Exploring Deep Learning Techniques for Seasonal Prediction of Autumn Precipitation in South America

Matheus Corrêa Domingos[1], **Valdivino Alexandre de Santiago Júnior**[1], **Juliana Aparecida Anochi**[1], **Elcio Hideiti Shiguemori**[2]

[1]Laboratório de Inteligência ARtificial para Aplicações AeroEspaciais e Ambientais (LIAREA), Coordenação de Pesquisa Aplicada e Desenvolvimento Tecnológico (COPDT), Programa de Pós-Graduação em Computação Aplicada (CAP), Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP, Brasil

[2]Laboratório de Inteligência ARtificial para Aplicações AeroEspaciais e Ambientais (LIAREA), Instituto de Estudos Avançados, São José dos Campos, SP, Brasil

matheus.domingos@inpe.br, valdivino.santiago@inpe.br, juliana.anochi@inpe.br, elcio.shiguemori@ieav.cta.br

Abstract. *Seasonal precipitation forecasting is essencial for climate planning and disaster risk reduction. This study used Convolutional Neural Network 1D (CNN 1D), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Graph Convolutional Long Short-Term Memory (GConvLSTM) to forecast autumn 2023 precipitation. CNN 1D had the best performance (MSE: 1.8827, Correlation Coefficient: 0.9137), excelling in accuracy but losing spatial details. GConvLSTM captured spatial patterns better, while LSTM and GRU underperformed. Despite CNN 1D's superiority, integrating spatial components, as in GConvLSTM, could improve predictions.*

Resumo. *A previsão de precipitação sazonal é essencial para o planejamento climático e a redução de riscos de desastres naturais. Este estudo utilizou Convolutional Neural Network 1D (CNN 1D), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) e Graph Convolutional Long Short-Term Memory (GConvLSTM) para prever a precipitação do outono de 2023. CNN 1D teve o melhor desempenho (MSE: 1.8827, Coeficiente de Correlação: 0.9137), destacando-se em precisão, mas perdendo detalhes espaciais. GConvLSTM capturou melhor os padrões espaciais, enquanto LSTM e GRU tiveram desempenho inferior. Apesar da superioridade da CNN 1D, a integração de componentes espaciais, como no GConvLSTM, poderia melhorar as previsões.*

1. Introduction

Precipitation is one of the most challenging meteorological variables to predict, primarily due to the complexity of cloud microphysics, which involves processes such as condensation, coalescence, and ice crystal formation. These processes are strongly influenced by a variety of other atmospheric variables [Holton and Hakim 2013]. With increasingly intense global environmental climate changes, precipitation forecasting becomes increasingly difficult to perform accurately.

Traditionally, numerical weather/climate prediction models, based on physical equations, remain the primary approach to address this issue. These models aim to identify atmospheric behavior by solving complex systems of differential equations that govern atmospheric dynamics. To optimize the computational process, discretization methods and parallelism techniques are employed, although this approach still incurs high computational costs [Reboita et al. 2018, CPTEC 1988].

With advancements in artificial intelligence, especially in Machine Learning (ML) and its subfield Deep Learning (DL), promising alternatives have emerged for precipitation forecasting. While ML is effective at recognizing patterns in large datasets, it can struggle with very large data or complex patterns. In contrast, DL techniques excel in handling these challenges, making them a powerful and effective approach for accurate precipitation forecasting. These methods are particularly suited for capturing intricate temporal and spatial dependencies in climate data, offering significant improvements in forecasting precision [DataCamp 2023]

Based on data from the Global Precipitation Climatology Project (GPCP), a seasonal precipitation forecast for South America was made using Extreme Gradient Boosting (XGBoost) and Multilayer Perceptron (MLP) models [Monego et al. 2022]. Also using GPCP data, monthly precipitation forecasts were made, including an analysis of droughts, using an MLP model and comparing the results with data from the North American Multi-Model Ensemble (NMME) [Anochi and Shimizu 2024]. However, the use of DL techniques for precipitation forecasting still needs to be properly explored, by evaluating several approaches and perceiving their performances.

