

DAVID BRUNO DE SOUSA TEIXEIRA

RAINFALL EROSIVITY IN BRAZIL

Thesis submitted to the Applied Meteorology Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Roberto Avelino Cecílio

Co-advisers: Elpídio Inácio Fernandes Filho
Gabrielle Ferreira Pires
Michel Castro Moreira

**VIÇOSA - MINAS GERAIS
2023**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

T266r
2023
Teixeira, David Bruno de Sousa, 1993-
Rainfall erosivity in Brazil / David Bruno de Sousa
Teixeira. – Viçosa, MG, 2023.
1 tese eletrônica (119 f.): il. (algumas color.).

Texto em inglês.

Orientador: Roberto Avelino Cecílio.

Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Engenharia Agrícola, 2023.

Referências bibliográficas: f. 102-119.

DOI: <https://doi.org/10.47328/ufvbbt.2023.154>

Modo de acesso: World Wide Web.

1. Solos - Erosão. 2. Hidrologia. 3. Meteorologia agrícola.
I. Cecílio, Roberto Avelino, 1976-. II. Universidade Federal de
Viçosa. Departamento de Engenharia Agrícola. Programa de
Pós-Graduação em Meteorologia Aplicada. III. Título.

CDD 22. ed. 631.45


DAVID BRUNO DE SOUSA TEIXEIRA

RAINFALL EROSIVITY IN BRAZIL


Thesis submitted to the Applied Meteorology
Graduate Program of the Universidade Federal de
Viçosa in partial fulfillment of the requirements
for the degree of *Doctor Scientiae*.

APPROVED: March 29, 2023.

Assent:

Documento assinado digitalmente
 DAVID BRUNO DE SOUSA TEIXEIRA
Data: 29/03/2023 17:29:14-0300
Verifique em <https://validar.iti.gov.br>

David Bruno de Sousa Teixeira
Author

Documento assinado digitalmente
 ROBERTO AVELINO CECILIO
Data: 29/03/2023 15:17:03-0300
Verifique em <https://validar.iti.gov.br>

Roberto Avelino Cecílio
Adviser

To God, Jesus, the Holy Spirit, and my family.

ACKNOWLEDGEMENTS

To God, for the gift of life and for his infinite grace that sustained me in difficult times and led me to get here.

To my beloved, Mariana, for all her love and for believing in me, even when I didn't. Thank you for making my days happier.

To my parents, Teixeira and Luciene, and my sister, Shelida, for all their love, encouragement, and unconditional support throughout my life. To my cousins, uncles and aunts, grandfathers and grandmothers, and to my brother-in-law for being happy with my achievements.

To my parents-in-law, Heraldo and Nadja, for all the support and encouragement.

To professor Roberto Cecílio, for the opportunity and trust, and for helping me to idealize and execute this thesis. To professors Elpídio, Gabrielle, and Michel for their co-supervision and for all the valuable contributions to this thesis. To João Paulo and Marcel for all suggestions for improving this thesis.

To my friends from UFV Laura, Fernanda, Mariana, and Arnaldo, for all the companionship and for making the routine in Viçosa more pleasant.

To my brothers in Christ in Viçosa, for their teachings and testimonies that strengthened my journey in faith. Especially Beth, for being an example of perseverance, and Rita, Rodrigo, and Pedro for being an example of family.

To Gustavo for helping me with the processing of the maps shown in this thesis.

To the Federal University of Viçosa for the structure granted, and to the graduate program in Applied Meteorology for the opportunity to pursue a doctorate. To Graça for her competence and willingness to help with whatever I needed.

This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

To all those who directly or indirectly contributed to the realization of this thesis.

*“Whoever believes in me, as Scripture has said,
rivers of living water will flow from within them”.*

John 7:38

ABSTRACT

TEIXEIRA, David Bruno de Sousa, D.Sc., Universidade Federal de Viçosa, March, 2023. **Rainfall erosivity in Brazil.** Adviser: Roberto Avelino Cecílio. Co-advisers: Elpídio Inácio Fernandes Filho, Gabrielle Ferreira Pires, and Michel Castro Moreira.

The phenomenon known as rainfall erosivity (RE) expresses the ability of rainfall to cause soil erosion. Thus, the estimation of RE magnitudes is relevant for understanding how the erosive processes vary in time and space. Considering this, the present thesis explores the main aspects of RE in Brazil. In Chapter 1, an in-depth review of scientific literature on the RE assessment in Brazil is shown. It was found that the EI_{30} has been the most employed erosivity index, while the use of pluviographic rainfall data and regression equations are the main methods for obtaining erosivity values. Kriging is the most widespread technique for obtaining RE maps in Brazil. Furthermore, the Southeast region accounts for the largest number of erosivity studies, while the North has a major lack of erosivity information. The advancements over the last decade are characterized by the use of synthetic series of rainfall and remote sensing products to estimate erosivity, as well as the use of machine learning techniques for its interpolation. In Chapter 2, a large national database was used to assess the RE patterns in time and space over the Brazilian territory. The results show that the mean annual RE value is $5,620 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$, with considerable spatial variation over the country. The RE values are more equitably distributed throughout the year in the southern region, while in some spots of the northeastern region, it is irregularly concentrated in specific months. Further analyses revealed that the annual RE gravity center for Brazil is in the Goiás State and that it presents a north-south migration pattern throughout the months. Complementarily, the erosivity density magnitudes allowed the identification of high-intensity rainfall spots. Additionally, the Brazilian territory was divided into eleven homogeneous regions regarding the RE patterns and for each defined region, a regression model was adjusted and validated. These models' statistical metrics were considered satisfactory and, thus, can be used to estimate RE values for the whole country using monthly rainfall depths. In Chapter 3 machine learning techniques were applied to obtain an annual RE map for Brazil. According to the accuracy metrics analyzed, Random Forest (RF) is considered the model with the best prediction performance for mapping the annual RE. The covariates with higher importance for the predictions were the total annual rainfall, rainfall depth for August, and rainfall of the coldest quarter. Further analysis revealed that the northeastern of the country as well as the Serra do Mar mountains

region are characterized as the areas with the highest uncertainties in the values mapped. The created map is considered an advancement regarding the availability of accurate RE values in the country. The present thesis shows the most complete panorama of the RE phenomenon in Brazil shown in the literature so far. Therefore, the values, maps, and analysis shown are relevant for improving the accuracy of soil loss estimates and for the establishment of soil conservation planning on a national scale.

Keywords: Erosivity index. Universal Soil Loss Equation. Soil and water conservation. Erosivity density. Machine learning.

RESUMO

TEIXEIRA, David Bruno de Sousa, D.Sc., Universidade Federal de Viçosa, março de 2023. **Erosividade da chuva no Brasil**. Orientador: Roberto Avelino Cecílio. Coorientadores: Elpídio Inácio Fernandes Filho, Gabrielle Ferreira Pires e Michel Castro Moreira.

O fenômeno conhecido como erosividade das chuvas (RE) expressa a capacidade das precipitações em provocar a erosão do solo. Assim, a estimativa de magnitudes de RE é relevante para entender como os processos erosivos variam no tempo e no espaço. Considerando isso, a presente tese explora os principais aspectos da RE no Brasil. No Capítulo 1, é apresentada uma revisão aprofundada da literatura científica sobre a avaliação da RE no Brasil. Verificou-se que o EI_{30} tem sido o índice de erosividade mais empregado, enquanto o uso de dados pluviográficos e de equações de regressão são os principais métodos para obtenção dos valores de erosividade. A krigagem é a técnica mais difundida para obtenção de mapas de RE no Brasil. Além disso, a região Sudeste concentra o maior número de estudos de erosividade, enquanto a região Norte apresenta grande carência de informações sobre erosividade. Os avanços na última década são caracterizados pelo uso de séries sintéticas de precipitação e produtos de sensoriamento remoto para estimar a erosividade, bem como o uso de técnicas de aprendizado de máquina para sua interpolação. No Capítulo 2, um grande banco de dados nacional foi utilizado para avaliar os padrões de RE no território brasileiro. Os resultados mostram que o valor médio anual de RE é $5.620 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ ano}^{-1}$, com considerável variação espacial. Os valores de RE são distribuídos de forma mais equitativa ao longo do ano na região sul, enquanto em alguns pontos da região nordeste concentram-se de forma irregular em meses específicos. Análises posteriores revelaram que o centro de gravidade anual da RE para o Brasil está localizado no estado de Goiás e que este apresenta um padrão de migração latitudinal ao longo dos meses. Complementarmente, as magnitudes de densidade de erosividade permitiram a identificação de pontos de chuva de alta intensidade. Adicionalmente, o território brasileiro foi dividido em onze regiões homogêneas quanto aos padrões de RE e para cada região, um modelo de regressão foi ajustado e validado. No Capítulo 3, técnicas de aprendizado de máquina foram aplicadas para obter um mapa anual de RE para o Brasil. De acordo com as métricas de acurácia analisadas, o Random Forest (RF) é considerado o modelo com melhor desempenho de previsão para mapear a RE anual. As covariáveis com maior importância para as previsões foram a precipitação total anual, a precipitação de agosto e a precipitação do trimestre mais frio. Análises mais aprofundadas revelaram que o nordeste do

país, assim como a região da Serra do Mar, caracterizam-se como as áreas com maiores incertezas nos valores mapeados. O mapa criado é considerado um avanço em relação à disponibilização de valores precisos de RE no país. A presente tese apresenta o panorama mais completo do fenômeno RE no Brasil apresentado na literatura até o momento. Portanto, os valores, mapas e análises apresentados são relevantes para melhorar a precisão das estimativas de perda de solo e para o estabelecimento de um planejamento para a conservação do solo em escala nacional.

Palavras-chave: Índice de erosividade. Equação Universal de Perda de Solo. Conservação do solo e da água. Densidade de erosividade. Aprendizado de máquina.

LIST OF FIGURES

CHAPTER 1: RECENT ADVANCEMENTS IN RAINFALL EROSION ASSESSMENT IN BRAZIL: A REVIEW

- Figure 1. Number of published studies related to the assessment of rainfall erosion in Brazil (a) per year and (b) per language.....24
- Figure 2. Number of published studies related to the assessment of rainfall erosion in Brazil regarding the erosion index used (a) in total and (b) per year.....26
- Figure 3. Number of published studies related to the assessment of rainfall erosion in Brazil according to the method for estimating the erosion, (a) in total and (b) per year.....30
- Figure 4. Number of published studies related to the assessment of rainfall erosion in Brazil per (a) spatial scale, (b) geographic region, and (c) State.....33
- Figure 5. (a) Number of published studies per rainfall erosion categories. (a) Quantity for the entire country and (b) percentage for each Brazilian region.....35
- Figure 6. (a) Number of published studies on rainfall erosion in Brazil that did or did not generate erosion maps and (b) most employed mapping techniques.....36

CHAPTER 2: ASSESSMENT, REGIONALIZATION, AND MODELING RAINFALL EROSION OVER BRAZIL: FINDINGS FROM A LARGE NATIONAL DATABASE

- Figure 1. Flowchart of the methodological steps adopted in the present study. RECI and RESI mean, respectively, rainfall erosion concentration index and rainfall erosion seasonality index. PCA and HCA denote, respectively, principal component analysis and hierarchical cluster analysis. R^2 means coefficient of determination.....48
- Figure 2. (a) Location of Brazil in South America and its altitude variation (SRTM-DEM 30-m resolution). (b) Köppen (1936)'s climate classification presented by Alvares et al. (2013). (c) Spatial distribution of the normal annual rainfall 1981-2010 (INMET, 2022) and location of the rainfall gauges considered in this study. (d) Land use and cover provided by the MapBiomass Project (2021). (e) Histogram of the number of years presented in the series of the gauges considered per region.....49
- Figure 3. (a) Point distribution of the annual RE values for Brazil and (b) its classification according to Carvalho (2008).....60
- Figure 4. Annual RE values per (a) climate according to the climatic classification of Köppen (1936) carried out by Alvares et al. (2013), (b) biome (IBGE, 2019), and (c) hydrographic region (IBGE, 2021), for Brazil.....61
- Figure 5. Correlogram presenting Pearson's correlation coefficient (r) between the variables altitude (m), latitude (degrees south), longitude (degrees west), annual RE ($\text{MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$), annual rainfall depth (R, mm), annual ED ($\text{MJ ha}^{-1} \text{h}^{-1}$), RECI (dimensionless), and RESI (dimensionless), as well as their scatter and density distribution plots, for the entire country and

the north (N), northeast (NE), center-west (CW), southeast (SE), and south (S) regions. Significant correlation was found if p-value < 0.001 (***), < 0.01 (**), < 0.05 (*), and < 0.10 (.).63

Figure 6. Point distribution of the monthly RE values for Brazil.....64

Figure 7. Classification for the (a) RECI and (b) RESI indexes over the Brazilian territory. (c) Monthly distribution of the RE values for Maraújo-CE, Buenópolis-MG, Rancharia-SP, and Montenegro-RS, including the coefficient of variation (CV).....66

Figure 8. Gravity center location of monthly, seasonal, and annual RE. Federal District (DF) and Brazilian states: Goiás (GO), Minas Gerais (MG), Mato Grosso (MT), Mato Grosso do Sul (MS), São Paulo (SP), and Tocantins (TO). Seasons: December-January-February (DJF), March-April-May (MAM), June-July-August (JJA), and September-October-November (SON).....68

Figure 9. (a) Annual ED classified according to Dash et al. (2019). (b) Monthly distribution of the RE, ED, and rainfall depth values. (c) Spatial distribution of the lower and higher monthly ED magnitudes, represented respectively by the 5th and 95th percentiles.....70

Figure 10. (a) Homogeneous regions regarding the RE for Brazil, as well as the (b) annual and (c) mean monthly RE values for each region.....73

Figure 11. Scatter plots between the observed monthly RE values and the predicted by the models established for each homogeneous region.....75

Figure 12. Percentage error map for the annual RE values predicted by the models established for each homogeneous region. n expresses the number of rainfall gauges in each percentage error interval.....76

CHAPTER 3: GRIDDED MAP OF ANNUAL RAINFALL EROSIVITY FOR BRAZIL: A MACHINE LEARNING APPROACH

Figure 1. Flowchart of the methodological steps adopted in the present study. RFE means recursive feature elimination. LM, KkNN, RF, and SVM denote, respectively, Linear Model, Weighted k-Nearest Neighbors, Random Forest, and Support Vector Machine Radial Sigma. LCCC, MAE, RMSE, R², SD, and CV means Lin’s concordance correlation coefficient, mean absolute error, root mean square error, coefficient of determination, standard deviation, and coefficient of variation.....83

Figure 2. (a) Altitude variation in Brazil (WorldClim-DEM 30 arc-second resolution). (b) Köppen (1936)’s climate classification presented by Alvares et al. (2013). (c) Spatial distribution of the normal annual rainfall 1981-2010 (INMET, 2022) and location of the rainfall gauges considered as database in this study.....84

Figure 3. Comparison of the MAE and RMSE values obtained by the analyzed models and those by the Null Model (values obtained considering the mean value of all observations of the rainfall gauges selected). The values refer to the test subset.....94

Figure 4. LCCC values for the predictions by the analyzed models, considering the test subset. The different letters following the mean values correspond to statistical difference occurrence.....95

Figure 5. (a) Mean annual RE map predicted by the RF model for Brazil. Maps of (b) standard deviation (SD) and (c) coefficient of variation (CV) of the annual RE maps. These maps were created considering the 100 runs of the RF model's test.....96

Figure 6. Number of times that the covariates selected by the RFE process are included in the five (Top5, red bars) and ten (Top10, blue bars) most important ones for predicting the annual RE map, considering the 100 RF model's runs.....98

LIST OF TABLES

CHAPTER 1: RECENT ADVANCEMENTS IN RAINFALL EROSION ASSESSMENT IN BRAZIL: A REVIEW

Table 1. List of information analyzed in the review process of the papers.23

CHAPTER 2: ASSESSMENT, REGIONALIZATION, AND MODELING RAINFALL EROSION OVER BRAZIL: FINDINGS FROM A LARGE NATIONAL DATABASE

Table 1. Qualitative classification for the RECI and RESI indexes.....53

Table 2. Descriptive statistics of annual and monthly RE and ED values for Brazil expressed respectively in MJ mm ha⁻¹ h⁻¹ time unit⁻¹ and MJ ha⁻¹ h⁻¹59

Table 3. Established rainfall erosion estimation models for each homogeneous region (HR), and their respective accuracy annual metrics. MAE and RMSE values are expressed in MJ mm ha⁻¹ h⁻¹ year⁻¹. RE_i, R_i, and MFI_i stand for rainfall erosion, rainfall depth, and modified Fournier index, respectively, for the month i.....74

CHAPTER 3: GRIDDED MAP OF ANNUAL RAINFALL EROSION FOR BRAZIL: A MACHINE LEARNING APPROACH

Table 1. Climatic covariates considered for the spatial prediction of the annual RE in Brazil.....86

Table 2. Terrain covariates considered for the spatial prediction of the annual RE in Brazil.....87

Table 3. Machine learning models used, their learning approaches, and R packages for application.....88

Table 4. Covariates that presented variance near zero, and that were eliminated from the prediction process.....93

Table 5. Covariates that presented more than 95% of Spearman's correlation with other covariates, and that were eliminated from the prediction process.....93

Table 6. Models' training and test statistical metrics obtained for the prediction of the annual RE map for Brazil. The best metrics are shown in bold.....94

Table 7. Annual RE range shown in the present study and in the previous ones, as well as the mapping technique used and the number of rainfall gauges considered for mapping.....97

LIST OF ACRONYMS AND ABBREVIATIONS

ANA	National Water and Sanitation Agency
ANNs	Artificial neural networks
CanESM2	Second Generation Canadian Earth System Model
CHIRPS	Climate Hazards Group Infrared Precipitation with Stations
CV	Coefficient of variation
DEM	Digital elevation model
DJF	December-January-February
ED	Erosivity density
EWD	Easterly Wave Disturbance
GPRH-UFV	Research Group in Water Resources of the Federal University of Viçosa
HadCM3	Hadley Centre Coupled Global-Ocean Model
HadGEM2	Hadley Centre Global Environment Model version 2
HCA	Hierarchical clustering analysis
HR	Homogeneous region
IBGE	Brazilian Institute of Geography and Statistics
IDW	Inverse distance weighting
INMET	National Institute of Meteorology
ITCZ	Intertropical Convergence Zone
JJA	June-July-August
KE	Kinetic energy
KNN	<i>k</i> -Nearest Neighbors
KKNN	Weighted <i>k</i> -Nearest Neighbors
LCCC	Lin's concordance correlation coefficient
LM	Linear model
MAE	Mean absolute error
MAM	March-April-May
MFI	Modified Fournier index
MIROC5	Model for Interdisciplinary Research on Climate version 5
ML	Machine learning
PCA	Principal component analysis
PCI	Precipitation concentration index

PCs	Principal components
PERSIANN	Precipitation Estimation from Remotely Sensed Information
R^2	Coefficient of determination
RCP	Representative Concentration Pathway
RE	Rainfall erosivity
RECI	Rainfall erosivity concentration index
RESI	Rainfall erosivity seasonality index
RF	Random Forest
RFE	Recursive feature elimination
RMSE	Root mean square error
RUSLE	Revised Universal Soil Loss
SACZ	South Atlantic Convergence Zone
SAMS	South American Monsoon System
SD	Standard deviation
SON	September-October-November
SRTM	Shuttle Radar Topography Mission
SS	Synthetic series
SVM	Radial Support Vector Machine Sigma
TRMM	Tropical Rainfall Measuring Mission
USLE	Universal Soil Loss Equation

SUMMARY

GENERAL INTRODUCTION	18
CHAPTER 1: RECENT ADVANCEMENTS IN RAINFALL EROSION ASSESSMENT IN BRAZIL: A REVIEW	20
Abstract.....	20
1. Introduction.....	21
2. Material and methods.....	22
3. Results and discussion	24
3.1. Erosion indexes used in Brazil	25
3.2. Methods for obtaining rainfall erosion values in Brazil	29
3.3. Spatial distribution of rainfall erosion studies in Brazil	32
3.4. Rainfall erosion values over Brazil	34
3.5. Mapping rainfall erosion in Brazil.....	36
3.6. Projected rainfall erosion over Brazil	39
3.7. Rainfall erosion products available for Brazil	40
3.8. Advancements and prospects.....	41
4. Conclusion	43
Data availability	44
CHAPTER 2: ASSESSMENT, REGIONALIZATION, AND MODELING RAINFALL EROSION OVER BRAZIL: FINDINGS FROM A LARGE NATIONAL DATABASE ..	45
Abstract.....	45
1. Introduction.....	46
2. Material and methods.....	47
2.1. Study area and database	48
2.2. Rainfall erosion estimation	50
2.3. Rainfall erosion concentration	52
2.4. Rainfall erosion gravity center model	53
2.5. Erosion density estimation.....	54
2.6. Establishing regionalized models to estimate rainfall erosion	54
2.6.1. Dataset obtention	54
2.6.2. Definition of homogeneous regions.....	55
2.6.3. Establishment of the estimation models	56
3. Results and discussion	58
3.1 Rainfall erosion values	58

3.2. Erosivity density values	68
3.3. Regionalized models to estimate rainfall erosivity	72
4. Conclusion	77
Data availability	79
CHAPTER 3: GRIDDED MAP OF ANNUAL RAINFALL EROSIVITY FOR BRAZIL: A MACHINE LEARNING APPROACH.....	80
Abstract.....	80
1. Introduction.....	81
2. Material and methods.....	82
2.1 Study area	84
2.2 Erosivity database	84
2.3 Prediction covariates database	85
2.4 Prediction models	88
2.5 Selection of the prediction covariates	89
2.6 Training, test, and selection of the models	91
2.7 Final map prediction and uncertainty analysis	92
3. Results and discussion	93
3.1 Elimination of covariates	93
3.2 Performance of the models	93
3.3 Rainfall erosivity map.....	95
3.4 Uncertainty maps	98
4. Conclusion	99
GENERAL CONCLUSION.....	100
REFERENCES	102

GENERAL INTRODUCTION

Water erosion causes environmental, economic, and social problems worldwide. Among its major impacts, it can be mentioned a decrease in agricultural yields and an increase in production costs (ALEWELL et al., 2020; CHALISE et al., 2020). In addition, water quality and availability are compromised due to pollution and sediment accumulation in the water sources. According to Borelli et al. (2020), soil erosion is projected to increase over the 21st century in many countries, including Brazil, due to climate change. Thus, high soil loss rates constitute a challenge regarding the promotion of current and future food security (WUEPPER; BORRELLI; FINGER, 2019).

Due to the need for a methodology capable of assessing the factors that cause water erosion and estimating soil losses, predictive models such as the Universal Soil Loss Equation (USLE) and its revised version (RUSLE) (RENARD et al., 1997; WISCHMEIER; SMITH, 1978) were developed. Among the components included in these models, the climatic factor, known as rainfall erosivity (RE), expresses the ability of rainfalls to cause soil erosion. As stated by Castro et al. (2022), rainfall, due to the direct impact of the drops and the generated runoff water, is the primary agent which causes erosion in the Brazilian soils, which evidences the need for accurate RE estimates for the country.

For obtaining reliable RE values, historical series of rainfall data is necessary on a sub-daily scale with a minimum length of 20 years of consistent and uninterrupted data (RENARD et al., 1997; WISCHMEIER; SMITH, 1978). However, the difficulty in acquiring data with large time length in Brazil, both in terms of quantity and quality, is characterized as a barrier to the availability of more accurate RE estimates. Considering this, many authors have used empirical models that correlate rainfall data with erosivity magnitudes to estimate RE (BRITO et al., 2021; CASTAGNA et al., 2022; MOREIRA et al., 2020; NEVES; DI LOLLO, 2022).

Although the use of empirical models is characterized as an easy way to estimate RE, so far, a restricted number of models have been adjusted for Brazil (OLIVEIRA; WENDLAND; NEARING, 2012). This has led to its indistinct and generalized application, resulting in RE values with large deviations in their magnitude, as demonstrated by Oliveira et al. (2018). As an alternative, the use of weather generators that simulate, from daily rainfall data, sub-daily series of synthetic rainfall data, has been a feasible option to estimate RE values in Brazil as shown by Cecílio et al. (2013), Moreira et al. (2016), and Teixeira et al. (2022a). The effectiveness of using weather generators for this purpose was proven by Oliveira et al. (2018),

which constitutes an advancement in the methods for expanding RE estimates over the Brazilian territory.

Regardless of the estimation method, the effective use of the RE estimates often depends on the availability of maps for this variable. In Brazil, RE maps have been widely used to obtain soil loss estimates at local (CUNHA et al., 2022; SILVA et al., 2016), State (LENSE et al., 2021; MEDEIROS et al., 2016), and regional (GOMES et al., 2019a) scales. This mapping process has usually been done using conventional interpolation methods, such as the use of deterministic (CASTAGNA et al., 2022; NEVES; DI LOLLO, 2022) and geostatistical (SOUZA et al., 2020b; TERASSI et al., 2020) techniques. In this context, the use of machine learning techniques appears as a promising alternative for obtaining RE maps with higher accuracy, as exemplified by Lee et al. (2021) and Souza et al. (2022).

