Transfer Learning for PV Generation Forecasting under Data Scarcity: A Case Study in Brazil

Fernando Vasconde de Arruda[1], Marcelo Almeida Pinho[1] and Fernando Ramos Martins[2]

[1] Institute of Energy And Environment, University of São Paulo (IEE-USP), São Paulo (Brazil)

[2] Institute of Marine Science, Federal University of São Paulo (Unifesp), Santos (Brazil)

Abstract

This study investigates the use of Transfer Learning (TL) techniques to improve photovoltaic (PV) power forecasting in Brazilian regions characterized by heterogeneous climatic conditions and limited data availability. An Encoder–Decoder Long Short-Term Memory (EDLSTM) model was initially trained using a multi-year dataset from São Paulo and subsequently adapted to three target sites, Tacaratu (PE), Itacarambi (MG), and Florianópolis (SC), through five distinct TL strategies. The experiments evaluated model performance under varying data availability scenarios (1 to 12 months) and were benchmarked using a persistence model. Among the tested configurations, TL1, which reuses pretrained weights without layer freezing, consistently achieved the lowest normalized Root Mean Square Error (pRMSE) and demonstrated superior stability across all climate contexts. These findings confirm that knowledge transfer from data-rich domains significantly enhances predictive performance in data-scarce regions, reducing both training effort and computational cost. The proposed approach provides a scalable and data-efficient framework for PV forecasting, contributing to the reliable integration of solar energy into the Brazilian power system.

Keywords: Transfer Learning, Photovoltaic Power Forecasting, LSTM, Encoder Decoder, Data Scarcity

1. Introduction

The global transition towards low-carbon energy systems has accelerated significantly over the past decade, driven by the urgent need to mitigate climate change and enhance energy security. According to the International Renewable Energy Agency (IRENA, 2023), achieving the 1.5°C climate target necessitates a substantial transformation of the global electricity generation sector, with renewable sources constituting over 90% of the power mix by the year 2050. Solar photovoltaic (PV) technology has emerged as a cornerstone of this transition due to its modularity, declining costs, and scalability. From 2020 to 2023, the global installed capacity of solar PV more than doubled, making it one of the fastest-growing renewable technologies (REN21, 2024; IEA, 2024).

In the context of Brazil, solar energy has undergone a substantial expansion in recent years, propelled by advantageous policy frameworks and the presence of abundant solar resources. According to the 2024 National Energy Balance (EPE, 2024), solar PV generation reached 50.6 TWh in 2023, with an installed capacity of 37.8 GW, corresponding to an annual growth rate of 54.8%. The nation currently exhibits one of the highest percentages of renewable energy in the global power mix, with a percentage exceeding 89% of total generation (IEA, 2023). Despite this progress, integrating intermittent PV generation into the electrical grid poses significant operational challenges, particularly in the context of short-term planning and dispatching. Accurate forecasting is imperative for ensuring grid reliability and economic efficiency.

The development of reliable PV forecasting models remains limited by data scarcity, particularly for new installations and remote regions lacking monitoring infrastructure. A significant number of newly commissioned PV systems, particularly in remote or under-monitored regions, lack long-term historical records of power generation and meteorological conditions. This limitation restricts the training of data-driven models. Conventional deep learning architectures, including Long Short-Term Memory (LSTM) networks, generally necessitate substantial and uninterrupted datasets to attain optimal performance. In such contexts, the predictive accuracy of models frequently declines when applied to data-limited sites.

10.18086/swc.2025.04.48 Available at http://proceedings.ises.org

To overcome these challenges, recent advances in machine learning, particularly Transfer Learning (TL), have shown promise in reusing knowledge from data-rich PV systems. The utilization of TL techniques facilitates the reapplication of knowledge acquired by a model trained on a data-rich source domain to enhance learning in a related, data-scarce target domain. In the domain of solar forecasting, research conducted by Sarmas et al. (2022) and Depoortere et al. (2024) has demonstrated that TL can significantly improve forecast accuracy in situations where data is limited by leveraging pretrained models across various locations and climatic conditions. However, despite the growing international literature on the subject, applications of transfer learning to PV forecasting remain scarce in the Brazilian context. This is because climate diversity and regional variability pose unique modeling challenges in Brazil.

