by high solar radiation and lower socio-economic development, underscoring the potential for targeted solar investments to advance energy justice and sustainable development goals in the region.

5.1. Recommendations

Based on these findings, several recommendations emerge. Policymakers are encouraged to **focus investments** on municipalities identified with high priority scores to strategically address energy poverty and promote equitable regional development. Further **research** incorporating more granular socio-economic data, local infrastructure assessments, and potentially climate resilience factors could refine the prioritization process. Crucially, **community engagement** throughout the planning and implementation phases of any solar project is vital to ensure local acceptance, sustainability, and alignment with community needs. Lastly, continuous efforts towards **data enhancement**, particularly improving the availability and quality of data for remote areas, will strengthen the accuracy and reliability of future analyses.

5.2. Future Work

Opportunities for future research include exploring the integration of other renewable energy potentials (e.g., hydro, biomass) for a more holistic energy planning perspective, modeling the temporal dynamics of socio-economic indicators, and conducting longitudinal studies to assess the long-term impacts of solar investments on regional development trajectories. Furthermore, deepening the engagement with local communities and policymakers through participatory research methods could yield richer insights and more contextually appropriate strategies.

5.3. Closing Remarks

The methodology presented herein highlights the significant potential of data-driven approaches to tackle complex developmental challenges like energy access in sensitive regions. By effectively synthesizing environmental and socio-economic data, stakeholders can make more informed, targeted decisions that foster sustainable development and promote energy equity throughout the Amazon region.

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

References

Alegre-Bravo, A. and Lindsay Anderson, C. (2023). Exploring the influence of multidimensional variables on access to electricity in rural areas of the global south. *Appl. Energy*, 333(120509): 120509.

Anderson, J., Rainie, L., and Vogels, E. (2021). Experts say the 'new normal' in 2025 will be far more tech-driven, presenting more big challenges. https://www.pewresearch.org/wp-content/uploads/sites/20/2021/02/PI_2021.02.18_New-Normal-2025_FINAL.pdf. [Accessed 04-12-2024].

Arantes, C. C., Laufer, J., Pinto, M. D. d. S., Moran, E. F., Lopez, M. C., Dutka-Gianelli, J., Pinto, D. M., Chaudhari, S., Pokhrel, Y., and Doria, C. R. (2022). Functional responses of fisheries to hydropower dams in the amazonian floodplain of the madeira river. *Journal of Applied Ecology*, 59(3):680–692.

Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T., and Gamboa, H. (2020). TSFEL: Time series feature extraction library. *SoftwareX*, 11(100456): 100456.

Benes, K. J., Porterfield, J. E., and Yang, C. (2024). Ai for energy: Opportunities for a modern grid and clean energy economy. Technical report, U.S. Department of Energy.

Facchinelli, F., Pappalardo, S. E., Della Fera, G., Crescini, E., Codato, D., Diantini, A., Jimenez, D. R. M., Mendoza, P. E. F., Bignante, E., and De Marchi, M. (2022). Extreme citizens science for climate justice: linking pixel to people for mapping gas flaring in amazon rainforest. *Environmental Research Letters*, 17(2):024003.

Ferreira, G. W., Reboita, M. S., and Ribeiro, J. (2024). Assessment of the solar energy potential over south america estimated by cmip6 mod-els in the present and future climate. *Journal of Environmental & Earth Sciences*, 6(2):110–143.

IBGE. Malha Municipal — IBGE — ibge.gov.br. https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html. [Accessed 04-12-2024].

IPEA. Ipeadata — ipeadata.gov.br. http://www.ipeadata.gov.br/Default.aspx. [Accessed 04-12-2024].

Jannuzzi, G. M., Gadgil, A., Silva, E., and Leonardi, M. L. (1999). Maximally-renewable and energy-efficient cost-neutral energy scenario for manaus, brazil. *Applied Energy*, 63(4):237–253.

Jean, W., Arcela, A., van Els, R. H., Brasil Junior, A. C., Echeverry, S. M. V., de Miranda, A. R. S., and de Souza, J. S. (2021). A gis for rural electrification strategies in the brazilian amazon. *Papers in Applied Geography*, 7(3):239–255.

Kapil, S. and Chawla, M. (2016). Performance evaluation of k-means clustering algorithm with various distance metrics. In *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, p. 1–4.

Kirchner, K., Zec, J., and Delibašić, B. (2016). Facilitating data preprocessing by a generic framework: a proposal for clustering. *Artif. Intell. Rev.*, 45(3):271–297.

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br

Linguet, L., Joubert, A., Asso, A., and Adouki, F. (2009). Analysis of the productivity of solar home systems for the electrification of isolated dwellings in french guiana. *Renewable and Sustainable Energy Reviews*, 13(9):2447–2456.

Lubba, C. H., Sethi, S. S., Knaute, P., Schultz, S. R., Fulcher, B. D., and Jones, N. S. (2019). catch22: CAnonical time-series CHaracteristics. *Data Min. Knowl. Discov.*, 33(6):1821–1852.

Miranda, D. d. S., Araujo, J. B. d., and Kulay, L. (2024). Replacing diesel with biogas in decentralized electricity generation systems: a feasibility assessment from workers' perspective with a modified human development index (hdi) approach. *Socio-Ecological Practice Research*, 6: 263–281.

Penteado, I. M., do Nascimento, A. C. S., Corrêa, D., Moura, E. A. F., Zilles, R., Gomesa, M. C. R. L., Piresa, F. J., Britoa, O. S., da Silvaa, J. F., Reisa, A. V., et al. (2018). Among people and artifacts: Actor-network theory and the adoption of solar ice machines in the brazilian amazon. *Energy Research & Social Science*, 45:124–132.

PVGIS. Photovoltaic Geographical Information System (PVGIS) — joint-research-centre.ec.europa.eu. https://joint-research-centre.ec.europa.eu/photovoltaic-geographical-information-system-pvgis_en. [Accessed 04-12-2024].

Şahin, M. (2021). A comprehensive analysis of weighting and multicriteria methods in the context of sustainable energy. *Int. J. Environ. Sci. Technol. (Tehran)*, 18(6):1591–1616.

Velastegui-Montoya, A., de Lima, A., and Hidalgo-Monroy, V. (2022). Sustainability of hydroelectric plants in the brazilian amazon based on socio-economic indicators: The case of tucuruí hydroelectric power plant. *Sustainability*, 14(3):1630.

Venkatesh, B. and Anuradha, J. (2019). A review of feature selection and its methods. *Cybern. Inf. Technol.*, 19(1):3–26.

Williams, A. and Cai, Y. (2025). Insights into weighted sum sampling approaches for multi-criteria decision making problems. URL https://arxiv.org/abs/2410.03931.

DOI: 10.59254/sbpo-2025-212546
https://proceedings.science/p/212546?lang=pt-br