This thesis explores the main aspects of the rainfall erosivity in Brazil. The first section of this thesis contains a general introduction. In the sequence, three chapters are shown with their own abstract, introduction, material and methods, results and discussion, and conclusion sections. Finally, a general conclusion section followed by the references list is presented. In Chapter 1, an in-depth review of scientific literature on the RE assessment in Brazil is shown. In Chapter 2, a large national database was used to assess the RE patterns in time and space over the Brazilian territory. In Chapter 3, machine learning techniques were applied to obtain an annual RE map for Brazil.

CHAPTER 1: RECENT ADVANCEMENTS IN RAINFALL EROSION ASSESSMENT IN BRAZIL: A REVIEW

Abstract

In this study, we report the results of an in-depth review of scientific literature on the assessment of rainfall erosivity in Brazil and published in peer-reviewed journals over the last three decades. In total, 123 articles regarding this topic in Brazil were published. These studies were analyzed and filtered into categories regarding the bibliographic information, methodology scope, and main results. It was found that the EI_{30} has been the most employed erosivity index in the country, while the use of pluviographic rainfall data and regression equations are the main methods for obtaining erosivity values. The magnitudes of annual rainfall erosivity reported for the Brazilian territory range from 59 to 26,891 MJ mm ha⁻¹ h⁻¹ year⁻¹. The lowest values are found in the Northeast region and the highest in the North. Kriging is the most widespread technique for obtaining rainfall erosivity maps in Brazil. Furthermore, the Southeast region accounts for the largest number of erosivity studies, while the North has a major lack of erosivity information. The advancements over the last decade are characterized by the use of synthetic series of rainfall and remote sensing products to estimate erosivity, as well as the use of machine learning techniques for its interpolation. For the next years, an increase in the use of these methodologies is expected, as well as an intensification of the assessment of the future patterns of rainfall erosivity over the country. The present review updates the findings regarding the assessment of rainfall erosivity in Brazil and brings a wider overview of the erosivity studies in the country. The information here summarized contributes to the establishment of a conservationist planning of soil and water management on a national scale. Finally, the complete database with all information retrieved from the literature was made available.

Keywords: Erosivity index. R-factor. Soil erosion. Universal Soil Loss Equation. Soil and water conservation.

1. Introduction

Water erosion causes environmental, economic, and social problems worldwide (WUEPPER; BORRELLI; FINGER, 2020). Among its main impacts are the decrease in the potential yield of agricultural lands and the reduced quality and availability of water resources. Also, water erosion changes the chemical, physical and biological properties of the soil, reducing its fertility (DUAN et al., 2016; NOVARA et al., 2018) and compromising current and future food security (PANAGOS; BORRELLI; ROBINSON, 2020; WUEPPER; BORRELLI; FINGER, 2020). According to Sartori et al. (2019), water erosion is estimated to incur a global annual cost of eight billion US dollars along with a decline in agricultural and food production of 33.7 million tons.

Considering the problems caused by water erosion processes, the quantification of soil losses has practical relevance for the adoption of soil management and conservation program. For this purpose, several models for estimating soil losses have been developed and improved. Among these models, the Universal Soil Loss Equation (USLE) and its revised version (RUSLE) (RENARD et al., 1997; WISCHMEIER; SMITH, 1978) are the biggest contributors to the advancements in quantifying erosion processes worldwide (BORRELLI et al., 2021). They are easier to use than other models, especially regarding the degree of flexibility and data accessibility (ALEWELL et al., 2019). Thus, these models are characterized as important conservationist tools to support decision-making.

Included as a component of these models, the climatic factor, known as rainfall erosivity (or R-factor), expresses the ability of rainfall to cause soil erosion. Erosivity is represented by a numerical index and, as explained by Serio et al. (2019), is a function of the physical characteristics of the rain itself, such as its distribution, fall velocity, intensity, duration, frequency, and kinetic energy of the raindrops. Thus, since rainfall is the inducing agent of the phenomenon of water erosion, the estimation of the erosivity magnitude is relevant for understanding how the erosive processes vary in time and space. Also, this variable can be used to identify areas highly susceptible to erosion (MOSAVI et al., 2020; PANDEY et al., 2021; SENANAYAKE; PRADHAN, 2022).

Rainfall erosivity has been estimated in global (LIU et al., 2020b), continental (BALLABIO et al., 2017), national (PADULANO; RIANNA; SANTINI, 2021), regional (ZHU; XIONG; XIAO, 2021), State (SOUZA et al., 2022), and local (CARDOSO et al., 2020) scale. According to Panagos et al. (2017a), South America presents the highest annual erosivity

values. In addition, these same authors state that countries within tropical climate zone face the most expressive erosivity magnitudes of the globe.

Inserted in this high erosivity context, Brazil has expanded its agricultural land areas and is responsible for the production of many commodities. Along with it, deforestation rates have increased over the last years (ARAÚJO et al., 2019; DIAS et al., 2016; SOUZA et al., 2020a). Thus, quantifying erosivity can potentially contribute to identifying erosion hotspots in the country and minimize soil loss rates.

Despite the importance of research aimed at analyzing rainfall erosivity, only a few studies have been carried out to review erosivity values worldwide (NAZUHAN; RAHAMAN; OTHMAN, 2018; NEARING et al., 2017). This is also the case for Brazil, where only one paper was published for this purpose so far. This effort was carried out by Oliveira et al. (2012), with a review article whose goal was to verify the quality and representativeness of the research on erosivity in the country. Although the results and discussion presented by these authors have contributed to significant advancements in this theme, the number of studies published over the last decade highlights the need for a new and broader review that addresses the progress made in the assessment of rainfall erosivity over the Brazilian territory.

Therefore, this study aimed to review studies on rainfall erosivity in Brazil and published in peer-reviewed journals over the last three decades. The specific objectives were to describe and analyze quantitatively and qualitatively the main methodologies for obtaining and mapping erosivity values for Brazil; to identify areas lacking studies; and to present the current and climate change projections values for rainfall erosivity in Brazil.

2. Material and methods

In this study, we report the results of an in-depth review of scientific peer-reviewed literature on the assessment of rainfall erosivity in Brazil published in journals between 1992 and 2021. Data was collected from five distinct platforms: (1) Web of Science core collections (<https://apps.webofknowledge.com>); (2) Science Direct (<http://www.sciencedirect.com>); (3) Scopus (<http://www.scopus.com>); (4) Scientific Electronic Library Online – SciELO (<https://scielo.org>); and (5) Google Scholar (<http://scholar.google.com>). The key terms used for researching articles were “rainfall erosivity” and “Brazil” as well as their translations to Portuguese (“*erosividade da chuva*” and “*Brasil*”). Papers including any quantitative or qualitative method to assess rainfall erosivity in Brazil were selected and reviewed.

After reading the articles, they were analyzed and divided into categories to systematize the discussion of the results in the present study. Some categories included the paper's bibliographic information (e.g. title, year of publication, authors, journal), others included the methodology scope (e.g. study area, method for estimating erosivity) and main results (e.g. mean annual erosivity values). The complete list of information retrieved from the papers is shown in Table 1.

Table 1. List of information analyzed in the review process of the papers.

Group	Entry	Types of data
Bibliography	Title	Open (alphanumeric)
	Author(s)	Open (alphanumeric)
	Year of publication	Open (numeric)
	Journal	Open (alphanumeric)
	Language	Open (alphanumeric)
Study area	Spatial scale	Multiple choice
	Geographic region	Open (alphanumeric)
	State	Open (alphanumeric)
	Municipality or basin	Open (alphanumeric)
Methodology	Number of years of the database	Open (numeric)
	Erosivity index used	Open (alphanumeric)
	Erosivity estimation method	Open (alphanumeric)
Modeling	Did the study establish a regression equation?	Open (alphanumeric)
	Independent variable of the equation	Open (alphanumeric)
	Coefficient of determination (R^2) of the equation	Open (numeric)
Erosivity values	Mean annual EI_{30} values	Open (numeric)
	Range of annual EI_{30} values	Open (numeric)
Mapping	Did the study generate erosivity maps?	Multiple choice
	Mapping technique	Open (alphanumeric)
Others	Did the study classify the rainfall erosivity values?	Open (alphanumeric)
	Did the study relate the erosivity values to soil loss?	Multiple choice
	Is the study presented in the review of Oliveira et al. (2012)?	Multiple choice

The spatial scale from the analyzed articles was discretized as (1) Local: studies within municipalities, small watersheds and considering specific rainfall gauges; (2) State: studies considering a whole state; (3) Regional: studies within interstate watersheds, biomes (IBGE, 2019), and for a specific geographic region of Brazil; and (4) National: studies considering the entire country. Also, to analyze the spatial distribution of the rainfall erosivity studies throughout the Brazilian territory, the municipality, state, and geographic region were identified for all papers (if applicable).

3. Results and discussion

From 1992 to 2021, 123 articles related to the assessment of rainfall erosivity in Brazil were published, including the review article presented by Oliveira et al. (2012). We found that 69 papers were published after the review of Oliveira et al. (2012), which evidences the increase of research on this theme over the last decade (Figure 1a). The fact that in the last nine years the number of published studies is higher than in the previous 21 years (1992-2012) highlights the importance of the present review and shows that this still is a highly researched topic.

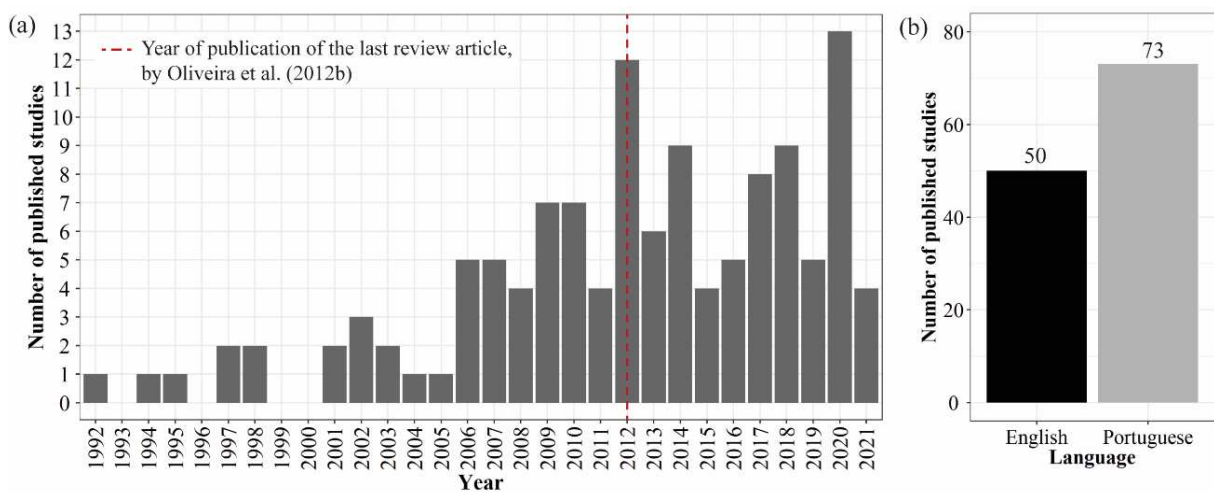


Figure 1. Number of published studies related to the assessment of rainfall erosivity in Brazil (a) per year and (b) per language.

An important aspect to consider is that most of the published papers related to the assessment of rainfall erosivity in Brazil over the last three decades were written in Portuguese (Figure 1b). In this sense, a higher number of articles published in English may improve the contribution of the erosivity values found in Brazil to international studies, such as in those by Liu et al. (2020b), Panagos et al. (2017a), and Riquetti et al. (2020).

Rainfall erosivity values are mainly used to estimate soil losses from models such as the Universal Soil Loss Equation (USLE) and its revised (RUSLE) version (RENARD et al., 1997; WISCHMEIER; SMITH, 1978). However, the relationship between rainfall erosivity and soil loss was analyzed in only 25 papers (about 20% of the total). This shows that the majority of the studies published for Brazil only obtained erosivity values but not related them to observed soil loss rates.

The correlation between measured soil loss and the estimated rainfall erosivity was established in studies for the Northeast (ALBUQUERQUE et al., 1998, 2002), Center-West (CARVALHO; HERNANI, 2001), Southeast (LOMBARDI NETO; MOLDENHAUER, 1992; MARQUES et al., 1997; SILVA et al., 2009), and South Brazil (BERTOL et al., 2007; SCHICK et al., 2014; SILVA et al., 2016). The coefficient of determination (R^2) of the rainfall erosivity and soil loss correlations varied from 0.37 to 0.98, found, respectively, by Silva et al. (2016) and Silva et al. (2009). In addition, some of the reviewed studies estimated rainfall erosivity values and used them as input for soil loss models, such as for USLE (FALCÃO; DUARTE; DA SILVA VELOSO, 2020; GUIMARÃES et al., 2019; SIMPLÍCIO et al., 2021) and RUSLE (COLMAN et al., 2019; DURÃES; MELLO, 2016; RODRIGUES et al., 2017).

Other aspects of the reviewed papers, such as main erosivity indexes, main methods for obtaining and mapping rainfall erosivity, current and projected erosivity values, the spatial distribution of rainfall erosivity studies, and erosivity products available for Brazil are discussed in the following subsections. A complete list of the reviewed articles in this study and a summary of the information retrieved from them are freely available for download at <http://dx.doi.org/10.17632/hgd2whtx55.1>.

3.1. Erosivity indexes used in Brazil

The main erosivity indexes used to express the erosive potential of rainfall in Brazil were the EI_{30} , $KE>25$, and $KE>10$. As discussed later, for specific studies, the modified Fournier index (MFI) and the erosivity density (ED) were also considered to represent the rainfall erosivity. Figure 2a shows that EI_{30} has been the most used erosivity index in the country. This index is present in 121 of the 123 reviewed papers and was used by Lombardi Neto and Moldenhauer (1992) in the first published study related to rainfall erosivity in Brazil analyzed in this paper (Figure 2b).

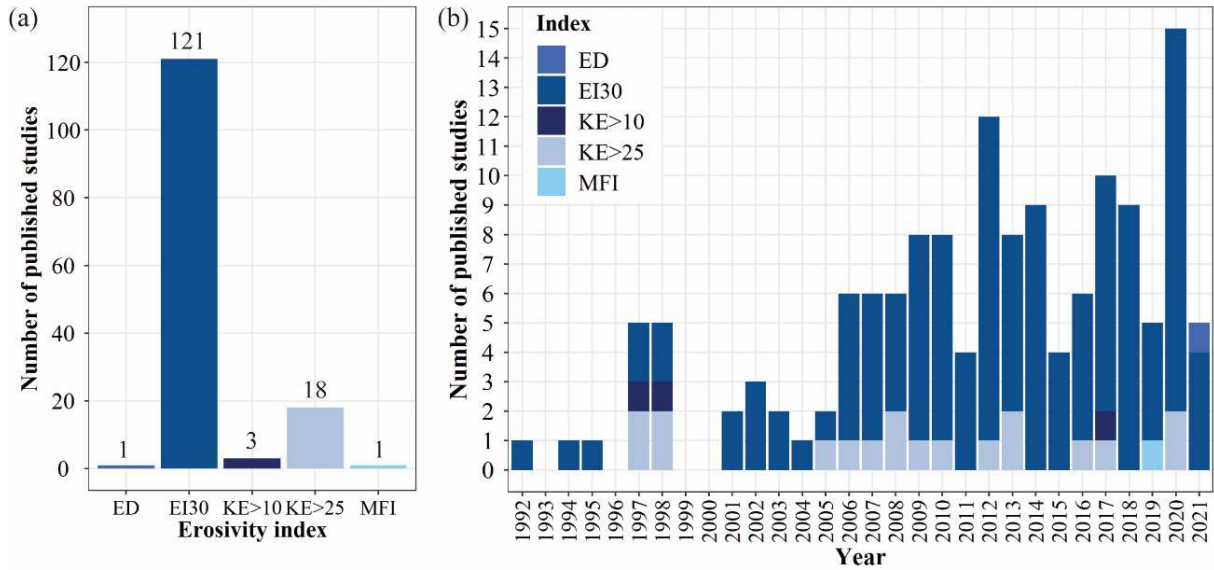


Figure 2. Number of published studies related to the assessment of rainfall erosivity in Brazil regarding the erosivity index used (a) in total and (b) per year.

The erosivity index EI_{30} , proposed by Wischmeier and Smith (1958), is the relationship between rainfall's kinetic energy and maximum intensity for 30 minutes. To determine the EI_{30} , it is necessary to identify all the rainfalls considered erosive in the data series. According to the criteria proposed by Wischmeier and Smith (1958) and Wischmeier (1959), and modified by Cabeda (1976), erosive rainfalls are the ones with 10 mm or more of total precipitation, or with less than 10 mm when the amount precipitated in 15 minutes is equal to or greater than 6 mm.

For erosive rainfalls, the associated kinetic energy (KE) is calculated as a function of the rainfall intensity, from minute to minute, using Equation (1), as suggested by Wischmeier and Smith (1958) and readjusted to the International System of Units by Foster et al. (1981). Foster et al. (1981) state that the KE adopted when the rainfall intensity is higher than 76 mm h^{-1} corresponds to 0.283 MJ $ha^{-1} mm^{-1}$. The KE associated with the rainfall of each day is calculated by adding the kinetic energies of each minute until the total duration of the rainfall.

$$KE = 0.119 + 0.0873 \log I \quad (1)$$

in which KE is the kinetic energy (MJ $ha^{-1} mm^{-1}$); I is the intensity of the rainfall (mm h^{-1}).

Using the daily KE values, the rainfall erosivity index EI_{30} , also daily, is determined through Equation (2), in MJ $mm ha^{-1} h^{-1} day^{-1}$. The monthly EI_{30} is determined by adding daily erosivity values, as shown in Equation (3). The annual EI_{30} values, also known as the R-factor

of the USLE model, are the sum of the rainfall erosivity for each month, as shown in Equation (4).

$$EI_{30} = KE \cdot I_{30} \quad (2)$$

$$RE_m = \sum_{j=1}^n (EI_{30})_j \quad (3)$$

$$RE_a \text{ or } R\text{-factor} = \sum_{i=1}^{12} (RE_m) \quad (4)$$

in which EI_{30} is the daily rainfall erosivity index in the day j ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ day}^{-1}$); I_{30} is the maximum rainfall intensity for 30 consecutive minutes (mm h^{-1}); n is the number of days in the month i ; RE_m is the rainfall erosivity in the month i ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$); RE_a or $R\text{-factor}$ is the annual rainfall erosivity ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$).

As mentioned before, the EI_{30} is used to represent most of the erosivity values over Brazil. This index was used in all five geographic regions of the country (AVANZI et al., 2019; RAIMO et al., 2018; SOUZA et al., 2020b; TEIXEIRA et al., 2022a; WALTRICK et al., 2015), and was used to estimate erosivity on a national scale (CECÍLIO et al., 2021; OLIVEIRA et al., 2018; TRINDADE et al., 2016).

The index $KE>25$ (also known as $KE>1$), in MJ ha^{-1} , was proposed by Hudson (1971) from studies on rainfall erosivity in Africa. As it was developed considering the typical conditions of African rainfalls, this index has been pointed out as the most suitable for regions with a tropical and subtropical climate, such as Brazil. The $KE>25$ is determined by the sum of the kinetic energy (Equation (1)) of rainfall segments whose intensity is greater than 1-inch h^{-1} or 25 mm h^{-1} . The $KE>10$ index is a variation of the $KE>25$ and is obtained considering rainfall segments whose intensity is greater than 10 mm h^{-1} .

In Brazil, both $KE>25$ and $KE>10$ have not yet been used to assess erosivity nationally, which is a gap to be filled. In local studies, Andrade et al. (2020a) and Andrade et al. (2020b) used the $KE>25$ index to estimate erosivity in the municipalities of Formosa and Aragarças, respectively, both in the state of Goiás. Carvalho et al. (2010), Cecílio et al. (2013), Moreira et al. (2008), and Moreira et al. (2016) used the $KE>25$ index to analyze the rainfall erosivity in the states of Rio de Janeiro, Espírito Santo, Minas Gerais, and Santa Catarina, respectively. Much less widespread, the $KE>10$ index was used only for the municipalities of Sete Lagoas

and Sumé, in the states of Minas Gerais and Pernambuco, respectively (ALBUQUERQUE et al., 1998; MARQUES et al., 1997).

Differently from the other reviewed papers, Back et al. (2019) studied the rainfall erosive potential for the entire South region of Brazil exclusively as a function of MFI. The authors introduced the term rainfall aggressiveness as an indication of the degree of rainfall erosivity, so this variable was also considered as an erosivity index in the present review. The MFI, obtained from Equation (5), relates the mean monthly rainfall and the mean annual rainfall (LOMBARDI NETO, 1977; LOMBARDI NETO; MOLDENHAUER, 1992) and, as pointed out by Oliveira et al. (2012), is widely used as an independent variable for obtaining erosivity estimation models in Brazil.

$$MFI_i = \frac{(R_i)^2}{R_a} \quad (5)$$

in which MFI_i is the modified Fournier index for the month i ; R_i is the mean monthly rainfall in the month i (mm); R_a is the mean annual rainfall (mm).

Back et al. (2019) concluded that the rainfall aggressiveness calculated for South Brazil has different monthly patterns depending on the State (Santa Catarina, Paraná, and Rio Grande do Sul). Also, these authors state that the high aggressiveness of rainfalls may not coincide with high amounts of annual rainfall, evidence of the influence of rainfall seasonality on the high erosive potential of specific months. Thus, the rainfall aggressiveness index (MFI) can be an easy tool to help assess soil erosion in Brazil.

Finally, ED was also considered as an index to infer the erosive potential of rainfalls in Brazil. ED is expressed by the ratio between monthly rainfall erosivity and monthly rainfall (FOSTER, 2008), and therefore shows the erosivity content per unit of rainfall, as shown in Equation (5).

$$ED_i = RE_i \cdot R_i^{-1} \quad (5)$$

in which ED_i is the mean erosivity density in the month i ($\text{MJ ha}^{-1} \text{ h}^{-1}$); RE_i is the mean monthly rainfall erosivity in the month i ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$).

The ED index concept was shown by Foster (2008) and was used in some studies in China (LI; YE, 2018; XU et al., 2019), Greece (PANAGOS et al., 2016; VANTAS; SIDIROPOULOS; LOUKAS, 2019), India (DASH; DAS; ADHIKARY, 2019; SINGH; SINGH, 2020), and South Korea (SHIN et al., 2019). In Brazil, this index was only estimated by Teixeira et al. (2022a) as a complement to the analysis of the EI₃₀ values obtained for the state of São Paulo. According to these authors, the ED analysis made it possible to identify areas susceptible to the impacts of the most extreme rainfall events even in months with low EI₃₀ magnitudes. This occurs once the high ED estimates indicate a greater occurrence of high-intensity precipitation events (PANAGOS et al., 2016).

Although the KE>25, KE>10, MFI, and ED indexes have advanced the understanding of the impacts of rainfall on erosive processes, the literature that covers the theme of soil loss in Brazil is mostly composed of studies that used the EI₃₀ erosivity index. This fact may be related, as highlighted by Back et al. (2018) and Back and Poletto (2018), to the good correlations between the EI₃₀ and the soil loss found in studies carried out in Brazil (ALBUQUERQUE; CHAVES; VASQUES FILHO, 1994; CARVALHO et al., 2010; LOMBARDI NETO; MOLDENHAUER, 1992; SCHICK et al., 2014; SILVA et al., 2009).

Furthermore, the lack of rainfall erosivity maps obtained from the mentioned indexes on a national scale, makes it difficult to compare with other erosivity maps previously obtained for Brazil using the EI₃₀ index, such as the maps proposed by Mello et al. (2013), Oliveira et al. (2015), Silva (2004) e Trindade et al. (2016).

3.2. Methods for obtaining rainfall erosivity values in Brazil

We found that the use of pluviographic rainfall data, regression equations, synthetic series (SS) of rainfall data, and MFI were the main methods applied for obtaining rainfall erosivity values in Brazil (Figure 3a). As discussed previously, MFI was used only by Back et al. (2019), who considered MFI as an index of the aggressiveness of rainfalls. According to Figure 3b, over the last three decades, pluviographic rainfall data and regression equations were the methods used in the majority of the reviewed studies.