The objective of this study is to assess the efficacy of transfer learning methodologies in facilitating the short-term prediction of photovoltaic power generation across Brazilian regions, which exhibit a range of climatic characteristics and varying degrees of data accessibility. In this study, we employed an Encoder–Decoder LSTM (EDLSTM) architecture, trained on data from São Paulo, as the source model. We then systematically analyzed five TL strategies and assessed their capacity to adapt to semi-arid (Pernambuco), dry tropical (Minas Gerais), and humid subtropical (Santa Catarina) environments. The findings underscore the robustness and scalability of transfer learning as a practical and data-efficient approach to PV forecasting in heterogeneous conditions.

This paper is structured as follows: Section 2 describes the methodology, including datasets, model architecture, and TL strategies; Section 3 presents the results and discussion; and Section 4 concludes with the main findings and perspectives.

2. Methodology

This section presents the methodological framework developed to evaluate the performance of transfer learning strategies in PV power forecasting under data scarcity conditions. The workflow integrates five main stages: (i) data acquisition and preprocessing; (ii) development of the base model using an Encoder–Decoder Long Short-Term Memory (EDLSTM) architecture; (iii) adaptation of the pretrained model to new climatic regions through transfer learning; (iv) model evaluation using statistical metrics and a persistence-based benchmark; and (v) analysis of the results under different data availability scenarios.

The complete methodological process is illustrated in Figure 1, which summarizes the data flow and experimental stages. The diagram highlights the central role of the São Paulo dataset in training the base model, which subsequently serves as the knowledge source for adaptation to three locations, Tacaratu (PE), Itacarambi (MG), and Florianópolis (SC). Each stage of the flowchart represents a key methodological component: data preparation, model training, transfer learning implementation, evaluation, and performance comparison.

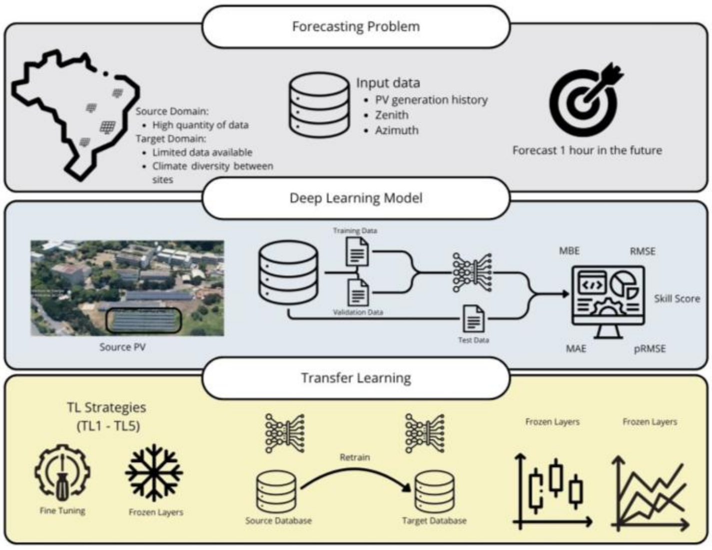

Figure 1: Workflow of the methodological framework, showing data preparation, model training, transfer learning adaptation, and evaluation steps.

2.1 Study area and dataset

The experimental analysis was conducted using PV systems located in distinct climatic regions of Brazil. These regions were selected to capture the influence of diverse meteorological conditions on solar power generation and to test the robustness of the proposed transfer learning approach. The source domain corresponded to a ground-mounted PV system installed at the Institute of Energy and Environment of the University of São Paulo (IEE-USP), located in São Paulo (SP).

In order to assess the transfer learning technique of the model to different operational contexts, three target systems were selected: Tacaratu, located in the semi-arid climate of Pernambuco; Itacarambi, in the dry tropical climate of Minas Gerais; and Florianópolis, in the humid subtropical climate of Santa Catarina (Figure 2 and Table 1). These sites exemplify contrasting atmospheric and irradiance regimes across Brazil, exhibiting substantial variations in cloud cover patterns, seasonal behavior, and temperature ranges. This variability enables a thorough evaluation of the model's generalizability when applied to regions with distinct climatic dynamics and limited data availability.

Table 1: Geographic and operational characteristics of PV systems used in the study.

