

Exploring Transfer Learning Techniques for Solar Irradiation Forecast across Geographically Diverse Locations in Brazil Using Reanalysis Data

Rodrigo Huber Marques Moreira Mendes

Department of Industrial Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
Marquês de São Vicente Street 225, Gávea, 22451-041 Rio de Janeiro, RJ, Brazil
rodrigo@aluno.puc-rio.br

Fernanda Araujo Baião

Department of Industrial Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
Marquês de São Vicente Street 225, Gávea, 22451-041 Rio de Janeiro, RJ, Brazil
fbaiao@puc-rio.br

Reinaldo Castro Souza

Department of Industrial Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
Marquês de São Vicente Street 225, Gávea, 22451-041 Rio de Janeiro, RJ, Brazil
reinaldo@puc-rio.br

ABSTRACT

This research explores the effectiveness of Transfer Learning (TL) for day-ahead solar forecasting across diverse locations in Brazil. Hourly data from 2017 to 2020 of air temperature, ground-level and top-of-the-atmosphere solar irradiation, and cloud coverage were collected from the MERRA-2 Reanalysis database for ten sites within Rio de Janeiro and ten others distributed across the country. A Random Forest Regressor (RFR) was trained using data from Galeão Airport (GIG) to forecast solar irradiation a day ahead, with evaluation over one year. This pre-trained model was then applied via TL, without fine-tuning, to all other sites. Results show that the model's forecast accuracy performs very well, but declines with increasing divergence between source and target, emphasizing the importance of domain-specific modeling in solar forecasting. The findings highlight the need for robust models that consider distance thresholds, climatic variability, and geographical features to ensure reliable TL performance across different regions.

KEYWORDS. Transfer Learning. Solar Irradiation Forecast. Renewable Energy. Machine Learning.

Paper topics (EN&OG – OR in Energy, Oil&Gas; EST&AM – OR Analytics in Statistics and Machine Learning; AS&DS – Data Analysis and Science)

1. Introduction

Recently, there has been a great body of research on Renewable Energies towards reducing carbon emissions and contributing to a more conscious energy consumption [EPE, 2023]. Based on the International Energy Agency, solar photovoltaic energy should exceed the coal capacity by the year 2027 [IEA, 2023].

In Brazil, the National Interconnected System (SIN) is responsible for the production and transmission of electricity throughout the country [EPE, 2023]. It is managed by the National Electricity System Operator (ONS). It comprises various power plants (such as hydroelectric, thermoelectric, wind, and solar) interconnected by a high-voltage transmission network [EPE, 2023]. However, acting as a real-time energy transmission coordinator in a country of such dimensions as Brazil is an extremely difficult task. To optimize the system, SIN forecasts the energy that should be produced the following day to obtain some guidance for dispatch coordination, allowing the flow of energy through the network to be balanced more efficiently.

To better coordinate the energy generation and the transmission lines in a given country, it is well stated that one should look to forecast with high confidence levels the output for each generation plant to correctly dispatch energy through the power transmission lines. This is a matter of safety and risk mitigation to avoid power outages or damage to the power lines.

Given the intrinsic fact that solar energy is intermittent (based heavily on the local climate and meteorology), it is extremely important to have a model in which its intermittencies are accounted for to make accurate forecasts with high confidence levels. These forecasts support energy agencies with day-ahead planning and reliable dispatch scheduling. Usually, forecasting weather data throughout the day is used to support the stability of the electricity grid, while forecasting data for a day or more ahead is used to support market participation [IEA, 2013]. Accurate short-term and day-ahead forecasting of solar irradiation is essential for the reliable integration of solar energy into power systems, supporting grid stability and operational efficiency.

Although locally trained models are capable of capturing site-specific meteorological dynamics, developing and maintaining a distinct model for every location is both data-intensive and computationally costly. In addition, many regions still lack historical data of sufficient quality and quantity to train robust forecasting models. In this context, TL [Wang and Chen, 2023] emerges as a promising alternative, enabling the reuse of models trained in data-rich environments for application in data-scarce regions.

This research explores the potential of TL, in which a model trained on data from one location is transferred to make predictions in other places with similar characteristics, to find the limitations and opportunities that TL presents to achieve the best results to help the country's energy management concerning solar energy forecasts.

Transfer Learning offers several advantages that make it particularly suitable for meteorological forecasting tasks. First, it addresses the challenge of limited data availability, a common constraint in many regions, by allowing knowledge learned from one location to be applied to another with minimal additional data. Second, it significantly reduces the computational cost associated with training models from scratch for each new site, thus improving scalability. Third, it enables faster deployment of forecasting models across broad geographic areas, accelerating the operational use of solar predictions. Finally, well-trained source models can capture generalized physical and atmospheric patterns that can be effectively transferred to target locations, especially when there are similarities in climatic conditions or seasonal behavior.

