Revista Brasileira de Geografia Física

ISSN:1984-2295

Homepage: https://periodicos.ufpe.br/revistas/rbgfe

Mapping the Effects of Climate Change on Reference Evapotranspiration in Future Scenarios in the Brazilian Semi-arid Region - South America

Lucas Augusto Pereira Silva[1], Cristiano Marcelo Pereira de Souza[2], Claudionor Ribeiro Silva[3],Roberto Filgueiras[4], João Paulo Sena-Souza[5], Elpídio Inácio Fernandes-Filho[6], Marcos Esdras Leite[7]

[1]Universidade Federal de Uberlândia – UFU. Doutorando em Geografia. E-mail: lucaskaio1605@gmail.com (correspondente).
[2]Universidade Estadual de Montes Claros - UNIMONTES. E-mail: cmpsgeografia@gmail.com
[3]Universidade Federal de Uberlândia – UFU, Professor do Programa de Pós-graduação em Geografia. E-mail: crs@ufu.br
[4]Pós-doutorando em Engenharia Agrícola - Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba, CODEVASF. E-mail: roberto.filgueiras@ufv.br
[5]Professor do Programa de Pós-graduação em Geografia - UNIMOTES. E-mail: jpsenasouza@gmail.com
[6]Professor da Universidade Federal de Viçosa. E-mail: elpidio@ufv.br
[7]Professor do Programa de Pós-graduação em Geografia - UNIMOTES. E-mail: marcosesdrasleite@gmail.com

Artigo recebido em 15/10/2022 e aceito em 13/03/2023

ABSTRACT

Brazil has the world's most populous semi-arid region and climate change represents significant ecological and socioeconomic challenges for this area. To better understand the impact of these changes, it is crucial to analyze the dynamics of climate variables and evapotranspiration (ETo), a critical climate variable. This study aimed to model ETo rates considering climate change scenarios in the Brazilian Semi-arid region (BSR). The modeling was based on tests of five machine learning algorithms: Bayesian Regularized Neural Networks (BRNN), Cubist, Earth, Linear Regression (LM), and Random Forest (RF). A dataset with 20 covariates was used to represent the current scenario. In the future prediction, covariates from two shared socio-economic pathways were used (SSPs 126 and 585). The best statistical performance was achieved by Cubist ($R^2 = 0.98$ and RMSE = 0.08 mm day^{-1} in the holdout-test). The current daily average ETo is 4.77 mm day^{-1}, while in future scenarios, it can increase by 3.56% in SSP 126 and 15.51% in SSP 585. ETo rates are expected to expand territorially; ranges from > 0.60 mm day^{-1} should increase 8% in SSP 126 and 40% in SSP 585. The applied model suggests that ETo may increase in future scenarios in the BSR, which could affect biodiversity levels and intensify social conflicts.

Keywords: Machine Learning; Spatial Prediction; Cubist; Semi-arid Zone; Climate changes.

Mapeamento dos Efeitos das Mudanças Climáticas na Evapotranspiração de Referência em Cenários Futuros no Semiárido Brasileiro - América do Sul

RESUMO

O Brasil possui a região semiárida mais populosa do mundo e as mudanças climáticas impõem desafios ecológicos e socioeconômicos significativos para esta área. Para entender melhor o impacto dessas mudanças, é crucial analisar a dinâmica das variáveis climáticas e a evapotranspiração (ETo), uma importante variável do clima. Este estudo teve como objetivo modelar taxas de ETo considerando cenários de mudanças climáticas no Semiárido Brasileiro (BSR). A modelagem foi baseada em testes de cinco algoritmos de aprendizado de máquina: Bayesian Regularized Neural Networks (BRNN), Cubist, Earth, Linear Regression (LM) e Random Forest (RF). Um conjunto de dados com 20 covariáveis foi usado para representar o cenário atual. Na predição futura, foram usadas covariáveis de dados dos Caminhos Socioeconômicos Compartilhados (SSPs 126 e 585). Na modelagem, o melhor desempenho estatístico foi com modelo Cubist ($R^2 = 0{,}98$ e RMSE = 0,08 mm dia^{-1} no holdout-test). A média diária atual de ETo é de 4,77 mm dia^{-1}, enquanto em cenários futuros a média pode aumentar 3,56% no SSP 126 e 15,51% no SSP 585. As taxas de ETo devem se expandir territorialmente; intervalos de > 0,60 mm dia^{-1} devem expandir 8% no SSP 126 e 40% no SSP 585. O modelo aplicado sugere que a ETo pode aumentar em cenários futuros na BSR, o que pode afetar os níveis de biodiversidade e intensificar os conflitos sociais.

Palavras-Chave: Aprendizado de máquina; Predição espacial; Cubist; Zona Semiárida; Mudanças climáticas.

Introduction

Anthropogenic activities have boosted greenhouse gas emissions (CO_2, CH_4, N_2O, etc.), leading to global warming. The Intergovernmental Panel on Climate Change (IPCC) has been projecting several climate change scenarios for future generations, pointing out the influence of human actions on climate dynamics (Kim et al., 2013). These scenarios are called Shared Socioeconomic Pathways (SSP). The SSP is a set of concentrations and emissions of greenhouse gases modeled for the future to assist research on the impacts and effects of public policies on climate change (van Vuuren et al., 2017). The SSPs are divided into five scenarios (i.e., 1.9, 126, 245, 370, and 585), which vary considering the intensity of climate change (Pielke Jr et al., 2022).

Increasing temperatures can cause climate changes in different regions worldwide, potentially transforming semi-arid regions into fully arid ones (Chen et al., 2017). This intensification directly affects the dynamics of components of the hydrological cycle, such as wind speed, air humidity, vapor pressure deficit, and evapotranspiration (Jiang et al., 2021). Notably, reference evapotranspiration (ETo) is a crucial component of the hydrological cycle, resulting from a complex interaction of several climatological variables.

In recent decades, ETo has followed an increasing trend in various parts globally (Fan et al., 2016; Li et al., 2022). This uncontrolled increase in ETo can lead to excessive moisture loss in the soil, intensifying droughts (Yoo et al., 2016; Shi et al., 2020; Nooni et al., 2021; Zhang et al., 2022). The increase in ETo may be more dramatic for semi-arid regions of the world, which are already affected by a water deficit, and support 14% of the world population (Huang et al., 2016).

The most densely populated semi-arid zone is located in South America, specifically in the Brazilian territory, with 28 million inhabitants (Costa et al., 2021). The region is called Brazilian semi-arid (BSR) and strongly depends on water resources for irrigated agricultural activity. Ecologically, it sustains an exclusively Brazilian biome (Caatinga biome). In the climate change context, the Caatinga biome can increase and expand over other biomes, resulting in a loss of biodiversity and affecting the dynamics of ecosystem services to maintain the economy of the BSR (Castro Oliveira et al., 2021). Therefore, monitoring and establishing scenarios for response variables to climate change, such as ETo, is an essential demand for the BSR to assist in formulating guidelines for management and planning.