Generator	Location	Latitude	Longitude	Nominal (kW)	Climatology	From date	To date
PV0	São Paulo / SP	-23.55	-46.73	256	Subtropical Humid	2018	2022
PV1	Tacaratu / PE	-9.07	-38.15	5000	Semi-arid	2021	2022
PV2	Itacarambi / MG	-15.13	-44.14	3375	Tropical Semi-arid	2021	2022
PV3	Florianopolis / SC	-27.58	-48.51	930	Subtropical Humid	2021	2022

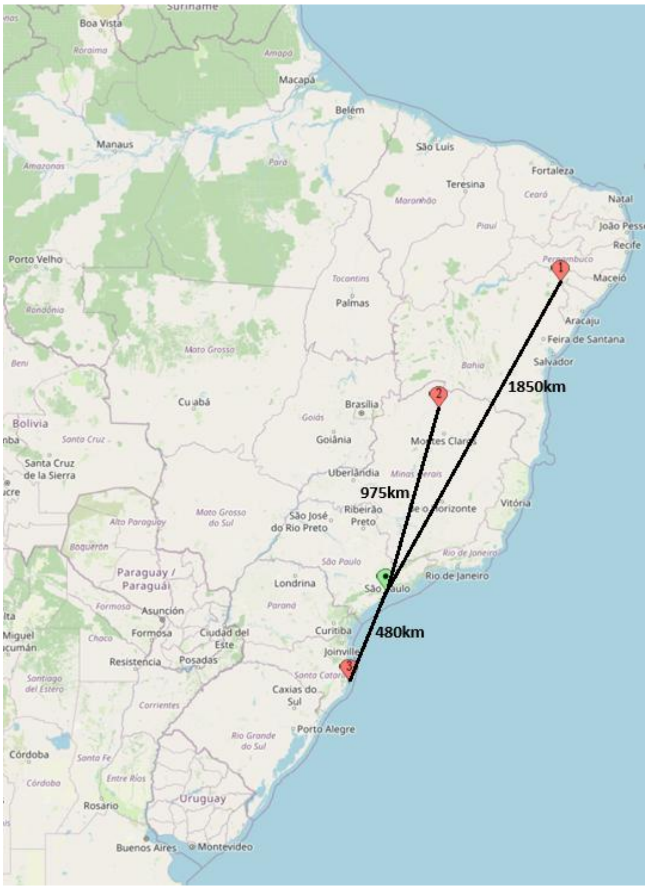

Figure 2: Map showing the locations and distances of the photovoltaic generators. The source photovoltaic generator is shown in green, and the generators in which transfer learning was used are shown in red.

For model training, the historical photovoltaic power generation data were used together with solar zenith and azimuth angles. The normalization of the data was achieved by dividing the historical power series by its nominal power. In the context of solar angles, the cosine of the zenith angle and the sine of the azimuth angle were employed.

Since São Paulo served as the source domain for network training, the dataset was divided as follows: data from 2018 to 2021 were used for model training, with 80% of this period dedicated to training and 20% reserved for validation, while the 2022 data were used exclusively for testing to ensure temporal independence of the evaluation set.

The application of TL strategies used 2021 data exclusively for training purposes. In contrast, the initial 15 days of each month in 2022 were allocated for validation, while the subsequent 15 days were designated for testing. This configuration was designed to prevent data leakage between the different stages, thereby preserving the statistical integrity of the experimental results.

2.2 Deep Learning Model

The predictive model is based on an Encoder–Decoder Long Short-Term Memory (EDLSTM) architecture (Figure 3), which is particularly suitable for sequence-to-sequence forecasting. The selection of this architecture was made based on its capacity to encompass both short-term and long-term temporal dependencies present within multivariate time series. This property is imperative for the effective modeling of the nonlinear dynamics of PV generation, which is, in turn, influenced by meteorological variability.

As illustrated in Figure 3, the proposed model is composed of several interconnected components. The diagram under consideration elucidates the two primary components: the encoder, which is responsible for processing the input sequence and summarizing it into a latent feature representation (context vector), and the decoder, which is tasked with the reconstruction of the predicted output sequence from that latent representation (Cho et al., 2014; Sutskever et al., 2014). A magnified view of the LSTM cell is provided to illustrate the internal gating mechanisms (input, forget, and output gates) that regulate information flow and preserve temporal dependencies across time steps (Hochreiter and Schmidhuber, 1997).