Despite its advantages, the effectiveness of TL in solar forecasting across geographically and climatically diverse regions remains underexplored. Investigating its limitations and potential, particularly with regard to spatial variability and domain specificity, is crucial to building robust and

scalable forecasting systems.

This research aims to contribute to the forecasting of photovoltaic energy generation by developing models to forecast solar irradiation with great reliability. Specifically, we focus on day-ahead forecasts using machine learning, comparing traditional location-specific training with a TL strategy across diverse geographic regions in Brazil. The transferability of a model trained at GIG in Rio de Janeiro to several other locations is assessed, with no fine-tuning, meaning: considering a model with no retraining and without using any data on the target locations, in contrast to the locally trained models, which are data-intensive and with high computational cost [Wang and Chen, 2023]. We analyze how predictive accuracy changes based on the distance between the source and target sites, contributing insights into when and how TL is viable for solar irradiation forecasting. In particular, we evaluate the performance of the RFR [Breiman, 1996, 2001] in a TL setting, seeking to determine under which spatial or statistical similarities the transfer is effective. By identifying conditions that support accurate model transfer, this study seeks to assist the SIN in optimizing solar energy planning in a scalable and data-efficient manner.

The text is structured as follows: Section 2 presents the data used in this study. Section 3 describes the methodology employed. Section 4 evaluates the results of TL, both within the state of Rio de Janeiro and across Brazil, and investigates which similarity/distance metrics for time-series best correlate with the obtained transfer losses. Section 5 provides a critical discussion of the findings. Finally, Section 6 outlines the main conclusions and suggests directions for future research based on the results obtained in this work.

2. Data Collection

The data set used was composed of data acquired from the MERRA-2 [Rienecker et al., 2011] database through the Renewables Ninja API interface [Pfenninger and Staffell, 2024].

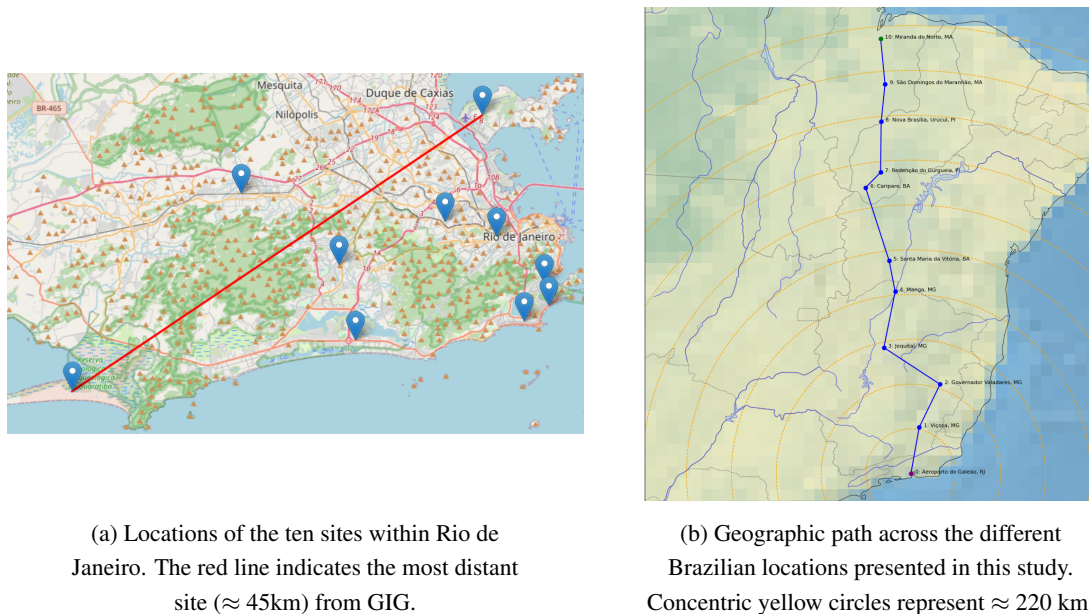


Figure 1: Geophysical maps of Rio de Janeiro and Brazil with all the sites presented in this research

Four years of data (2017-2020) were acquired with an hourly granularity, and the features used were: Air temperature ($^{\circ}\text{C}$), Ground-level solar irradiation (W/m^2), Cloud coverage (%), and Top of the atmosphere solar irradiation (W/m^2).

The features were obtained for twenty locations, comprising ten randomly chosen sites within Rio de Janeiro and ten additional, geographically distant locations distributed across south-eastern and northeastern Brazil. The intra-urban locations in Rio de Janeiro include: Bangú, Barra da Tijuca, Botafogo, Copacabana, GIG, Ipanema, Jacarepaguá, Maracanã, Marambaia, and Méier, as shown in Figure 1a.