There is an ongoing effort in ETo modeling for large areas, such as Brazil. Xavier et al. (2016) applied ETo prediction by interpolation and supported several climate studies nationally (Dias and Sentelhas, 2021; Santos et al., 2021). However, interpolation can be problematic in studies of large areas due to spatial dependence requirements (Oliver and Webster, 1990), and variables strictly related to ETo can be neglected. Therefore, more robust methods based on machine learning (ML) algorithms and covariate data sets have been evaluated in Brazil, which aligns with other studies worldwide (Althoff et al., 2020; Del Cerro et al., 2021; Wu et al., 2021). Notably, the insertion of covariates represents a significant advantage in predicting ETo; for example, climate data from the Worldclim dataset has been widely used in studies of this scope (Dias et al., 2021). In Brazil, several factors create an opportunity for studying ETo in the BSR under future scenarios: (i) previous studies with ETo (current scenario); (ii) availability of robust methods of prediction; (iii) freely available spatial environmental data (covariates). This study aimed to establish scenarios for ETo, usingmachine learning algorithms and a data set of covariates, to understand the spatial-temporaldynamics of ETo in the climate change context.

Material and methods

Study area

The Brazilian semi-arid is located in the tropical zone between latitude 0° and 20° S and represents 13% of the Brazilian territory (Figure 1a). It comprises states in the Northeast region and part of the north of Minas Gerais State. In total, 1,262 municipalities are included in the semi-arid zone, which comprises ~28 million inhabitants. All municipalities included in the semi-arid zone followed specific climatic criteria, such as mean annual precipitation ≤ 800 mm, Aridity Index ≤ 0.5, and daily percentage of water deficit ≥ 60% considering all days of the year (Castro Oliveira et al., 2021).

The semi-arid region is characterized by an irregular rainfall distribution throughout the year (Cunha et al., 2015). Rainfall ranges from 390 mm year^{-1} to 1990 mm year^{-1}, with an average of 840 mm ± 230 mm (1970 - 2000 period) (Figure 1b). In addition, the region is influenced by elements that drive its climatic dynamics, such as cold fronts originating in southern Brazil, the Intertropical Convergence Zone (ITCZ), and the amplitudes between the ocean and coastal temperatures

(Cunha et al., 2015). This climatic context influences the composition of the vegetation of the semi-arid region, containing mainly vast and continuous Tropical Dry Forests.

The BSR has vegetation from the Caatinga (predominantly), Cerrado, and Atlantic Forest biomes. These three environments are sometimes in contact, forming ecotone zones (~ 22% of the territory) (Castro Oliveira et al., 2021). This condition was given in past contexts of climate change, mainly due to the displacement of theITCZ towards the south of the Brazilian Northeastregion during the Heinrich Stadials in the northern hemisphere, which increased precipitation (Wendt et al., 2019), favoring the expansion of areas from the Cerrado to the semi-arid. Therefore, we emphasize the importance of studies that analyze the influence of future climate change scenarios in the BSR.

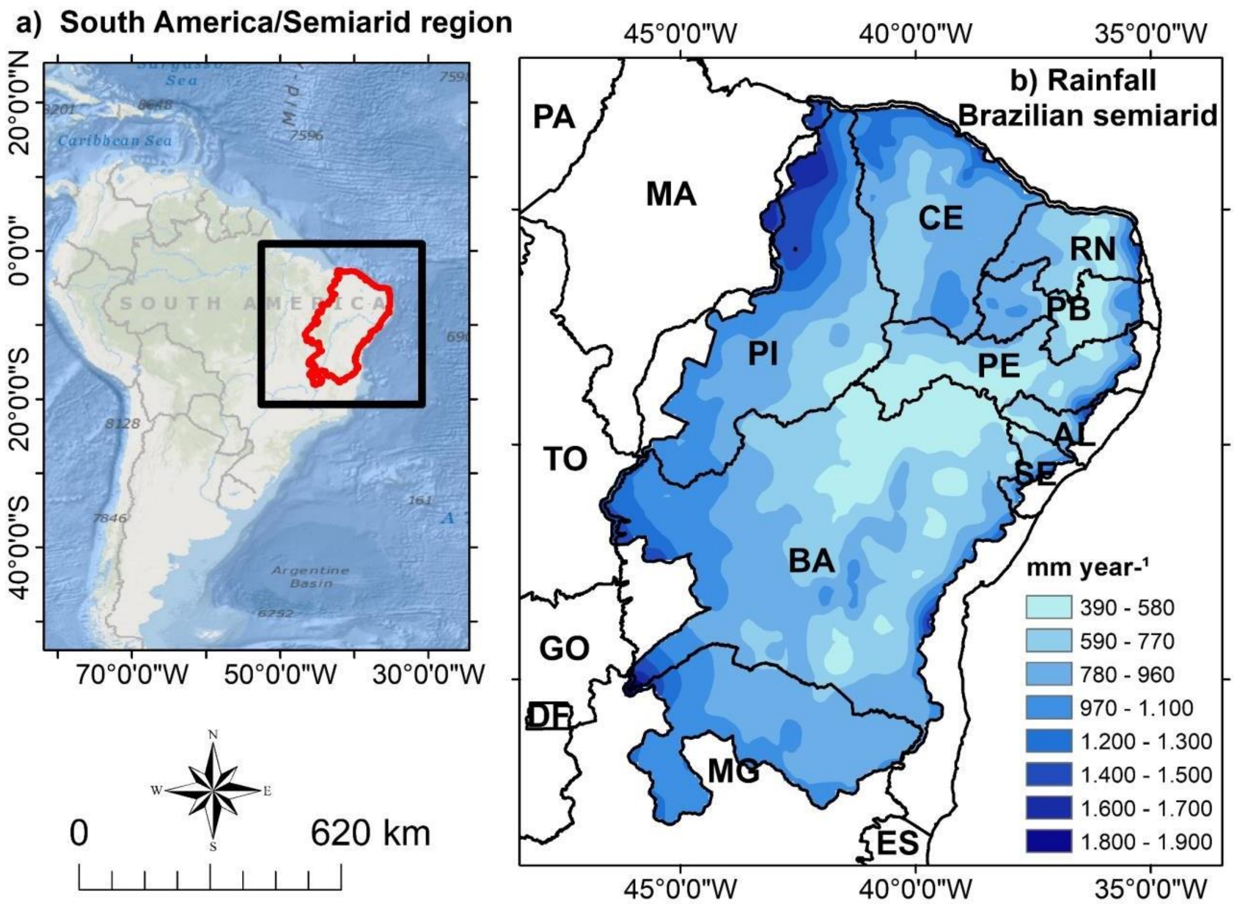

Figure 1 - A) Location of the Semi-arid region in the context of South America; B) Spatial Distribution of Rainfall in the Semi-arid Region (1970 – 2000). Source: (Hijmans et al., 2005).

