Subseasonal Precipitation Forecasting in Southeastern Brazil Using Machine Learning with Explainable Artificial Intelligence

João Vyctor Ferreira da Costa[1], **Matheus Corrêa Domingos**[1], **Valdivino Alexandre de Santiago Júnior**[1], **Jorge Luís Gomes**[1]

[1] Laboratório de Inteligência ARtificial para Aplicações AeroEspaciais e Ambientais (LIAREA)
Programa de Pós-Graduação em Computação Aplicada (PGCAP)
Instituto Nacional de Pesquisas Espaciais (INPE)
São José dos Campos, SP - Brazil

joao.ferreira@inpe.br, matheus.domingos@inpe.br,

valdivino.santiago@inpe.br, jorge.gomes@inpe.br

Abstract. *Subseasonal precipitation forecasting is essential for mitigating climate-related risks, but prediction remains challenging for traditional numerical weather prediction models. This study evaluates machine learning (ML) models—MLP, LSTM, GRU, and CNN-1D—for subseasonal precipitation forecasting in Southeastern Brazil. ERA5 reanalysis data (1991–2018) were used as the observational reference, with 1991–2016 for training and 2017–2018 for testing. The ML models were compared to the Brazilian Global Atmospheric Model (BAM), and the results show that all ML models outperform BAM, with MLP achieving the best overall performance. SHAP values were used to explain the MLP predictions, highlighting the importance of the input variables.*

1. Introduction

Rainfall is one of the most important climatic variables, directly influencing sectors such as agriculture, water resource management, transportation, tourism, urban planning, and disaster risk management [Sales et al. 2018]. In Brazil, precipitation exhibits strong spatial and seasonal variability, leading to both droughts and intense rainfall events with significant socio-economic impacts. Therefore, improving precipitation forecasting capabilities is essential for reducing risks and supporting decision-making in multiple sectors [Yan et al. 2025].

Subseasonal forecasting bridges the gap between short-range numerical weather prediction and seasonal climate forecasts. In operational subseasonal prediction systems, forecasts typically extend from about two weeks to three months ahead, which makes accurate predictions at this timescale challenging [NOOA 2017].

At the subseasonal time scale, forecasts are typically produced using numerical weather prediction (NWP) models. These models are based on physical laws that describe atmospheric dynamics and have been applied for decades to simulate weather and climate processes [NOOA 2017]. However, they still face significant limitations, particularly in representing processes related to cloud formation and precipitation distribution.

These processes involve complex interactions among multiple atmospheric variables and microphysical cloud mechanisms that occur at spatial scales smaller than model resolution [Morrison et al. 2020]. Even with recent advances, accurately predicting rainfall remains challenging because of this complexity and the high computational cost associated with numerical simulations [Lam et al. 2023].

In the context of seasonal climate forecasting, the work by [Monego et al. 2022] stands out, producing precipitation forecasts (mm/day) for each season in South America during 2018–2019 using Global Precipitation Climatology Project (GPCP) data interpolated with National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR), with results compared to the Brazilian Global Atmospheric Model (BAM) [Figueroa et al. 2016], employing algorithms such as Gradient Boosting and Multilayer Perceptron (MLP) [Rumelhart et al. 1986]. Similarly,[Domingos et al. 2025] forecasts precipitation for autumn 2023 using ERA5 reanalysis data [Hersbach et al. 2020], employing architectures like a one-dimensional Convolutional Neural Network (CNN-1D) [LeCun et al. 1998], Long Short-Term Memory (LSTM) [Hochreiter and Schmidhuber 1997], Gated Recurrent Unit (GRU) [Chung et al. 2014], and Graph Convolutional Long Short-Term Memory (GConvLSTM), with results compared to the NCEP-CFSv2 model from the North American Multi-Model Ensemble (NMME) dataset; both studies show strong performance, even at broader spatial scales than those used here. In this scenario, artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a promising approach to complement numerical models [Anochi et al. 2021, Anochi and Shimizu 2025], as these methods can capture complex non-linear relationships and process large datasets, while offering lower computational cost and faster inference [Lam et al. 2023].