The ten most distant sites were selected to approximate a north-northwestward trajectory from GIG to Miranda do Norte, with intervals of approximately 220 km; for each step, a close city to the ideal straight-line path was chosen, resulting in the following sequence: Viçosa, Governador Valadares, Jequitaiá, Manga, Santa Maria da Vitória, Cariparé, Redenção do Gurguéia, Nova Brasília, São Domingos, and Miranda do Norte, as shown in Figure 1b.

This spatial distribution was designed to assess the impact of increasing geographical and climatic divergence on the effectiveness of TL for solar irradiation forecasting.

GIG was chosen as the source for TL due to its role as Rio de Janeiro's main port of entry and the availability of consistent, high-quality meteorological data, making it a practical reference point for evaluating model generalization.

A 4-year period with a total of 2,805,120 data points were collected across 20 locations. For each location, hourly data were obtained for four meteorological variables (air temperature, ground-level solar irradiation, top-of-the-atmosphere solar irradiation, and cloud coverage), resulting in 140,256 values per location. This comprehensive dataset provides a robust foundation for training and evaluating machine learning models and TL for day-ahead solar irradiation forecasting.

The features were pre-processed to identify and handle outliers using the Interquartile Range (IQR) method. For each feature, values outside the range defined by 1.5 times the IQR below the first quartile (Q1) or above the third quartile (Q3) were considered outliers. Depending on their frequency and impact, these anomalous values were either removed or replaced to ensure data quality, reduce noise, and improve the robustness of the model training process.

Following outlier treatment, daily mean solar irradiation values were computed for each location, considering only the daylight period. Additionally, daily mean values were calculated for all other features. This aggregation aligns with the research objective, which focuses on day-ahead forecasting of solar irradiation, and ensures that the input data accurately represents daily meteorological conditions relevant to solar energy modeling.

3. Modeling

Each of the 20 selected locations was independently modeled to assess local modeling performance using an RFR model for day-ahead solar irradiation forecasting.

Each model was trained using data from 2017 to 2019 and validated on the subsequent 365 days of 2020. This structure ensured temporal separation between training and testing phases, mimicking real-world forecasting conditions. Model performance was evaluated using the Mean Absolute Percentage Error (MAPE), defined as:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (1)$$

Equation 1: Where y_t and $\hat{y}_t$ represent the observed and predicted daily solar irradiation values on day t , respectively, and $n = 365$ corresponds to the number of days in the validation period.

During the model development phase, a walk-forward validation strategy was employed to assess model stability and guide hyperparameter tuning. This approach was implemented using the *TimeSeriesSplit* method [Benson et al., 2020; Hasanov et al., 2022], which preserves the temporal ordering of the data, a critical requirement in time series forecasting tasks. Unlike traditional cross-validation, *TimeSeriesSplit* incrementally expands the training window while validating on the immediately following time segment, thereby avoiding data leakage and more realistically simulating operational conditions.

TimeSeriesSplit (with three splits) was used; the data were divided into four contiguous folds, resulting in three sequential training/testing pairs. The first split trains on the initial fold (2017) and tests on the second (2018). The second split trains on the first two folds (2017–2018) and tests on the third (2019). This progressive expansion of the training window mirrors the real-world scenario of model retraining as new data becomes available.

While this walk-forward strategy was used to support model development and selection, the final evaluation used a hold-out approach. Each model was trained on the complete dataset from 2017 to 2019 and validated on the subsequent 365 days of 2020. This setup ensures a temporally consistent and unbiased estimate of forecasting performance.

Every time the RFR was trained (including every *TimeSeriesSplit*), it was necessary to adjust its hyperparameters, which were first manually adjusted based on the literature, following a Bayesian search with an exhaustive search for a maximum of 100 iterations to select the best hyperparameters.

Transfer Learning models can follow two main strategies: (i) pretraining followed by fine-tuning, where the model is exposed to some data from the target location to adapt its performance, or (ii) direct transfer without fine-tuning, where the pretrained model is applied to new locations without any additional training or data. This research adopts the latter approach (TL without fine-tuning) to assess the generalization capacity of the TL model under data-scarce conditions.

To explore the viability of TL for solar forecasting across diverse geographical and climatic conditions, a RFR model was trained on data from GIG, located in Rio de Janeiro. This source RFR model was then saved and transferred *as-is* without any fine-tuning, retraining, or exposure to data from the target locations. The TL model's performance was evaluated solely on the holdout validation set (year 2020) from their respective target locations.

In this context, TL used no data from the target locations for learning purposes, unlike locally trained models, which rely on location-specific data and demand significant computational resources [Wang and Chen, 2023]. This strict approach ensures no data leakage and allows for a fair assessment of the model's generalization capacity.