The Chinese summer precipitation forecast between 2019 and 2021 was investigated using climate data from 1991 to 2021 with a spatial resolution of $1° \times 1°$ and observed precipitation data from Merged Analysis of Precipitation (CMAP) with a resolution of $2.5° \times 2.5°$ [Yang et al. 2022]. In Brazil, a seasonal precipitation forecast was performed with the Support Vector Machine (SVM) model and climate indices as input variables [Torres et al. 2024].

The seasonal precipitation forecast in South America for the years 2018 and 2019 was evaluated using data from the Global Precipitation Climatology Project (GPCP) with a spatial resolution of $2.5° \times 2.5°$. A Multilayer Perceptron (MLP) optimized by the Multiparticle Collision Algorithm (MPCA) was compared with the standard MLP and the Brazilian Atmospheric Model (BAM). The results indicated that the MPCA improved the accuracy of the MLP, although limitations remained in some regions, as evidenced by the MSE (Mean Squared Error) and other evaluation metrics [Anochi et al. 2021].

The literature still presents a lack of more comprehensive investigations on the application of DL models in climate precipitation forecasting, indicating a promising field to be explored.

This article aims to make a climate seasonal precipitation forecast for the autumn season in South America in the year of 2023 using DL techniques, such as CNN 1D [Kiranyaz et al. 2019], LSTM [Staudemeyer and Morris 2019], GRU [Chung et al. 2014], and GConvLSTM [Seo et al. 2016]. The choice of autumn as the focus of the research is due to it being a transitional season between summer and winter, presenting atmospheric characteristics that can significantly influence climate forecasts. The

forecast will be based on atmospheric behaviors from the previous summer season as well as historical data from the autumn season itself. Additionally, the results obtained from the DL models will be compared with predictions from the North American Multi-Model Ensemble (NMME) to assess their performance and reliability [Kirtman et al. 2014].

2. Materials and methods

Figure 1 presents the flowchart of the proposed methodology. To perform the autumn climate forecasting, tabular and gridded data from ECMWF Reanalysis v5 (ERA5) were used. These data have a resolution of 0.25° x 0.25°, which is considered high for climate studies, and are part of a reanalysis dataset [Hersbach et al. 2023, Service 2023]. Reanalysis combines global observations with model data to produce a globally comprehensive and consistent dataset, based on the laws of physics.

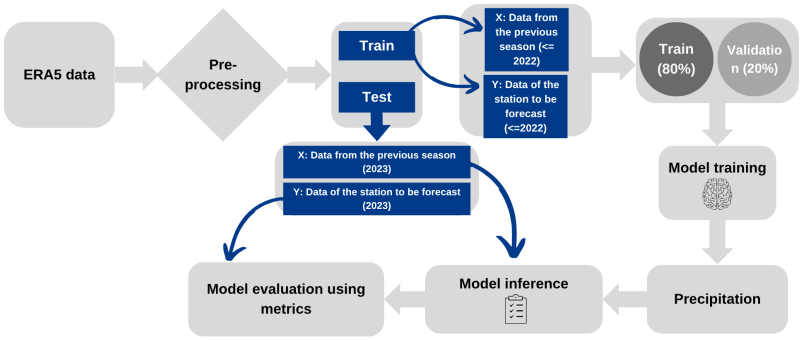

Figure 1. Flowchart of the proposed methodology

The climatic variables contained in this dataset are detailed in Table 1, where the variables precipitation and mean precipitation rate were originally in meters per day (m/day) but were converted to mm/day during preprocessing.