This study investigates the potential of ML techniques for subseasonal precipitation forecasting in Southeastern Brazil. Four models were analyzed: MLP, LSTM, GRU, and CNN-1D. The models were trained with ERA5 atmospheric variables using the previous four weeks to predict precipitation for the subsequent four weeks. The performance of the ML techniques was also compared to forecasts from the BAM, allowing an assessment of their ability to complement traditional numerical prediction systems in subseasonal precipitation forecasting. In addition, SHAP (SHapley Additive exPlanations) values [Lundberg and Lee 2017] were used to explain the behavior of the best-performing ML model, identifying the most influential atmospheric variables for its predictions.

2. Materials and methods

Figure 1 illustrates the workflow of the proposed methodology considered in this study.

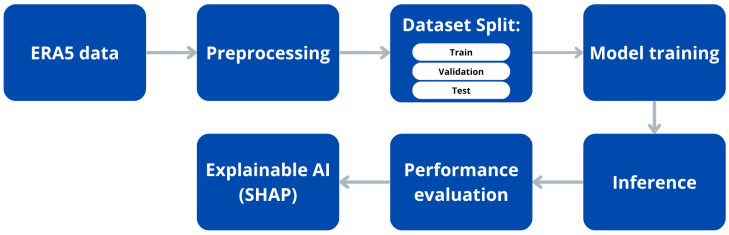


Figure 1. Workflow of the proposed methodology.

2.1. Data

The research focuses on the Southeast region of Brazil, which includes the states of São Paulo, Minas Gerais, Rio de Janeiro, and Espírito Santo. This region stands out for its economic relevance, population density, and diversity of industrial and urban activities, making it an important area for weather studies.

As for this study, weekly mean values from 1991 to 2018 were obtained from the ERA5 reanalysis product, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA5 reanalysis provides global historical time series of atmospheric variables at hourly resolution since 1940, covering multiple pressure levels (from 1000 to 1 hPa) with a spatial resolution of $0.25° \times 0.25°(\approx 28\,\text{km} \times 28\,\text{km})$ in latitude and longitude. The dataset combines global observations with numerical weather prediction models via data assimilation techniques, resulting in a physically consistent and globally complete representation of atmospheric conditions [Copernicus Climate Change Service (C3S) 2017].

The variables used in this study are presented in Table 1, including their names, units, and corresponding atmospheric levels. The temporal division of the dataset was defined as follows: the period from 1991 to 2016 was used for training and validation, while the years 2017 and 2018 were reserved for testing. As input, the weekly mean values observed during the previous four weeks were used to predict the weekly mean precipitation for the subsequent four weeks.

Variable	Unit	Atmospheric Level
Total precipitation (tp)	mm	Surface
2 m air temperature (t2m)	K	Surface
2 m dewpoint temperature (d2m)	K	Surface
Surface pressure (sp)	Pa	Surface
Mean sea level pressure (msl)	Pa	Surface
Total column water vapour (tcwv)	kg/m²	Surface
Specific humidity (q)	kg/kg	850 hPa
10 m u-component of wind (u10)	m/s	Surface
10 m v-component of wind (v10)	m/s	Surface
U-component of wind (u)	m/s	850 hPa
V-component of wind (v)	m/s	850 hPa
Geopotential (z)	m²/s²	500 hPa

Table 1. ERA5 reanalysis variables used in the study.

In total, the dataset comprises 3,708,018 samples distributed over 2,538 grid points with 12 features. The final dataset occupies approximately 328.87 MB. A temporal holdout strategy was adopted. The training and validation set contains 3,433,914 samples (92.6%), while the test set comprises 274,104 samples (7.4%). The test period spans from 2016-12-07 to 2018-12-26.

2.2. ML Models

The ML models used in this study were selected to capture nonlinear relationships and temporal dependencies in precipitation data. Four techniques were evaluated: MLP, LSTM, GRU, and CNN-1D.