Database

Due to the region's low coverage of weather stations, alternative methodologies are needed for modeling climate variables. We obtained the ETo variable from the data set ETo-Brazil (Althoff et al., 2020), which was prepared for all of Brazil from a dense network of meteorological stations using a model of machine learning (Cubist). This dataset covers Brazil with ETo values showing low error rates (RMSE = 0.65 mm day-¹) (available at: data.mendeley.com/datasets/sstjw74ryh/1). To cover the temporal range satisfactorily, we obtained 6.827 ETo products daily between 2000 and 2020.

We organized our methodological structure in two stages, i) training the models for the current scenario (1970 - 2000) and ii) spatial prediction for future scenarios (2061 - 2080). In both stages, we used bioclimatic (19) and topographic (1) covariates from Worldclim 2.1 (Fick and Hijmans, 2017) (Table 1). We selected these variables because they influence the spatial behavior of ETo (Althoff et al., 2020; Liu et al., 2021).

Table 1 - Covariates used for training the Machine Learning models and ETo spatial prediction in current (1970 - 2000) and future (2061 - 2080) scenarios.

Abbreviation	Covariates of current and future scenarios	Abbreviation	Covariates of current and future scenarios
Bio 01	Mean annual temperature	Bio 11	Mean temperature of coldest quarter
Bio 02	Mean diurnal range	Bio 12	Annual precipitation
Bio 03	Isothermality	Bio 13	Precipitation of wettest month
Bio 04	Temperature seasonality	Bio 14	Precipitation of driest month
Bio 05	Maximum temperature of warmest month	Bio 15	Precipitation seasonality
Bio 06	Minimum temperature of coldest month	Bio 16	Precipitation of wettest quarter
Bio 07	Temperature annual range	Bio 17	Precipitation of driest quarter
Bio 08	Mean temperature of wettest quarter	Bio 18	Precipitation of warmest quarter
Bio 09	Mean temperature of driest quarter	Bio 19	Precipitation of coldest quarter
Bio 10	Mean temperature of warmest quarter	SRTM	Altitude

Source: Hijmans et al.(2005).

The bioclimatic covariates for the current scenario were elaborated by interpolating global meteorological stations (Fick and Hijmans, 2017). The altitude is interpolation to 10 km from the Shuttle Radar Topography Mission - SRTM digital elevation model. For the future context, weselected covariates considering global climate change scenarios from the Intercompared Project 6Coupled Model (CMIP6), specifically for The Shared Socioconomic Pathways (SSPs) 126 and 585, as they are two climatic extremes and socioeconomic. In the SSP 126 scenario, a substantial decline in carbon emissions is expected by 2050, with temperature stability at 1.8°C, considered an optimistic scenario. In SSP 585, considered a pessimistic scenario, the expected behavior is opposite to SSP 126, where CO_2 emission levels tend to rise until 2050 and the average temperature will increase by 4.4°C (PielkeJr et al., 2022).

The CMIP6 project provides a wide range of global climate models (GCM) with future simulations. However, there are uncertainties in these models, especially considering the sensitivity to climate change. For this study, following recommendations from the literature (Hausfather et al., 2022), we use the average of the five GCMs with the lowest uncertainty in the projections, i.e., INM-CM4-8, INM-CM5-0, MIROC6, GISS-E2-1-H, MIROC-ES2L (available at www.worldclim.org/data/cmip6/cmip6climate.html).

Sampling, pre-processing and tested models

We calculated a daily average image from all 6,827 ETo images. We generated 2056 random points with a minimum spacing of 10 km to be compatible with the ETo image, covering ~20% of the study area. Finally, the values of the dependent and independent variables were extracted for each point to create the regression matrix.

Prioritizing the parsimony, we performed a correlation matrix (with the Spearman method) from the *findcorrelation* tool. We established a limit of 0.95 for correlations between pairs of covariates, that is, pairs with Spearman coefficient values greater than 0.95 were removed from the database. We use the *caret* R package for this process (Kuhn et al., 2020).

We selected five machine learning models to predict the effects of climate change on ETo: Bayesian Regularized Neural Networks (BRNN) (Rodriguez and Gianola, 2016), Cubist (Kuhn and Quinlan, 2018), Earth (Milborrow and Tibshirani, 2019), linear regression (LM) (Kuhn et al., 2020), and Random Forest (RF) (Breiman, 2001). These models are recurrently used in spatial predictions of evapotranspiration in Brazil (Althoff et al., 2020; Dias et al., 2021).

Calibration, training and validation

We extracted the values from thecovariates of the current scenario and altitude to train the models. The data set was randomly divided into 75% for training and 25% for holdout-test. We applied 10-fold Cross-Validation with 5 repetitions to obtain the training metrics and optimize the hyperparameters of each model (Kuhnand Johnson, 2013). Each training and holdout-test was performed 10 times to reduce prediction uncertainties. After the training process, we applied the *varImp* function to obtain the covariates that best explained the spatial distribution of ETo. The function calculates the percentage increase in the average standard error of each covariate (Kuhn et al., 2020). Sequentially, we performed an exploratory analysis between the most explanatory covariates and the target variable (i.e., ETo). For this, we applied simple linear regressions to observe the trend (slope of the line) and the significance (*p-value* < 0.001) of each covariate concerning the behavior of ETo.

The model's performance was by external validation (holdout-test). The validations provided RMSE (Root Mean Squared Error) and R^2 (R-squared) as statistical metrics for comparing the models. For the final prediction, we selected the model with the best metrics in the holdout-test (25%). The same procedures were applied for predictions of future ETo scenarios, however, replacing current covariates with covariates of future scenarios. The SRTM covariate was inserted in all modeling (present and future).

Comparative Analysis between Spatial Prediction of Scenarios

From the spatial modeling of the prediction of climate change scenarios for the BSR, we performed a comparative analysis to observe the areas of ETo increases and decreases in the region. We apply Equation 01:

$$\mathbf{\Delta ETo = ETo_{\mathit{FS}} - ETo_{\mathit{CS}}}$$

Where ΔETo represents the difference between the scenarios; ETo_{FS} is the future climate change scenarios (SSPs 126 and 585); ETo_{CS} is the current scenario of ETo (1970 – 2000). We obtained two ΔETo maps to observe the spatial variations of ETo considering climate change scenarios.

Results

Models performance for ETo prediction

Among the 20 covariates selected topredict ETo in the semi-arid region, four wereremoved (Bio2: Mean diurnal range; Bio11: Mean temperature of coldest quarter; Bio16:Precipitation of wettest quarter; Bio17: Precipitation of driest quarter), as they had a coefficient of Spearman above 0.95. Therefore, wetrained five ML models with 16 explanatory covariates.