Variables	Representation	Units
Precipitation	tp	mm/day
Air temperature at 2 meters	t2m	K
Horizontal wind speed component	u	m/s
Vertical wind speed component	v	m/s
Surface pressure	sp	Pa
Mean precipitation rate	mtpr	mm/day
Geopotential height	z	m
Specific humidity	q	kg/kg
Air temperature	t	K
Divergence	d	s^{-1}

Table 1. Variable representations and their respective units.

Initially, the data were organized based on the seasons of the year. Then, two sets were created: the input set, consisting of historical data from 1989 to 2022, and the test set, containing data from 2023. Before proceeding, a sensitivity analysis was

conducted, focusing on the correlation matrix between the variables, as it helps to assess the importance of the variables based on their linear correlations. After this analysis, the input set was further subdivided into training and validation sets, with 80% of the data allocated for training and 20% for validation.

Based on the correlation matrix (see Figure 2), some variables were removed to avoid redundancy and reduce collinearity in the model. The year, month and pressure level variables were excluded, as they did not present a significant correlation with other variables, making it irrelevant for prediction. The mtpr variable was discarded due to its 100% correlation with tp (total precipitation; this is the variable that needs to be predicted), indicating that both carry the same information. Similarly, sp was removed due to its strong correlation with other meteorological variables, while d was excluded for its low predictive power and high correlation with existing variables. These removals help optimize the model by preserving only the most relevant information and reducing data complexity. After this analysis, the final set of input variables contains q, t, t2m, tp, u, v, z, latitude, and longitude.

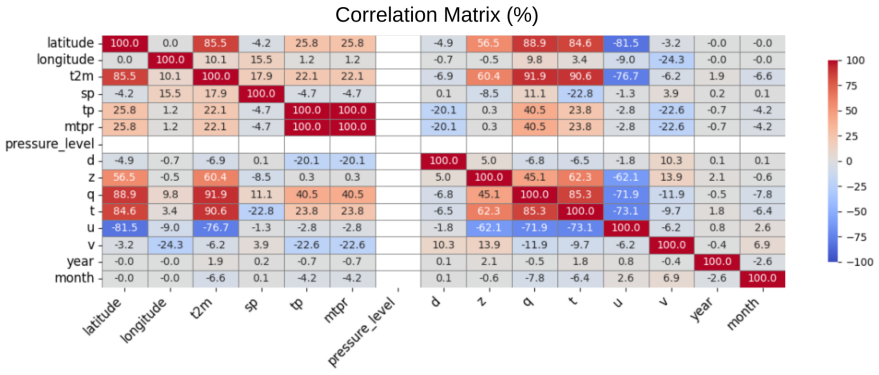

Figure 2. Correlation matrix

After the sensitivity analysis, the model hyperparameters were fine-tuned to optimize its performance. The adjusted values, which showed the best combination of results for precipitation forecasting, are detailed in Table 2. This table presents the key configurations, such as the number of hidden layers, the activation function used, the batch size, the learning rate, among others, which were essential for the success of all models during training.

Parameters	**Value**
Number of hidden layers	5
Epochs	200
Activation function	ReLU
Batch size	1024
Patience	20
Neurons	[256, 128, 64, 32, 64, 128, 256]
Optimizer	AdamW
Learning Rate	0.001

Table 2. Hyperparameters of all DL models.

As for the GConvLSTM model, the number of connections was determined using a Multi-Layer Perceptron (MLP) network with 2 layers. The MLP was trained for 20 epochs with a learning rate of 0.001 to optimize the adjustment of the connections.

In the data extracted from NMME for the autumn of 2023, total precipitation data was used, and bilinear interpolation was applied to align it with the ERA5 grid. This interpolation aimed to ensure spatial compatibility between the two data sources, enabling a more consistent and comparable analysis between the model results and the NMME forecasts. By aligning the precipitation data to the same grid, this process minimized spatial discrepancies and allowed for a more accurate assessment of precipitation patterns and forecasting performance across both datasets.