MLP is a feedforward neural network composed of multiple fully connected layers that learn nonlinear relationships between input variables through the application of activation functions. By propagating information through hidden layers, the model is able to approximate complex mappings between inputs and outputs. In contrast, LSTM and GRU are recurrent neural network architectures specifically designed to model sequential data and capture temporal dependencies. These models incorporate gating mechanisms that regulate the flow of information across time steps, enabling the learning of patterns that evolve over time. Additionally, the CNN-1D architecture applies convolutional filters along the temporal dimension, allowing the extraction of local temporal patterns from the input sequences.

All models were trained using the same input configuration, consisting of weekly averages of atmospheric variables from the previous four weeks to predict precipitation for the subsequent four weeks. This setup enables a consistent comparison of the predictive performance of the different architectures.

Hyperparameter optimization was performed using the Optuna framework [Akiba et al. 2019] for automated hyperparameter search, where the loss computed on the validation set was used as the objective function for model selection. The final configurations selected for each model are presented in Table 2 (MLP), Table 3 (GRU), Table 4 (LSTM), and Table 5 (CNN-1D).

Hyperparameter	Range	Selected
Hidden Layers	[1, 5]	4
Hidden Size	[32, 512]	480
Dropout	[0.1, 0.5]	0.134
Learning Rate	$[10^{-4}, 10^{-2}]$	0.0003
Weight Decay	$[10^{-6}, 10^{-2}]$	2.88
Epochs	200	200
Patience	10	10
Optimizer	[Adam, SGD]	Adam

Table 2. Selected hyperparameters of the MLP Model.

Hyperparameter	Range	Selected
Hidden Layers	[1, 5]	2
Hidden Size	[32, 512]	64
Dropout	[0.1, 0.5]	0.2
Learning Rate	$[10^{-4}, 10^{-2}]$	0.001
Weight Decay	$[10^{-6}, 10^{-2}]$	0.0001
Epochs	200	200
Patience	10	10
Optimizer	[Adam, SGD]	Adam

Table 3. Selected hyperparameters of the GRU Model.

Hyperparameter	Range	Selected
Hidden Layers	[1, 5]	2
Hidden Size	[32, 512]	64
Dropout	[0.1, 0.5]	0.2
Learning Rate	$[10^{-4}, 10^{-2}]$	0.001
Weight Decay	$[10^{-6}, 10^{-2}]$	0.0001
Epochs	200	200
Patience	10	10
Optimizer	[Adam, SGD]	Adam

Table 4. Selected hyperparameters of the LSTM Model.

Hyperparameter	Range	Selected
Hidden Layers	[1, 5]	4
Number of Filters	[32, 128]	128
Kernel Size	2	2
Dropout	[0.1, 0.5]	0.125
Learning Rate	$[10^{-4}, 10^{-2}]$	0.0002
Weight Decay	$[10^{-6}, 10^{-2}]$	9.76
Epochs	200	200
Patience	10	10
Optimizer	[Adam, SGD]	Adam

Table 5. Selected hyperparameters of the CNN-1D Model.

The hyperparameters epochs, patience, and kernel size were fixed and kept constant across all experiments.

2.3. BAM Model

In addition to the ML models, forecasts from the BAM model were used as a reference for comparison. BAM is a global atmospheric model developed by the Center for Weather

Forecasting and Climate Studies (CPTEC) of the National Institute for Space Research (INPE). The model is designed for both weather and climate prediction and can operate on time scales ranging from days to seasonal forecasts, with horizontal resolutions ranging from approximately 10 to 200 km.

BAM is based on the physical equations that describe atmospheric dynamics and includes numerical and physical parameterization schemes to represent atmospheric processes. In this study, BAM forecasts are used as a baseline to evaluate the performance of the ML models in subseasonal precipitation prediction.