In the model performance phase (holdout-test), Cubist and RF showed better performance for ETo spatialization with low errors (0.07 and 0.08 mm day^{1}) and higher R^2 (0.98 and 0.97) (Figure 2). The BRNN showed an intermediate behavior (R^2 = 0.95 and RMSE = 0.11 mm day^{-1}). Earth and LM models presented the worst performances. We considered Cubist to predict ETo spatially as it was slightly superior to RF.

From the learning process of the Cubist model, we obtained a ranking of importance (Overall %) among the 16 covariates selected for the prediction of ETo: Annual Precipitation (Bio12), Temperature Annual Range (Bio7), Annual Mean Temperature (Bio1), Temperature Seasonality (Bio4) and Precipitation Seasonality (Bio15) (Figure 3).

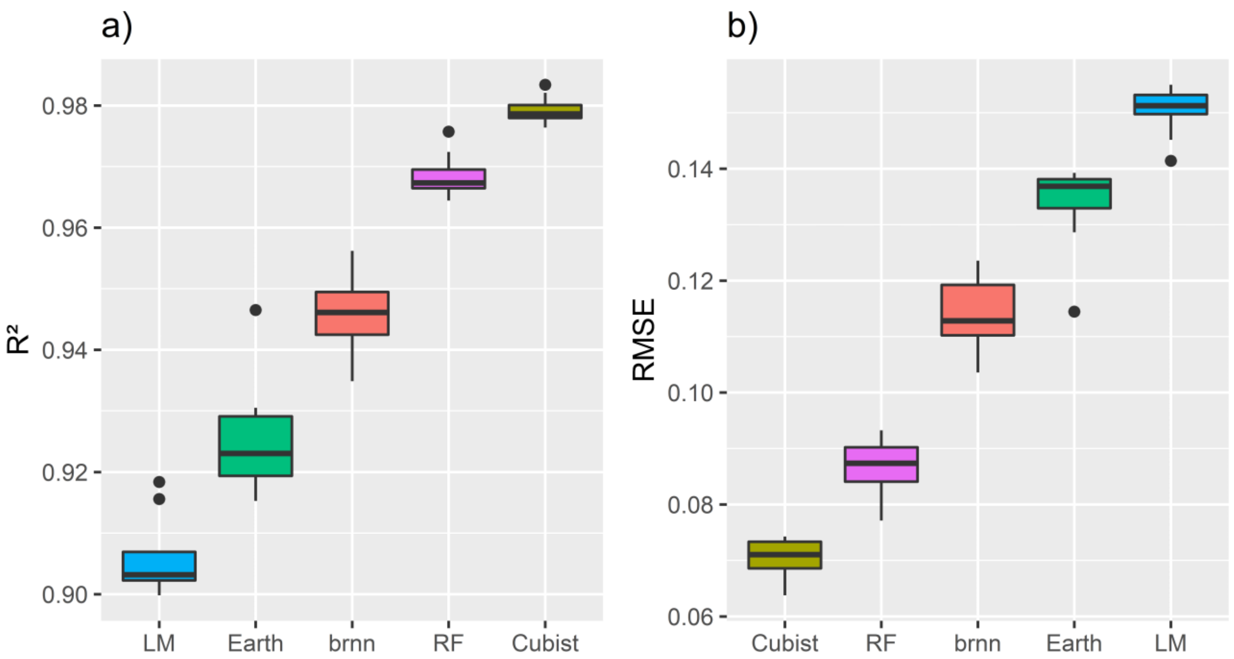

Figure 2 - Holdout-test for BRNN (Bayesian regularized neural networks) Cubist, Earth, LM (linear regression), and RF (Random Forest). Figure 2a) Boxplot of the R^2 (R-Squared) of the models and Figure 2b) Boxplot of the RMSE (Root Mean Squared Error) for the models.

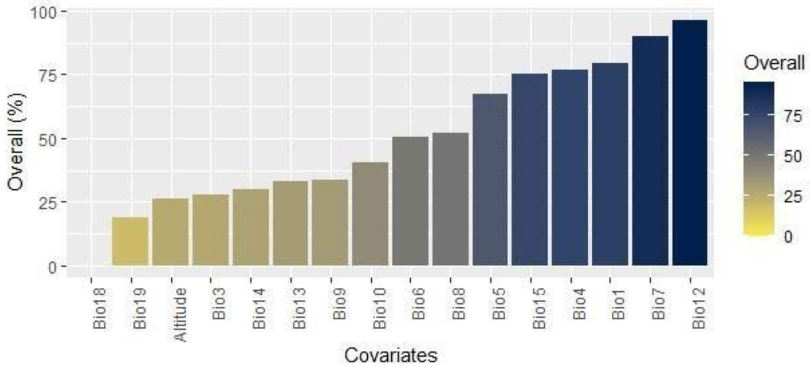

Figure 3 - Importance Ranking (Overall %) for covariates of Cubist model. Bio1: Mean Annual Temperature; Bioc3: Isothermality; Bio4: Temperature Seasonality; Bio5: Maximum Temperature of the Warmest Month; Bio6: Minimum Temperature of the Coldest Month; Bio7: Temperature Annual Range; Bio8: Mean Temperature of Wettest Quarter; Bio9: Mean Temperature of Driest Quarter; Bioc12: Annual Precipitation; Bioc13: Precipitation of Wettest Month; Bioc14: Precipitation of Driest Month; Bioc15: Precipitation Seasonality; Bio18: Precipitation of Warmest Quarter; Bio19: Precipitation of Coldest Quarter.

The main explanatory covariates selected in the ranking of importance showed significant effects (*p-value* <0.001) concerning the behavior of ETo (Figure 4). ETo has an inverse relationship with Annual Precipitation (Figure 4a). The Mean Annual Temperature presented a positive relationship with ETo rates. For the Annual Temperature Range there was no significant correlation. Temperature seasonality is more intense in regions with low ETo, while the seasonal intensity decreases in areas with high vapor flow ascendancy. While seasonality and variation in rainfall distribution intensify ETo rates.