4. Transfer Learning Evaluation Results

The TL model's performance was evaluated using the same validation window used for local models (year 2020). The impact of TL was quantified using a metric referred to as *Transfer Loss*, which captures the relative degradation in forecasting accuracy when using TL model instead of a locally trained one.

$$\text{Transfer Loss} = \text{MAPE}_{\text{TL}} - \text{MAPE}_{\text{local}} \quad (2)$$

Equation 2: Where MAPE_{TL} is the MAPE of the TL model, and $\text{MAPE}_{\text{local}}$ is the MAPE of the locally trained model.

A positive transfer loss indicates worse performance of the transferred model compared to

the local model. This evaluation aims to measure how forecast accuracy deteriorates or is enhanced with the use of the TL technique, and as a way to investigate the relationship of the increasing geographical distance from the source to the target model and its performance.

4.1. Assessing Transfer Learning of a Solar Forecasting Model within Rio de Janeiro

The performance of the locally trained RFR models, measured by MAPE, varied across the locations within Rio de Janeiro. The highest local MAPE was observed at Barra, with a value of 13.41%, followed by Copacabana 13.03% and Maracanã 12.88%. These locations show relatively higher forecast errors, possibly due to their specific local weather patterns. On the other hand, locations like Bangu 12.29%, Meier 12.66%, and Marambaia 12.19% exhibited the lowest MAPE values, indicating better forecast accuracy in these regions. The remaining locations, such as Jacarepagua 13.19% and Ipanema 13.19%, showed intermediate performance. These results emphasise there is variability in forecasting accuracy across Rio de Janeiro, indicating that the values fall within a relatively narrow delta range of variability, suggesting the potential benefit of exploring TL as an alternative to training models locally in each area.

A precise analysis of the results presented in Figure 2 and Table 1 shows that the transfer loss tends not to increase with geographical distance for the sites within Rio de Janeiro, indicating the effectiveness of TL for those sites, since the transfer loss is so minute, with the greater transfer loss being 0.24%.

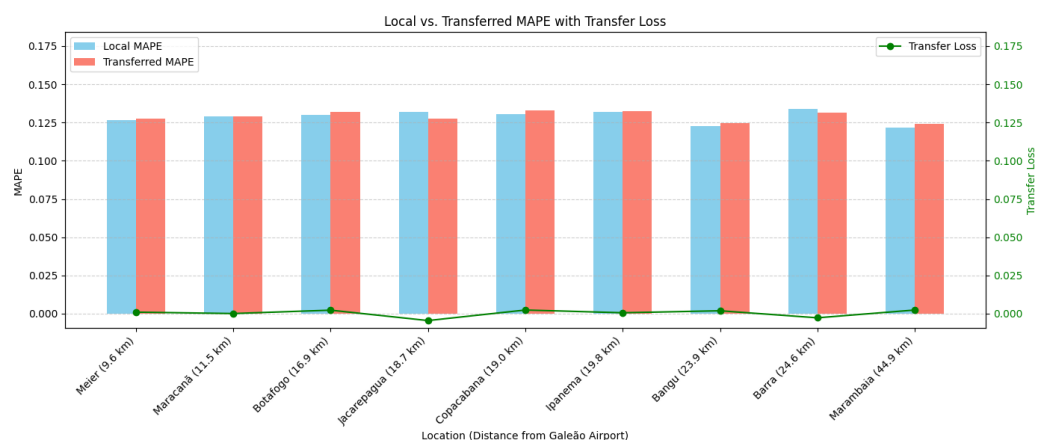


Figure 2: Comparison between local and TL RFR models across ten locations. The blue bars indicate the MAPE for locally trained models, while the red bars show the MAPE of the TL model trained on GIG data.

The green line represents the transfer loss, defined as the difference between TL and local MAPE. The x-axis lists the target locations and their distance (kilometers) from the source model GIG.

Based on the observed results, we can define a spatial threshold for effective model transferability. Locations within a radius of approximately 45 km from the source model GIG exhibited low transfer loss, with all values remaining below 0.24% increase in MAPE. This suggests that TL within this spatial range is very well acceptable for the RFR model trained on GIG data, with only a minor degradation in forecasting performance. Thus, a 45 km radius could be a practical guideline for geographic transferability, assuming an acceptable forecast degradation of 0.24% increase in MAPE for closer locations, we observe that all the locations inside Rio de Janeiro presented in this research could be suited to be used in a TL scenario.