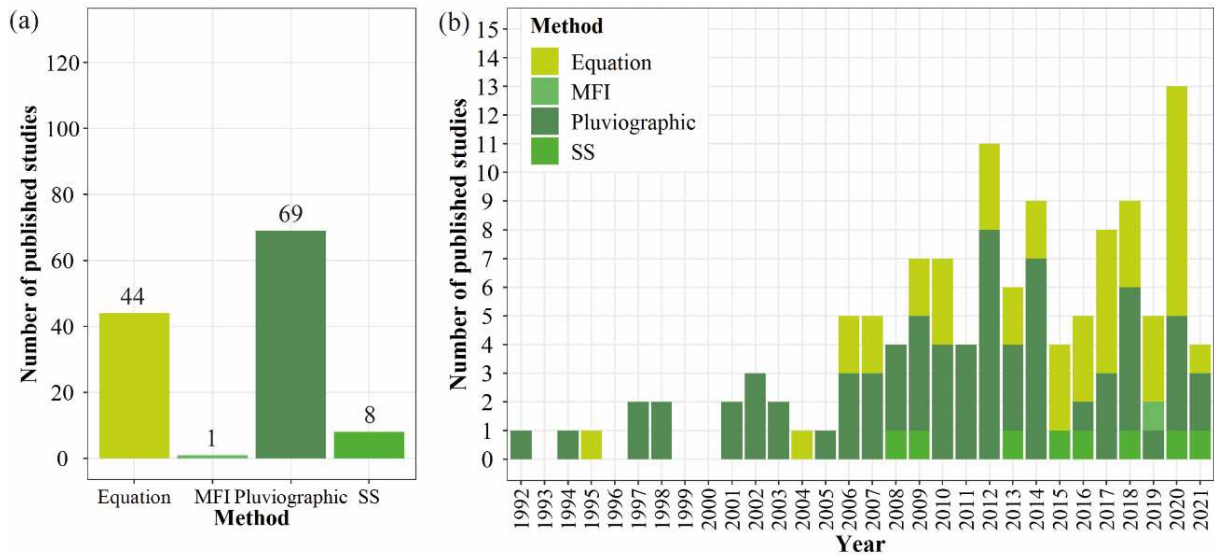


Figure 3. Number of published studies related to the assessment of rainfall erosivity in Brazil according to the method for estimating the erosivity, (a) in total and (b) per year.

The use of pluviographic rainfall data is the most recommended estimation method in literature and allows a more realistic characterization of rainfall erosivity (WISCHMEIER; SMITH, 1978). Historical series of rainfall data on a sub-daily scale with a minimum of 22 years of consistent and uninterrupted data is required to obtain reliable erosivity values (WISCHMEIER, 1959). However, as stated by Montebeller et al. (2007) and Oliveira et al. (2018), the difficulty of acquiring data with a time horizon of this magnitude in the Brazilian territory is an obstacle to estimating erosivity more accurately.

Despite this difficulty, pluviographic rainfall data was used for rainfall erosivity estimates in most of Brazil (ANDRADE et al., 2020b; BACK; ALBERTON; POLETO, 2018; MACHADO et al., 2017; PINHEIRO et al., 2018). Considering this, some examples of databases with rainfall data in a sub-daily resolution available for Brazil that can be used for obtaining rainfall erosivity values are those presented by the National Water and Sanitation Agency – ANA (<https://www.snirh.gov.br/hidroweb/>), National Institute of Meteorology – INMET (<https://bdmep.inmet.gov.br/>), and the National Center for Monitoring and Natural Disaster Alerts – CEMADEN (<http://www2.cemaden.gov.br/mapainterativo/>).

Authors such as Back et al. (2017), Valvassori and Back (2014), and Back (2018) obtained erosivity values using sub-daily series with 31, 33, and 45 years of data, respectively, well above the minimum recommended. On the other hand, studies from Guimarães et al. (2019), Montenegro et al. (2018), and Lima et al. (2013) used historical series with less than 5 years of data, which constitutes a risk of not capturing the long-term patterns of the erosive

rainfalls. As explained by Nearing et al. (2017), the erosivity values obtained using short records have the potential to be biased by unusual wet or dry periods.

Considering the low availability of pluviographic rainfall data, simplified methods for predicting rainfall erosivity using monthly and yearly data from rainfall gauges are the most widely used alternative for estimating erosivity in Brazil. Several regression equations to obtain erosivity values have been established, for the entire country, based on the relationship between this variable and MFI (OLIVEIRA; WENDLAND; NEARING, 2012). This index relates the mean monthly rainfall with the mean annual. These variables are available for most locations with good spatial and temporal coverage, allowing the establishment of estimation models with satisfactory accuracy.

Given the good reliability and easy use of these equations to determine erosivity, several authors have established regression models for this purpose in Brazil. Based on the literature presented by Oliveira et al. (2012), most of the proposed equations used MFI to estimate erosivity. Other variables can also be used, such as the topographic aspects of altitude, latitude, and longitude, used as independent variables in the models presented by Avanzi et al. (2019), Mello et al. (2013), and Riquetti et al. (2020).

Regression equations allow estimating erosivity values in several places where pluviographic rainfall records are not available. However, since these equations are empirical models, their use is limited to locations with climatic characteristics, especially rainfall, similar to those for which they were developed. As explained by Oliveira et al. (2018), their widespread use can lead to incorrect erosivity estimates. Examples of the inappropriate use of these equations can be seen in the studies proposed by Falcão et al. (2020), Sousa et al. (2019), and Souza et al. (2020b). Although these studies were developed for the states of Pernambuco, Ceará, and Alagoas, respectively, whose climate is semi-arid, the authors used the equation from Lombardi Neto and Moldenhauer (1992) for the municipality of Campinas, in the state of São Paulo, whose climate is humid subtropical. Thus, the use of the erosivity values obtained by the mentioned examples is probably not adequate and may imply inaccurate erosion rates.

Silva (2004) divided the Brazilian territory into eight homogeneous regions in terms of total annual rainfall and identified regression equations to estimate erosivity for each one. This division was used to obtain erosivity values in some studies, such as those presented by Almeida et al. (2017), Brito et al. (2021), and Silva et al. (2020c). Despite this attempt to regionalize the equations, Silva (2004) simply defined an existing model to represent a region and did not establish a model considering the specific data of each region. Teixeira et al. (2022a) did this last procedure for the state of São Paulo. Differently from Silva (2004), these authors

firstly identified homogeneous regions concerning rainfall erosivity, and then obtained and validated models for each region, which is characterized as a good method for establishing regionalized estimation models. In addition, Mello et al. (2013) proposed to use the Thiessen polygon's method as well as the precipitation concentration index (PCI) to identify a region to use a given erosivity model already established for Brazil. This procedure was also applied by Almagro et al. (2017).

Another alternative method used to obtain erosivity estimates in Brazil has been the use of synthetic series of rainfall from climatic generators, as shown by Lobo et al. (2015), Yu (2002), and Zhang et al. (2008). From the association of real rainfall data series with random numbers produced by computational algorithms, these generators provide sequences of numbers that resemble real sub-daily rainfall data, which means that possible non-stationarity in the series is not captured.

In this sense, the software ClimaBR, a stochastic weather generator (BAENA et al., 2005; OLIVEIRA; ZANETTI; PRUSKI, 2005a, 2005b; ZANETTI et al., 2005), was developed for the Brazilian climatic conditions. Zanetti et al. (2005) validated the software regarding the number of rainy days and the total daily rainfall. The synthetic series generated by this software show information that characterizes the daily rainfall profile such as the daily amount and its duration, maximum instantaneous rainfall intensity, and its time of occurrence.

The synthetic series generated by ClimaBR were used for estimating rainfall erosivity values by Cecílio et al. (2013), Moreira et al. (2009), and Teixeira et al. (2022a) for the states of Espírito Santo, Minas Gerais, and São Paulo, respectively, all in the Southeast region of Brazil. Oliveira et al. (2015) used the ClimaBR for determining erosivity values for the entire Brazilian territory. The use of the synthetic series generated by ClimaBR was validated for Brazil by Oliveira et al. (2018) and consists of a method with great potential to expand the understanding of the erosive dynamics of rainfall in the country.

3.3. Spatial distribution of rainfall erosivity studies in Brazil

In Brazil, most published papers studied rainfall erosivity at a local scale (municipalities, small watersheds, and specific rainfall gauges), as shown in Figure 4a. The amount of research to understand erosivity patterns throughout Brazil is still scarce (only ten studies), which makes it difficult to establish national guidelines for soil conservation planning.

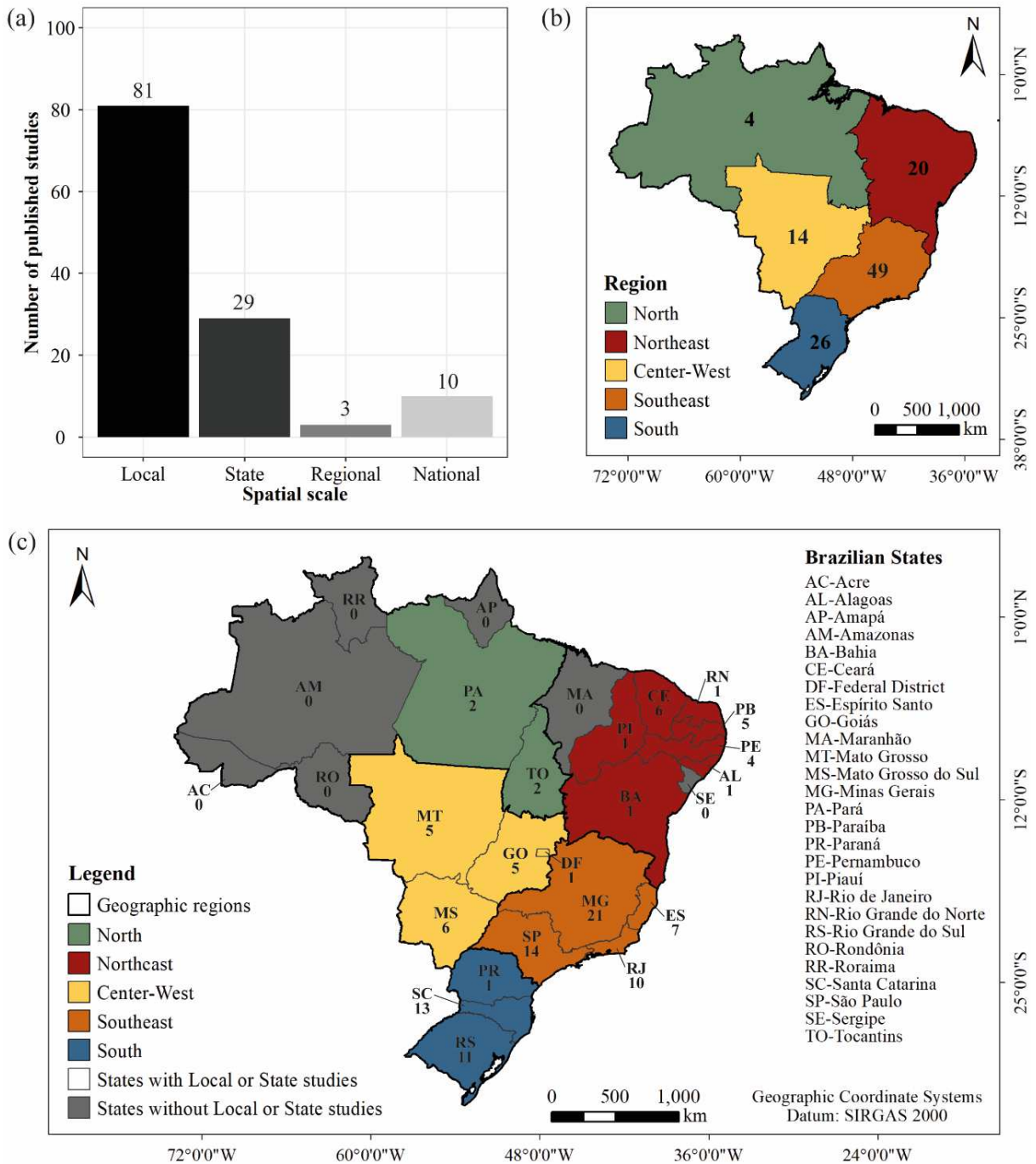


Figure 4. Number of published studies related to the assessment of rainfall erosivity in Brazil per (a) spatial scale, (b) geographic region, and (c) State.

Considering the five geographic regions in the country, the Southeast region accounts for most of the erosivity studies, around 43% of all Local, State, and Regional studies (Figures 4b and 4c), something also pointed by Oliveira et al. (2012). On the other hand, the North and Center-West regions have the lowest number of studies. According to Dias et al. (2016), some of the areas in these two regions (Amazonia and Cerrado biomes agricultural frontiers) suffered

an increase in cropland and pastureland areas in recent years. Also, for the state of Mato Grosso (Center-West) an intensification of pasture and cattle ranching has been reported (GARRETT et al., 2018; GOLLNOW; LAKES, 2014). In this context, an increase in the studies on the erosive potential of rainfall in these two regions may highlight the need for the adoption of conservation practices to mitigate changes in land use.

It is also important to point out that the reduced number of erosivity studies in the states of the North and Center-West regions can reflect the lower number of climatic monitoring stations in these regions than in others. Based on the inventory of the Brazilian National Water and Sanitation Agency (ANA), Silva et al. (2021) stated that the North and Center-West regions account for, respectively, only 7 and 8.6% of the Brazilian rainfall gauges, minimizing the amount and availability of data for hydrological studies. As an alternative, the use of satellite data such as those presented by Brito et al. (2021) and Moreira et al. (2020) can spatially improve the understanding of the dynamics of rainfall erosivity in areas with low rainfall gauge data, despite the problem with the temporal resolution that will still persist.

3.4. Rainfall erosivity values over Brazil

Since EI_{30} was the most used erosivity index in the published rainfall erosivity studies in the country, the erosivity magnitudes discussed in this section refer only to this index. According to the results reported in the reviewed literature, considering values for a specific year, the magnitudes of annual rainfall erosivity over the Brazilian territory range from 59 to 26,891 MJ mm ha⁻¹ h⁻¹ year⁻¹, in the states of Pernambuco (SANTOS; MONTENEGRO, 2012) and Tocantins (AVANZI et al., 2019), respectively.

The Northeast region has the lowest erosivity values, from 59 to 11,469 MJ mm ha⁻¹ h⁻¹ year⁻¹ (SANTOS; MONTENEGRO, 2012; SOUZA et al., 2020b). Oppositely, the highest erosivity values are in the North region, from 6,599 to 26,891 MJ mm ha⁻¹ h⁻¹ year⁻¹ (AVANZI et al., 2019; VIOLA et al., 2014). Additionally, expressive erosivity values (higher than 15,000 MJ mm ha⁻¹ h⁻¹ year⁻¹) were also found in the states of Minas Gerais (OLIVEIRA et al., 2009, 2014; SILVA et al., 2010a), Rio de Janeiro (CARVALHO et al., 2012; GONÇALVES et al., 2006; MACHADO et al., 2013; MONTEBELLER et al., 2007), and São Paulo (SILVA et al., 2010b; TEIXEIRA et al., 2022a), in the Southeast region; in the state of Mato Grosso (ALMEIDA et al., 2012), in the Center-West region; and in the state of Rio Grande do Sul (BAZZANO; ELTZ; CASSOL, 2007; CASSOL et al., 2008; HICKMANN et al., 2008; MAZURANA et al., 2009; PAULA et al., 2010), in the South region.

Many studies in Brazil classify the erosivity values in the following categories: Low (below 2,452 MJ mm ha⁻¹ h⁻¹ year⁻¹), Moderate (higher than 2,452 and below 4,905 MJ mm ha⁻¹ h⁻¹ year⁻¹), Moderate-Strong (higher than 4,905 and below 7,357 MJ mm ha⁻¹ h⁻¹ year⁻¹), Strong (higher than 7,357 and below 9,810 MJ mm ha⁻¹ h⁻¹ year⁻¹), and Very strong (higher than 9,810 MJ mm ha⁻¹ h⁻¹ year⁻¹). Despite its popularity, no pattern is observed concerning the reference of the original study that proposed this classification for the country. The most cited authors for this purpose are Carvalho (1994), Carvalho (2008), and Santos (2008).

According to this classification, the country is mostly classified with Moderate-Strong and Strong rainfall erosivity (Figure 5a). Considering the different Brazilian regions, the highest percentages of Low and Moderate values are reported for the Northeast region, while the North and Center-West regions have the highest percentages of papers that found Very strong erosivity values (Figure 5b). Despite this, annual erosivity higher than 7,357 MJ mm ha⁻¹ h⁻¹ year⁻¹ (Strong and Very strong categories) are found in all Brazilian territory which evidences the existence of several areas of water erosion susceptibility in the country. Therefore, an increase in the number of studies aiming to estimate local rainfall erosivity values may help identify areas at high erosion risk and support soil and water conservation plans.

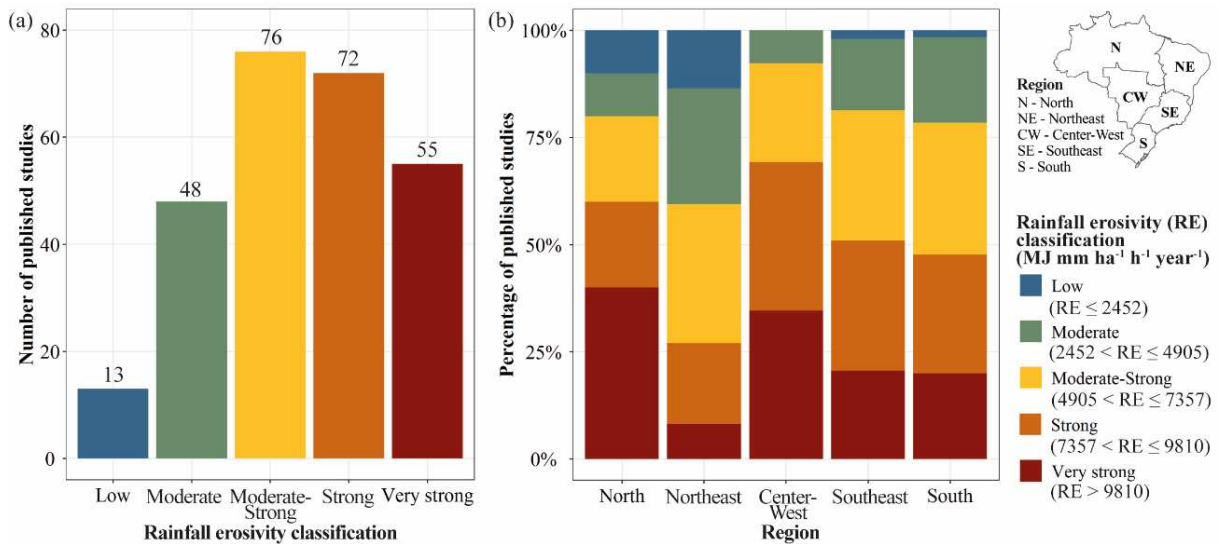


Figure 5. (a) Number of published studies per rainfall erosivity categories. (a) Quantity for the entire country and (b) percentage for each Brazilian region.

3.5. Mapping rainfall erosivity in Brazil

According to the reviewed literature, about 42% of the published papers on rainfall erosivity in Brazil produced erosivity maps (Figure 6a). The most employed interpolation techniques were kriging, artificial neural networks (ANNs), and inverse distance weighting (IDW) (Figure 6b). As discussed later, a small number of studies used other methods such as satellite and climate models' products.

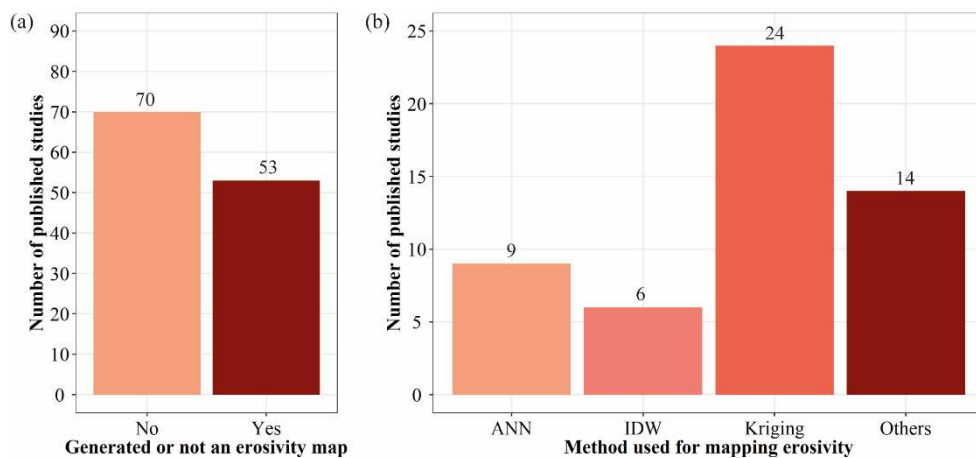


Figure 6. (a) Number of published studies on rainfall erosivity in Brazil that did or did not generate erosivity maps and (b) most employed mapping techniques.

Kriging is the most widespread method for obtaining rainfall erosivity maps in Brazil. This method consists of geostatistical interpolation whose linear weights follow an unbiased constraint and the minimum square error condition. The resulting system of linear equations is solved to determine the estimator's weights considering the stochastic dependence between the data sampled in space (VAROUCHEKIS, 2019). The spatial variation in kriging interpolation is quantified by a semivariogram, which consists of a scatter plot of semivariance versus the distance between the sampled points.

In Brazil, kriging was first used for obtaining rainfall erosivity maps for the state of São Paulo (VIEIRA; LOMBARDI NETO, 1995). Then, this method was also employed to interpolate erosivity in many other studies in state scale, such as for Alagoas (SOUZA et al., 2020b), Espírito Santo (MELLO et al., 2012; SAITO et al., 2009; SILVA et al., 2010c), Mato Grosso (RAIMO et al., 2018), Mato Grosso do Sul (OLIVEIRA et al., 2012), Minas Gerais (MELLO et al., 2007), Paraná (WALTRICK et al., 2015), Rio de Janeiro (MONTEBELLER et

al., 2007), Santa Catarina (BACK; POLETO, 2018), and Tocantins (AVANZI et al., 2019; VIOLA et al., 2014). Local studies also used kriging, but only for locations in southeastern Brazil (AQUINO et al., 2012; MELLO et al., 2015a, 2020; PONTES et al., 2017; SILVA et al., 2010a; TERASSI et al., 2020).

On a national scale, kriging was used by Mello et al. (2013) for obtaining an annual erosivity map. These authors used regression-kriging considering the geographical coordinates and altitude as predictive variables. Also for the entire Brazilian territory, Mello et al. (2015b) tested many interpolation techniques and concluded that the regression-kriging method provided the greatest prediction accuracy and was considered the most reliable for mapping annual rainfall erosivity in the country. Additionally, kriging also showed satisfactory performance for mapping the erosivity patterns in Brazil on a monthly time scale (TRINDADE et al., 2016).

The IDW, another interpolation technique, is a deterministic method in which the value of the estimated variable in any position is calculated by the linear combination of values recorded by the n nearest neighbor points weighted by the inverse of their distance raised to a power (ATTORRE et al., 2007; CARUSO; QUARTA, 1998), as shown in Equation (6). In Brazil, this method was applied for mapping erosivity in the states of Paraíba (SILVA et al., 2013, 2020c), Pernambuco (PINHEIRO et al., 2018), Rio de Janeiro (MACHADO et al., 2013), and São Paulo (TEIXEIRA et al., 2022a).

$$Z_i = \frac{\sum_{j=1}^n \left(\frac{Z_j}{d_{ij}^p} \right)}{\sum_{j=1}^n \left(\frac{1}{d_{ij}^p} \right)} \quad (6)$$

in which Z_i is the rainfall erosivity interpolated at the point i ($\text{MJ mm ha}^{-1} \text{ h}^{-1}$); d_{ij} is the distance between the points i and j ; Z_j is the rainfall erosivity calculated at the point j ($\text{MJ mm ha}^{-1} \text{ h}^{-1}$); p is the power used; n is the number of neighboring points considered for the interpolation.

Differently from the results presented by Mello et al. (2015b), Oliveira et al. (2015) showed that the IDW performed better than the kriging technique to interpolate rainfall erosivity in Brazil. For this, Oliveira et al. (2015) considered erosivity values estimated using synthetic pluviographic series for 142 locations in the country. As stated by these authors, the good performance of IDW is an advantage in modeling erosivity, as it is a method that is well

suiting to irregularly distributed data samples and is a fast and simple technique (it requires few decisions to be made about the model parameters).

The use of machine learning algorithms, more specifically the technique named artificial neural networks (ANNs), is another method employed for mapping erosivity in Brazil. The ANNs are a set of computational elements, called artificial neurons. These neurons are layered and interconnected by weights to form a network and are adjusted as model training progresses. Although there are numerous variant parameters, in ANNs they can be summarized in terms of the characteristics of the activation function, the network architecture, and the training algorithm (KUBAT, 2021). In this way, ANNs provide a learning rule to modify their weights and neurons based on patterns in the input data to result in a specific output estimate.

In Brazil, the use of ANNs to obtain rainfall erosivity maps is limited to the Center-West, Southeast, and South regions. In summary, ANNs were developed for the states of Mato Grosso do Sul (SOBRINHO et al., 2011), Espírito Santo (CECÍLIO et al., 2013; MOREIRA et al., 2012), Minas Gerais (MOREIRA et al., 2008), Rio de Janeiro (CARVALHO et al., 2012), São Paulo (MOREIRA et al., 2006a, 2006b; SILVA et al., 2010b), and Rio Grande do Sul (MOREIRA et al., 2016). As discussed later, for most of these states, computational software was developed to obtain erosivity values using ANNs.

Erosivity maps were obtained using ANNs specifically for the coastal region of the São Paulo state using latitude, longitude, and total rainfall as predictive variables (SILVA et al., 2010b). The results demonstrate the effectiveness of this methodology for the interpolation of annual rainfall erosivity. In this context, Moreira et al. (2006a) also used ANNs to map erosivity, but for the whole state of São Paulo.