2.4. Explainable artificial intelligence (XAI)

To better understand the contribution of the input variables to the precipitation predictions, a XAI technique, SHAP values, was used to quantify the contribution of each predictor variable to the model outputs. This technique helps to explain the behavior of the model. SHAP is based on concepts from cooperative game theory and assigns an importance value to each feature according to its contribution to the prediction. This approach allows the identification of the most influential atmospheric variables in the forecasting process and provides additional insights into the relationships learned by the models.

2.5. Evaluation metrics

To evaluate the performance of the models, six metrics were used, including continuous and categorical metrics. Continuous metrics quantify the differences between observed and predicted precipitation values, whereas categorical metrics assess the ability of the models to detect precipitation events. The metrics and their corresponding equations are presented in Table 6, where y_i represents the observed value, $\hat{y}_i$ the predicted value, and n is the number of samples. For the categorical metrics, TP represents True Positives, FP False Positives, and FN False Negatives.

Metric	Equation	Description
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	Root Mean Squared Error. Measures the square root of the average squared difference between observed and predicted values. The lower its value, the better.
BIAS	$\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$	Bias. Indicates whether the model systematically overestimates (positive values) or underestimates (negative values) the observations. Values closer to zero indicate better performance.
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$	Mean Absolute Error. Represents the average magnitude of the absolute differences between observed and predicted values. The lower its value, the better.
POD	$\frac{TP}{TP+FN}$	Probability of Detection. Measures the proportion of observed events that were correctly predicted. The higher its value, the better.
FAR	$\frac{FP}{TP+FP}$	False Alarm Ratio. Represents the proportion of predicted events that did not occur. The lower its value, the better.
CSI	$\frac{TP}{TP+FP+FN}$	Critical Success Index. Evaluates the model's ability to correctly capture events considering true positives, false positives, and false negatives. The higher its value, the better.

Table 6. Evaluation metrics used to assess model performance.

All experiments were performed on a workstation equipped with an Intel Core Ultra 5 245KF CPU (4.20 GHz), 64 GB of RAM, and an NVIDIA GeForce RTX 4080 SUPER GPU with 16 GB of VRAM.

3. Results

3.1. Model performance across forecast horizons

Table 7 presents the evaluation metrics for the precipitation forecasts produced by the different models across the four forecast weeks. In general, the ML models outperform the BAM model in terms of RMSE and MAE for all forecast horizons, indicating higher predictive accuracy. Among the models, the MLP achieves the lowest errors in most weeks (weeks 1–3), while LSTM presents the lowest RMSE in week 4.

A gradual increase in RMSE and MAE is observed from week 1 to week 4, which is expected in subseasonal forecasting due to the growth of predictive uncertainty with lead time. Even so, the ML models maintain relatively stable performance, indicating their ability to capture persistent atmospheric patterns affecting precipitation variability. In contrast, BAM consistently shows higher errors, suggesting limitations in representing local precipitation dynamics at the subseasonal scale.

Bias values are generally closer to zero for the ML models, indicating a better representation of precipitation magnitude. BAM shows a systematic positive bias in most weeks, suggesting a tendency to overestimate precipitation. Among the ML approaches, LSTM often presents the smallest bias values, indicating a more balanced representation of precipitation intensity.

For the categorical metrics, the ML models also achieve higher POD values, indicating improved detection of rainfall events. The MLP model reaches the highest POD values in several weeks, including 0.977 in week 1 and 0.976 in week 4. Although BAM shows the lowest FAR values, this occurs alongside its larger RMSE and MAE, suggesting a more conservative behavior that reduces false alarms but limits event detection and intensity representation.

CSI values remain relatively similar among the models, although the ML approaches generally maintain competitive performance. Considering all metrics and forecast horizons, MLP shows the best overall performance, followed by LSTM, particularly at longer lead times. CNN-1D and GRU also achieve competitive results, though with slightly higher errors in some weeks.

Overall, these results indicate that ML approaches can effectively complement traditional numerical weather prediction models for subseasonal precipitation forecasting, as they are able to capture complex nonlinear relationships among atmospheric variables that are difficult to fully represent in conventional dynamical models such as BAM.