4.2. Assessing Transfer Learning of a Solar Forecasting Model across Brazil

The performance of the locally trained RFR models, as measured by the MAPE, varied across the ten selected locations. The highest local MAPE was observed in Miranda do Norte,

Location	Transfer Loss %
Meier (9.6 km)	0.10%
Maracanã (11.5 km)	0.01%
Botafogo (16.9 km)	0.23%
Jacarepagua (18.7 km)	-0.45%
Copacabana (19.0 km)	0.24%
Ipanema (19.8 km)	0.06%
Bangu (23.9 km)	0.19%
Barra (24.6 km)	-0.27%
Marambaia (44.9 km)	0.24%

Table 1: Transfer loss between the locally trained and TL RFR model from GIG across locations within Rio de Janeiro

with a value of 15.71%, followed by São Domingos 15.04% and Nova Brasília 13.91%. These higher errors may reflect increased variability in local atmospheric conditions or data limitations. On the other hand, Santa Maria 9.42%, Manga 8.80%, and Cariparé 10.70% exhibited the lowest MAPE values, indicating relatively better forecast accuracy in these regions. The remaining locations showed intermediate performance: Redenção 10.98%, Jequitai 10.91%, Gov. Valadares 10.84%, and Viçosa 11.25%. These results emphasize the spatial heterogeneity in forecasting accuracy and support the need to evaluate the potential of transfer learning relative to strong local baselines.

The graphical interpretation of the results shown in Table 2 and Figure 3 clearly illustrates that transfer loss tends to increase with geographical distance, highlighting the diminishing effectiveness of transfer learning as spatial divergence grows.

Based on the observed results, defining a spatial threshold for effective model transferability is possible. Locations situated within a radius of approximately 630 kilometers from the source (GIG) exhibited low transfer loss, below 1.02% increase in MAPE. This suggests that, for the RFR model trained on GIG data, transfers within this spatial range can be considered acceptable to use TL, if a minor degradation in forecasting performance is tolerated. Thus, a 630 km radius may be a practical guideline for geographic TL, assuming a predefined threshold for acceptable forecast degradation (e.g., < 1.02% absolute increase in MAPE for distant places).

TL models performed comparably well in locations closer to GIG, as shown in Table 2, where transfer loss and geographical distances can be compared:

Location	Transfer Loss %
Viçosa (230.5 km)	-0.31%
Gov. Valadares (458.4 km)	-0.41%
Jequitai (630.2 km)	1.02%
Manga (894.7 km)	2.18%
Santa Maria (1047.5 km)	1.23%
Cariparé (1416.8 km)	2.16%
Redenção (1482.0 km)	0.32%
Nova Brasília (1728.6 km)	2.39%
São Domingos (1910.2 km)	3.79%
Miranda do Norte (2134.2 km)	3.67%

Table 2: Transfer loss and distance from GIG for locations across Brazil

These values represent the performance discrepancy between the locally trained models

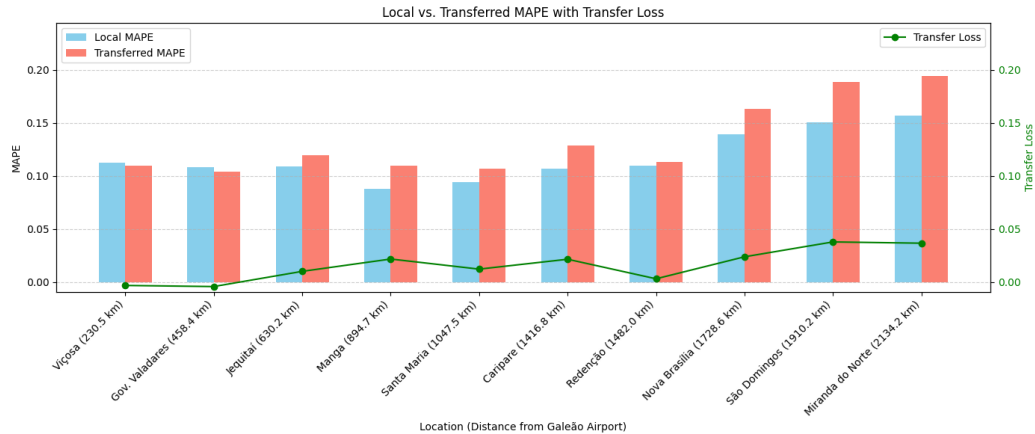


Figure 3: Comparison between local and TL RFR models across ten locations. The blue bars indicate the MAPE for locally trained models, while the red bars show the MAPE of the TL model trained on GIG data.

The green line represents the transfer loss, defined as the difference between the transferred and local MAPE. The x-axis lists the target locations and their distance (kilometers) from the source model (GIG).

and those transferred from GIG in Rio de Janeiro. Negative transfer loss values indicate an improvement in forecasting accuracy when applying the pre-trained model, while positive values suggest a deterioration in forecast performance. The results highlight the inverse relationship between transfer performance and geographical distance from the source location, with accuracy generally decreasing as the distance increases. This underscores the importance of domain-specific modelling in solar irradiation forecasting, especially for locations distant from the training dataset.