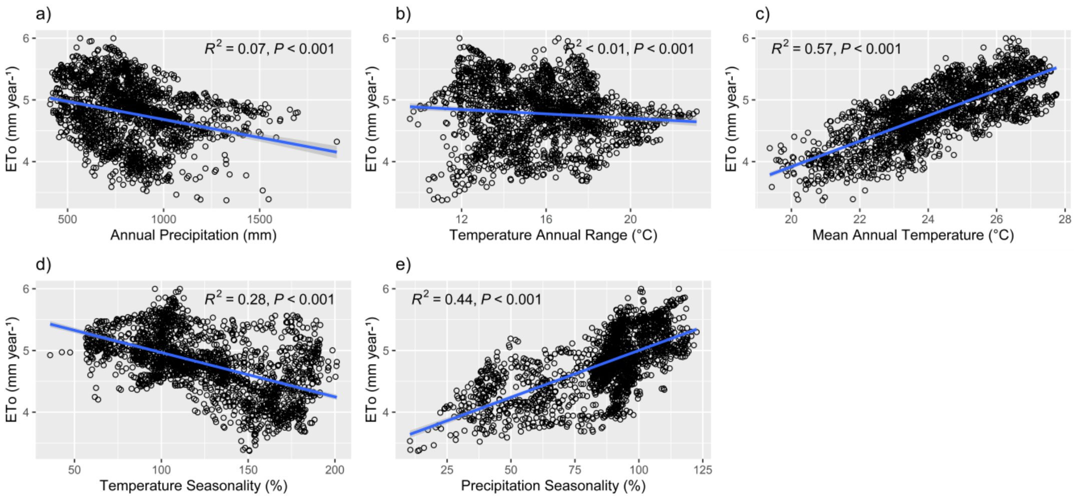

Figure 4 - Influence of the most important covariates to explain ETo. A simple linear regression was applied to the five covariates vs ETo, we obtained the R-squared (R^2) p-value and slope. Slope for regressions: ETo vs Annual Precipitation = -0.0005; ETo vs Temperature Annual Range = -0.02; ETo vs Mean Annual Temperature = 0.21; ETo vs Temperature Seasonality = -0.007; ETo vs Precipitation Seasonality = 0.02.

Spatial Distribution of ETo - Current (1970 - 2000) and Future Scenarios (2061- 2080)

Under two future climate change scenarios ETo in the BSR will increase (Figure 5). The mean ETo in the current scenario is 4.77 mm day^{-1}. However, it shows an increase of +0.17 mm day^{-1} in the SSP 126 and +0.51 mm day^{-1} on the SSP 585. These changes are also manifested through the configuration of spatial patterns. In the current scenario, the maximum value (6 mm day^{-1}) comprises 1.09% of the region, while in the future scenarios (SSPs 126 and 585) represent 11.36%

and 25.74%, respectively. ETo values > 5.71 mm day^{-1} that predominate in the northern and central portion of the BSR should expand to the southern part, reaching latitudes of 15° S in future scenarios. We also noticed that the minimum values (< 4.2 mm day^{-1}) in the current scenario were more expressive (13.61% of the territory), decreasing in the future scenarios (SSP 126 = 11%; SSP 585 = 2.85%).

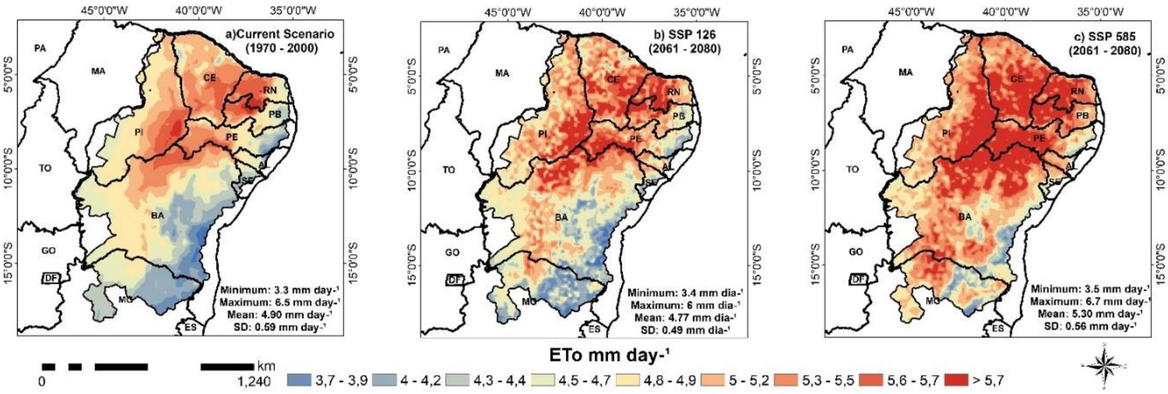

Figure 5 - Spatial Distribution of ETo (mean) in the Brazilian Semi-arid Region for the Current and Climate Change Scenario. Figure a) ETo Prediction for the Current Scenario (1970 – 2000); Figure b) ETo Prediction for the SSP 126 (2061-2080); Figure c) Spatial Prediction for the SSP 585 (2061-2080).

From the spatial patterns, the most extreme changes in ETo may occur in the worst climate change scenario, SSP 585. Considering the ΔETo in SSP 126, by 2070 about 8% of the study area may experience an increase between 0.60 and 1 mm day^{-1}. In SSP 585, ~40% of the semi-arid willincrease ETo rates between 0.60 and 1.8 mm day^{-1}, reaching ~650 mm-year^{1} in some areas.

The changes in ETo in both scenarios reflect an increase in air temperature, an abrupt decrease in precipitation, and climatic seasonalityvariations compared to the present context (Table 2). The mean temperature in the current scenario is24.16°C and will increase by 1.13 °C and 2.73°C in SSPs 126 and 585, respectively. The Annual Temperature Range also shows a boost in the face of climate change of 0.72 °C and 1.28 °C comparedto the current mean (15.74 °C). While, adversely, the rainfall input will decrease until 2070 in the order of 25.09 mm and 50.13 mm in SSPs 126 and585.

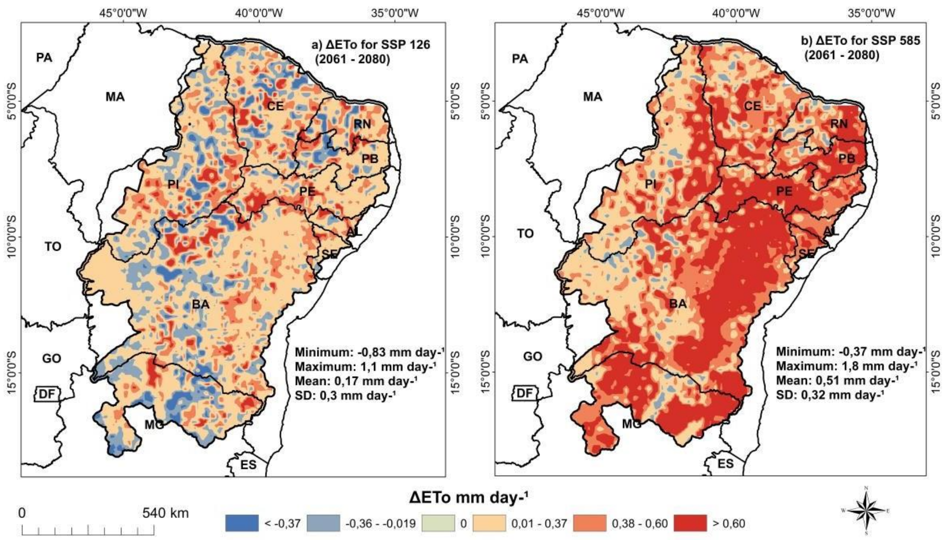

Figure 6 – ETo difference maps between scenarios (ΔETo = ETo_{FS} – Eto_{CS}). a) ETo difference for the SSP 126 scenario compared to the current one; b) ETo difference for the SSP 585 scenario compared to the current one. Statistical summary with maximum, minimum, mean, and standard deviation (SD).