Regarding the selected DL models, LSTM, GRU, 1D CNN, and GConvLSTM are neural networks used to process sequential and structured data. LSTM captures long-term dependencies through three gates: forget, input, and output [Staudemeyer and Morris 2019]. GRU is a more compact version with reset and update gates [Chung et al. 2014]. 1D CNN learns patterns in one-dimensional data and uses them in MLP models for tasks like classification and prediction [Kiranyaz et al. 2019]. GConvLSTM combines graph convolutions with LSTM, where each grid point is a node, and the connections between them can be adjusted using a distance measure, enabling the modeling of spatial and temporal data [Seo et al. 2016].

2.1. Assessment metrics

To assess the performance of the models, we employed the following evaluation metrics: R^2, MSE, MAE, and the Correlation Coefficient. The corresponding equations are provided below:

Coefficient of Determination (R^2): This metric quantifies the proportion of the variance in the observed data explained by the model. A value closer to 1 indicates a better model fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

Where y_i is the observed value for the i-th sample, $\hat{y_i}$ is the predicted value for the i-th sample, and $\bar{y}$ is the mean of the observed values.

Mean Squared Error (MSE): MSE computes the average of the squared differences between observed and predicted values, with larger errors being penalized more.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{2}$$

Where y_i is the observed value for the i-th sample, $\hat{y_i}$ is the predicted value for the i-th sample, and n is the total number of samples.

Mean Absolute Error (MAE): MAE calculates the average of the absolute differences between observed and predicted values, offering a straightforward indication of model accuracy.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}| \tag{3}$$

Where y_i is the observed value for the i-th sample, $\hat{y}_i$ is the predicted value for the i-th sample, and n is the total number of samples.

Correlation Coefficient (r): The correlation coefficient quantifies the linear relationship between observed and predicted values, with a range from -1 to 1. A value of 1 indicates perfect positive correlation, while -1 indicates perfect negative correlation.

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}} \tag{4}$$

Where y_i represents the observed values for the i-th sample, $\hat{y}_i$ represents the predicted values for the i-th sample, $\bar{y}$ and $\bar{\hat{y}}$ are the means of the observed and predicted values, respectively, and n is the total number of samples.

3. Results

During autumn in the Southern Hemisphere, which serves as a transitional period between summer and winter, one of the most notable characteristics is the persistence of rainfall in the northern and northeastern regions of Brazil, often influenced by the Intertropical Convergence Zone (ITCZ). This period is also marked by the occurrence of the first typical adverse weather phenomena, such as fog in the South, Southeast, and Central-West regions, frosts in the South, Southeast, and Mato Grosso do Sul, snow in the mountainous areas and plateaus of the South region, and cold spells in the southern part of the North region, as well as in Mato Grosso do Sul and Mato Grosso. Another typical phenomenon of autumn is the intrusion of cold air masses from the southern part of the continent, which cause a significant drop in temperatures, especially in the South and Southeast regions. These cold air masses result in colder nights and, in some areas, frosts, marking the gradual transition between warmer and cooler seasons, as well as contributing to a more unstable climate in various parts of Brazil [INSTITUTO NACIONAL DE METEOROLOGIA 2024].