4.3. Assessing distinct Distance Metrics for time-series Transfer Learning

Four statistical similarity/divergence metrics for time-series were used to evaluate possible correlations with transfer loss, namely Catch22 Distance [Lubba et al., 2019; Ehrig et al., 2024], Cosine Similarity, Euclidean Distance, Dynamic Time Warping Distance (DTW), and a Geographical Distance metric (measured in kilometers). The similarity/divergence metric that shows the strongest and most consistent correlation with transfer loss should be adopted as the basis for defining the boundaries of effective transfer learning. This ensures that a meaningful measure of compatibility between source and target locations guides model transfer.

For sites within Rio de Janeiro, the Critical Pearson r value is 0.765 for a confidence level of 99% and $n = 10$. This means that any Pearson correlation coefficient greater than or equal to 0.765 (in absolute value) is considered statistically significant at a significance level (α) of 0.010. In the context of the correlation tests, correlations with an absolute value greater than or equal to 0.765 indicate significant relationships, while those below this threshold are not considered statistically significant, as shown in Figure 4.

For the sites across Brazil, the Critical Pearson r is 0.612 for a confidence level of 95.45% and $n = 11$. This means that any Pearson correlation coefficient greater than or equal to 0.612 (in absolute value) is considered statistically significant at an alpha level of 0.045. In the context of your heatmap, correlations with an absolute value greater than or equal to 0.612 indicate significant relationships, while those below this threshold are not considered statistically significant, as shown in Figure 4.

Among the variable pairs analyzed within Rio de Janeiro, three show statistically significant correlations at a confidence level of 99% ($\alpha = 0.010$, $|r| \geq 0.765$). No significant correlations were found between Transfer Loss % and any of the tested metrics, indicating that, under this

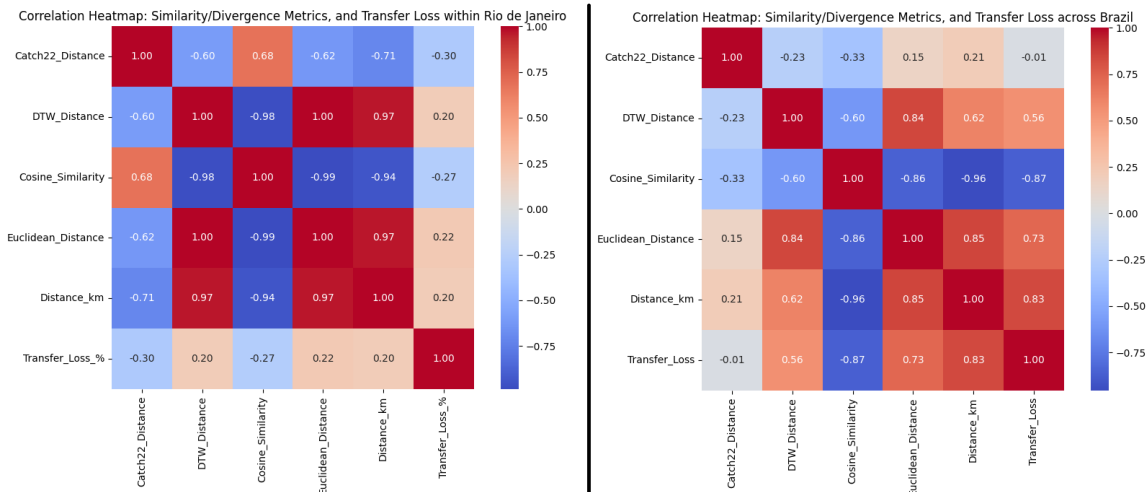


Figure 4: Heatmap showing Pearson correlation coefficients between Transfer Loss, Geographical Distance (measured in kilometers), and multiple statistical similarity measures across locations (Catch22 Distance, Cosine Similarity, Euclidean Distance, and DTW). Darker shades indicate stronger correlations (positive or negative). Significant correlations (based on critical r values and p -values) highlight the most representative metrics for capturing relationships among locations in the context of model transferability.

stricter threshold, neither feature-based nor geographical similarity metrics show a strong association with model performance degradation.

However, `Distance_km` presents very strong and significant correlations with several similarity measures. Specifically, it is strongly and positively correlated with both `DTW Distance` ($r = 0.969$) and `Euclidean Distance` ($r = 0.968$), implying that greater physical separation corresponds to greater divergence in time-series patterns. Moreover, a strong negative correlation with `Cosine Similarity` ($r = -0.942$) reinforces that spatially distant locations tend to be less similar in feature space. This suggests that while transfer loss may remain relatively stable within the region, spatial distance strongly influences structural divergence among locations.