Cecílio et al. (2013), using rainfall erosivity values estimated from synthetic rainfall series, evaluated the performance of ANNs, IDW, and kriging techniques for the spatial interpolation of erosivity in the state of Espírito Santo. According to these authors, ANNs had the best values of the Willmott (1981) index. Also, for the Southeast region, Carvalho et al. (2012) used ANNs to map erosivity in the state of Rio de Janeiro. The results were satisfactory using latitude, longitude, and altitude as predictive covariates.

For the Brazilian Center-West region, Sobrinho et al. (2011) developed an ANN capable of mapping with satisfactory accuracy the rainfall erosivity in any location in the state of Mato Grosso do Sul. Finally, for the southern region, Moreira et al. (2016) used ANNs for spatial interpolation of monthly and annual erosivity in the state of Rio Grande do Sul. Using data from 103 rainfall gauges, these authors obtained erosivity maps not only for the EI_{30} index but also for the $KE > 25$. Although the use of the ANN method has intensified in recent years,

the use of other machine learning techniques to interpolate erosivity values in Brazil is still a gap to be filled.

In addition to the previously discussed methods, other studies proposed to obtain maps using satellite products, such as those shown by Silva et al. (2020a) and Moreira et al. (2020), who employed images from the Tropical Rainfall Measuring Mission (TRMM), and those shown by Brito et al. (2021), who used data from the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) products. Spatial data from regionalized global climate models were employed to map projected erosivity patterns over the 21st-century (ALMAGRO et al., 2017; COLMAN et al., 2019; RIQUETTI et al., 2020). As the use of these technology becomes more popular for mapping environmental and climatic variables, the use of satellite and climate models' products for mapping rainfall erosivity is expected to increase over the next years.

3.6. Projected rainfall erosivity over Brazil

Studies proposing to analyze future patterns of rainfall erosivity over the Brazilian territory are scarce. In total, only five papers were published with this objective for Brazil (ALMAGRO et al., 2017; COLMAN et al., 2019; MELLO et al., 2015a; RIQUETTI et al., 2020; ROSA et al., 2016).

The first attempt to estimate future erosivity values in the country was carried out by Mello et al. (2015a). These authors used the products of the HadCM3 model (Hadley Centre Coupled Global-Ocean Model) to estimate rainfall erosivity from 2011 to 2098 for the Grande River Basin (Southeastern Brazil). The future erosivity values considered an intermediary climate scenario (A1B) and were obtained by the MFI, which was calculated using the rainfall projected patterns from the statistical relationship between annual rainfall erosivity obtained by local rainfall gauges and the respective MFI. Mello et al. (2015a) evidenced a perspective of increasing the erosive potential of rainfalls in the studied area, with a significant increase in the MFI values for December and January.

Also considering an intermediate climate change scenario (RCP 4.5), Rosa et al. (2016) projected the rainfall erosivity for the municipality of Rondon do Pará (Northern Brazil) from 2016 to 2035. These authors used projected data from the HadGEM2 model (Hadley Centre Global Environment Model version 2) as input in the erosivity estimation equation established by Morais et al. (2006) and presented by Silva (2004). As result, a decrease in the annual rainfall

erosivity is projected. Although the results obtained by Rosa et al. (2016) contribute to advancing the understanding of how climate change can impact rainfall erosivity in the future, it is important to note that the empirical equation used to calculate future erosivity was established considering the stationarity of rainfalls. In addition, this equation was adjusted for a State in another region of the country and using obsolete data (before 1991).

Almagro et al. (2017) projected rainfall erosivity values for the entire Brazilian territory from 2007 to 2099. For this, data from the HadGEM2 and MIROC5 (Model for Interdisciplinary Research on Climate version 5) models were used, considering the climate scenarios RCP4.5 and RCP8.5. These authors suggest that Northeastern and Southern Brazil will be the most affected regions due to the increasing rainfall erosivity over the century. On the other hand, a decrease in erosivity values is projected for the Southeastern, Central, and Northwestern parts of the country. The results presented by Almagro et al. (2017) seem to be useful for soil and water conservation planning, as evidenced by Colman et al. (2019), who used these projections to evaluate the potential effects of climate change in soil erosion for the Pantanal biome.

Finally, Riquetti et al. (2020) used the projections from the HadGEM2, MIROC5, and CanESM2 (Second Generation Canadian Earth System Model) models to analyze the impact of climate change on the long-term average annual rainfall erosivity in South America considering the RCP8.5 climate scenario. The results obtained by these authors for Brazil are in agreement with several results shown by Almagro et al. (2017). Furthermore, Riquetti et al. (2020) point out a strong reduction trend in the annual erosivity values for the Amazon Forest throughout the century.

3.7. Rainfall erosivity products available for Brazil

In the reviewed literature, products are cited as tools for estimating rainfall erosivity values for Brazil. Some of the most popular ones are from the group of computational software called *netErosividade*, developed using ANNs (MOREIRA et al., 2006a, 2009). The softwares *netErosividadeES* (MOREIRA et al., 2012), *netErosividadeMG* (MOREIRA et al., 2008), *netErosividadeRS* (MOREIRA et al., 2016), and *netErosividadeSP* (MOREIRA et al., 2006b) were developed, respectively, for the states of Espírito Santo, Minas Gerais, Rio Grande do Sul, and São Paulo. By using them, it is possible, in a practical way, to obtain rainfall erosivity values for any location within the respective states. This group of software is freely available

online at the website of the Research Group in Water Resources of the Federal University of Viçosa (GPRH-UFV) (<http://www.gprh.ufv.br/?area=softwares>).

Another tool used for estimating rainfall erosivity values over the Brazilian territory is the software *Chuveros*, developed by professor Elemar Antonino Cassol (Federal University of Rio Grande do Sul-UFRGS). This software was widely used in southern Brazil, as shown by Bazzano et al. (2007), Bazzano et al. (2010), Cassol et al. (2008), Cogo et al. (2006), Eltz et al. (2011), Hickmann et al. (2008), and Mazurana et al. (2009) for some municipalities of the Rio Grande do Sul state, and Schick et al. (2014) for Lages, in Santa Catarina state. Despite the large number of studies that used *Chuveros*, this software is not available for download online, limiting its use for future studies.

Cardoso et al. (2020) developed the *RainfallErosivityFactor* package for the R software environment (R CORE TEAM, 2023) as a tool for the analysis of rainfall data and the calculation of erosivity values (<https://cran.r-project.org/web/packages/RainfallErosivityFactor/index.html>). This package consists of a routine for loading and classifying large rainfall datasets into erosive or non-erosive events, and then using the erosive events to compute rainfall erosivity. The methodology proposed in this package was tested and validated for the municipality of Pirassununga, in São Paulo state, and according to the authors, erosivity values were calculated fast and accurately.

Many papers have calculated erosivity values for Brazil, but most of them do not make the estimates available in online databases. In this context, the rainfall erosivity values database presented by Cecílio et al. (2021) for 141 locations in Brazil can contribute to soil conservation planning applications since they were estimated using the direct method considering pluviographic rainfall data. In addition, these authors also established regression models for obtaining erosivity for most of the locations studied. Thus, this dataset consists of a product with a high potential for use and it is freely available online on a monthly and yearly time scale (<https://ojs.datainscience.com.br/index.php/lads/article/view/37/26>).

3.8. Advancements and prospects

Over the last decade, an increase in the number of erosivity studies was observed in Brazil. With this, some methodologies different from the usual ones are considered advancements in this study field as well as are expected to be more common in the next years. As mentioned before, despite the limitation of not capturing non-stationarity, the use of stochastic weather generators to obtain sub-daily data on rainfall has the potential to increase.

A validated weather generator that considers the specific aspects of the Brazilian rainfall patterns, the ClimaBR, as well as its free availability and ease of use, can enhance the access of rainfall erosivity values in the country.

Another methodological advancement has been the use of machine learning algorithms to estimate rainfall erosivity. In summary, for Brazil, its use has been limited to artificial neural networks. However, as the use of other machine learning techniques becomes more and more common (LEE et al., 2021, 2022), testing a larger group of algorithms is expected to increase. An example can be cited in the study presented by Souza et al. (2022) which analyzed the Random Forest, Cubist, Support Vector Machine, Earth, and Linear Model techniques, associated with topographic, climatic, and vegetation covariates for spatial prediction of rainfall erosivity in Southeastern Brazil. This approach was considered promising as it is a method capable of estimating erosivity in unsampled areas using information available from significant spatial covariates.

A prospect is an increase in the use of remote sensing products to obtain erosivity estimates. As previously discussed, the use of these products for this purpose in Brazil is still limited to some studies, such as those presented by Brito et al. (2021), Moreira et al. (2020), and Silva et al. (2020a) over the last decade. Despite this, the use of remote sensing data products become more popular for estimating rainfall erosivity values around the world (CHEN et al., 2021; MELVILLE; WUDDIVIRA; SUTHERLAND, 2022), so an intensification in the employment of these technologies in Brazil is also probable.

For the next years, an increase in the application of rainfall erosivity values different from the normal use in USLE/RUSLE models is also expected. As an example, it can be mentioned the use of daily erosivity as an indicator to identify areas more susceptible to natural disasters related to rainfall. The concept of rainfall erosivity is based on the kinetic energy of rainfall, rainfall intensity, and maximum rainfall intensity, so it is a climatic index that can be related to damages caused by erosion, landslides, and flooding.

Considering this, Mello et al. (2020) propose the index named R_{maxday} , which is defined as the maximum daily rainfall erosivity and is determined from the maximum daily rainfall. These authors calculated the R_{maxday} for the Mantiqueira Range Region, a mountainous region in Southeastern Brazil, and found that $R_{\text{maxday}} > 2000 \text{ MJ ha}^{-1} \text{ mm h}^{-1} \text{ day}^{-1}$ is a threshold associated with rainfall events that caused fatalities in the region. Also, the areas and months most vulnerable to natural disasters originating from heavy rainfalls were identified for the region using the proposed index.

Similarly, Alves et al. (2022) analyzed the daily erosivity values for a mountainous region of the Rio de Janeiro state. These authors evidenced that the use of the daily estimates proved to be a promising indicator of rainfall disasters, and is considered more effective than those normally used that are only based on the quantity (mm) and/or intensity (mm h^{-1}) of the rainfalls. Thus, Alves et al. (2022) assessed the daily rainfall erosivity as an early warning index for natural disasters. Also for the Rio de Janeiro (municipality), Terassi et al. (2020) found that the sectorial classification of greater and reduced risk areas to the erosive potential of rainfalls in the city is essential in socioenvironmental evaluations and for the prediction of human, social and economic losses. This fact shows that obtaining erosivity values can also be useful not only for rural locations but also for urban areas, especially those located in mountainous regions.

4. Conclusion

This review updated the findings regarding the assessment of rainfall erosivity in Brazil, especially over the last decade, when an increase in the amount of research on this topic was observed. The EI_{30} has been the most employed erosivity index in the country. Pluviographic rainfall data and regression equations are the main methods for obtaining erosivity values. Additionally, the use of synthetic series of rainfall for this purpose has expanded and has great potential for improving the availability of erosivity data in Brazil.

The magnitudes of annual rainfall erosivity reported in the literature for Brazil range from 59 to 26,891 $\text{MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$. The lowest erosivity values are found in the Northeast region and the highest in the North. In Brazil, most published papers assessed erosivity at a local scale, which minimizes the advances in the establishment of soil conservation planning on a national scale. Considering the five geographic regions in the country, the Southeast accounts for the largest number of erosivity studies, while the North constitutes the region with a major lack of erosivity information.

Kriging is the most widespread technique for obtaining rainfall erosivity maps in Brazil. The use ANNs have also substantial importance for mapping erosivity in the country. Through this technique, there was the development of software to obtain erosivity values for some Brazilian States. Despite this, the use of other machine learning techniques for the interpolation of erosivity in Brazil is still a gap to be filled and its use can enable more accurate erosivity maps.

The assessment of the future patterns of rainfall erosivity over the Brazilian territory is still scarce. An increase in the erosivity magnitudes is expected for Northeastern and Southern

Brazil, while a decrease is projected for the Southeastern, Central, and Northwestern parts of the country. Throughout the century, a strong reduction trend in the annual erosivity values for the Amazon Forest is also expected. For future studies, the use of updated climate and land-use scenarios is recommended as well as the quantification of uncertainties to support decision-making on a national scale.

In total, 123 articles on the assessment of rainfall erosivity in Brazil were published. Considering this, the present review complements the review article presented by Oliveira et al. (2012) and brings a wider overview of the erosive dynamics of rainfalls in the country. The information summarized in this paper contributes to understanding erosivity patterns over the Brazilian territory and is relevant for subsidizing the establishment of a conservationist soil and water management.

Data availability

The data produced in this study is freely available online. A complete list of the reviewed articles and a summary of the information retrieved from them is available for download at <http://dx.doi.org/10.17632/hgd2whtx55.1>.

CHAPTER 2: ASSESSMENT, REGIONALIZATION, AND MODELING RAINFALL EROSION OVER BRAZIL: FINDINGS FROM A LARGE NATIONAL DATABASE

Abstract

In this study, we used a large national database to assess the rainfall erosivity (RE) patterns in time and space over the Brazilian territory. Thereby, RE and erosivity density (ED) values were obtained for 5,166 rainfall gauges. Also, the concentration of the RE throughout the year and the RE's gravity center locations were analyzed. Finally, homogeneous regions regarding RE values were delimited and estimative regression models were established. The results show that Brazil's mean annual RE value is $5,620 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$, with considerable spatial variation over the country. The highest RE magnitudes were found for the north region, while the northeast region shows the lowest values. Regarding the RE's distribution throughout the year, in the southern region of Brazil, it is more equitable, while in some spots of the northeastern region, it is irregularly concentrated in specific months. Further analyses revealed that for most of the months, the RE's gravity centers for Brazil are in the Goiás State and that they present a north-south migration pattern throughout the year. Complementarily, the ED magnitudes allowed the identification of high-intensity rainfall spots. Additionally, the Brazilian territory was divided into eleven homogeneous regions regarding the RE patterns and for each defined region, a regression model was established and validated. These models' statistical metrics were considered satisfactory and, thus, can be used to estimate RE values for the whole country using monthly rainfall depths. Finally, all the database produced is available for download. The results here discussed show the most complete panorama of the RE phenomenon in Brazil present in the literature so far. Therefore, the values and maps shown in this study are relevant for improving the accuracy of soil loss estimates and for the establishment of soil and water conservation planning on a national scale.

Keywords: Erosivity index. R-factor. Soil erosion. Universal Soil Loss Equation. Soil and water conservation.

1. Introduction

Rainfall erosivity (RE) is defined as the potential capacity to cause soil erosion. It is a parameter of great importance for soil and water conservation planning. This variable is included in the main soil loss prediction models, such as the Universal Soil Loss Equation (USLE) and its revised version (RUSLE) (RENARD et al., 1997; WISCHMEIER; SMITH, 1978). Its assessment, in time and space, is the basis for detecting the effects of climate on long-term erosion rates in a region. As demonstrated by Jia et al. (2022) and Shi et al. (2021), high RE values are strongly linked to high soil erosion rates. Thus, the estimation of this variable on a national scale may help to identify areas with great erosion potential (CHALISE; KUMAR; KRISTIANSEN, 2019; LIU et al., 2020a).

RE has been assessed all over the world (LIU et al., 2020b; PANAGOS et al., 2022). In Brazil, the estimation of RE values was first reported in the literature by Lombardi Neto et al. (1992), for the municipality of Campinas, in São Paulo state. Since then, many studies with this purpose have been developed, mostly on a local scale (TEIXEIRA et al., 2022b). Some national estimates of RE values were done for Brazil, such as those presented by Cecílio et al. (2021), Oliveira et al. (2018), and Trindade et al. (2016). Although the results presented in these papers constitute an advance regarding the availability of RE magnitudes for the country, these studies considered a limited database, especially regarding the continental dimension of Brazil. Consequently, the use of large databases for obtaining RE values for the entire country is a gap to be filled.

The erosivity density (ED) is defined as the erosivity content per rainfall unit and can express the variations in the rainfall intensity of a particular gauge (FOSTER, 2008). The ED concept is relatively new and was used to complement the assessment of RE in some studies in Austria (JOHANNSEN et al., 2022), China (ZHU; XIONG; XIAO, 2021), Greece (VANTAS; SIDIROPOULOS; LOUKAS, 2019), India (DAS; JAIN; GUPTA, 2022), Italy (DIODATO et al., 2021), and South Korea (SHIN et al., 2019). In Brazil, this variable was only estimated by Teixeira et al. (2022a), for the São Paulo state. These authors state that ED analysis can identify areas susceptible to the impacts of the most extreme rainfall events even in months with low RE magnitudes. As a result, obtaining ED values on a national scale could be useful for establishing more efficient soil conservationist planning in Brazil.

In the last few years, some indexes were used to complement the assessment of RE (DI LENA; CURCI; VERGNI, 2021). Among them, concentration indexes aiming to analyze the distribution of the RE magnitudes throughout the year were applied in China (HUANG et al.,

2019; ZHU; XIONG; XIAO, 2021) and Italy (DI LENA; CURCI; VERGNI, 2021). As shown by Zhu et al. (2021), these indexes can help to identify regions with strong irregular distribution of RE values. Furthermore, some studies employed the gravity center model concept to assess how the RE phenomenon occurs in a location (GUO et al., 2019; ZHU; XIONG; XIAO, 2021). Therefore, the consideration of concentration indexes and the definition of gravity centers have the potential to improve the understanding of the RE patterns over the Brazilian territory.

As explained by Teixeira et al. (2022b), using regression models to estimate RE values has been the most employed alternative to the lack of sub-daily rainfall data in Brazil. These models are considered easy to use since they usually require only rainfall depths data, however, due to the unavailability of these equations for the entire country many studies have used them inappropriately (e.g. Falcão et al. (2020) and Sousa et al. (2019)). To overcome this problem, Teixeira et al. (2022a) suggest that regionalized models can be helpful, and proposed homogeneous regions regarding RE values for the state of São Paulo, as well as validated regression models for each of them. Thus, considering that the amount of erosivity models for Brazil is still scarce, a regionalized approach may improve the availability of RE estimates in the country.

This paper aims to expand the availability of RE values in Brazil as well as to assess its patterns in time and space, using a large national database. The specific objectives of this study were: i) to estimate and analyze RE values for 5,166 rainfall gauges over the Brazilian territory; ii) to calculate ED values for each gauge analyzed; iii) to assess the distribution of the RE throughout the year using erosivity concentration indexes; iv) to locate the gravity centers for RE; v) to delimitate homogeneous regions regarding RE; vi) to establish models to estimate RE for each defined region.

2. Material and methods

Figure 1 shows the flowchart of the methodological steps adopted in the present study which is better described in the following sections.

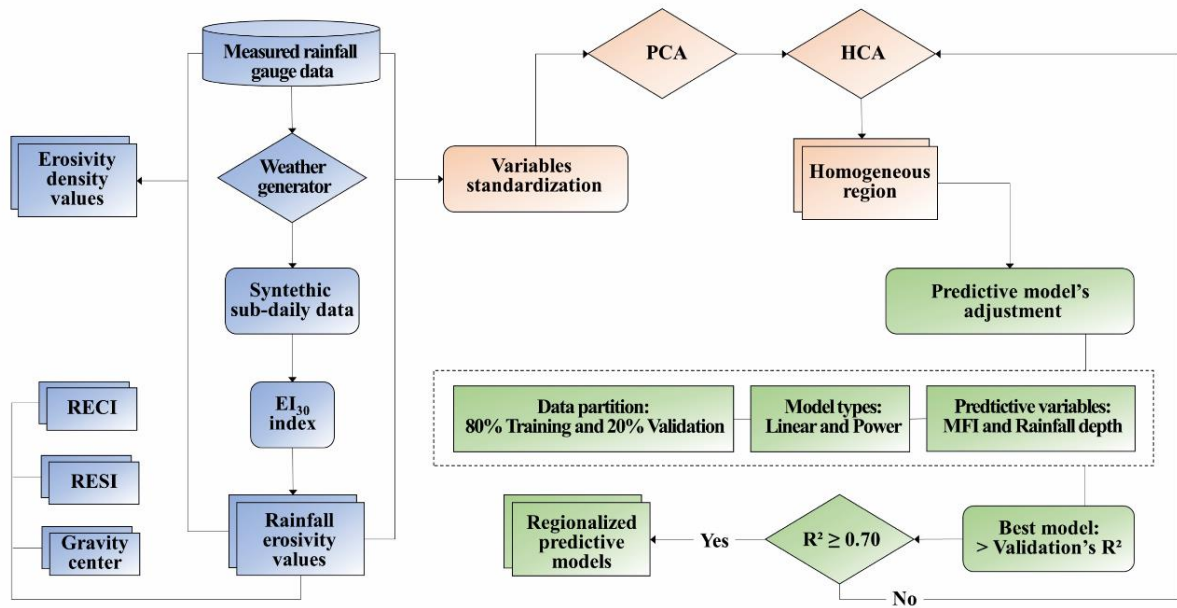


Figure 1. Flowchart of the methodological steps adopted in the present study. RECI and RESI mean, respectively, rainfall erosivity concentration index and rainfall erosivity seasonality index. PCA and HCA denote, respectively, principal component analysis and hierarchical cluster analysis. R^2 means coefficient of determination.

2.1. Study area and database

Covering about 8,511,000 km², the Brazilian territory occupies 47.3% of South America, and has altitudes ranging from zero, in most coastal areas, to 2,800 m, in the region of the Guianas Plateau, north of the state of Amazonas at the border with Venezuela. As shown in Figure 2a, Brazilian altitudes range mostly from 200 to 800 m.

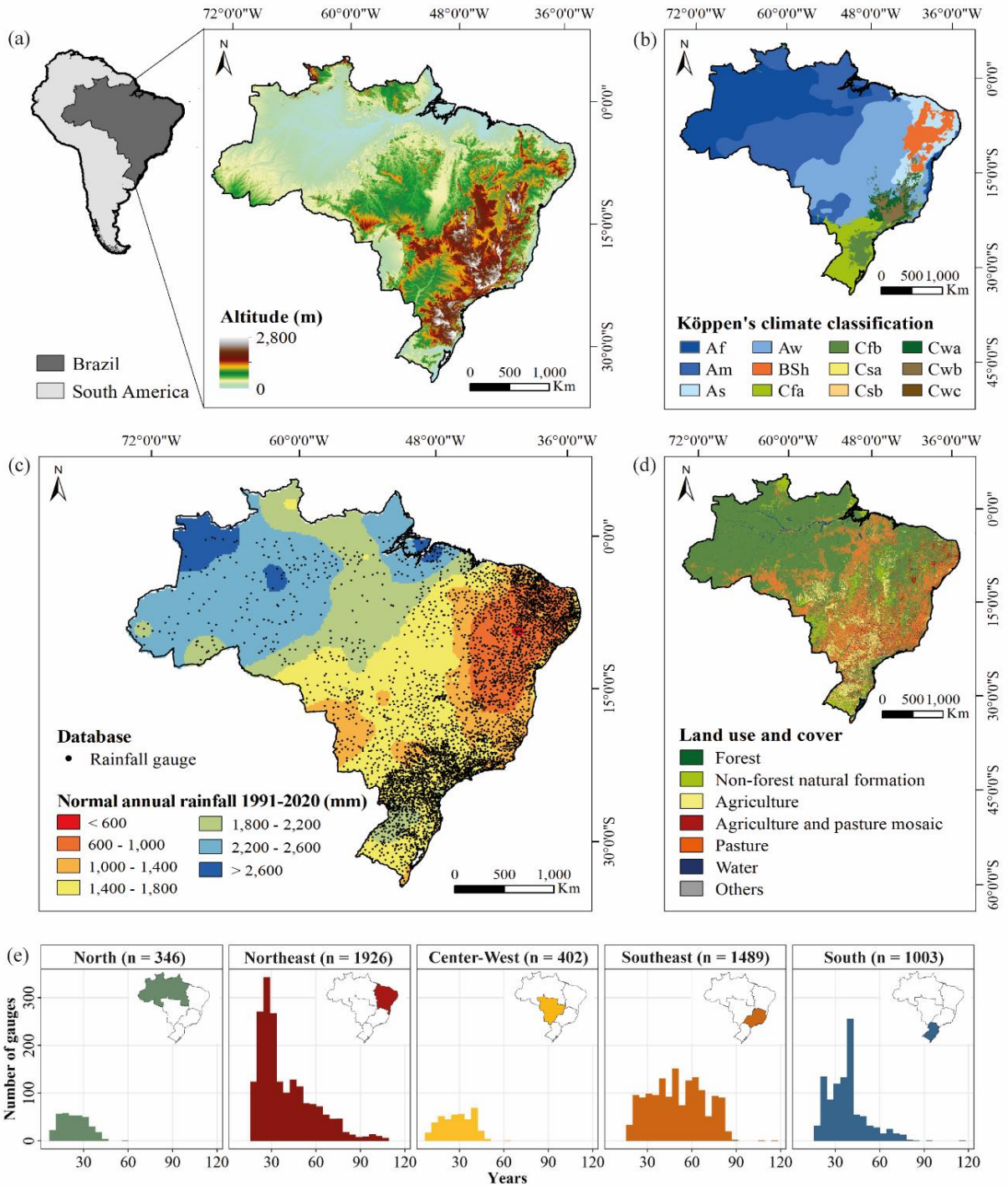


Figure 2. (a) Location of Brazil in South America and its altitude variation (SRTM-DEM 30-m resolution). (b) Köppen (1936)’s climate classification presented by Alvares et al. (2013). (c) Spatial distribution of the normal annual rainfall 1981-2010 (INMET, 2022) and location of the rainfall gauges considered in this study. (d) Land use and cover provided by the MapBiomas Project (2021). (e) Histogram of the number of years presented in the series of the gauges considered per region.