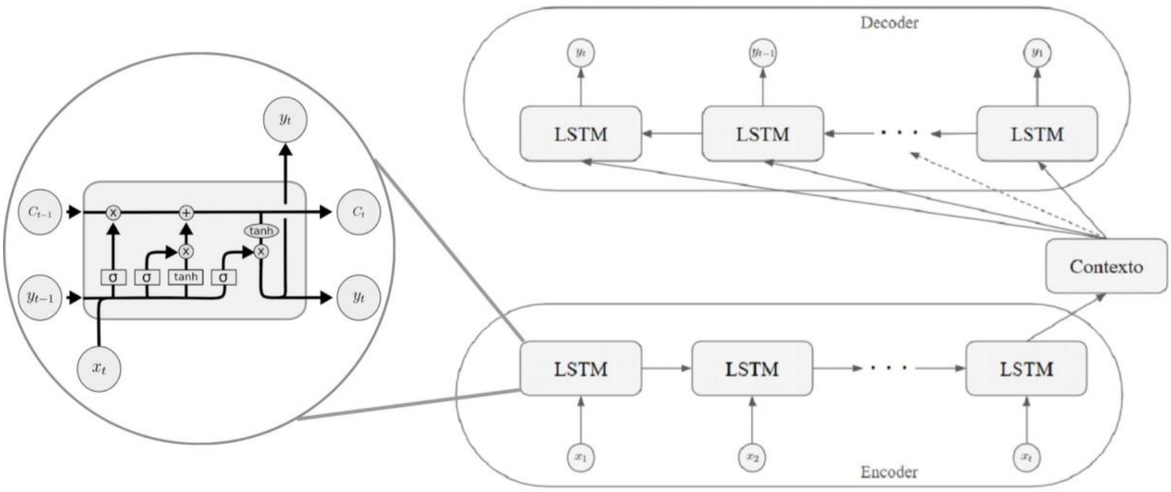

Figure 3: Structure of the Encoder–Decoder LSTM architecture used for short-term PV forecasting, highlighting the encoder, decoder, and LSTM gating mechanisms.

Equations (1)–(6) describe the internal dynamics of an LSTM unit, where σ denotes the sigmoid activation and tanh the hyperbolic tangent function.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad \text{(eq. 1)}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \text{(eq. 2)}$$

$$\tilde{C}_t = \tanh\,(W_C \cdot [h_{t-1}, x_t] + b_C) \quad \text{(eq. 3)}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad \text{(eq. 4)}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \text{(eq. 5)}$$

$$h_t = o_t * \tanh\,(C_t) \quad \text{(eq. 6)}$$

where f_t, i_t, and o_t represent the forget, input, and output gates, respectively; $C_t$is the cell state; and $h_t$is the hidden state at time t. These mechanisms enable the LSTM to selectively retain or discard information, effectively learning long-term dependencies in solar generation sequences affected by temporal lag and weather variability.

2.3 Transfer Learning Strategies

Transfer Learning is a machine learning technique that facilitates the utilization of knowledge acquired during the training of a model in a source domain to enhance performance in a different but related target domain. In the context of photovoltaic power forecasting, TL is beneficial in overcoming challenges such as the scarcity of historical data and significant climatic variability between locations. This strategy enables the adaptation of models pre-trained in data-rich environments to new locations with limited meteorological and generation data, thereby maintaining high predictive performance with reduced training time. (Pan and Yang, 2010; Zhuang et al., 2021).

In this framework, the Encoder–Decoder LSTM (EDLSTM) model trained in São Paulo (SP) serves as the source model. At the same time, the systems in Tacaratu (PE), Itacarambi (MG), and Florianópolis (SC) represent the target domains. The pretrained model weights from the source domain were reused as the initial condition for fine-tuning in each target domain. This procedure enables the network to maintain the general spatiotemporal characteristics of PV generation while adapting to local meteorological and operational patterns.

To systematically analyze the influence of knowledge reuse on forecasting performance, five transfer learning strategies (TL1–TL5) were implemented. The distinguishing characteristics of these strategies include variations in layer freezing, weight initialization, and the extent of fine-tuning applied to the network.

- **TL1:** All layers are retrained using the target data, starting from the pretrained weights. This configuration provides complete adaptation to the new environment while preserving initial feature representations from the source model.
- **TL2:** The encoder layers are frozen, while the output layer is replaced and trained from scratch, enabling adaptation of the output mapping to new data without altering the learned temporal representations.
- **TL3:** Similar to TL2, but the output layer weights are retained and fine-tuned instead of being reinitialized.
- **TL4:** Starting from the TL2 configuration, all layers are progressively unfrozen during training, with a reduced learning rate ($\alpha = 10^{-6}$) to prevent abrupt weight updates and catastrophic forgetting.
- **TL5:** Similar to TL4 but starting from the TL3 configuration, combining gradual adaptation with full reuse of pretrained weights.