Among the variable pairs analyzed across Brazil, five show statistically significant correlations at a confidence level of 95.45% ($\alpha = 0.045$, $|r| \geq 0.612$). A strong negative correlation was found between `Transfer Loss %` and `Cosine Similarity` ($r = -0.869$), suggesting that greater similarity in feature space is associated with lower transfer loss. Additionally, `Transfer Loss %` showed strong positive correlations with `Euclidean Distance` ($r = 0.726$) and `Distance_km` ($r = 0.835$), indicating that larger geographic or feature-space distances are associated with higher transfer losses. Regarding spatial relationships, `Distance_km` is strongly and positively correlated with both `DTW Distance` ($r = 0.617$) and `Euclidean Distance` ($r = 0.851$), and strongly negatively correlated with `Cosine Similarity` ($r = -0.956$). This pattern confirms that more distant locations diverge in statistical and geometric terms. As shown, a higher `Cosine Similarity` is strongly associated with lower transfer loss, and as geographical distance increases, `Cosine Similarity` decreases sharply (distant locations are less similar in feature space).

5. Discussions

The Maximum Delta Difference (greatest absolute difference between the highest and lowest transfer losses) observed in the dataset across the nine locations within the city of Rio de Janeiro, excluding GIG, was 0.0069. This relatively low variation suggests a more homogeneous transfer performance when the model is applied to nearby, geographically contiguous areas.

Conversely, the Maximum Delta Difference for locations spread across Brazil is markedly larger, amounting to 0.0420. This substantial difference underscores the heightened variability in transfer loss as the geographical span increases, reflecting the greater divergence in model performance when applied to regions located at considerable distances from the source location. The pronounced disparity in transfer losses across such extensive distances emphasizes the challenges associated with applying a single model to diverse and spatially heterogeneous locations, thereby highlighting the importance of considering region-specific factors in forecasting accuracy.

Transfer loss increased with both geographic and climatic distance, highlighting the limitations of applying static models across diverse environments.

These results align with previous findings by Pan and Yang [2010]; Weiss et al. [2016], emphasizing that TL is most effective when the source and target domains are similar. Poor performance in distant locations underscores the role of environmental divergence. Methods such as fine-tuning and domain adaptation as suggested by Blitzer et al. [2007]; Arase and Tsujii [2019] need to be taken into account to develop a more broader TL considered to develop analysis of MAPE for both local and TL models, as well as the calculated transfer loss, provides important insights into the implications of geographical distance on model accuracy.

For locations within Rio de Janeiro, the results indicate that local models tend to perform well in regions geographically close to the source data. The transfer loss for these locations were low, reflecting a minor degradation in performance when applying TL to these areas. These findings are consistent with the observation that local models can more effectively capture region-specific features, which leads to more accurate predictions [Pan and Yang, 2010]. In these areas, the TL model was able to generalize very well, with a lower transfer loss compared to more distant locations. This supports the findings of Weiss et al. [2016], who emphasized that TL can be effective when the source and target domains are similar.

In contrast, for locations farther from the source, TL models struggled to conform to these regions' unique environmental characteristics and data patterns. This finding aligns with the work of Yosinski et al. [2014], who demonstrated that the performance of TL diminishes when the source and target domains diverge significantly. Furthermore, the transfer loss for these locations was significantly higher, reflecting a marked degradation in model performance when applying TL from the source data to geographically distant areas. This discrepancy highlights the importance of location-specific features, such as climate, land cover, and other environmental factors, which are likely to vary significantly between Rio de Janeiro and other regions in Brazil. Pan and Yang [2010] and Pereira and Torezzan [2024] have similarly noted that models trained in one context (e.g., geographical) may fail to generalize well to significantly different contexts due to variations in domain-specific characteristics.

The results of this study underscore the importance of geographical proximity in determining the performance of TL models. For locations closer to Rio de Janeiro, TL performed very well, exhibiting relatively small transfer losses. This suggests that the model could still provide reasonable predictions even for other locations within the same geographical constraints. However, for more distant locations, the performance of TL significantly declined, with larger transfer losses observed as distances increase. This suggests that geographical factors, such as distance, regional climate differences, land use, and other local characteristics, substantially impact the model's ability to generalize effectively, as pointed out by Weiss et al. [2016].

The results observed within the state of Rio de Janeiro demonstrate that the sites are geographically and statistically close to one another, which is reflected in the lack of significant correlation between transfer loss and any of the distance or similarity metrics tested. This spatial compactness likely reduces variability in meteorological patterns, making model transfer more sta-

ble regardless of the metric used. In contrast, when evaluating sites across Brazil, covering diverse climates and regions, distance and similarity metrics such as Cosine Similarity, Euclidean Distance, and geographic Distance in km show strong correlations with transfer loss. This underscores their importance in representing meaningful divergences or affinities between locations. These findings highlight the relevance of choosing appropriate similarity measures when scaling the transfer learning approach to broader or more heterogeneous regions for this study and future applications or methodological improvements.