Brazil has an expressive climate variability (Figure 2b), among which the tropical subtypes (Af, Am, As, and Aw) stand out, as shown by Alvares et al. (2013). In the northeast region, the semi-arid climate (Bsh) is more representative, while the subtropical humid climate subtypes (Cfa, Cfb, Csa, Csb, Cwa, and Cwb) stand out in the south and southeast. Regarding annual rainfall, it ranges from 380 to 4,000 mm (ALVARES et al., 2013; INMET, 2022). Annual rainfall above 2,000 mm is observed mostly in northern Brazil, while annual totals below 1,000 mm occur in the central strip of the northeast region (Figure 2c).

Concerning land use and cover, forest and pasture areas predominate over the Brazilian territory (Figure 2d). They occupy about 59.7 and 18.2% of the country's total area, respectively (MAPBIOMAS PROJECT, 2021). Agricultural areas correspond to 6.5% of Brazil, in which soybean, corn, sugarcane, and cotton stand out due to their economic importance (BRASIL, 2022). Regarding other uses and covers, areas with non-forest natural formations stand out, covering 6.6% of the country.

Daily historical series of rainfall data measured at 5,166 rainfall gauges (Figure 2c) were used for the present study. These data were obtained from the Hidroweb portal (<http://www.snirh.gov.br/hidroweb/serieshistoricas>). As shown in Figure 2e, the minimum time length of the series used was defined as 20 years of data for the Northeast, Southeast, and South regions. For the North and Center-West regions, the minimum time length was defined as 10 years, since the availability of rainfall gauges is still scarce (SILVA, 2021). Finally, the mean time length considering all rainfall gauges was 40.5 years.

2.2. Rainfall erosivity estimation

On ClimaBR stochastic weather generator (BAENA et al., 2005; OLIVEIRA; ZANETTI; PRUSKI, 2005a, 2005b; ZANETTI et al., 2005), version 2.0, the measured daily data from each rainfall gauge was individually inserted. ClimaBR generates synthetic series of pluviographic data on a sub-daily scale. The use of the synthetic series generated by ClimaBR for estimating RE values was proposed and validated by Oliveira et al. (2018) for the entire Brazilian territory. This software was developed considering specifically the climatic conditions of Brazil and was validated regarding the number of rainy days and the total daily precipitation by Zanetti et al. (2005).

The synthetic series were generated to contain a total of 100 years in length, as presented by Teixeira et al. (2022a). These series show information that characterizes the daily rainfall

profile such as the daily amount and its duration, the maximum instantaneous precipitation intensity, and its time of occurrence.

From the synthetic series of pluviographic data obtained for each rainfall gauge, RE was estimated based on the criteria proposed by Wischmeier and Smith (1958) and Wischmeier (1959), and modified by Cabeda (1976). For the pluviographic series, the rainfall events that were considered erosive were identified day by day. Erosive rainfall is considered as rainfall events with 10 mm or more of total volume, or with a total of less than 10 mm when the amount precipitated in 15 minutes is equal to or greater than 6 mm.

Once identified the erosive rainfalls, the kinetic energy (KE) associated with them was calculated as a function of the rainfall intensity, from minute to minute, using Equation (1), suggested by Wischmeier and Smith (1958) and readjusted to the International System of Units by Foster et al. (1981). These authors state that the KE corresponds to $0.283 \text{ MJ ha}^{-1} \text{ mm}^{-1}$ when the rainfall intensity is higher than 76 mm h^{-1} . The KE associated with the rainfall of each day was calculated by adding the kinetic energies of each minute until the total duration of the rainfall.

$$KE = 0.119 + 0.0873 \log I \quad (1)$$

in which KE is the kinetic energy ($\text{MJ ha}^{-1} \text{ mm}^{-1}$); I is the intensity of the rainfall (mm h^{-1}).

To represent the erosivity of rainfalls, the EI_{30} erosivity index was used, as proposed by Wischmeier and Smith (1958). The daily EI_{30} was calculated by the product of the kinetic energy of each rainfall of the day and the maximum intensity of precipitation that occurred in 30 minutes (I_{30}), as shown in Equation (2).

$$(EI_{30})_j = KE \cdot I_{30} \quad (2)$$

in which $(EI_{30})_j$ is the rainfall erosivity index in day j ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ day}^{-1}$); I_{30} is the maximum rainfall intensity for 30 consecutive minutes (mm h^{-1}).

The RE on the monthly scale was determined by the sum of the daily EI_{30} values as shown in Equation (3). Next, the annual RE values were obtained from the sum of the RE for each month, as shown in Equation (4). Finally, the long-term mean values of monthly and

annual RE were obtained for each rainfall gauge from the average of the values considering all years of the generated series (100 years).

$$RE_m = \sum_{j=1}^n (EI_{30})_j \quad (3)$$

$$RE_a = \sum_{m=1}^{12} (RE_m) \quad (4)$$

in which RE_m is the rainfall erosivity in the month m ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$); n is the number of days in the month m ; RE_a is the annual rainfall erosivity ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$).

2.3. Rainfall erosivity concentration

To characterize the non-uniformity of the RE distribution throughout the year, we used the rainfall erosivity concentration index (RECI) and the rainfall erosivity seasonality index (RESI), as shown respectively in Equations (6) and (7).

$$RECI = \frac{\sum_{m=1}^{12} \overline{RE}_m^{-2}}{(\sum_{m=1}^{12} \overline{RE}_m)^2} \times 100 \quad (6)$$

$$RESI = \frac{1}{\overline{RE}_a} \sum_{m=1}^{12} \left| \overline{RE}_m - \frac{\overline{RE}_a}{12} \right| \quad (7)$$

in which \overline{RE}_m is the mean monthly rainfall erosivity in the month m ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$); \overline{RE}_a is the mean annual rainfall erosivity ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$).

The RECI (DI LENA; CURCI; VERGNI, 2021; ZHU; XIONG; XIAO, 2021) is an adaptation of the rainfall concentration index proposed by Oliver (1980). Theoretically, this index ranges from 8.3 for equal monthly increments (uniform distribution) to 100 for extreme monthly distribution (strongly irregular distribution). Similarly, the RESI (DI LENA; CURCI; VERGNI, 2021) was adapted from the rainfall seasonality index proposed by Walsh and Lawler (1981) to represent the RE distribution within a year in the present study. This index can in theory vary from zero (if all the months have equal amounts) to 1.83 (if all the amount occurs

in a single month). The description of the qualitative classification for the RECI and RESI values is shown in Table 1.

Table 1. Qualitative classification for the RECI and RESI indexes.

Index	Class	Description
RECI	<10	Uniform distribution
	10–15	Moderate distribution
	15–20	Irregular distribution
	>20	Strongly irregular distribution
RESI	≤0.19	Very equable
	0.20–0.39	Equable but with a definite wetter (erosive) season
	0.40–0.59	Rather seasonal with a short drier (non-erosive) season
	0.60–0.79	Seasonal
	0.80–0.99	Markedly seasonal with a long drier (non-erosive) season
	1.00–1.19	Most precipitation (erosivity) in 3 months
	≥1.20	Extreme seasonality, with almost all precipitation (erosivity) in 1–2 months

The description between parenthesis represents the adaptation of the original interpretation proposed by Oliver (1980) and Walsh and Lawler (1981), to the rainfall erosivity context.

2.4. Rainfall erosivity gravity center model

The concept of gravity center derives from the field of physics and refers to the point at which the force of gravity is exerted equally on each part of an object (GUO et al., 2019). In the present study, this object is considered as the entire Brazilian territory. The spatial variation characteristics of a gravity center can reflect the degrees of variation and trends of a geographical phenomenon. Considering this, the gravity center concept was employed to analyze the RE phenomenon in some studies, such as the ones from Guo et al. (2019) and Zhu et al. (2021). Here, the gravity center of the RE values on monthly, seasonal, and annual scales were calculated using Equations (8) and (9) to characterize the spatial distribution of this phenomenon in different time scales.

$$X = \frac{\sum_{g=1}^n X_g \overline{RE}_g}{\sum_{g=1}^n \overline{RE}_g} \quad (8)$$

$$Y = \frac{\sum_{g=1}^n Y_g \overline{RE}_g}{\sum_{g=1}^n \overline{RE}_g} \quad (9)$$

in which X and Y are, respectively, the longitude and latitude of the gravity center on the monthly, seasonal, and annual scale (decimal degrees); X_g and Y_g are, respectively, the

longitude and latitude of the rainfall gauge g (decimal degrees); \overline{RE}_g is the mean rainfall erosivity in the rainfall gauge g ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ time unit}^{-1}$); n is the number of rainfall gauges considered (total of 5,166 gauges).

2.5. Erosivity density estimation

To express the erosivity content per unit of rainfall, the variable erosivity density (ED) was calculated by the ratio between mean monthly RE and mean monthly total rainfall (FOSTER, 2008), as shown in Equation (5). High ED values suggest that rainfall is characterized by high-intensity events of short duration in a specific period (PANAGOS et al., 2015).

$$ED_i = RE_i \cdot R_i^{-1} \quad (5)$$

in which ED_i is the mean erosivity density in the month i ($\text{MJ ha}^{-1} \text{ h}^{-1}$); R_i is the mean monthly rainfall depth in the month i (mm); RE_i is the mean monthly rainfall erosivity in the month i ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$).

2.6. Establishing regionalized models to estimate rainfall erosivity

In summary, to establish regionalized models to estimate RE in Brazil, three steps were followed: 1) the obtention of new variables from the initial dataset using principal component analysis (PCA); 2) the definition of homogeneous regions regarding RE using hierarchical clustering analysis (HCA); 3) the adjustment of regression models for each defined region based on the correlation between RE and rainfall depths. These procedures are explained in detail in the following subsections. This methodology was partially used by Teixeira et al. (2022a) to establish RE regionalized models for the São Paulo state, in southeastern Brazil.

2.6.1. Dataset obtention

For the initial dataset, the mean monthly and annual RE values and mean monthly and annual rainfall depths were used, totaling 26 variables for each rainfall gauge. These variables were standardized regarding their scale using Equation (6). This procedure was applied to eliminate interferences in the results due to the different magnitudes of the data.

$$Xn' = \frac{Xn - \text{mean}(X)}{\text{SD}(X)} \quad (6)$$

in which Xn' is the standardized value of Xn ; Xn is a value observed in an array of values for a given variable X ; $\text{mean}(X)$ is the mean value of the variable X ; $\text{SD}(X)$ is the standard deviation of the variable X .

Using the standardized variables, the PCA was performed to obtain a new set of variables from the contribution of the initial variables. As explained by Hair et al. (2009) and Kassambara (2017), PCA is a statistical procedure used to extract relevant information from a multivariate database using orthogonal transformation to convert a set of correlated observations into linearly uncorrelated variables, called principal components (PCs). This conversion is conducted such that the first PC explains the largest portion of the variability of the data, and each of the next PCs explains less variability than the previous one.

Considering this, only PCs with an eigenvalue greater than the unity (>1.0) were considered for selection. This criterion is based on the fact that any component must explain a variability higher than that presented by a single standardized variable (BERTOSSO et al., 2013; HAIR; BLACK; SANT'ANNA, 2009). Thus, each PC selected consisted of a new variable obtained from the contribution of the initial variables. Therefore, this new dataset presented a greater explanatory power concerning the total variability of the initial data since the firsts PCs retain the majority of the data information (DEMŠAR et al., 2013).

These analyses were performed in the R software version 4.0.2 (R CORE TEAM, 2023), using the *FactoMineR* (LÊ; JOSSE; HUSSON, 2008) and *factoextra* (KASSAMBARA; MUNDT, 2017) packages.

2.6.2. Definition of homogeneous regions

The similarity between the RE patterns observed in the 5,166 rainfall gauges was assessed through clustering analysis. For this, HCA was employed considering the dataset composed of the PCs generated by PCA. Additionally, the latitude and longitude values of each rainfall gauge were considered for grouping. In the HCA, Euclidean distance was used as a measure of the similarity, and Ward's method (WARD JR, 1963) was used as an agglomerative

hierarchical technique, according to Dehghan et al. (2018), Terassi and Galvani (2017), and Rodriguez et al. (2016). The HCA was performed using the *Stat* (BOLAR, 2019) and *Dendextend* (GALILI, 2020) packages in the R software environment.

As a result of these procedures, each rainfall gauge was classified as belonging to a homogeneous region (cluster). Subsequently, the influence area of each rainfall gauge was obtained using the Thiessen polygons' method to generate maps with the spatial distribution of each homogeneous region defined, as performed by Shirin and Thomas (2016).

2.6.3. Establishment of the estimation models

For each defined region, regression models to estimate RE values were established. So, only the data of the rainfall gauges inserted within the respective region was considered for the model's adjustment.

For this, two variables were tested as explanatory variables in the equations: the rainfall depth (R) and the modified Fournier index (MFI) (LOMBARDI NETO, 1977; LOMBARDI NETO; MOLDENHAUER, 1992). The MFI, obtained from Equation (7), relates the mean monthly rainfall and the mean annual rainfall and, as pointed out by Oliveira et al. (2012), is widely used as an independent variable for obtaining erosivity estimation models in Brazil.

$$MFI_i = \frac{(R_i)^2}{R_a} \quad (7)$$

in which MFI_i is the modified Fournier index for the month i ; R_i is the mean monthly rainfall in the month i (mm); R_a is the mean annual rainfall (mm).

Furthermore, two model types were adjusted: the linear and the power model, as shown in Equations (8) and (9).

$$RE_m = \alpha + \beta(R_m) \quad \text{or} \quad RE_m = \alpha + \beta(MFI_m) \quad (8)$$

$$RE_m = \alpha(R_m)^\beta \quad \text{or} \quad RE_m = \alpha(MFI_m)^\beta \quad (9)$$

in which RE_m is the rainfall erosivity in the month m ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$); R_m is the rainfall depth in the month m (mm); MFI_m is the modified Fournier index in the month m ; α and β are the regression coefficients that describe the relationship between RE_m and R_m or MFI_m .

The coefficient of determination (R^2) of the models' adjustment was used as criteria to stop the homogeneous regions' subdivisions. The HCA was employed to obtain clusters that the rainfall gauges within it have an adjustment's R^2 between the RE and the explanatory variables equal to or greater than 0.7. Thus, if a cluster (region) presented a model with an R^2 value equal to or greater than 0.7, the model was considered satisfactory, otherwise (R^2 lower than 0.7) the cluster was subdivided and a new model was adjusted. This process was repeated until all regions had a model considered satisfactory.

The R^2 values were analyzed for two datasets: the training and the validation dataset. For this, the rainfall gauges within each homogenous region were divided, randomly, in which 80% of the gauges were used for training the model, and 20% were used for its validation. Considering the combination of the two explanatory variables (R and MFI) and the two model types (linear and power) tested, the chosen model for each region was the one with the higher R^2 value.

Finally, the monthly and annual RE values estimated (E_i) by the models established for each region were compared to the observed RE values (O_i), estimated using the synthetic sub-daily data (section 2.2). For this comparison, the statistical evaluators used were: the mean absolute error; the mean absolute percentage error; the root mean square error; and Willmott's agreement index (WILLMOTT, 1981), as shown respectively in Equations (10) to (13). Also, a percentage error (Equation (14)) map was produced for the estimated annual RE values.

$$MAE = \frac{\sum_{i=1}^n (|O_i - E_i|)}{n} \quad (10)$$

$$MAPE = \frac{\sum_{i=1}^n \left(\frac{|O_i - E_i|}{O_i} \right)}{n} 100 \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (E_i - O_i)^2}{n}} \quad (12)$$

$$d = 1 - \left[\frac{\sum_{i=1}^n (E_i - O_i)^2}{\sum_{i=1}^n (|E_i - \bar{O}_i| + |O_i - \bar{O}_i|)^2} \right] \quad (13)$$

$$PE = \left(\frac{O_i - E_i}{O_i} \right) 100 \quad (14)$$

in which *MAE* is the mean absolute error ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ time unit}^{-1}$); *MAPE* is the mean absolute percentage error (%); *RMSE* is the root mean square error ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ time unit}^{-1}$); *d* is Willmott's agreement index (dimensionless); *PE* is the percentage error (%); *O_i* is the monthly or annual RE value estimated from the synthetic sub-daily rainfall series ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ time unit}^{-1}$); *E_i* is the monthly or annual RE value estimated using the models established for each homogeneous region ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ time unit}^{-1}$); *n* is the number of observations.

3. Results and discussion

3.1 Rainfall erosivity values

The annual and monthly RE values obtained in the present study for each of the 5,166 rainfall gauges used are available for download at <http://dx.doi.org/10.17632/hzxfvvmr6p.1> and are useful for checking RE values for a specific location or region of Brazil.

Considering the entire Brazilian territory, a mean annual RE value of $5,620 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$ was observed (Table 2). The annual values ranged from 252 to $23,916 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$, with a coefficient of variation (CV) of 52.5%, which expresses the large variation of the annual RE values over the country (Figure 3a). Also, the minimum and maximum values found in the present study, expanded the range of annual RE values for Brazil found in national studies, such as those found by Cecílio et al. (2021) with a minimum value of $338 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$, and by Mello et al. (2013) with a maximum value of $23,400 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$. The obtention of values that were not previously found in the literature is probably due to the larger number of rainfall gauges analyzed. In previous studies on a national scale, the number of gauges considered was limited from only 141 (CECÍLIO et al., 2021) to 1,600 (SILVA, 2004) gauges.

Table 2. Descriptive statistics of annual and monthly RE and ED values for Brazil expressed respectively in MJ mm ha⁻¹ h⁻¹ time unit⁻¹ and MJ ha⁻¹ h⁻¹.

Variable	Statistics	Annual	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
RE	Mean	5620	820	762	766	538	368	213	167	138	222	394	536	698
	SD	2950	498	437	433	353	276	181	157	128	182	282	363	456
	Min.	252	1	5	16	14	0	0	0	0	0	0	0	0
	Max.	23916	3260	3469	4402	4050	4256	2488	2543	2496	3563	2581	2464	3204
	Range	23664	3260	3464	4387	4036	4256	2488	2543	2496	3563	2581	2464	3204
	Median	6027	887	771	675	388	259	127	77	72	165	404	525	682
	CV (%)	52.5	60.7	57.3	56.5	65.5	75	85.3	93.8	93.1	82.1	71.7	67.7	65.4
ED	Mean	4.1	4.4	4.6	4.5	4.2	3.7	3.1	2.8	2.8	3.0	3.7	4.1	4.3
	SD	1.9	2.0	2.0	2.0	2.0	1.9	1.9	1.8	1.8	1.8	2.0	2.1	1.9
	Min.	0.9	0.5	0.5	0.8	0.7	0.3	0	0	0	0	0	0	0
	Max.	9.7	12.4	12.0	11.1	10.4	11.2	10.9	12.3	12.7	11.3	12.0	13.4	11.4
	Range	8.8	11.9	11.5	10.3	9.6	10.9	10.9	12.3	12.7	11.3	12.0	13.4	11.4
	Median	4.6	5.2	5.4	5.2	4.6	3.9	3.0	2.8	2.8	3.2	3.9	4.6	4.8
	CV (%)	45.7	44.1	43.4	45.4	47.4	52.1	60.2	64.9	64.0	59.1	54.1	50.3	45.4

SD stands for standard deviation and CV stands for coefficient of variation.

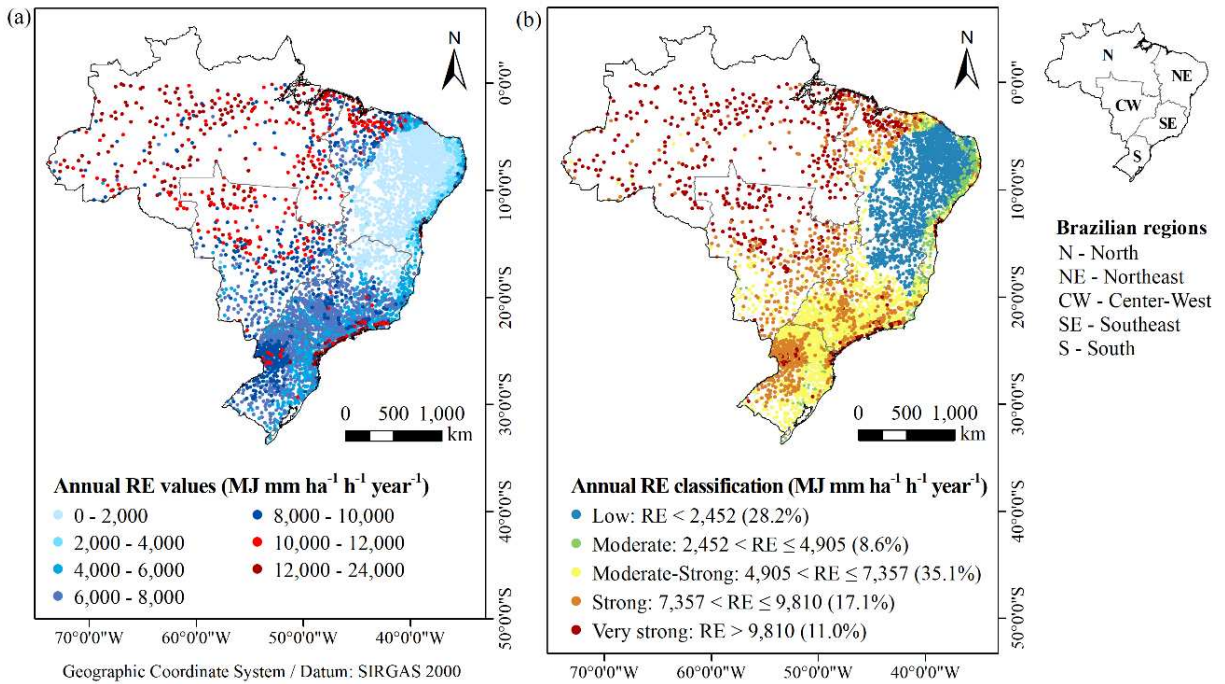


Figure 3. (a) Point distribution of the annual RE values for Brazil and (b) its classification according to Carvalho (2008).

As shown in Figure 3a, the highest annual RE values are in north and center-west regions, where magnitudes above $12,000 \text{ MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$ are common. On the other hand, the northeast region has the lowest RE values, around $2,000 \text{ MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$ or less. According to the classification proposed by Carvalho (2008) and presented by Oliveira et al. (2012), most of the analyzed rainfall gauges (35.1%) have RE values considered “Moderate-Strong” (Figure 3b). Gauges inserted in this category are mostly in the southeastern and southern regions of Brazil. Considering this same classification method, Silva (2004) did not identify any area in the country with annual RE values considered “Low”, but for the present study, this category represents 28.2% of the analyzed data. This can be explained by the greater number of gauges considered for the northeast region in the present study when compared to Silva (2004).

Considering the different climate types over the Brazilian territory, the tropical subtypes Af (without dry season) and Am (monsoon) had the most expressive average magnitudes (Figure 4a). The mean values found, around $10,000 \text{ MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$, are superior to those presented by Panagos et al. (2017a) for areas with these same climates. The Af and Am subtypes, which are characterized by the absence of months without rainfall (ALVARES et al., 2013), could increase the chances of erosive rainfall occurrence. On the other hand, the lowest mean values in the present study were found for the semiarid climate (Bsh), whose annual

rainfall depths are usually below 800 mm. Finally, for most of the other climatic subtypes in Brazil (Aw, Cfa, Cfb, Cwa, and Cwb) the mean annual RE values varied around 7,000 MJ mm ha⁻¹ h⁻¹ year⁻¹.