2.4 Evaluation Method

The performance of the transfer learning models was evaluated using a set of statistical metrics commonly adopted in the literature on PV forecasting. These include the Root Mean Square Error (RMSE), the Mean Bias Error (MBE), and the Mean Absolute Error (MAE), as well as the normalized RMSE (pRMSE) based on the mean power output. Additionally, a skill score (SS) was computed to compare the proposed models against a reference persistence model, providing a relative measure of improvement in forecasting accuracy.

The Root Mean Square Error (RMSE) quantifies the quadratic mean of deviations between predicted ($\hat{P}_t$) and observed (P_t) PV power at time t. This approach penalizes larger errors more heavily. It is expressed as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N} (P_t - \hat{P}_t)^2} \quad \text{(eq. 7)}$$

where N is the number of forecasted instances. Lower RMSE values indicate higher predictive accuracy.

The Mean Bias Error (MBE) provides information about the systematic tendency of the model to overestimate or underestimate the PV generation. It is calculated as:

$$MBE = \frac{1}{N}\sum_{t=1}^{N} (P_t - \hat{P}_t) \quad \text{(eq. 8)}$$

A positive MBE suggests underestimation, while a negative value indicates overestimation by the model.

The Mean Absolute Error (MAE) measures the average magnitude of errors without considering their direction, offering a more interpretable indication of typical forecast deviation. It is defined as:

$$MAE = \frac{1}{N}\sum_{t=1}^{N} \mid P_t - \hat{P}_t \mid \qquad \text{(eq. 9)}$$

To account for variations in power output magnitude between different PV systems, the normalized RMSE (pRMSE) was used. This metric scales the RMSE by the mean observed PV power ($\bar{P}$), facilitating comparisons across sites with distinct generation capacities:

$$pRMSE = \frac{RMSE}{\bar{P}} \qquad \text{(eq.10)}$$

The Skill Score (SS) evaluates the relative improvement of a given model over a reference model, typically the persistence model. It is expressed as:

$$SS = 1 - \frac{RMSE_{model}}{RMSE_{pers}} \qquad \text{(eq.11)}$$

where $RMSE_{model}$ and $RMSE_{pers}$ correspond to the RMSE values of the evaluated model and the persistence model, respectively. Positive values of SS indicate that the model outperforms the persistence benchmark, while negative values denote inferior performance.

The persistence model serves as a baseline for short-term forecasting and assumes that the future PV power will be equal to the most recent measured value, expressed as:

$$\hat{P}_{t+1} = \left(\frac{P_t}{P_{t,clr}}\right) \times P_{t+1,clr} \qquad \text{(eq.12)}$$

where P_t is the measured PV power at time t, $P_{t,clr}$ is the expected clear-sky power estimated from the clear-sky irradiance model, and $\hat{P}_{t+1}$ is the predicted power at time $t+1$ according to the persistence assumption.

3. Results

3.1 Performance in the Source Domain

Before evaluating the transfer learning strategies, it is essential to establish the baseline performance of the EDLSTM model in the source domain, São Paulo/SP. This model served as the foundation for all subsequent transfer operations, and its performance provides the reference against which the efficiency and adaptability of the transfer processes were measured.

Overall, the EDLSTM achieved strong results in the source domain, with a pRMSE of approximately 25.3% and a Skill Score of 15.16%. Figure 4 presents the comparison between the predicted and observed power for São Paulo.

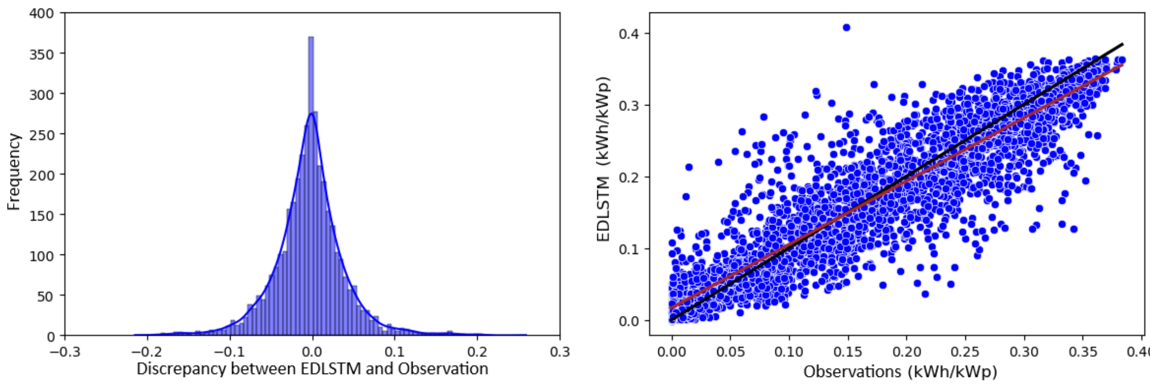

Figure 4: Comparison between predicted and observed one-hour-ahead PV power for the EDLSTM model in São Paulo/SP.