6. Conclusion

Our findings indicate that TL models trained on data from one geographical region (e.g., Rio de Janeiro) show a diminishing capacity to generalize as the distance to the training site increases. While the model performs very well in regions that are geographically similar to the training data, its performance deteriorates significantly when applied to more distant locations. Therefore, it is crucial to consider the geographical context and specific environmental features when transferring models across different locations, as TL is most effective when the training and application sites share similar characteristics. This underscores the importance of domain adaptation in TL, as suggested by Blitzer et al. [2007], where domain-specific differences must be addressed for successful model transfer and underscores, as suggested by Arase and Tsujii [2019], the importance of pre-training a model to acquire general knowledge, while fine-tuning it to adapt to the specific characteristics of a target domain, ensuring better performance despite domain-specific variations.

While TL without fine-tuning performs very well across nearby locations with similar conditions, performance tends to degrade across broader geographic or climatic boundaries. This highlights the importance of domain-specific features and motivates the use of more advanced TL strategies, such as pretraining and fine-tuning, to ensure a reliable and adaptable solar forecasting across diverse and distant regions.

Building on these findings, future work includes the development of a more robust and reliable TL model capable of generalizing across wider geographic and climatic boundaries and not confined to the limitations encountered in this study. We are also investigating domain-specific similarities/differences metrics to better correlate different sites.

7. Limitations and Future Perspectives

This study assessed TL without fine-tuning to simulate data-scarce conditions and isolate the model's intrinsic generalization ability. While fine-tuning may improve performance by adapting to local features, its exclusion allowed for a clear baseline evaluation. Future work should incorporate domain adaptation and fine-tuning to enhance scalability and robustness.

Daily mean values were used to align with one of the needs of the SIN: day-ahead forecasting, but this aggregation may overlook seasonal influences on model accuracy. The lack of seasonal segmentation may limit understanding of when TL is more effective. Future studies should analyze seasonal patterns to refine model deployment and improve forecasting reliability under varying climatic-seasonal conditions.

Acknowledgements

The authors gratefully acknowledge the CAPES, CNPq, and FAPERJ research agencies for their financial support.

References

Arase, Y. and Tsujii, J. (2019). Transfer fine-tuning: A BERT case study. In Inui, K., Jiang, J., Ng, V., and Wan, X., editors, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, p. 5393–5404, Hong Kong, China. Association for Computational Linguistics. URL <https://aclanthology.org/D19-1542/>.

- Benson, B., Pan, W., Prasad, A., Gary, G., and Hu, Q. (2020). Forecasting solar cycle 25 using deep neural networks. *Solar Physics*, 295(5):65.
- Blitzer, J., Dredze, M., and Pereira, F. (2007). Biographies, bollywood, boom-boxes, and blenders: Domain adaptation for sentiment classification. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, p. 440–447. URL <https://www.aclweb.org/anthology/P07-1045/>.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24:123–140.
- Breiman, L. (2001). Random forests. *Machine learning*, 45:5–32.
- Ehrig, C., Sonnleitner, B., Neumann, U., Cleophas, C., and Forestier, G. (2024). The impact of data set similarity and diversity on transfer learning success in time series forecasting. *arXiv preprint arXiv:2404.06198*.
- EPE (2023). Matriz energética e elétrica. URL <https://www.epe.gov.br/pt/abcdenergia/matriz-energetica-e-eletrica>.
- Hasanov, M., Wolter, M., and Glende, E. (2022). Time series data splitting for short-term load forecasting. In *PESS+ PELSS 2022; Power and Energy Student Summit*, p. 1–6. VDE.
- IEA (2013). Photovoltaic and solar forecasting: State of the art. URL <https://iea-pvps.org/>.
- IEA (2023). Solar pv still dominates renewable energy capacity additions. Technical report, International Energy Agency.
- Lubba, C. H., Sethi, S. S., Knaute, P., Schultz, S. R., Fulcher, B. D., and Jones, N. S. (2019). catch22: Canonical time-series characteristics: Selected through highly comparative time-series analysis. *Data mining and knowledge discovery*, 33(6):1821–1852.
- Pan, S. J. and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359.
- Pereira, J. P. d. L. and Torezzan, C. (2024). Transferência de aprendizado em modelos de predição de turnover voluntário. In *Anais do LVI Simpósio Brasileiro de Pesquisa Operacional*, Fortaleza. Galoá.
- Pfenninger, S. and Staffell, I. (2024). Renewables ninja. URL <https://www.renewables.ninja/about>.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., et al. (2011). Merra: Nasa’s modern-era retrospective analysis for research and applications. *Journal of climate*, 24(14):3624–3648.
- Wang, J. and Chen, Y. (2023). *Introduction to transfer learning: algorithms and practice*. Springer Nature.
- Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1):9.
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems*, p. 3320–3328. URL <https://arxiv.org/abs/1411.1792>.