Considering all metrics and weeks, the MLP model exhibited the best overall performance. Following the MLP, the LSTM model also demonstrated strong performance. As previously noted, all ML models, in general, outperformed the traditional BAM model.

Week	Model	RMSE	MAE	Bias	POD	FAR	CSI
W1	BAM	32.09	19.56	2.34	0.933	**0.153**	**0.798**
	MLP	**26.87**	**16.57**	-1.62	**0.977**	0.198	0.788
	LSTM	27.34	17.25	**-0.48**	0.952	0.184	0.784
	CNN	27.57	17.29	-1.19	0.972	0.201	0.781
	GRU	28.00	18.20	1.88	0.955	0.191	0.779
W2	BAM	31.58	19.57	2.81	0.905	**0.160**	0.772
	MLP	**26.83**	**16.58**	-1.34	0.970	0.203	**0.778**
	LSTM	26.99	17.26	**-0.09**	0.940	0.203	0.758
	CNN	27.72	17.58	-0.19	**0.975**	0.210	0.775
	GRU	28.66	18.70	1.84	0.930	0.202	0.753
W3	BAM	31.79	20.26	3.42	0.928	**0.166**	**0.783**
	MLP	**26.38**	**16.29**	-1.54	**0.971**	0.206	0.775
	LSTM	26.41	16.97	**-0.19**	0.962	0.219	0.758
	CNN	27.05	17.01	-1.89	0.964	0.209	0.769
	GRU	27.03	17.43	0.71	0.958	0.210	0.763
W4	BAM	31.33	20.13	4.24	0.911	**0.179**	0.760
	MLP	27.51	**17.43**	-1.89	**0.976**	0.212	**0.773**
	LSTM	**27.17**	17.62	**-0.27**	0.952	0.215	0.755
	CNN	27.75	17.67	-2.30	0.956	0.206	0.766
	GRU	28.32	18.83	-0.19	0.942	0.193	0.768

Table 7. Evaluation metrics of precipitation forecasts for the models.

3.2. Spatial analysis of precipitation forecasts

To illustrate the spatial performance of the models, a representative case from March 2018 was selected. This month highlights the differences between ML models and the traditional BAM forecast.

The monthly mean precipitation map (Figure 2) shows that MLP, LSTM and CNN-1D capture the spatial distribution of rainfall more accurately than GRU and BAM. Observed rainfall reveals intense precipitation peaks along the coast of Rio de Janeiro and Espírito Santo, with additional areas in the interior regions of São Paulo and Minas Gerais. MLP, LSTM and CNN-1D reproduce the coastal gradients and inland decrease reasonably well. In contrast, the GRU tends to overestimate precipitation across most regions, while the BAM shows greater difficulty in representing the spatial location and intensity of precipitation.

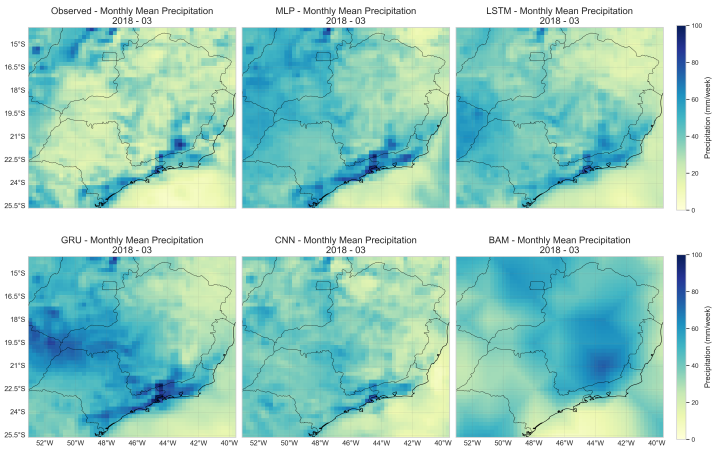


Figure 2. Monthly mean precipitation for March 2018: observed vs. predicted by MLP, CNN-1D, LSTM, GRU, and BAM.