These outcomes validate the EDLSTM as a reliable forecasting model and confirm that its learned representations encode meaningful physical and temporal relationships. Consequently, São Paulo served as the ideal source domain for knowledge transfer to other Brazilian regions with different climatic profiles. The following sections (3.2 and 3.3) explore how the trained model's internal knowledge was adapted and fine-tuned through TL strategies to locations with reduced data availability.

3.2 Performance vs. Data Availability

The relationship between data availability and model performance is a crucial aspect when assessing the applicability of forecasting methods in real-world PV systems. As illustrated in Figure 5, the normalized Root Mean Square Error (pRMSE) demonstrates a clear trend across varying training data durations. The data sets, ranging from 1 to 12 months, are examined for three distinct target locations. As anticipated, the augmentation of the available data resulted in incremental enhancements in the performance of all models. However, the magnitude of improvement and the rate of convergence exhibited significant variation between the baseline and the TL strategies.

In regions with limited historical data, such as Tacaratu/PE, the baseline model demonstrated elevated initial error values (pRMSE ≈ 0.28), signifying challenges in capturing the local generation dynamics with a mere month of training data. In contrast, all TL-based models, particularly TL1, exhibited markedly superior results, achieving reductions in pRMSE even under conditions of extreme data scarcity. This suggests that the reuse of pre-trained weights enables the model to leverage learned representations of the PV generation process, effectively compensating for the lack of local information. Analogous patterns were observed for Itacarambi/MG, where TL1 exhibited consistent superiority over the baseline across all data lengths, underscoring its notable capacity for knowledge generalization.

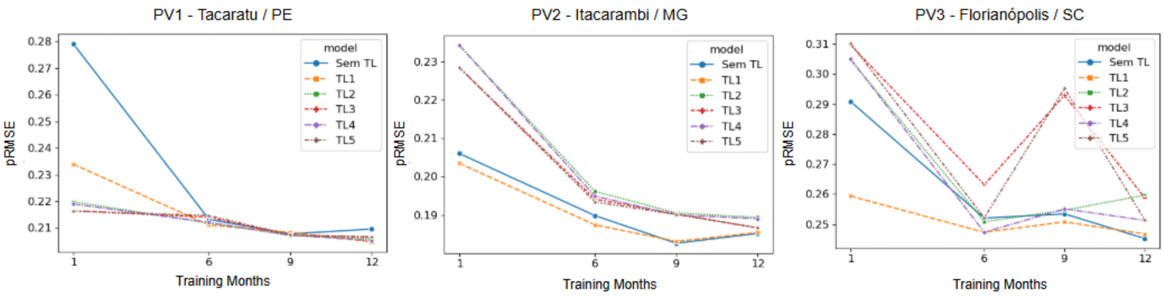

Figure 5: One-hour-ahead forecasting performance (pRMSE) as a function of training data availability for the three target locations (PE, MG, SC).

In the case of Florianópolis/SC, a region characterized by high atmospheric variability, model performance exhibited significant instability. Notwithstanding, TL1 exhibited a competitive pRMSE, indicative of its resilience to climatic variations. The comparative advantage of TL1 across all case studies indicates that full weight reuse without layer freezing allows the model to adapt flexibly to the target site's specific meteorological and operational features while preserving previously learned patterns related to solar power generation dynamics.

From a pragmatic standpoint, these findings have direct consequences for PV forecasting in settings with limited data, such as novel installations, small-scale distributed systems, and remote regions with inadequate monitoring infrastructure. The capacity of TL to generate precise forecasts with limited local data significantly reduces the temporal and financial costs associated with model implementation, thereby enabling the quicker integration of predictive intelligence into PV operations and grid management. Furthermore, these findings underscore the potential of TL-based frameworks to facilitate energy transition strategies in emerging solar markets, where data scarcity persists as a significant impediment to the implementation of advanced forecasting solutions.

3.3 Performance Stability with 12 Months of Data

Beyond predictive accuracy, the stability and reproducibility of forecasting results are critical for evaluating a model's operational robustness. To assess this, we performed 20 independent training repetitions using a full 12-month dataset for each target location. The resulting performance metrics are detailed in the provided Table 2, while the distribution of the pRMSE values from these repetitions is visualized in boxplots (Figure 6).