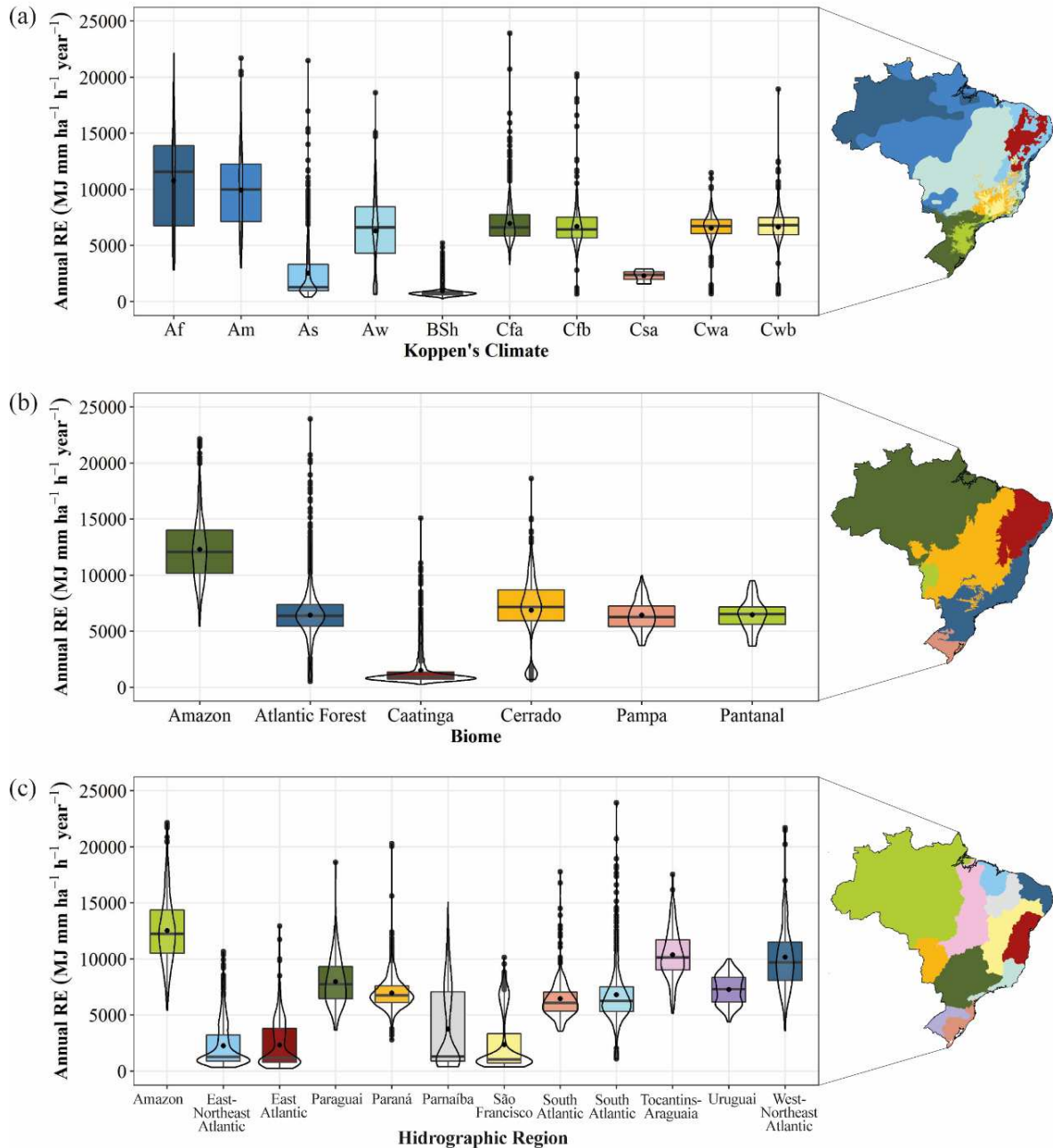


Figure 4. Annual RE values per (a) climate according to the climatic classification of Köppen (1936) carried out by Alvares et al. (2013), (b) biome (IBGE, 2019), and (c) hydrographic region (IBGE, 2021), for Brazil.

In the Amazon biome, which mostly corresponds to the Amazon hydrographic region, the annual RE can reach values over 20,000 MJ mm ha⁻¹ h⁻¹ year⁻¹ (Figures 4b and 4c), classified

as “Very Strong” RE values. Such magnitudes are also found in the Atlantic Forest biome as well as in Paraná, South Atlantic, and West-Northeast Atlantic hydrographic regions. For the Amazon and Pantanal biomes, the studies from Silva et al. (2020b) and Machado et al. (2014), respectively, found annual RE maximum values superior to the ones found in the present paper. Differently, Castagna et al. (2022) analyzed the RE values for the Cerrado biome and the magnitudes corroborate the ones found in the present paper. Studies like this are important to soil conservationist programs since the Cerrado biome has most of the Brazilian agricultural production (GOMES et al., 2019a; SPERA, 2017).

As shown in Figure 5, considering Pearson’s correlation coefficient (r) between the annual RE values found for all rainfall gauges, and the altitudes and latitudes, a very slight correlation was found (-0.06 and 0.21, respectively). However, a closer relationship between the longitudes and annual RE values was evidenced, with an $r = 0.70$. Thus, a direct relation to these variables was found, which shows that the more continental (far from the coast) the rainfall gauge is, the higher the RE magnitudes. This was also found by Silva (2004) and Oliveira et al. (2012). Furthermore, the annual erosivity values are strongly correlated to the annual rainfall depths ($r = 0.91$). The correlation between the annual RE values and the variables altitude, latitude, longitude, and rainfall depth were considered significant at the 0.001 level.

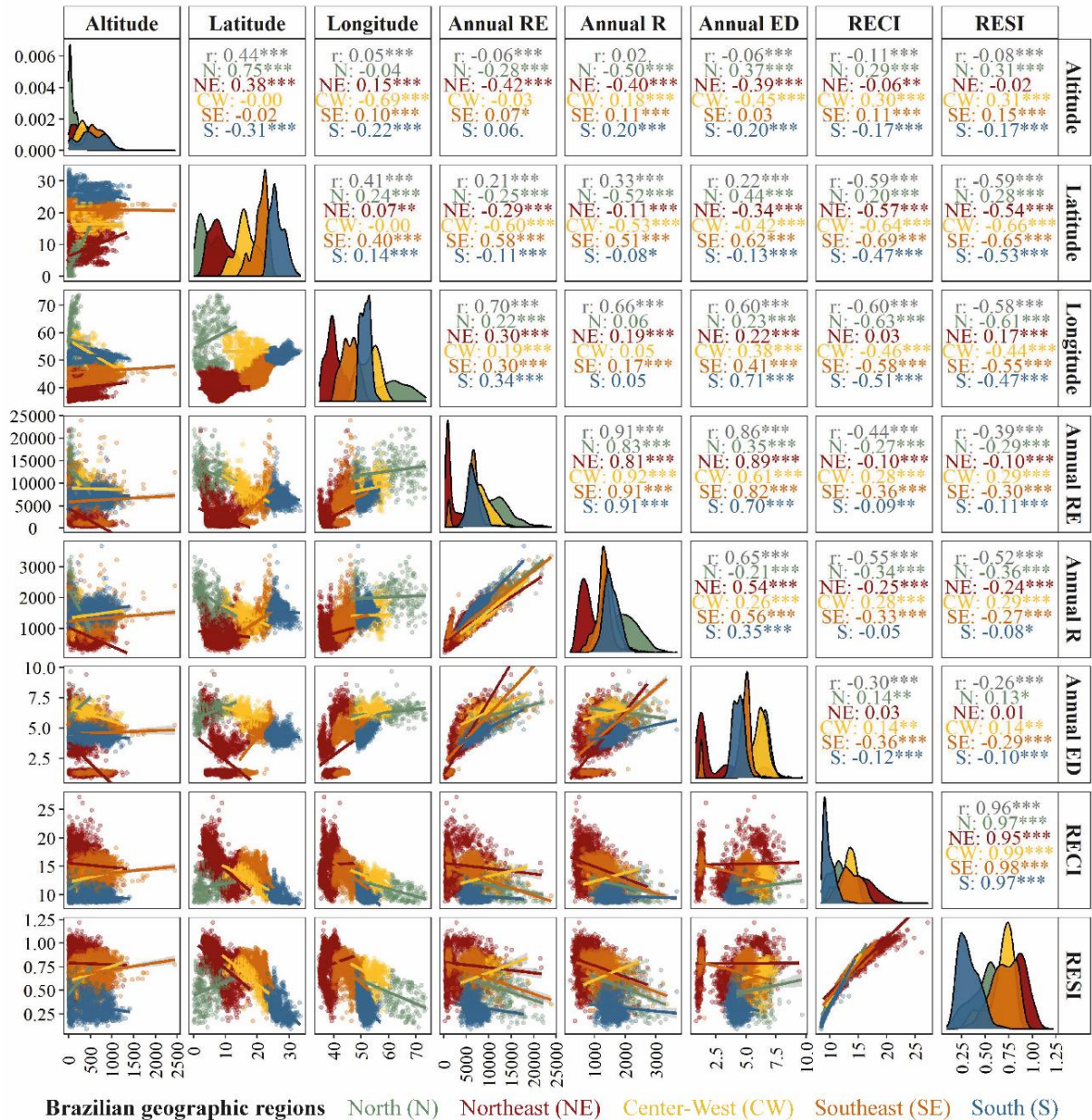


Figure 5. Correlogram presenting Pearson's correlation coefficient (r) between the variables altitude (m), latitude (degrees south), longitude (degrees west), annual RE ($\text{MJ mm ha}^{-1} \text{h}^{-1} \text{year}^{-1}$), annual rainfall depth (R, mm), annual ED ($\text{MJ ha}^{-1} \text{h}^{-1}$), RECI (dimensionless), and RESI (dimensionless), as well as their scatter and density distribution plots, for the entire country and the north (N), northeast (NE), center-west (CW), southeast (SE), and south (S) regions. Significant correlation was found if p -value < 0.001 (***), < 0.01 (**), < 0.05 (*), and < 0.10 (.)

The mean monthly RE values over the Brazilian territory range from 138 to 820 $\text{MJ mm ha}^{-1} \text{h}^{-1} \text{month}^{-1}$ respectively in August and January (Table 2). The months with the highest maximum values range from March to May when magnitudes can be over 4,000 MJ mm ha^{-1}

$\text{h}^{-1} \text{ month}^{-1}$. Furthermore, RE values significantly reduce from June, whose the highest magnitudes are 41.5% lower than the maximum recorded in the previous month.

From June to August, it is observed the highest contrast in the monthly RE values over the country (CV higher than 85%). As shown in Figure 6, it can be explained by the low erosivity values ($245 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$) in most of the gauges, while in the Amazonas, Amapá, Pará, and Roraima states as well as in the northeastern coast and the states of the south region, values higher than $490 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$ are found.

v

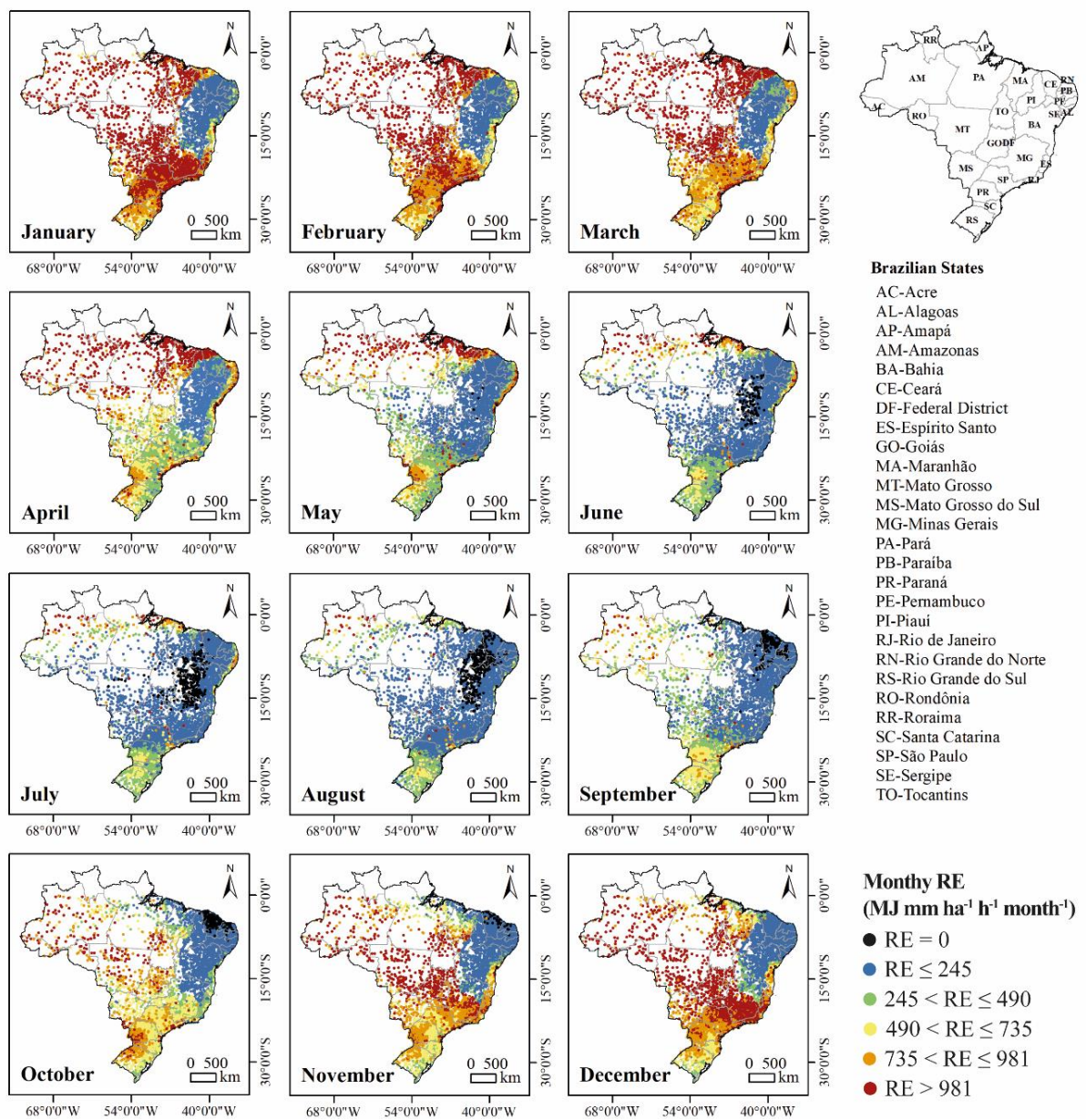


Figure 6. Point distribution of the monthly RE values for Brazil.

Null erosivity values were found in at least one of the gauges analyzed from May to December. Two spots with monthly erosivity equal to zero were observed in the northeast region of Brazil (black dots in Figure 6). The first is mainly at the central strip of this region (southern Piauí, western Bahia, and northern Minas Gerais states) from June to August. The second can be observed mainly in the Ceará state from September to November. Also, for most of the northeast region, especially in the Brazilian semiarid region, the RE values remain below $245 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$ throughout the year.

Regarding the spatial dynamics of the highest RE values, magnitudes over $981 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$ are found in northwestern Brazil in all months. From January to May, the north region of the Ceará, Maranhão, and Piauí states is considered another high erosivity spot. In these states, rainfall is greatly influenced by the Intertropical Convergence Zone (ITCZ), which is responsible for high rainfall volumes, especially in March and April (OLIVA, 2019; SILVA; GALVÍNCIO; COSTA, 2017). This happens due to ITCZ being at its southernmost position from the Equator in these months. The ITCZ also influences the formation of Easterly Wave Disturbances (EWD), which occur on the coast of northeastern Brazil during the austral autumn and winter seasons (Figure 6). This atmospheric system is important from an erosive perspective since it is associated with high-intensity rainfall events as shown by Alves et al. (2013), Machado et al. (2012), Neves et al. (2016), and Santos et al. (2012).

Another atmospheric system of great relevance, the South Atlantic Convergence Zone (SACZ), is responsible for increasing the monthly RE values over the northwest-southeast (NW-SE) direction of the country. For the center-west and southeast regions of Brazil, the precipitation climatology presents the SACZ, which operates during the austral summer and follows the annual rainfall cycle as one of the most important components (ALVARENGA, 2012; NIELSEN et al., 2016). As shown in Figure 6, the SACZ accounts for RE magnitudes higher the $981 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$ from November to March, especially in the states of Mato Grosso, Goiás, São Paulo, Minas Gerais, Rio de Janeiro, and Espírito Santo. The SACZ is a band of nebulosity that may cause intense or persistent rainfall and, as shown by Aguiar and Cataldi (2021), its occurrence is closely related to records of natural hazards.

The erosivity patterns in southern Brazil are characterized by magnitudes over $245 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ month}^{-1}$ for all months. In this region, the presence of cold fronts and Mesoscale Convective Complexes are the main sources of rainfall (DURKEE; MOTE; SHEPHERD, 2009; MORAES et al., 2020; NERY; CARFAN, 2014). As shown in Figure 7, the uniformity of the monthly RE values throughout the year is shown in both of the indexes used to assess the concentration of the erosivity, the RECI and the RESI. These indexes show that for southern

Brazil the RE occurs equitably in the months. As demonstrated for the municipality of Montenegro, in the eastern portion of the Rio Grande do Sul state (Figure 7c). The low CV, considering the monthly values (14.8%), shows small variations of RE magnitudes throughout the year (RECI = 8.50 and RESI = 0.11).

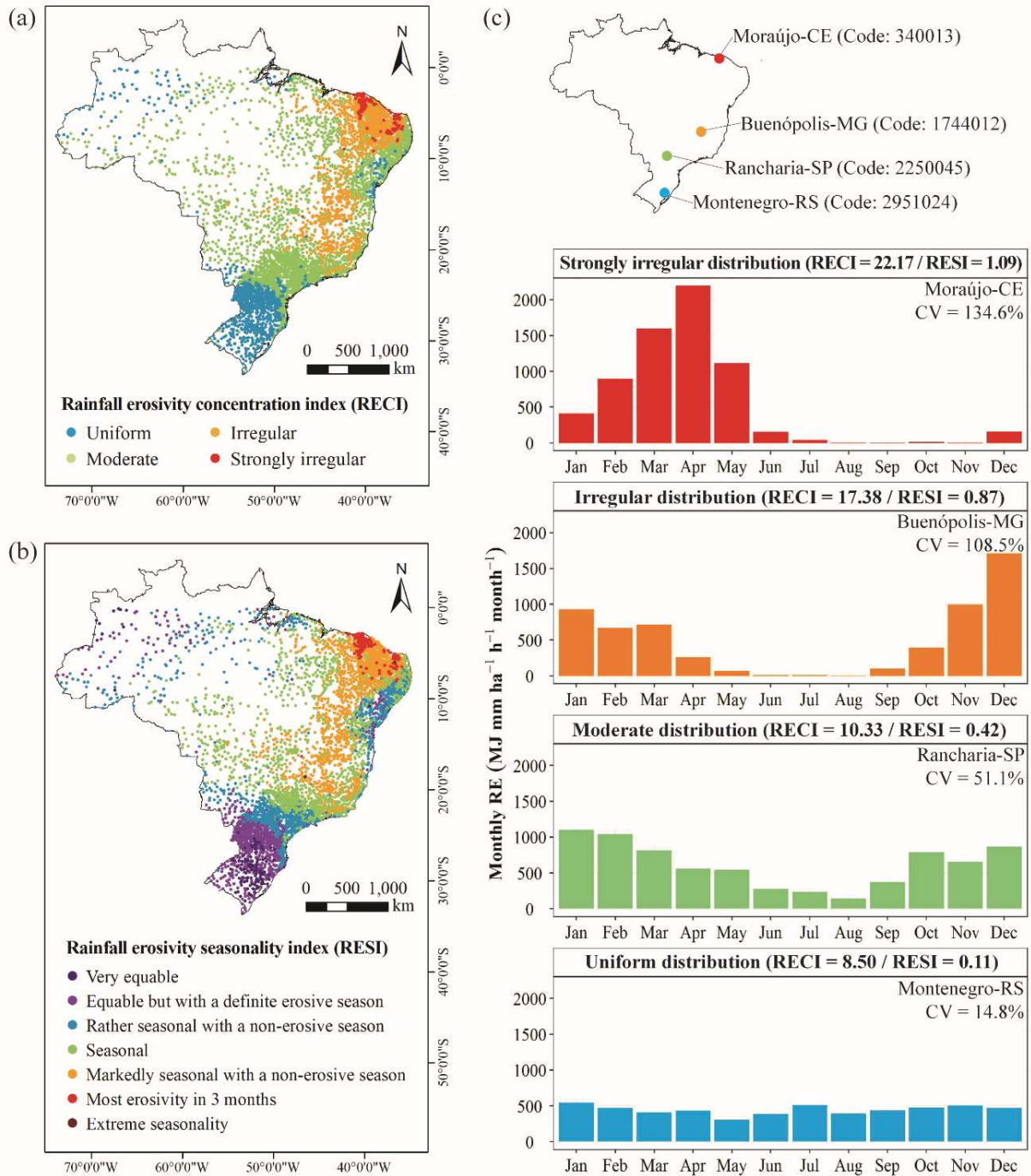


Figure 7. Classification for the (a) RECI and (b) RESI indexes over the Brazilian territory. (c) Monthly distribution of the RE values for Maraujo-CE, Buenópolis-MG, Rancharia-SP, and Montenegro-RS, including the coefficient of variation (CV).

Overall, the RECI and RESI values and their classification agreed for the entire country (Figures 7a and 7b) and had a strong correlation ($r = 0.96$, Figure 6). The use of these indexes identified that the western Ceará and central Rio Grande do Norte states are areas in Brazil where RE is mostly irregularly distributed. Also, for most of the northeastern region of the country, where the annual RE values are considered “Low” (Figure 3b), the RECI and RESI values classify the erosivity as “irregular” or “markedly seasonal”, which shows that despite the low magnitudes, the RE is concentrated in individual months. These results are relevant since they can be used to direct soil and water conservationist practices for specific months of the year, increasing the efficiency of the conservation programs. It is also important to highlight that the RE distribution for the municipalities in Figure 7c is not representative of the distribution for all gauges in the same RECI or RESI category. Thus, the months in which the highest RE magnitudes are concentrated should be identified for each location.

As mentioned before, southern Brazil has the most uniform RE distribution throughout the year. For this region, the analysis of the RESI and RECI demonstrates that the central strip of the Paraná, Santa Catarina, and Rio Grande do Sul states have very equitable distribution. In this case, despite the RECI and RESI values corroborating each other, the use of the RESI gave more detailed results. Thus, the authors recommend a joint analysis of both of these indexes. Finally, we also highlight that the RECI and RESI indexes were never calculated for the Brazilian territory before and constitute an advance in the assessment of RE values in the country.

The gravity center for the annual RE values in Brazil is in the municipality of Hidrolândia, in Goiás state (Figure 8). This means that considering all of the gauges used, this municipality represents the location where the spatial phenomenon of RE is equally distributed throughout the country. The gravity center here presented was obtained using long-term mean RE values, however, as shown by Guo et al. (2019) and Zhu et al. (2021), gravity center locations can be used to assess the evolution of the distribution of RE values in different periods (e.g., over the years). The authors emphasize that, obviously, the RE gravity centers here presented are influenced by the irregular distribution of the rainfall gauges analyzed. However, these results constitute a novelty since this methodology had never been used to assess the spatial dynamics of erosivity in the country.

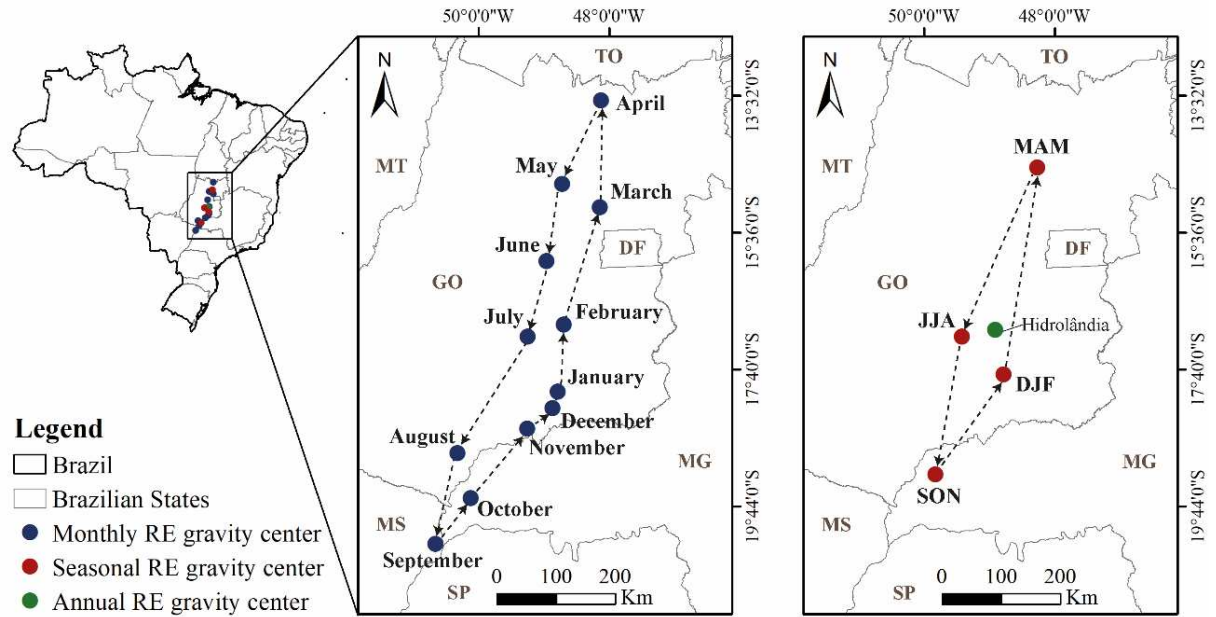


Figure 8. Gravity center location of monthly, seasonal, and annual RE. Federal District (DF) and Brazilian states: Goiás (GO), Minas Gerais (MG), Mato Grosso (MT), Mato Grosso do Sul (MS), São Paulo (SP), and Tocantins (TO). Seasons: December-January-February (DJF), March-April-May (MAM), June-July-August (JJA), and September-October-November (SON).