The corresponding monthly mean error maps (Figure 3) reveal the magnitude and spatial distribution of forecast errors. In this analysis, the error is computed as the difference between predicted and observed precipitation ($\hat{y} - y$), where positive values indicate overestimation and negative values indicate underestimation. GRU exhibits the largest overestimation in the interior (red patches), while BAM shows widespread positive bias across Minas Gerais. CNN-1D and LSTM present more balanced errors, with CNN-1D slightly underestimating rainfall along the coast. MLP displays moderate overestimation across the continental area. These error patterns confirm the ability of MLP, LSTM and CNN-1D to capture the spatial structure of precipitation, albeit with small magnitude biases in certain regions.

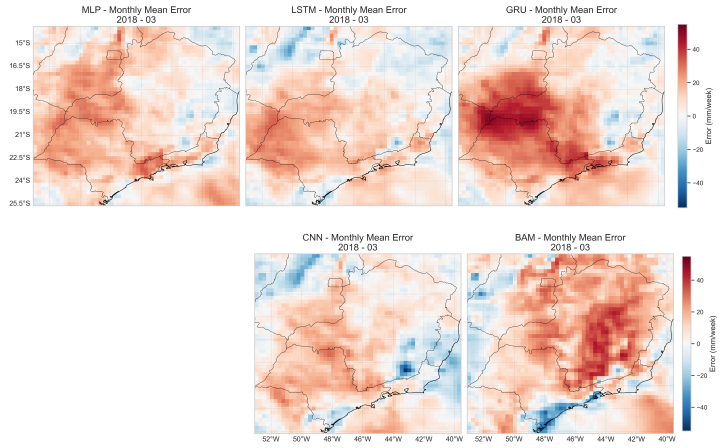


Figure 3. Mean error of monthly precipitation forecasts for March 2018, highlighting overestimation (red) and underestimation (blue) by each model.

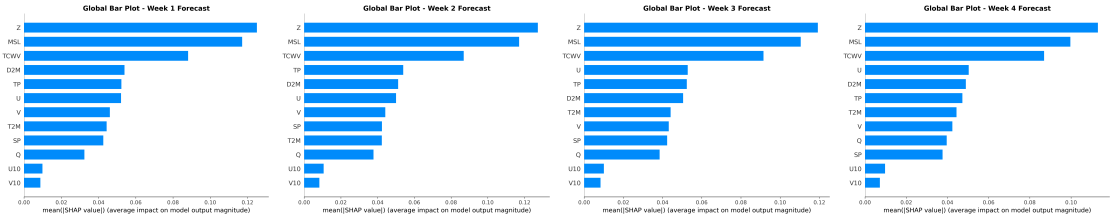
3.3. Model explainability using SHAP

SHAP (SHapley Additive exPlanations) values were used to analyze how the atmospheric variables influence the forecasts generated by the MLP model across the four subseasonal forecast horizons (week 1 to 4). Since the MLP achieved the best predictive performance, it was selected for the explainability analysis.

The global SHAP importance results reveal a consistent dominance of large-scale atmospheric variables. Geopotential appears as the most influential predictor in all forecast weeks, followed by mean sea level pressure and total column water vapor. This indicates that the model primarily relies on the large-scale atmospheric structure and moisture availability to generate predictions.

Although the leading predictors remain stable, small variations are observed among intermediate variables as the forecast horizon increases. In Week 1 (Fig. 4a), wind components and 2-m dewpoint temperature show similar importance. At Week 4 (Fig. 4d), the zonal wind component becomes slightly more relevant, suggesting a stronger influence of large-scale circulation patterns for longer lead times.

Variables associated with near-surface wind and 2-m air temperature consistently present lower contributions. Overall, the SHAP results indicate a stable hierarchy of predictor importance across the four weeks, suggesting that the MLP model captures physically consistent relationships within the atmospheric system.