Table 2 - Mean values of the most important covariates to explain ETo in the current and future scenarios (SSPs 126 and 585). The table also shows the difference between the covariates in the future and current scenarios.

Covariates	Current	SSP 126	SSP 585	Difference SSP 126 to Current	Difference SSP 585 to Current
Annual Precipitation (mm)	839.10	814.01	788.97	-25.09	-50.13
Temperature Annual Range (°C)	15.74	16.46	17.02	+0.72	+1.28
Mean Annual Temperature (°C)	24.16	25.29	26.89	+1.13	+2.73
Temperature Seasonality (%)	126.27	134.67	143.94	+8.4	+17.67
Precipitation Seasonality (%)	85.54	86.93	87.22	+1.39	+1.68

Discussion

Model performance and uncertainty

Considering the statistical analyzes and our dataset, Cubist was the most suitable model for spatially predicting ETo variations in the Brazilian semi-arid region (Figure 2). Compared to our study, works highlighted the high capacity of Cubist in predicting ETo for Brazil (Althoff et al., 2020; Dias et al., 2021). Cubist is a model based ontree logic (as well as RF); models of this nature have shown superiority over others inenvironmental predictions (Gomes et al., 2019;Althoff et al., 2020; Souza et al., 2022). In this model, the predictions are performed in sequence; if the first prediction has low metrics, it is adjustedand improved in the second, and so on (Kuhn e Quinlan, 2018). The RF model presented the second-best performance, and several studies havepointed out this pattern for ETo prediction in regions of Brazil (Dias et al., 2021). Due to the proximity of the RF performance to the Cubist, studies have proposed formulating hybrid models (Houborg and McCabe, 2018). The BRNN had an intermediate performance, a similar behavior to studies of Dias et al. (2021) for Brazil. However, when used without comparing to other models, it has shown a potential to predict ETo (Althoff et al.,2018; Silva et al., 2020). The two worst models (Earth and LM) are based on linear regression, andgenerally, due to their simplicity, they cannot capture ETo variations, as it is a complex and non- linear phenomenon (Althoff et al., 2020).

Our results suggest that for analysis of complex phenomena such as ETo, more robust models such as Cubist should be used. However, itis important to note that a model's efficacy may be influenced by its robustness, the complexity of the problem, and the quality of the data set utilized.

Influence of Climate Change Scenarios on ETo

Our results showed that in the BSR, under climate change scenarios (SSPs 126 and 585), ETo will increase on average (0.17 and 0.51 mm year-¹). Under these conditions, some areas may experience water loss of 600 mm per year through increased ETo rates. Spatial change of ETo rates in the BSR is also expected with climate changes. Maximum values of the ETo (6 mm year-¹) expand towards the southern portion of the BSR (Figure 6). Corroborating our study, the increase in ETo rates in climate change scenarios has been recurrent in several semi-arid zones worldwide. The intensification of water loss has expanded spatially, leading to aridity conditions (Nooni et al., 2021; Shi et al., 2020; Yoo et al., 2016; Zhang et al., 2022).

The main drivers that control the behavior of ETo in the BSR are Annual Precipitation, Air Temperature, and Climatic Seasonality (Figures 3 and 4). Annual Precipitation presented a negative relationship with ETo (Figure 4a), probably because areas with a higher concentration of clouds may inhibit radiation incidence to the surface (Althoff et al., 2020). In future scenarios, precipitation will decrease by 3 to 6% (Table 2), which may decrease the cloud cover, leading to more significant radiation inputs and increasing water loss.

In the context of climate variability, we found that temperature seasonality has a negative relationship with ETo (Figure 4d). In regions with more significant thermal variation in the year, ETo decreases, while portions with temperature stability and homogeneity positively influence water loss. Although the seasonality of precipitation positively impacts ETo, regions with high precipitation variability experience increased evapotranspiration rates (Figure 4e).

The intensification of ETo translates into water deficiency, which has practical implications in the ecosystem, economic and social aspects.There is a direct link between water availability and ecosystem conditioning (Gao et al., 2022). For example, vegetation quickly responds to variations in soil water content, especially considering its metabolic activities (Xing et al., 2022). With an abrupt decline in water, the vegetation loses levels of primary productivity, directly affecting the carbon fixation capacity (Kang et al., 2022). As vegetation decreases carbon storage rates, it becomes an emitting source. In a global context, semi-arid zones are already becoming a source of carbon emissions, being susceptible regions concerning climate change (Scott et al., 2015).

In addition, the amplification of water deficit levels due to the increase in ETo can change the patterns of the spatial distribution of vegetation (Burrell et al., 2020; Castro Oliveira et al., 2021; Grünzweig et al., 2022). Modeling of future scenarios has shown that by the end of the century, vegetation with xerophytic characteristics will extend into the BSR (Castro Oliveira et al., 2021), mainly on the southern edge (latitude 15° S), where maximum ETo values may extend (Figures 5 and 6).

The expansion of maximum ETo values (6 mm day-¹) may increase the frequency andintensity of droughts in the BSR, which historically is a phenomenon that has affected the socioeconomic context of the region (Marengo et al., 2022). In several parts of the globe, the intensification of droughts is one of the main concerns in the agricultural production sector (Salas-Martínez et al., 2021; Orimoloye et al., 2022). Specifically for semi-arid regions, significant crop losses have been observed (Núñez-López et al., 2022). Therefore, increases in ETo rates may jeopardize the productive capacity of the agricultural sector of the BSR, affecting the economic dynamics of several municipalities.

Furthermore, water shortages in future scenarios can amplify the context of social conflicts, as observed in several parts of the world (Leal Filho et al., 2022). In the BSR, conflicts are due to the multiple uses of water, mainly due to the appropriation of large volumes of water for the production of large commodities in the agro-industrial sector, as well as the formulation ofpublic policies aimed at large landowners (Silveira and Silva, 2019). Surveys indicate that in a five- year interval (2012 - 2017), there was a growth of +40% in conflicts involving ~13 thousand families in the semi-arid region (Silveira and Silva, 2019).

Therefore, conflicts over water use may intensify in future water loss scenarios due to increased evapotranspiration rates.

Conclusions

The modeling of the reference evapotranspiration (ETo) dynamics was evaluated for the Brazilian semi-arid region in the current (1970 - 2000) and future scenarios (until 2070). The adopted methodological structure indicated high performance for ETo prediction, with the cubist model reaching 98% of the explanation of this phenomenon.