The results of seasonal and monthly RE gravity centers evidence a latitudinal path variation (north-south migration) of the erosivity, despite the slight correlation between latitudes and annual RE magnitudes ($r = 0.21$, Figure 5). From April onwards a southward migration is observed, while a return towards the north is observed after September. This variation may be related to the South American Monsoon System (SAMS) dynamics during the summer (DJF) that increases rainfall magnitudes in southeastern Brazil (CARVALHO; CAVALCANTI, 2016). From the northernmost and westernmost gravity center (April) to the southernmost and easternmost (September), a distance of approximately 800 km (planar straight line) is observed. Furthermore, for most of the months and seasons, the RE gravity center is within the Goiás state.

3.2. Erosivity density values

The annual ED ranged from 0.9 to 9.7 MJ ha⁻¹ h⁻¹, with a mean value of 4.1 MJ ha⁻¹ h⁻¹ (Table 2). As shown in Figure 9a, most of the gauges analyzed have annual ED values considered “Low” and “Very low” (84.6% in total) according to the ED classification

proposed by Dash et al. (2019). These values ($< 6 \text{ MJ ha}^{-1} \text{ h}^{-1}$) characterize ED for most of northeastern, southeastern, and southern Brazil. ED values considered “Moderate” were mostly found for the north and center-west regions, while magnitudes considered “High” (> 6 and $\leq 15 \text{ MJ ha}^{-1} \text{ h}^{-1}$) were found only for two of the 5,166 gauges analyzed, the municipalities of Mirador-MA and Manicoré-AM. Finally, Dash et al. (2019) state that ED values considered “Very high” are higher than $15 \text{ MJ ha}^{-1} \text{ h}^{-1}$. However, magnitudes like this were not found in the Brazilian territory.

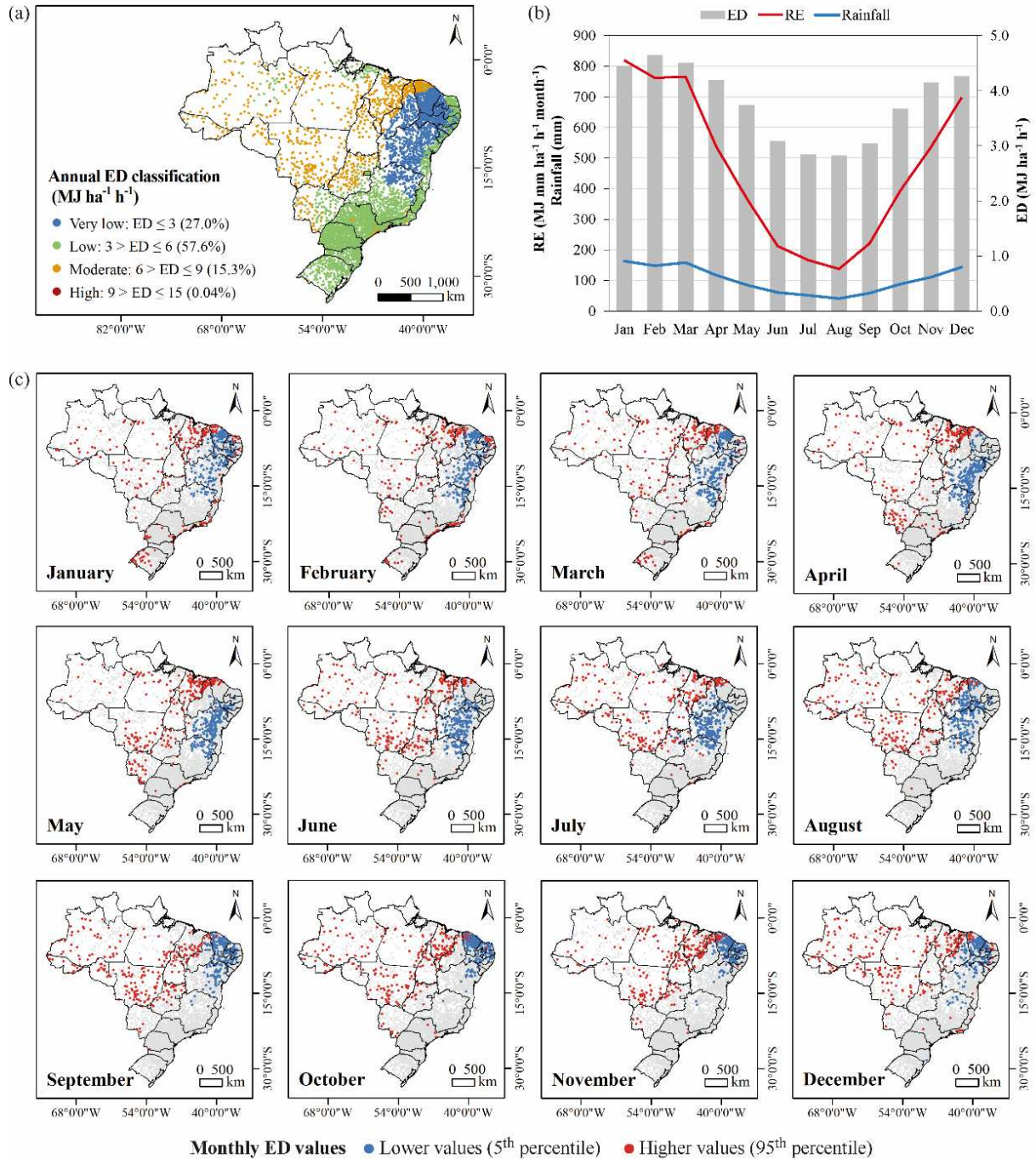


Figure 9. (a) Annual ED classified according to Dash et al. (2019). (b) Monthly distribution of the RE, ED, and rainfall depth values. (c) Spatial distribution of the lower and higher monthly ED magnitudes, represented respectively by the 5th and 95th percentiles.

As shown in Figure 5, a strong correlation was found between the annual ED and RE values ($r = 0.86$). This correlation was expected since ED is obtained by the ratio of RE and rainfall depth values. However, only a moderate correlation of 0.65 was found between annual ED and rainfall depth values. Additionally, the ED had a weak correlation with the altitude (r

= -0.06), latitude ($r = 0.22$), RECI ($r = -0.30$), and RESI ($r = -0.26$) values, similar to the RE variable. Finally, compared to the annual RE (CV of 52.5%), the annual ED values showed less spatial variation (CV of 45.7%).

In summary, the monthly ED values followed the seasonal patterns presented by the monthly RE and rainfall depth (Figure 9b). The mean monthly ED magnitudes were higher for February and March, and lower for July and August. The highest maximum ED was in November ($13.4 \text{ MJ ha}^{-1} \text{ h}^{-1}$, Table 2). However, ED magnitudes higher than $10 \text{ MJ ha}^{-1} \text{ h}^{-1}$, considered by Dash et al. (2019) as “High” ED values, are observed for at least one gauge analyzed in all months, which evidence the existence of high-intensity precipitation spots in Brazil in all year long.

As shown in Figure 9c, the highest monthly ED values are mostly concentrated in the north and center-west regions of Brazil. Additionally, some high ED spots are found for all regions. Some examples are the western Rio Grande do Sul state from January to March, the coast of the São Paulo and Rio de Janeiro states from December to April (due to the orographic effect of the Serra do Mar mountains, as explain Luiz-Silva and Oscar-Júnior (2022)), and the northern Ceará, Maranhão, and Piauí states for most of the months (due to the ICTZ influence on the local rainfall patterns, as previously discussed). As explained by Zhu et al. (2021), a high ED value indicates a high rainfall intensity and a strong soil erosion potential over a short period. Thus, the identification of the mentioned spots can be useful for finding areas most susceptible to water erosion and may contribute to a better implementation of soil conservation practices. Finally, as expected, the lowest ED values found for Brazil are in the northeast region, mainly in the central strip of this region from January to August, as well as for the states of Ceará, Rio Grande do Norte, Paraíba, and Pernambuco for October and November.

Results of ED magnitudes are relatively new for the soil erosion scientific literature, and were only obtained for China (LI; YE, 2018; XU et al., 2019; ZHU; XIONG; XIAO, 2021), Greece (PANAGOS et al., 2016; VANTAS; SIDIROPOULOS; LOUKAS, 2019), India (BAGWAN, 2020; DASH; DAS; ADHIKARY, 2019; SINGH; SINGH, 2020), Indonesia (SUPRIYONO et al., 2021), Italy (DIODATO et al., 2021), South Korea (SHIN et al., 2019), as well as the entire Europe (BALLABIO et al., 2017; PANAGOS et al., 2015). For Brazil, ED values estimation was limited to the Tocantins-Araguaia basin (SANTOS et al., 2022) and the São Paulo state (TEIXEIRA et al., 2022a). Therefore, the ED results here presented constitutes a novelty and fill the gap of ED values for most of the Brazilian territory.

As shown by Dash et al. (2019), the ED variable constitutes a relevant perspective of the erosive potential of rainfalls and can be used to identify floods, erosion, and landslide-prone

areas. Therefore, the authors recommend a joint analysis of the ED maps here presented (Figure 9) with those for RE (Figures 3 and 6), RECI, and RESI (Figure 7), which may contribute to a more accurate soil conservation planning for the country.

3.3. Regionalized models to estimate rainfall erosivity

The joint application of the multivariate statistical techniques PCA and HCA divided the Brazilian territory into eleven homogeneous regions (HRs) regarding the RE patterns, as presented in Figure 10a. As shown in Figure 10b, Regions 1, 2, and 3 have the highest RE values, while the lowest are found for Regions 5, 7, and 8. For each defined region, a regression model to estimate RE values was adjusted (Table 3). For most of the HRs, the rainfall depth (R) was considered the best explanatory variable, while the MFI is the best predictive variable only for Region 9. For all regions, R^2 higher than 0.7 was found for both the training and validation datasets.

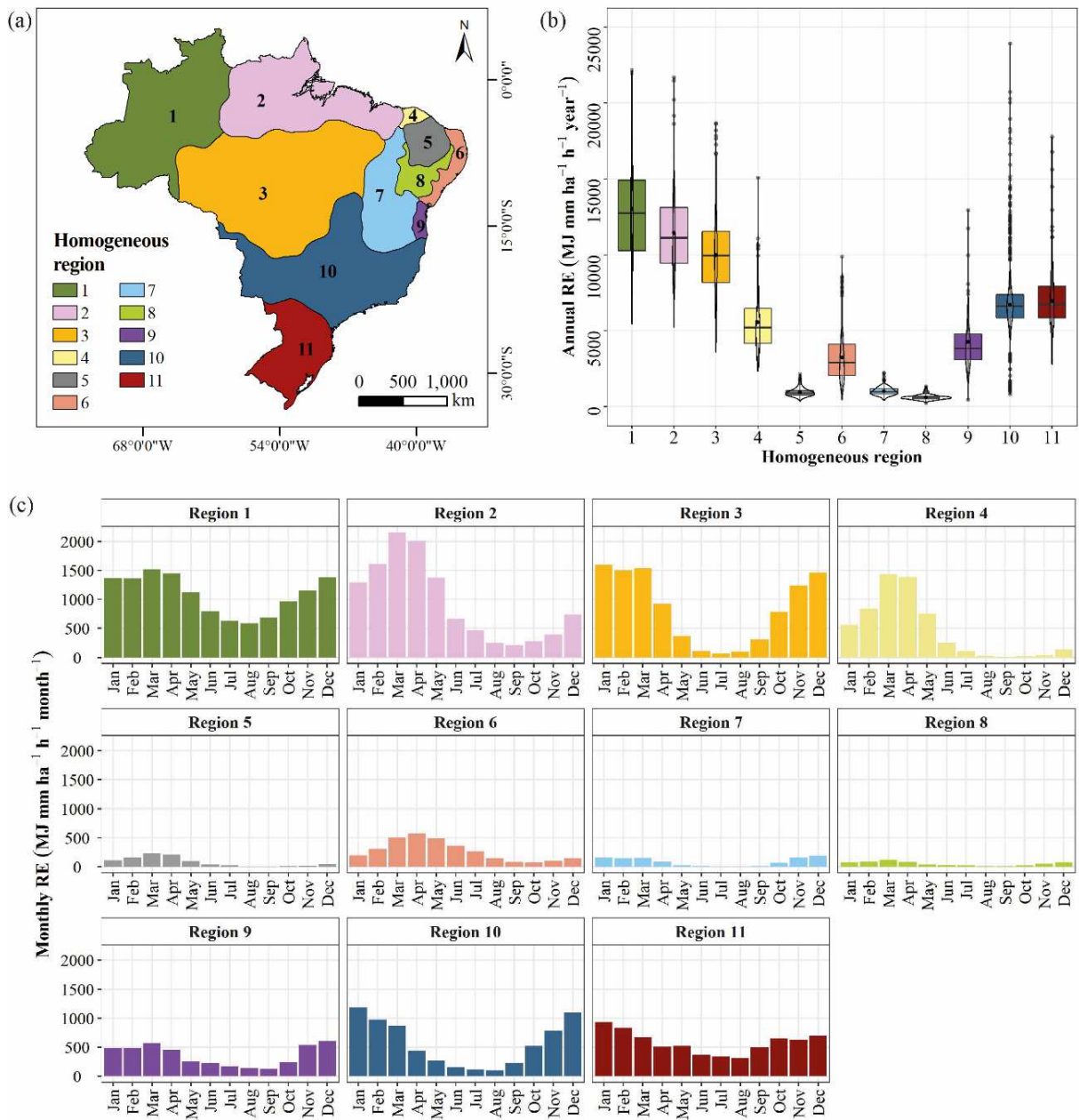


Figure 10. (a) Homogeneous regions regarding the RE for Brazil, as well as the (b) annual and (c) mean monthly RE values for each region.

Table 3. Established rainfall erosivity estimation models for each homogeneous region (HR), and their respective accuracy annual metrics. *MAE* and *RMSE* values are expressed in MJ mm ha⁻¹ h⁻¹ year⁻¹. *RE_i*, *R_i*, and *MFI_i* stand for rainfall erosivity, rainfall depth, and modified Fournier index, respectively, for the month *i*.

HR	Fitted model	Training	Validation	All dataset			
		<i>R</i> ²	<i>R</i> ²	<i>MAE</i>	<i>MAPE</i> (%)	<i>RMSE</i>	<i>d</i>
Region 1	$RE_i = 4.1233(R_i)^{1.0818}$	0.953	0.955	1002.52	7.87	1346	0.963
Region 2	$RE_i = 4.1374(R_i)^{1.0755}$	0.973	0.977	1546.38	13.39	1955.4	0.884
Region 3	$RE_i = 4.4688(R_i)^{1.0778}$	0.981	0.986	524.41	5.43	708.84	0.981
Region 4	$RE_i = 6.7578(R_i) - 31.927$	0.979	0.976	353.09	6.60	501.35	0.983
Region 5	$RE_i = 1.3571(R_i) - 2.7382$	0.983	0.984	46.73	4.82	63.65	0.985
Region 6	$RE_i = 1.4386(R_i)^{1.1776}$	0.828	0.826	488.69	19.56	649.18	0.956
Region 7	$RE_i = 1.4015(R_i) - 5.2407$	0.980	0.983	51.34	5.30	71.62	0.980
Region 8	$RE_i = 0.6605(R_i)^{1.1418}$	0.933	0.939	59.91	9.91	72.06	0.959
Region 9	$RE_i = 31.946(MFI_i)^{1.0581}$	0.784	0.853	584.71	16.49	1018.6	0.913
Region 10	$RE_i = 5.7928(R_i) - 89.449$	0.927	0.923	597.41	13.46	933.29	0.930
Region 11	$RE_i = 0.48(R_i)^{1.4473}$	0.783	0.770	588.93	8.45	787.64	0.940

As mentioned before, northeastern Brazil has the lowest RE values. Also, it is the most subdivided area in the country (Regions 4 to 9), which means that the HCA technique was sensitive enough to differentiate these regions regarding the monthly values despite their low magnitudes (Fig 10c). This occurs since the percentage differences between adjacent regions' RE values are high even with the small differences in the magnitudes. Furthermore, the authors highlight that Region 4, which has the highest RE and ED values for the Brazilian northeastern region (Figures 3, 6, 9a, and 9c) and is highly influenced by the ITCZ dynamics, presents an interesting border pattern. This region showed RE magnitudes similar to Region 2 (neighbor on the left) for the first half of the year, while for the second half, it was similar to Region 5 (neighbor on the right), hence requiring its own estimation model.

Considering both the training and validation datasets, *R*² values higher than 0.90 were obtained for most regions, except Regions 6, 9, and 11. This means that for most regions the monthly rainfall depths explain more than 90% of the RE variance. Complementarily, the values obtained for the *d* index were above 0.80 for all regions. For this index, the closer to 1 the *d* value is, the greater the agreement between the predicted and the observed values (WILLMOTT, 1981). This agreement can also be evidenced in the scatter plots between the observed monthly RE values and the predicted by the models for each HR, presented in Figure 11. In these plots, the dots and the regression line are mainly close to the 1:1 line. Also, the *r* and *R*² values are higher than 0.7. Therefore, we consider the established models for all HRs satisfactory.

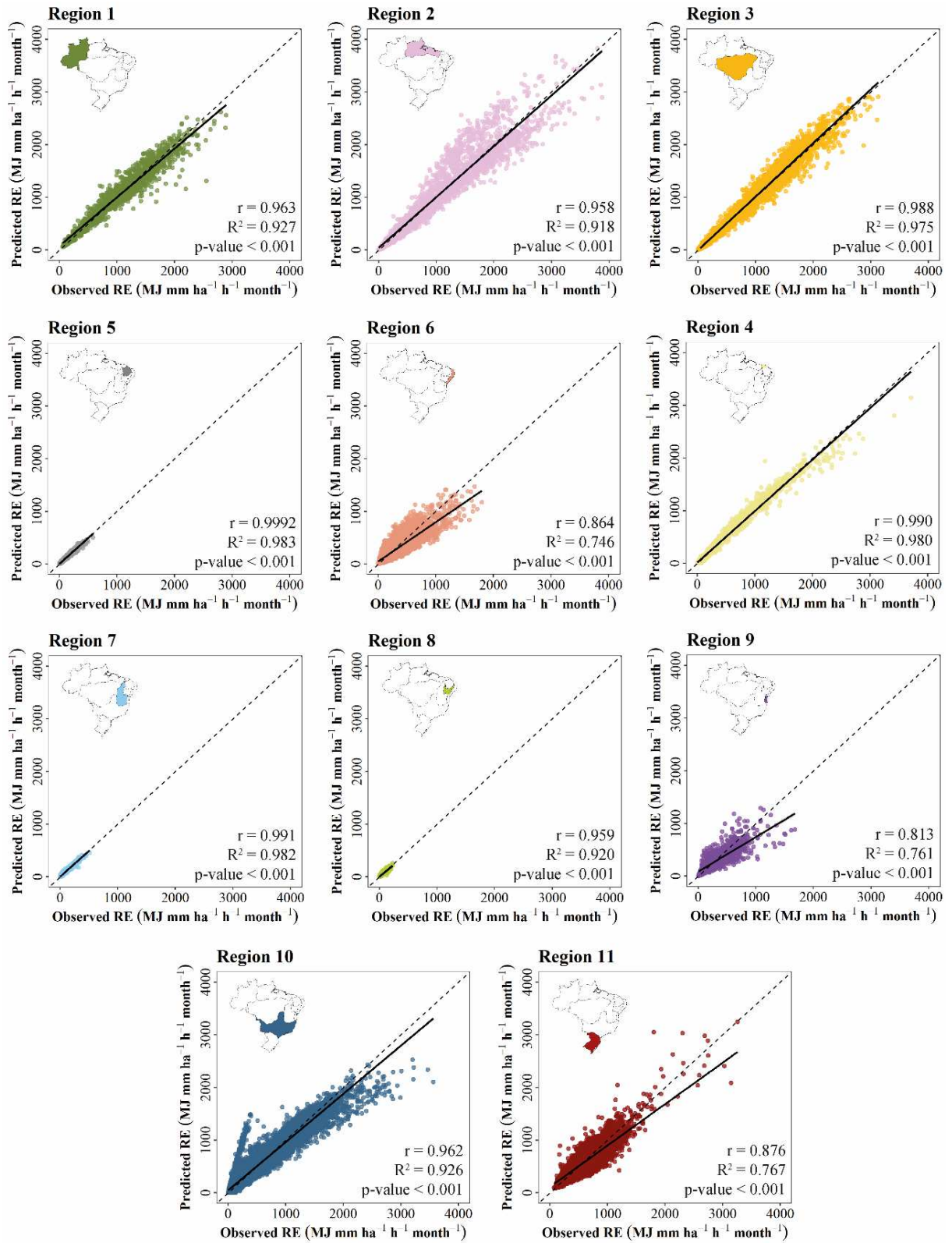


Figure 11. Scatter plots between the observed monthly RE values and the predicted by the models established for each homogeneous region.

As shown in Figure 12, the annual RE values predicted by the established models presented percentage errors below 10% for most of the rainfall gauges over Brazil. However, spots with underestimates higher than 50% were observed mainly in the eastern part of Region 10. Underestimates higher than 10% were also found in northern Region 2, in the coastal area of Regions 10 and 11, and at the border of these two regions. On the other hand, the northern part of Region 6 concentrated the gauges with higher overestimates in the country. This region also shows the highest *MAPE* values (about 19%), which makes this region one of those with the greatest uncertainty in predictions by the established models. Other overestimated spots are found in the western part of Regions 10 and 11.

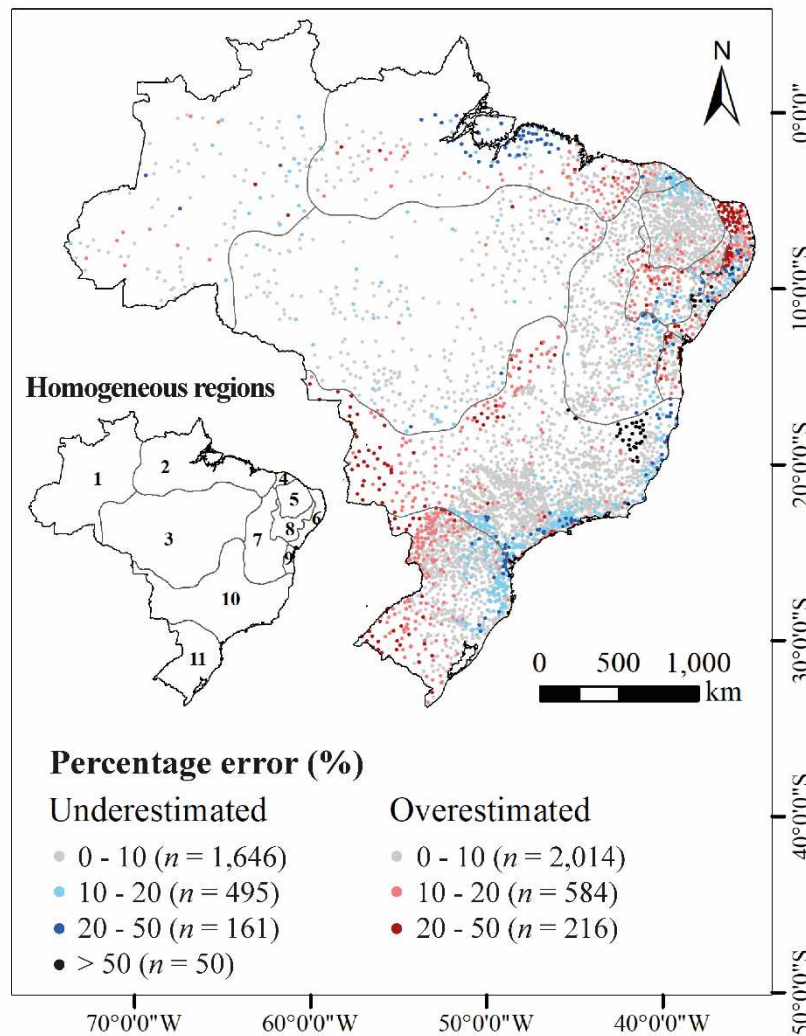


Figure 12. Percentage error map for the annual RE values predicted by the models established for each homogeneous region. n expresses the number of rainfall gauges in each percentage error interval.

Despite the uncertainty in the RE estimates for specific areas over the Brazilian territory using the established models, the equations here presented are an advance regarding the availability of regression models with this purpose. As stated by Teixeira et al. (2022b), the lack of models to obtain RE values for the entire country leads to the inadequate use of these empiric equations in regions where they were not developed. Also, the inequality in the number of studies in the different Brazilian states results in areas with few or no models while others have many estimation models. As mentioned before, in areas without models, equations developed for other regions have been used (e.g. the studies by Falcão et al. (2020) and Sousa et al. (2019)). On the other hand, for regions with greater availability of models, there may be a misunderstanding about choosing the best model to be used. Thus, regionalized models can solve these problems.

An attempt to obtain RE regionalized models for Brazil was done by Silva (2004). However, this study is limited to identifying a region to use a given erosivity model already established for Brazil so this author did not develop new models considering the RE patterns of each region. Also, Mello et al. (2013) proposed multivariate models for the entire country. Although these models consider only the latitude, longitude, and altitude values to estimate RE, many of them are adjusted considering more than eight regression coefficients, which makes its use difficult. Therefore, the regionalized models proposed in the present study can be considered an important tool for improving the availability of RE values over the Brazilian territory since they require only monthly rainfall depths to be used.