(a) Week 1 **(b) Week 2** **(c) Week 3** **(d) Week 4**

Figure 4. Mean SHAP values of weekly precipitation forecasts for the MLP model.

4. Conclusions

Subseasonal precipitation forecasting remains challenging due to the complex interactions among atmospheric dynamics, convective processes, and regional climate variability. This study investigated the performance of ML models, i.e., MLP, CNN-1D, LSTM, and GRU, for subseasonal precipitation forecasting in Southeastern Brazil. ML models consistently outperformed the traditional BAM forecasting model, with MLP presenting the best overall performance. In terms of RMSE and MAE, LSTM presented the second-best performance. LSTM and GRU produced smoother rainfall patterns, capturing general trends, while BAM exhibited the largest bias.

The explainability analysis using SHAP values highlighted geopotential, mean sea level pressure, and total column water vapor as the most influential predictors, emphasizing the importance of large-scale atmospheric structure and moisture availability in driving subseasonal precipitation forecasts. These results indicate that, the ML models demonstrate the capability to enhance subseasonal precipitation forecasts, contributing to better decision-making in water management, agriculture, and disaster preparedness across the region.

Future work includes extending the analysis to the entire Brazilian territory, investigating subseasonal precipitation forecasts at daily resolution, and further improving the performance of the ML models through architectural refinements, and the incorporation of additional atmospheric variables as input predictors.

5. Acknowledgements

This research was developed within the project *Classificação de imagens e dados via redes neurais profundas para múltiplos domínios* (IDeepS; available online: `https://github.com/vsantjr/IDeepS`), supported by LNCC, and in the *Laboratório de Inteligência ARtificial para Aplicações AeroEspaciais e Ambientais* (LIAREA; available online: `https://liarealab.github.io/liarea_page/`). This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES) – Finance Code 001. This research was also supported by CNPq (Finance Code 130463/2025-6).

Referências

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *CoRR*, abs/1907.10902.

Anochi, J. A., de Almeida, V. A., and de Campos Velho, H. F. (2021). Machine learning for climate precipitation prediction modeling over south america. *Remote Sensing*.

Anochi, J. A. and Shimizu, M. H. (2025). Precipitation forecasting and drought monitoring in south america using a machine learning approach. *Meteorology*.

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Copernicus Climate Change Service (C3S) (2017). Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate.

Domingos, M., Júnior, V. S., Anochi, J., and Shiguemori, E. (2025). Exploring deep learning techniques for seasonal prediction of autumn precipitation in south america. In *Anais do XVI Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais*, pages 157–166. SBC.

Figueroa, S. N., Bonatti, J. P., Kubota, P. Y., et al. (2016). The brazilian global atmospheric model (bam): Performance for tropical rainfall forecasting and sensitivity to convective scheme and horizontal resolution. *Weather and Forecasting*, 31(5):1547–1565.

Hersbach, H. et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., and Battaglia, P. (2023). Learning skillful medium-range global weather forecasting.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, volume 30.

Monego, V. S., Anochi, J. A., and de Campos Velho, H. F. (2022). South america seasonal precipitation prediction by gradient-boosting machine-learning approach. *Atmosphere*, 13(2).

Morrison, H., van Lier-Walqui, M., Fridlind, A. M., et al. (2020). Confronting the challenge of modeling cloud and precipitation microphysics. *Journal of Advances in Modeling Earth Systems*, 12.

NOOA (2017). Subseasonal and seasonal forecasting innovation: Plans for the twenty-first century - annotated outline. Technical report, NOAA.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323:533–536.

Sales, R. A., de Oliveira, E. C., Delgado, R. C., Leite, M. C. T., Ribeiro, W. R., and Berilli, S. S. (2018). Sazonal and interanual rainfall variability for colatina, espírito santo, brazil.

Yan, Z., Lu, X., Wu, L., Liu, F., Qiu, R., Cui, Y., and Ma, X. (2025). Evaluation of precipitation forecasting based on graphcast over mainland china.