We found that due to the increase in air temperature, decrease in precipitation, and climatic seasonality, ETo will increase under climatechange scenarios in the region (SSP 126 = 3.56%, SSP 585 = 15.51%). Considering the spatial configurations, it has been observed that higher ETo rates have the potential to expand geographically. In the most severe scenario (SSP 585), there is a notable increase in the areas with ETo rates reaching up to 600 mm yr-¹.

Given this, we draw attention to the need for decision-making guided by robust scientific methods. Furthermore, we emphasize the critical scenario for the coming decades in the Brazilian semi-arid region with the increase in evapotranspiration because if they materialize, many segments can be altered, such as biodiversity patterns, changes in the spatial structure of biomes, provision of ecosystem services and agricultural production.

Acknowledgment

The first author thanks the Research Support Foundation of the Federal University of Minas Gerais (FAPEMIG) for the doctoral scholarship. The second author thanks the Coordination for the Improvement of Higher Education Personnel (CAPES) for the postdoctoral scholarship. The last author thanks the National Council for Scientific and Technological Development (CNPq) for the productivity grant.

Data availability

https://github.com/lucasagus/Reference-evapotranspiration-data-

References

Althoff, D., Bazame, H.C., Filgueiras, R., Dias, S.H.B., 2018. Heuristic methods applied in reference evapotranspiration modeling. Ciênciae Agrotecnologia 42, 314–324.

Althoff, D., Dias, S.H.B., Filgueiras, R., Rodrigues, L.N., 2020. ETo-Brazil: A Daily Gridded Reference Evapotranspiration Data Setfor Brazil (2000–2018). Water Resources Research 56, 1–24. https://doi.org/10.1029/2020WR027562.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Burrell, A.L., Evans, J.P., De Kauwe, M.G., 2020. Anthropogenic climate change has driven over 5 million km2 of drylands towards desertification. Nat Commun 11, 3853. https://doi.org/10.1038/s41467-020-17710-7

Castro Oliveira, G., Arruda, D.M., Fernandes Filho, E.I., Veloso, G.V., Francelino, M.R., Schaefer, C.E.G.R., 2021. Soil predictors are crucial for modelling vegetation distribution and its responses to climate change. Science of The Total Environment 146680. https://doi.org/10.1016/j.scitotenv.2021.146680.

Costa, J.F.C.B., Silva, R.M., Santos, C.A.G., Silva, A.M., Vianna, P.C.G., 2021. Analysis of the response of the Epitácio Pessoa reservoir (Brazilian semi-arid region) to potential future drought, water transfer and LULC scenarios. Natural Hazards 1–25. https://dx.doi.org/10.1007/s11069-021-04736-3.

Cunha, A.P.M., Alvalá, R.C., Nobre, C.A., Carvalho, M.A., 2015. Monitoring vegetative drought dynamics in the Brazilian semi-arid region. Agricultural and forest meteorology 214, 494–505. https://doi.org/10.1016/j.agrformet.2015.09.010.

Del Cerro, R.T.G., Subathra, M.S.P., Kumar, N.M., Verrastro, S., George, S.T., 2021. Modelling the daily reference evapotranspiration in semi-arid region of South India: A case study comparing ANFIS and empirical models. Information Processing in Agriculture 8, 173–184. https://dx.doi.org/10.1016/j.inpa.2020.02.003.

Dias, H.B., Sentelhas, P.C., 2021. Assessing the performance of two gridded weather data for sugarcane crop simulations with a process-based model in Center-South Brazil. International Journal of Biometeorology 1–13. https://dx.doi.org/10.1007/s00484-021-02145-6.

Dias, S.H.B., Filgueiras, R., Fernandes Filho, E.I., Arcanjo, G.S., Silva, G.H. da, Mantovani, E.C., Cunha, F.F. da, 2021. Reference evapotranspiration of Brazil modeled with machine learning techniques and remote sensing. Plos one 16, e0245834. https://dx.doi.org/10.1371/journal.pone.0245834.

Fan, J., Wu, L., Zhang, F., Xiang, Y., Zheng, J., 2016. Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956–2015. Journal of Hydrology 542, 923–937. https://doi.org/10.1016/j.jhydrol.2016.09.060.

Fick, S.E. and R.J. Hijmans, 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37 (12): 4302-4315.

Gao, W., Zheng, C., Liu, X., Lu, Y., Chen, Y., Wei, Y., Ma, Y., 2022. NDVI-based vegetation dynamics and their responses to climate changeand human activities from 1982 to 2020: A casestudy in the Mu Us Sandy Land, China. Ecological Indicators 137, 108745.

Gomes, L.C., Faria, R.M., Souza, E., Veloso, G.V., Schaefer, C.E.G.R., Fernandes-Filho, E.I., 2019. Modelling and mapping soil organic carbon stocks in Brazil. Geoderma 340, 337–350. https://doi.org/10.1016/j.geoderma.2019.01.007

Grünzweig, J.M., Boeck, H.J., Rey, A., Santos, M.J., Adam, O., Bahn, M., Belnap, J., Deckmyn, G., Dekker, S.C., Flores, O., 2022. Dryland mechanisms could widely control ecosystem functioning in a drier and warmer world. Nature Ecology & Evolution 1–13.

Hausfather, Z., Marvel, K., Schmidt, G.A., Nielsen-Gammon, J.W., Zelinka, M., 2022. Climate simulations: Recognize the 'hot model' problem.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology: A Journal of the Royal Meteorological Society 25, 1965–1978. https://doi.org/10.1002/joc.1276.

Houborg, R., McCabe, M.F., 2018. A hybrid training approach for leaf area index estimation via Cubist and random forests machine- learning. ISPRS Journal of Photogrammetry and Remote Sensing 135, 173–188.

Huang, J., Ji, M., Xie, Y., Wang, S., He, Y., Ran, J., 2016. Global semi-arid climate change over last 60 years. Climate Dynamics 46, 1131–1150. https://doi.org/10.1007/s00382-015-2636-8.

Jiang, F., Xie, X., Liang, S., Wang, Y., Zhu, B., Zhang, X., Chen, Y., 2021. Loess Plateau evapotranspiration intensified by land surface radiative forcing associated with ecological restoration. Agricultural and Forest Meteorology 311, 108669.

Kang, Z., Qiu, B., Xiang, Z., Liu, Y., Lin, Z., Guo, W., 2022. Improving simulations of vegetation dynamics over the Tibetan Plateau: Role of atmospheric forcing data and spatial resolution. Advances in Atmospheric Sciences 39, 1115–1132.

Kim, J., Choi, J., Choi, C., Park, S., 2013. Impactsof changes in climate and land use/land cover under IPCC RCP scenarios on streamflow in theHoeya River Basin, Korea. Science of the Total Environment 452, 181–195. https://doi.org/10.1016/j.scitotenv.2013.02.005

Kuhn, M., Johnson, K., 2013. Applied predictive modeling. Springer. https://dx.doi.org/10.1007/978-1-4614-6849-3.