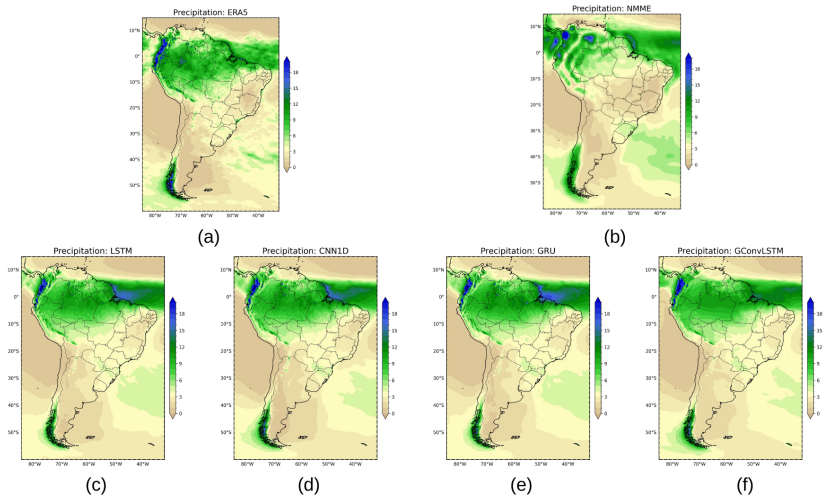

Figure 3. Observed precipitation and estimates. (a) ERA5 (reference), (b) NMME, (c) LSTM, (d) CNN 1D, (e) GRU, (f) GConvLSTM.

Figure 3 presents a comparison between observed precipitation (ERA5) and estimates generated by the NMME model and different neural network architectures (LSTM, CNN 1D, GRU, and GConvLSTM) for South America. In general, the models capture the seasonal distribution of precipitation, with the highest accumulations over the Amazon region and the Andes Mountains, which are typical characteristics of the region's climate regime.

However, differences in the intensity and location of precipitation maxima can be observed. The NMME model (b) exhibits general patterns similar to ERA5 but shows excessive intensification in northern South America and underestimates precipitation in southern Brazil. The LSTM (c) and GRU (e) display very similar patterns, but with variations in rainfall intensity, especially in the Amazon, where both seem to smooth out maximum values. The CNN 1D (d) tends to further smooth spatial patterns, losing local details, particularly in the Andes region and along the eastern coast of Brazil. The GConvLSTM (f), by incorporating graph convolutions, better captures certain geographical structures of precipitation, making it more similar to the spatial distribution observed in ERA5. When compared to the observed data, some models tend to underestimate precipitation in the Amazon and southern Brazil, while others overestimate it in coastal areas of the Andes Mountains.

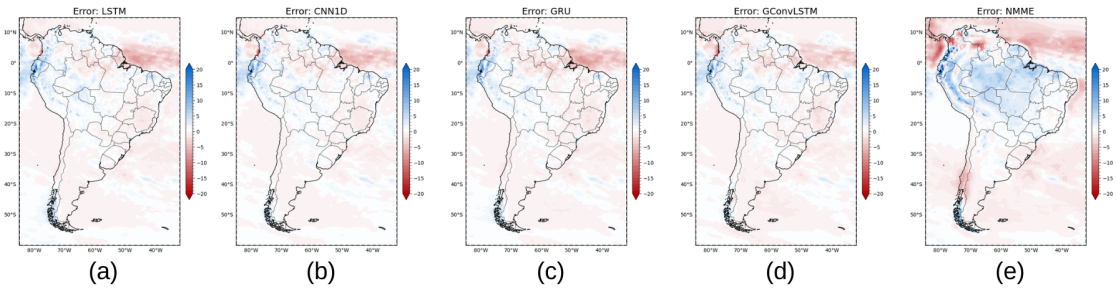

Figure 4. Error maps. (a) LSTM, (b) CNN 1D, (c) GRU, (d) GConvLSTM, (e) NMME.

Figure 4 presents precipitation error maps over South America during autumn, comparing the LSTM (a), CNN1D (b), GRU (c), GConvLSTM (d) and NMME (e) models. LSTM (a) underestimates precipitation over most of the continent, with some overestimations located in the Amazon and the Pacific coast. CNN1D (b) has a more balanced error distribution, but still presents underestimations in the eastern region. GRU (c) exhibits a similar behavior to CNN1D, but with smoother underestimations and more uniform patterns. GConvLSTM (d) demonstrates the smallest underestimation bias, with overestimations concentrated on the northwest coast. NMME (e) presents the largest overall error, with significant underestimations in the south and intense overestimations in the northern and western coastal regions.