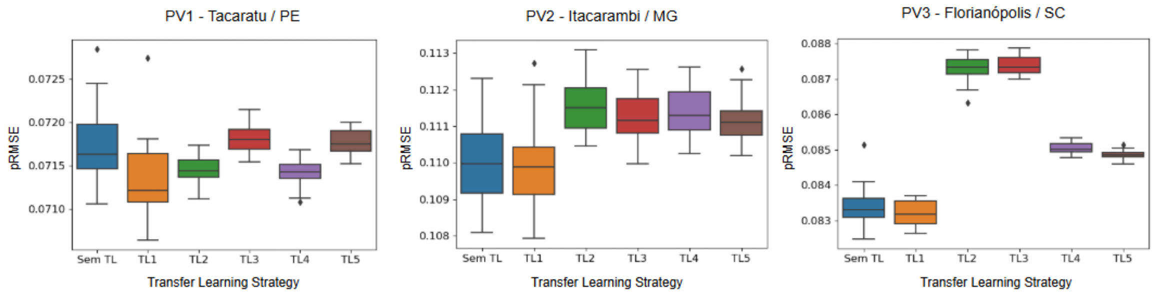

Figure 6: Distribution of pRMSE values across 20 training repetitions using 12 months of data for each target location.

In Tacaratu/PE, the table shows that the TL1 strategy achieved a slightly lower mean pRMSE (20.24%) compared to the baseline model without transfer learning (Sem TL, 20.36%). The boxplot further reveals that both TL1 and TL4 not only yield lower median errors but also exhibit visibly reduced dispersion. This finding indicates enhanced training stability, a crucial attribute for models intended for real-time forecasting pipelines.

For Itacarambi/MG, the performance of TL1 (18.21% pRMSE) was nearly identical to the baseline (18.23%), according to the table. However, the boxplot shows that TL1 consistently delivers this accuracy with lower variance, underscoring its reliability. In contrast, other strategies (TL2-TL5) led to a degradation in performance.

The analysis for Florianópolis/SC, a region with more complex weather dynamics, highlights the trade-off between median accuracy and consistency. While the baseline model showed a marginally lower pRMSE in the table (24.38% vs. 24.42% for TL1), the boxplot clearly illustrates that TL1's performance is significantly more stable, with much lower variability between runs. This consistency is a major advantage for operational reliability.

Table 2: Results of statistical metrics for TL.

PV	Métrica	Sem TL	TL1	TL2	TL3	TL4	TL5
PV1	RMSE (kWh/kWp)	0.071	0.071	0.071	0.072	0.071	0.072
	pRMSE (%)	20.36%	**20.24%**	20.38%	20.50%	20.37%	20.49%
	MBE (kWh/kWp)	-0.003	0.002	-0.003	-0.002	-0.001	0
	MAE (kWh/kWp)	0.051	0.051	0.051	0.051	0.051	0.051
	SS (%)	0.73	**0.732**	0.73	0.728	0.73	0.729
PV2	RMSE	0.108	0.108	0.11	0.11	0.11	0.11
	pRMSE	18.23%	**18.21%**	18.63%	18.55%	18.60%	18.59%
	MBE	0.001	-0.011	-0.001	-0.008	-0.003	-0.01
	MAE	0.074	0.076	0.076	0.076	0.075	0.077
	SS	0.527	**0.527**	0.516	0.518	0.517	0.517
PV3	RMSE	0.082	0.083	0.086	0.087	0.085	0.085
	pRMSE	**24.38%**	24.42%	25.51%	25.71%	25.06%	25.00%
	MBE	0.002	-0.003	-0.005	-0.002	0.001	0.002
	MAE	0.057	0.058	0.061	0.061	0.058	0.058
	SS	**0.663**	0.663	0.648	0.645	0.654	0.655

A comparative analysis across all sites reveals that strategies involving excessive layer freezing, particularly TL2 and TL3, often reduce the model's ability to adapt to local climatic features. This is evident from their higher error metrics in the table and the increased dispersion shown in the boxplots, especially for Florianópolis. Conversely, the TL1 approach, which retrains all layers using pre-trained weights as an initialization, proved most effective. This method preserves generalized knowledge while allowing for fine-tuned local adjustments, leading to more robust and reliable solutions.