Kuhn, M., Quinlan, R., 2018. Cubist: Rule-and instance-based regression modeling. R package version 0.2. 2.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R.C., 2020. Package 'caret.' The R Journal 223, 7.

Leal Filho, W., Totin, E., Franke, J.A., Andrew, S.M., Abubakar, I.R., Azadi, H., Nunn, P.D., Ouweneel, B., Williams, P.A., Simpson, N.P., 2022. Understanding responses to climate-related water scarcity in Africa. Science of the Total Environment 806, 150420

Li, M., Chu, R., Sha, X., Islam, A.R.M.T., Jiang, Y., Shen, S., 2022. How Has the Recent Climate Change Affected the Spatiotemporal Variation of Reference Evapotranspiration in a Climate Transitional Zone of Eastern China? ISPRS International Journal of Geo-Information 11,300.

Liu, Y., Yao, X., Wang, Q., Yu, J., Jiang, Q., Jiang, W., Li, L., 2021. Differences in reference evapotranspiration variation and climate-driven patterns in different altitudes of the Qinghai–Tibet plateau (1961–2017). Water 13, 1749.

Marengo, J.A., Galdos, M.V., Challinor, A.,Cunha, A.P., Marin, F.R., Vianna, M. dos S., Alvala, R.C., Alves, L.M., Moraes, O.L.,Bender, F., 2022. Drought in Northeast Brazil: A review of agricultural and policy adaptation options for food security. Climate Resilience and Sustainability 1, e17.

Milborrow, S., Tibshirani, R., 2019. Package 'earth': Multivariate Adaptive Regression Splines.

Nooni, I.K., Hagan, D.F.T., Wang, G., Ullah, W., Lu, J., Li, S., Dzakpasu, M., Prempeh, N.A., Lim Kam Sian, K.T., 2021. Future Changes in Simulated Evapotranspiration across Continental Africa Based on CMIP6 CNRM-CM6. International Journal of Environmental Research and Public Health 18, 6760.

Núñez-López, J.M., Cansino-Loeza, B., Sánchez-Zarco, X.G., Ponce-Ortega, J.M., 2022. Involving resilience in assessment of the water–energy–food nexus for arid and semi-arid regions. Clean Technologies and Environmental Policy 1–13.

Oliver, M.A., Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information System, 4(3), 313-332.

Orimoloye, I.R., Belle, J.A., Orimoloye, Y.M., Olusola, A.O., Ololade, O.O., 2022. Drought: A common environmental disaster. Atmosphere13, 111.

Pielke Jr, R., Burgess, M.G., Ritchie, J., 2022. Plausible 2005–2050 emissions scenariosproject between 2° C and 3° C of warming by 2100. Environmental Research Letters 17, 024027.

Rodriguez, P.P., Gianola, D., 2016. BRNN: Bayesian regularization for feed-forward neural networks. R package version 0.6.

Salas-Martínez, F., Valdés-Rodríguez, O.A., Palacios-Wassenaar, O.M., Márquez-Grajales, A., 2021. Analysis of the Evolution of Drought through SPI and Its Relationship with the Agricultural Sector in the Central Zone of the State of Veracruz, Mexico. Agronomy 11, 2099.

Santos, T.G., Battisti, R., Casaroli, D., Alves, J., Evangelista, A.W.P., 2021. Assessment of agricultural efficiency and yield gap for soybean in the Brazilian Central Cerrado biome.Bragantia 80, 1–11. https://doi.org/10.1590/1678-4499.20200352.

Scott, R.L., Biederman, J.A., Hamerlynck, E.P., Barron-Gafford, G.A., 2015. The carbon balance pivot point of southwestern US semi- arid ecosystems: Insights from the 21st century drought. Journal of Geophysical Research: Biogeosciences 120, 2612–2624.

Shi, L., Feng, P., Wang, B., Liu, D.L., Yu, Q., 2020. Quantifying future drought change and associated uncertainty in southeastern Australia with multiple potential evapotranspiration models. Journal of Hydrology 590, 125394. https://doi.org/10.1016/j.jhydrol.2020.125394

Silveira, S.M.B., Silva, M.G., 2019. Conflitos socioambientais por água no Nordeste brasileiro: expropriações contemporâneas e lutas sociais no campo. Revista Katálysis 22, 342–352.

Souza, C.M.P., Veloso, G.V., Mello, C.R., Ribeiro, R.P., Silva, L.A.P., Leite, M.E., Fernandes Filho, E.I., 2022. Spatiotemporal prediction of rainfall erosivity by machine learning in southeastern Brazil. Geocarto International 1– 19. https://doi.org/10.1080/10106049.2022.2060318

Silva, C., Teixeira, A.C., Manzione, R., 2020.

Utilização de Redes Neurais com Regularização Bayesiana na Modelagem de Evapotranspiração de Referência em AgroecossistemasSemiáridos. Revista Brasileira de Engenharia de Biossistemas 14, 73–84.

Van Vuuren, D.P., Riahi, K., Calvin, K., Dellink, R., Emmerling, J., Fujimori, S., Kc, S., Kriegler,E., O'Neill, B., 2017. The Shared Socio- economic Pathways: Trajectories for human development and global environmental change. Global Environmental Change 42, 148–152. https://doi.org/10.1016/j.gloenvcha.2016.10.009

Wendt, K.A., Häuselmann, A.D., Fleitmann, D., Berry, A.E., Wang, X., Auler, A.S., Cheng, H., Edwards, R.L., 2019. Three-phased Heinrich Stadial 4 recorded in NE Brazil stalagmites. Earth and Planetary Science Letters 510, 94–102. https://doi.org/10.1016/j.epsl.2018.12.025.

Wu, T., Zhang, W., Jiao, X., Guo, W., Hamoud, Y.A., 2021. Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration.Computers and Electronics in Agriculture 184, 106039. https://doi.org/10.1016/j.compag.2021.1060

Xavier, A.C., King, C.W., Scanlon, B.R., 2016. Daily gridded meteorological variables in Brazil (1980–2013). International Journal of Climatology 36, 2644–2659.https://doi.org/10.1002/joc.4518.

Xing, X., Qian, J., Chen, X., Chen, C., Sun, J., Wei, S., Yimamaidi, D., Zhanar, Z., 2022. Analysis of Effects of Recent Changes in Hydrothermal Conditions on Vegetation in Central Asia. Land 11, 327.

Yoo, J., Kwon, H.-H., Lee, J.-H., Kim, T.-W., 2016. Influence of evapotranspiration on future drought risk using bivariate drought frequency curves. KSCE Journal of Civil Engineering 20, 2059–2069. https://doi.org/10.1007/s12205-015-0078-9.

Zhang, G., Gan, T.Y., Su, X., 2022. Twenty-first century drought analysis across China under climate change. Climate Dynamics 59, 1665–1685.