The models struggle to accurately capture precipitation intensities, especially in regions with high climate variability like the Amazon and the Andes. The LSTM model shows the strongest dry bias, while the GConvLSTM model appears to better capture some spatial structures, reducing systematic errors. Quantitative evaluations are crucial to determine which model provides the most accurate seasonal precipitation predictions.

Model	MSE	R^2	MAE	Correlation Coefficient
CNN 1D	**1.8827**	**0.8264**	**0.8226**	**0.9137**
LSTM	2.1719	0.7998	0.9277	0.8976
GRU	2.3152	0.7866	0.9327	0.8979
GConvLSTM	1.9589	0.8194	0.9061	0.9083
NMME	6.9661	0.3586	1.7353	0.6479

Table 3. Performance metrics for each model in autumn.

The results shown in Table 3 indicate that the CNN 1D model performs best in predicting precipitation during the autumn of 2023, standing out with the lowest Mean Squared Error (MSE = 1.8827), the highest R^2 (0.8264), the lowest Mean Absolute Error (MAE = 0.8226), and the highest Correlation Coefficient (0.9137), suggesting a strong correspondence between the predicted and observed values. The GConvLSTM model also performs competitively, with an R^2 close to that of CNN 1D (0.8194), but with a slightly higher MSE. LSTM and GRU perform worse, with higher MSE and lower overall accuracy, evidenced by the higher MAE values. The NMME data, with an MSE of 6.9661, an R^2 of 0.3586, an MAE of 1.7353, and a Correlation Coefficient of 0.6479, shows relatively poorer performance in comparison. Overall, CNN 1D stands out as the most efficient and reliable model for predicting precipitation in the fall, while the other models, especially GRU, may require adjustments to improve their accuracy.

4. Discussion

The evaluation of deep learning models for seasonal precipitation forecasting in South America shows distinct strengths and weaknesses. All models capture the general seasonal distribution, with the Amazon and Andes experiencing the highest precipitation levels, but each model exhibits specific biases in spatial patterns and rainfall intensity.

LSTM and GRU models show similar precipitation patterns but tend to underestimate rainfall in areas like the Amazon and southern Brazil, despite capturing large-scale trends. The CNN 1D model, which achieves the best performance metrics (Table 3), smooths precipitation patterns, potentially losing localized details. This is evident in the error maps (Figure 4), where CNN 1D presents a balanced error distribution but underestimates rainfall in the eastern part of the continent.

The GConvLSTM model, incorporating graph convolutions, better captures geographic precipitation structures and reduces underestimation bias compared to other models. However, it does not surpass CNN 1D in accuracy, as shown by the performance metrics.

In comparison, the NMME data, bilinearly interpolated to the ERA5 grid, shows poorer performance with an MSE of 6.9661, an R^2 of 0.3586, an MAE of 1.7353, and a Correlation Coefficient of 0.6479. Although it offers valuable multi-model forecasts, its larger errors and lower correlation demonstrate its limitations in seasonal precipitation prediction, particularly when compared to the deep learning models.

Overall, while the CNN 1D model provides the best overall predictive performance, the integration of spatial components, like those in GConvLSTM, could further improve precipitation predictions.

5. Conclusion

In data-driven approaches, such as deep learning, it is challenging for a single solution to perform optimally across all domains. This study evaluated various DL models for seasonal precipitation forecasting in South America, identifying their strengths and weaknesses. The CNN 1D model showed the best predictive performance, although with some loss of spatial detail, while the GConvLSTM model demonstrated promise in capturing geographic precipitation structures, highlighting the potential of graph-based neural networks for meteorological predictions.

In contrast, the NMME model, despite offering valuable multi-model forecasts, displayed weaker predictive performance than the DL models, with higher MSE, lower R^2, and less favorable correlation. This underscores the limitations of traditional model ensembles in seasonal precipitation forecasting when compared to the spatially aware DL models.

This research improves seasonal precipitation forecasting, supporting water allocation, disaster preparedness, and conservation. It enhances decision-making for environmental management and natural resources. The findings contribute to more reliable climate predictions.