From an operational perspective, the reduced variance demonstrated by TL1 translates to higher forecast reliability, an essential characteristic for grid management and energy trading. The stability of the model's output enables better uncertainty quantification and supports the design of control strategies with reduced risk. Consequently, the consistency of the TL1 strategy across distinct climates strengthens its potential as a scalable and dependable solution for a nationwide PV forecasting framework in Brazil.

4. Conclusion

The results obtained in this study confirm the effectiveness of TL as a strategy to enhance PV power forecasting in data-scarce environments. Using the EDLSTM model initially trained in São Paulo as a knowledge source, the analysis demonstrated that pre-trained models can be successfully adapted to new regions with distinct climatic conditions and limited data availability.

Across all target sites, TL-based configurations consistently outperformed the baseline model trained from scratch. The TL1 strategy, which reuses all pre-trained weights without layer freezing, achieved the lowest normalized errors and the highest Skill Scores in every scenario tested. Even with only one month of local data, TL1 maintained accurate forecasts, evidencing its strong capacity to generalize and transfer learned temporal and radiative features. When trained with twelve months of data, TL1 also exhibited lower variance among repetitions, confirming its stability and robustness compared to other TL configurations.

These outcomes demonstrate that transfer learning not only improves forecast accuracy under data scarcity but also enhances model reproducibility and convergence efficiency. By leveraging prior knowledge from a data-rich domain, TL minimizes the dependency on extensive local datasets, allowing reliable forecasting in new or remote PV installations. This approach thus provides a technically feasible and scalable solution for accelerating the deployment of predictive intelligence in distributed solar generation systems.

In summary, the findings reinforce the potential of transfer learning to address one of the main challenges in PV power forecasting: the limited availability of site-specific data. This approach enables knowledge reuse and model adaptation across diverse climatic contexts. This methodology represents a practical advancement toward more accurate, stable, and data-efficient forecasting frameworks for large-scale solar integration in Brazil.

5. Acknowledgments

The authors would like to thank FAPESP for the master's scholarship provided in process 2022/10281-6. This research was also carried out within the scope of the National Institute of Science and Technology for the Development of Photovoltaic Solar Energy Applications (INCT-DAESF), supported by CNPq/MCTI/Brazil (Process 408486/2024-4), FAPESP (Process 2025/26902-8), and CAPES. Furthermore, the authors gratefully acknowledge the financial support provided by the INCT Climate Change Phase 2 (FAPESP 2014/50848-9, CNPq 465501/2014-1, CAPES/FAPS 16/2014). Fernando Martins also thanks FAPESP (grant 2025/15669-0) and CNPq (grant 311620-2023-0) for their financial support. Finally, the authors would like to thank the Federal University of São Paulo and the University of São Paulo for their infrastructure and institutional support.

6. References

ABSOLAR (2024). Dados do Setor de Geração Distribuída Solar Fotovoltaica. Associação Brasileira de Energia Solar Fotovoltaica. Disponível em: https://www.absolar.org.br/

Cho, K., van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259.

Depoortere, J., Driesen, J., Suykens, J., & Kazmi, H. S. (2024). SolNet: Open-source Deep Learning Models for Photovoltaic Power Forecasting Across the Globe. arXiv preprint arXiv:2405.14472. Disponível em: https://arxiv.org/abs/2405.14472

EPE (2024). Relatório Síntese do Balanço Energético Nacional 2024. Empresa de Pesquisa Energética. Disponível em: https://www.epe.gov.br/pt/imprensa/noticias/epe-publica-o-relatorio-sintese-do-balanco-

energetico-nacional-2024
Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8):1735–1780.
IEA (2024). Electricity Market Report 2024. International Energy Agency. Disponível em: https://www.iea.org/reports/electricity-market-report-2024
IRENA (2023). World Energy Transitions Outlook 2023: 1.5°C Pathway. International Renewable Energy Agency. Disponível em: https://www.irena.org/publications/2023/Jun/World-Energy-Transitions-Outlook-2023
REN21 (2024). Renewables 2024 Global Status Report. REN21 Secretariat. Disponível em: https://www.ren21.net/reports/global-status-report/
Sarmas, M., Chatzis, S. P., & Kleftakis, V. (2022). Transfer Learning for Solar Power Forecasting Under Data Scarcity Using Stacked LSTM Models. IEEE Transactions on Sustainable Energy, 13(4), 2102–2112. DOI: 10.1109/TSTE.2021.3135785
Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. In Advances in Neural Information Processing Systems (NIPS 2014).