Future work will involve using the Optuna hyperparameter optimization framework to improve model performance and extending the evaluation to other seasons, , and incorporating SHAP values to enhance the interpretability of the models. We also plan to explore graph-based neural networks further, aiming to enhance both accuracy and spatial representation of precipitation forecasts. Additionally, combining dynamic modeling with DL will be investigated to assess its feasibility for precipitation forecasting, potentially advancing reliable seasonal modeling and benefiting weather-sensitive sectors.

6. Acknowledgements

This research was developed within the project *Classificação de imagens e dados via redes neurais profundas para múltiplos domínios* (Image and data classification via Deep neural networks for multiple domains - IDeepS). The IDeepS project (available online: `https://github.com/vsantjr/IDeepS`, accessed on 25 February 2025) is supported by the Laboratório Nacional de Computação Científica (LNCC - National Laboratory for Scientific Computing, MCTI, Brazil) via resources of the SDumont supercomputer. This research was carried out within the scope of the Laboratório de Inteligência ARtificial para Aplicações AeroEspaciais e Ambientais (Artificial Intelligence Laboratory for Aerospace and Environmental Applications - LIAREA) (available online: `https://liarealab.github.io/liarea_page/index.html`, accessed on 25 February 2025). This research was also supported by CAPES (Finance Code 88887.951221/2024-00).

References

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*, 13(13).

Anochi, J. A. and Shimizu, M. H. (2024). Precipitation forecasting and drought monitoring in south america using a machine learning approach. *Meteorology*, 4(1):1.

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*.

CPTEC (1988). Eta model documentation. `http://etamodel.cptec.inpe.br/documentation/`. Accessed March 6, 2025.

DataCamp (2023). Machine learning and deep learning tutorial. Accessed on 06-Mar-2025.

Hersbach, H. et al. (2023). Era5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Data set, available online.

Holton, J. R. and Hakim, G. J. (2013). *An introduction to dynamic meteorology*, volume 88. Academic press.

INSTITUTO NACIONAL DE METEOROLOGIA (2024). Prognóstico climático de outono: Nota técnica conjunta inmet/inpe. INMET; INPE, Brasília, DF; São José dos Campos, SP. Disponível em: `https://portal.inmet.gov.br/notasTecnicas\#`. Accessed on: 20 Jan 2025.

Kiranyaz, S., Avcı, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2019). 1d convolutional neural networks and applications: A survey. *ArXiv*, abs/1905.03554.

Kirtman, B. P., Min, D., Infanti, J. M., Kinter III, J. L., Paolino, D. A., Zhang, Q., Van Den Dool, H., Saha, S., Mendez, M. P., Becker, E., et al. (2014). The north american multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, 95(4):585–601.

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):243.

Reboita, M. S., Dias, C. G., Dutra, L. M. M., Rocha, R. P. d., and Llopart, M. (2018). Previsão climática sazonal para o brasil obtida através de modelos climáticos globais e regional. *Revista Brasileira de Meteorologia*, 33:207–224.

Seo, Y., Defferrard, M., Vandergheynst, P., and Bresson, X. (2016). Structured sequence modeling with graph convolutional recurrent networks.

Service, C. C. C. (2023). Era5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed on 06-Mar-2025.

Staudemeyer, R. C. and Morris, E. R. (2019). Understanding lstm–a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

Torres, F. L. R., Kuki, C. A. C., Reboita, M. S., Lima, L. M. M., Lima, J. W. M., and Queiroz, A. R. d. (2024). Refining seasonal precipitation forecast in brazil using simple data-driven techniques and climate indices. *Revista Brasileira de Meteorologia*, 39:e39240052.

Yang, J., Xiang, Y., Sun, J., and Xu, X. (2022). Multi-model ensemble prediction of summer precipitation in china based on machine learning algorithms. *Atmosphere*, 13(9).