Finally, to ease the use of the models here proposed, the identification of the HR from each analyzed rainfall gauge and the spatial delimitation of the eleven homogenous regions defined for Brazil are available for download, respectively, at <http://dx.doi.org/10.17632/hzxfvvr6p.1> and <http://dx.doi.org/10.17632/dwgzpztc7.1>.

4. Conclusion

In this study, we update the findings regarding the assessment of RE values in Brazil, using a large national database. For this, RE and ED values were obtained for 5,166 rainfall gauges. Also, the concentration of the RE throughout the year was analyzed using the RECI and RESI index, and the RE's gravity center locations on the monthly, seasonal, and annual scales were defined. Finally, homogeneous regions regarding RE values were delimited and estimative regression models were established for each region.

The annual RE values found for Brazil range from 252 to 23,916 MJ mm ha⁻¹ h⁻¹ year⁻¹

¹, magnitudes never found in national studies before. The mean annual value was 5,620 MJ mm ha⁻¹ h⁻¹ year⁻¹, with high spatial variation over the country. The highest RE magnitudes were found for the north region, while the northeast region concentrates the lowest values. In addition, a strong correlation between annual erosivity values and annual rainfall depths was evidenced, while the altitudes and latitudes presented a very slight correlation with the annual RE magnitudes. Throughout the year, the RE patterns show a migration of the high erosivity spots through the different regions of the country. Considering this, the monthly RE values and maps presented in this study are relevant for improving the accuracy of soil loss estimates since they show a panorama of the RE phenomenon in the country.

The RECI and RESI indexes corroborated each other and evidenced that the southern region of Brazil has the most regular distribution of the RE values throughout the year. On the other hand, in western Ceará and central Rio Grande do Norte states the RE is mostly concentrated and, thus, requires caution for the months with higher RE magnitudes. Complementarily, the RE gravity center analyses showed a latitudinal path variation (north-south migration) of the erosivity throughout the year. Additionally, for most of the months, the RE gravity center of Brazil is in the Goiás state.

The ED magnitudes obtained for the country allowed the identification of high-intensity rainfall spots in Brazil all year long. Despite this, most of the gauges analyzed have annual ED values considered “Low” and “Very low”. In summary, the monthly ED values followed the seasonal patterns presented by the monthly RE and rainfall depth. Considering this, the ED magnitudes here presented fill the gap of ED estimates for the Brazilian territory and due to its relationship with high-intensity rainfall events can help understand the RE impacts on soil loss. Therefore, we encourage the consideration of these results by scientists, policymakers, and society in general, since they can be useful for finding areas highly susceptible to water erosion in the country.

The application of multivariate statistical techniques resulted in the division of the Brazilian territory into eleven homogeneous regions regarding the RE patterns. For each defined region, a regionalized regression model was established, and their performance is considered satisfactory for all regions. Thus, the models here presented constitute an advancement regarding the possibility of estimating RE magnitudes for the entire country since they can be used for predicting values for a specific location or period. Furthermore, the adjusted models can be considered easy to use, since only monthly rainfall depths are required for the estimation, so they are important tools for improving the availability of RE values in Brazil.

Moreover, we highlight that all the data produced by this study is available for download. This database includes RE and ED values on monthly, seasonal, and annual scales, as well as the RECI and RESI values for the 5,166 rainfall gauges analyzed. Also, it was specified which regionalized model should be used based on the homogeneous region each gauge is.

For future studies, we suggest the production of maps containing the spatial distribution of RE values for the entire country using the values here presented. For this, we encourage the use of techniques not widely used for mapping RE in Brazil so far, such as those based on machine learning. These maps have the potential to improve the understanding of the RE patterns in time and space, which may enhance planning soil and water management and conservation practices.

Data availability

The vectorial data produced in this study is freely available online. The spatial database containing the RE, ED, RECI, and RESI values for the 5,166 rainfall gauges are available for download at <http://dx.doi.org/10.17632/hzxfvvr6p.1>. The spatial delimitation of the eleven homogenous regions regarding RE defined for Brazil is available for download at <http://dx.doi.org/10.17632/dwgzpztc7.1>.

CHAPTER 3: GRIDDED MAP OF ANNUAL RAINFALL EROSIVITY FOR BRAZIL: A MACHINE LEARNING APPROACH

Abstract

In this study, a gridded annual map of rainfall erosivity (RE) was created for Brazil using machine learning techniques. For this, long-term mean RE values obtained from synthetic series of pluviographic data for 5,166 rainfall gauges over the Brazilian territory were used. In total, 93 covariates regarding climatic and terrain aspects were used as auxiliary variables for the predictions. For obtaining the map, four algorithms were tested in this study: linear model (LM), Weighted k-Nearest Neighbors (KKNN), Random Forest (RF), and Support Vector Machine Radial Sigma (SVM). The best subset of covariates was selected by removing variables with little variance, excluding correlated variables, and eliminating the less important ones. The training and test process was repeated 100 times. Results showed that all applied models statistically differed from each other in predicting annual RE values. According to the accuracy metrics analyzed, RF is considered the model with the best prediction performance. The annual RE map generated ranges from 633 to 14,288 MJ mm ha⁻¹ h⁻¹ year⁻¹. In addition to the total annual rainfall, the other covariates with higher importance for the predictions were the rainfall depth for August and rainfall of the coldest quarter, which evidence the relevance of the months with lower rainfall amounts in the annual RE mapping. Further analysis revealed that the northeastern of the country as well as the Serra do Mar mountains region are characterized as the areas with the highest uncertainties in the values mapped. Finally, the RF model presented satisfactory precision for the obtention of a gridded annual RE map for Brazil resolution grid (resolution of ~1 km²). This map is considered an advancement regarding the availability of accurate RE values for the entire Brazilian territory and can be used for improving the soil loss assessment in the country.

Keywords: Erosivity index. Linear model. Nearest neighbors. Random forest. Support vector machine.

1. Introduction

Inserted in the main soil loss prediction models, such as the Universal Soil Loss Equation (USLE) and its revised version (RUSLE) (RENARD et al., 1997; WISCHMEIER; SMITH, 1978), the rainfall erosivity (RE) is a parameter of great importance for soil and water conservation planning. The RE is defined as the potential capacity to cause soil erosion by rainfalls and can be useful to identify areas highly susceptible to erosion (MOSAVI et al., 2020; PANDEY et al., 2021; SENANAYAKE; PRADHAN, 2022), as well as to help policymakers planning and implementing erosion mitigation measures to avoid land degradation (GUDURU; JILO, 2023; WUEPPER; BORRELLI; FINGER, 2019).

The obtention of RE maps is helpful in quantifying accurate soil erosion rates (DISSANAYAKE; MORIMOTO; RANAGALAGE, 2019; STEFANIDIS et al., 2021). Considering this and other applications, RE maps have been created in different spatial scales throughout the world. Some examples are those presented by Panagos et al. (2017a, 2022) and Liu et al. (2020b) on a global scale, and those presented by Kim et al. (2020) and Zhu et al. (2021), respectively on national and regional scales. Independent of the spatial scale, RE maps are important for areas with large agricultural land and production (ŠARAPATKA; BEDNÁŘ, 2022), such as the Brazilian territory. This country has also faced an increase in deforestation rates over the last years (ARAÚJO et al., 2019; DIAS et al., 2016; SOUZA et al., 2020a), which increases the exposure of soils to rainfall impacts.

In Brazil, the RE maps are useful for assessing soil loss in many states (LENSE et al., 2021; MEDEIROS et al., 2016) as well as for the conservationist management of several water basins (CUNHA et al., 2022; SILVA et al., 2016). RE maps on a national scale were also developed, such as those shown by Silva (2004) and Mello et al. (2013). Despite these maps' availability, most of them were created using data from empirical models that were not developed for the regions they are applied (ANJOS et al., 2020; SOUSA; PAULA, 2019; SOUZA et al., 2020b). In addition, many of the maps for the country were generated using relatively sparse spatial data (low number of gauge stations) (OLIVEIRA et al., 2015; TRINDADE et al., 2016) as well as considering short time length records (less than 10 years of data) (CARVALHO et al., 2012; NEVES; DI LOLLO, 2022; SILVA et al., 2010a). These products may not represent accurately the erosivity dynamic of rainfalls. Therefore, using a large national database for obtaining RE maps for Brazil is still a gap to be filled.

As shown by Teixeira et al. (2022b), among the methods used for mapping the RE over the Brazilian territory, kriging is the most popular. In addition, the use of the inverse distance

weighting technique also contributed to the availability of RE maps in the country. Recently, studies have employed machine learning techniques to obtain RE maps (LEE et al., 2021, 2022; VANTAS; SIDIROPOULOS; LOUKAS, 2019). The use of these algorithms is considered a promising alternative for obtaining spatially interpolated maps, especially due to their predictive accuracy (DA SILVA JÚNIOR et al., 2019; KARIMI et al., 2020; SEKULIĆ et al., 2020). Another advantage of these techniques is the possibility of using a large set of covariates in the spatial modeling, which according to Souza et al. (2022) represents a methodological gain.

For many years, the use of machine learning techniques for mapping RE in Brazil was limited to the employment of artificial neural networks, that were developed for some states such as Espírito Santo (CECÍLIO et al., 2013), Minas Gerais (MOREIRA et al., 2009), Rio de Janeiro (CARVALHO et al., 2012), and São Paulo (MOREIRA et al., 2006a). The use of other models, such as random forest and support vector machines, was only assessed by Souza et al. (2022) for the state of Minas Gerais, which produced monthly and annual RE maps with good spatial resolution. Considering this, the use of different machine learning techniques shows great potential for obtaining RE maps that present good reliability, increasing the availability of estimates for this variable on a national scale.

This study aims to generate an annual RE gridded map for the Brazilian territory using machine learning techniques. The specific objectives of this study were: i) to define the best set of prediction covariates; ii) to identify the best model for obtaining the RE map; and iii) to create an annual map of RE for Brazil.

2. Material and methods

Figure 1 shows the flowchart of the methodological steps adopted in the present study which are better described in the following sections.

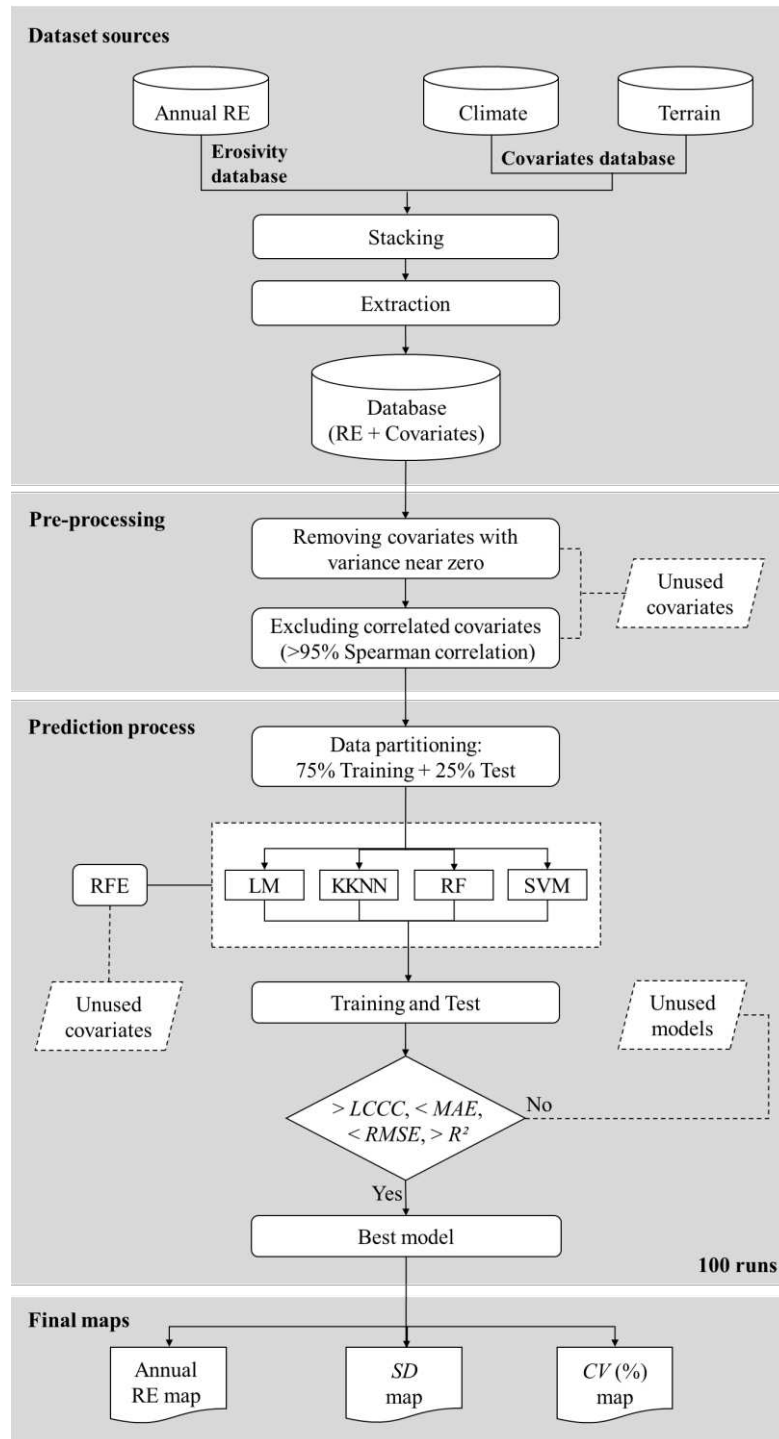


Figure 1. Flowchart of the methodological steps adopted in the present study. RFE means recursive feature elimination. LM, KKNN, RF, and SVM denote, respectively, Linear Model, Weighted k -Nearest Neighbors, Random Forest, and Support Vector Machine Radial Sigma. $LCCC$, MAE , $RMSE$, R^2 , SD , and CV means Lin's concordance correlation coefficient, mean absolute error, root mean square error, coefficient of determination, standard deviation, and coefficient of variation.

2.1 Study area

The Brazilian territory covers about 8,511,000 km² and has altitudes ranging from zero to 2,800 m (Figure 2a). In addition, Brazil has an expressive climate variability (Figure 2b), among which the tropical subtypes (Af, Am, As, and Aw) stand out (ALVARES et al., 2013). In the northeast region, the semi-arid climate (Bsh) is more representative, while the subtropical humid climate subtypes (Cfa, Cfb, Csa, Csb, Cwa, and Cwb) stand out in the south and southeast. Regarding rainfall patterns, it annually ranges from 380 to 4,000 mm (ALVARES et al., 2013; INMET, 2022), as shown in Figure 2c.

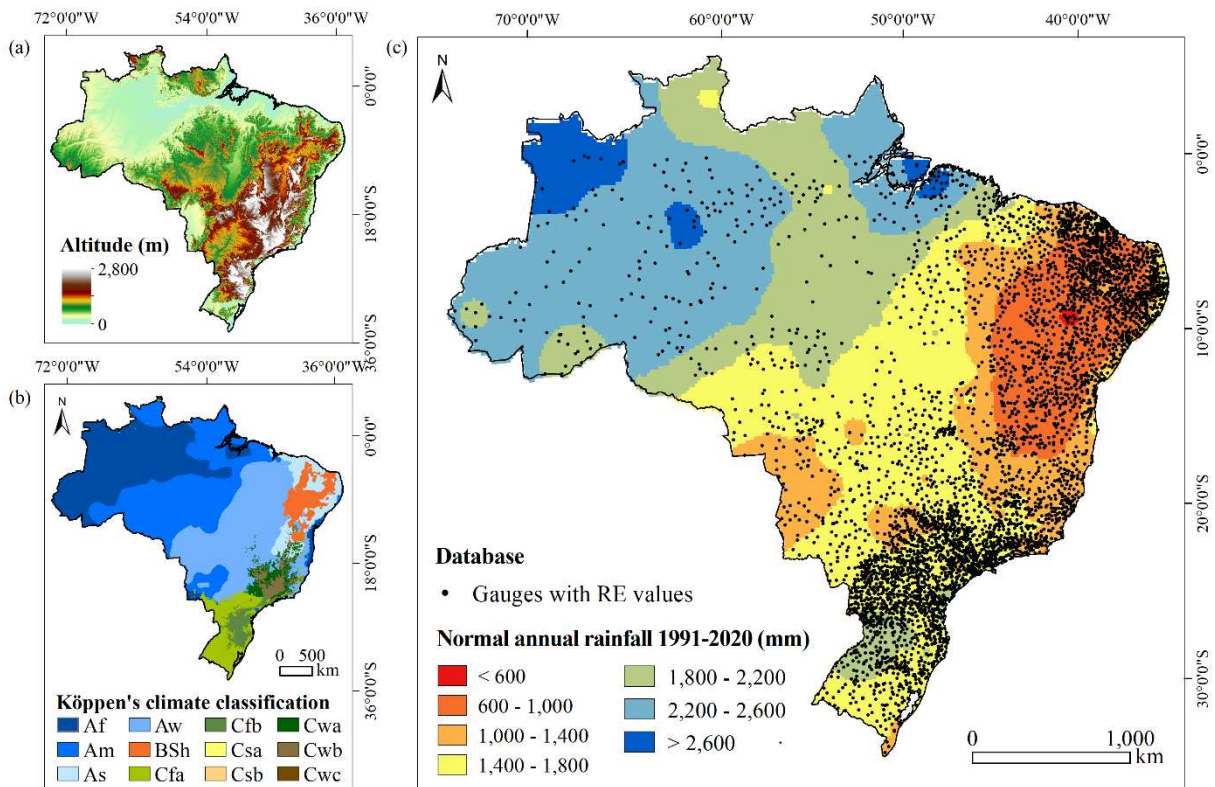


Figure 2. (a) Altitude variation in Brazil (WorldClim-DEM 30 arc-second resolution). (b) Köppen (1936)'s climate classification presented by Alvares et al. (2013). (c) Spatial distribution of the normal annual rainfall 1981-2010 (INMET, 2022) and location of the rainfall gauges considered as database in this study.

2.2 Erosivity database

The values of annual RE used for training the models were those presented in the Chapter 2 of this thesis. The RE magnitudes were estimated for 5,166 rainfall gauges over the

Brazilian territory (Figure 2c). For this, synthetic series of pluviographic data on a sub-daily scale were generated using the ClimaBR stochastic weather generator (BAENA et al., 2005; OLIVEIRA; ZANETTI; PRUSKI, 2005a, 2005b; ZANETTI et al., 2005), based on the measured daily data from each gauge (see Chapter 2).

From the rainfall synthetic series, RE was estimated based on the criteria proposed by Wischmeier and Smith (1958) and Wischmeier (1959), and modified by Cabeda (1976). Therefore, the rainfalls considered erosive were identified and the kinetic energy associated with them was calculated as a function of the rainfall intensity. Then, the EI_{30} erosivity index was applied, as proposed by Wischmeier and Smith (1958). The daily EI_{30} was calculated by the product of the kinetic energy of each rainfall of the day and the maximum intensity of precipitation that occurred in 30 minutes (I_{30}), as shown in Equation (1).

$$(EI_{30})_j = KE \cdot I_{30} \quad (1)$$

in which $(EI_{30})_j$ is the rainfall erosivity index in the day j ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ day}^{-1}$); KE is the kinetic energy ($\text{MJ ha}^{-1} \text{ mm}^{-1}$); I_{30} is the maximum rainfall intensity for 30 consecutive minutes (mm h^{-1}).

The RE on a monthly scale was obtained by summing the daily EI_{30} values and the annual RE values were obtained from the sum of the RE for each month. Finally, the average of the values considering all years of the data was calculated. Therefore, the RE values considered in the present study consist of long-term mean values of annual RE values obtained for each rainfall gauge, that is, the R-factor for USLE and RUSLE.

2.3 Prediction covariates database

In this study, two spatial databases were used as sets of covariates for predicting the annual RE map. The first one is composed of climatic variables, while the second one by terrain variables.

The climatic dataset was composed by 57 covariates (Table 1) based on rainfall precipitation and temperature, available at the WorldClim Data Portal (FICK; HIJMANS, 2017). This database was generated using information from satellites and weather stations, obtaining spatially interpolated climate variables, on a global scale, with high spatial resolution. In this study, the average historical climate data was used (1970 to 2000) at a 30 arc-second

resolution grid ($\sim 1 \text{ km}^2$). These data were also applied in other spatial studies as covariates to the spatial prediction of rainfall erosivity, as show Panagos et al. (2016, 2017b), Riquetti et al. (2020), and Souza et al. (2022).

Table 1. Climatic covariates considered for the spatial prediction of the annual RE in Brazil.

Covariate	N° of covariates	Abbreviation
Monthly rainfall precipitation	12	Prec1 to Prec12
Monthly minimum temperature	12	Tmin1 to Tmin12
Accumulated minimum temperature (sum of the 12 months)	1	TminAcc
Monthly maximum temperature	12	Tmax1 to Tmax12
Accumulated maximum temperature (sum of the 12 months)	1	TmaxAcc
Annual mean temperature	1	Bio1
Mean diurnal range (mean of monthly (max temp - min temp))	1	Bio2
Isothermality (Bio2/Bio7) ($\times 100$)	1	Bio3
Temperature seasonality (standard deviation $\times 100$)	1	Bio4
Max temperature of warmest month	1	Bio5
Min temperature of coldest month	1	Bio6
Temperature annual range (Bio5-Bio6)	1	Bio7
Mean temperature of wettest quarter	1	Bio8
Mean temperature of driest quarter	1	Bio9
Mean temperature of warmest quarter	1	Bio10
Mean temperature of coldest quarter	1	Bio11
Annual rainfall precipitation (sum of the 12 months)	1	Bio12
Precipitation of wettest month	1	Bio13
Precipitation of driest month	1	Bio14
Precipitation seasonality (coefficient of variation)	1	Bio15
Precipitation of wettest quarter	1	Bio16
Precipitation of driest quarter	1	Bio17
Precipitation of warmest quarter	1	Bio18
Precipitation of coldest quarter	1	Bio19
Total	57	

The terrain dataset was formed by 36 covariates derived from the digital elevation model available at the WorldClim Data Portal with the same spatial resolution that the climatic dataset variables. These terrain covariates represent different topographic attributes and were obtained using the packages *RSAGA* (BRENNING, 2008), *raster* (HIJMANS, 2023), and *rgrass7* (BIVAND, 2017), in the R software environment (R CORE TEAM, 2023). In Table 2 are presented the terrain covariates used in this study, which were also employed by Mello et al. (2022), Sena et al. (2021), Siqueira et al. (2023), and Souza et al. (2021, 2022).

Table 2. Terrain covariates considered for the spatial prediction of the annual RE in Brazil.

Covariate	Description	Abbreviation
Digital elevation model	Represents the elevation in each cell	DEM
Aspect	Slope orientation	ASP
Convergence index	Convergence/divergence index concerning runoff	CI
Cross sectional curvature	Measures the curvature perpendicular to the downslope direction	CSC
Diurnal anisotropic heating	Measures the relation between the aspect and the slope angle	DAH
Easternness	The property of being to the east	E
Flow line curvature	Represents the projection of a gradient line to a horizontal plane	FLC
General curvature	The combination of both plan and profile curvatures	GC
Hill	Demonstrates the hills	H
Hill index	Simulation of diffusive hillslope evolution using an Alternating-Direction-Implicit (ADI) method	HI
Longitudinal curvature	Measures the curvature in the downslope Direction	LC
Mass balance index	Balance index between erosion and deposition	MBI
Maximal curvature	Maximum curvature in local normal section	MAXC
Mid-slope position	Represents the distance from the top to the valley, ranging from 0 to 1	MSP
Minimal curvature	Minimum curvature for local normal section	MINC
Multiresolution index of ridge top flatness	Indicates flat positions in high altitude areas	MRRTF
Multiresolution index of valley bottom flatness	Indicates flat surfaces at bottom of valley	MRVBF
Normalized height	Vertical distance between base and ridge of normalized slope	NH
Northernness	The property of being to the north	N
Plan curvature	Described as the curvature of the hypothetical contour line passing through a specific cell	PLANC
Profile curvature	Describes surface curvature in the direction of the steepest incline	PROC
Real surface area	The actual calculation of cell area	RSA
Slope	Represents local angular slope	S
Slope height	The vertical distance between base and ridge of slope	SH
Standardized height	The vertical distance between base and the standardized slope index	STANH
Surface specific points	Indicates differences between specific surface shift points	SSP
Tangencial curvature	Measured in the normal plane in a direction perpendicular to the gradient	TANC
Terrain ruggedness index	Quantitative index of topography heterogeneity	TRI

Terrain surface convexity	The ratio of the number of cells that have positive curvature to the number of all valid cells within a specified search radius	TSC
Terrain surface texture	Splits surface texture into 8, 12, or 16 classes	TST
Topographic position index	Difference between a point elevation with surrounding elevation	TPI
Topographic wetness index	Describes the tendency of each cell to accumulate water as a function of relief	TWI
Total curvature	General measure of surface curvature	TC
Valley depth	Calculation of vertical distance at drainage base level	VD
Valley index	Calculation of the fuzzy valley index using the top-hat approach	VI
Vector ruggedness measure	Measures the variation in terrain roughness	VRM
Total		36

2.4 Prediction models

For predicting the annual RE map, four algorithms were tested in this study: linear model (LM), Weighted k -Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine Radial Sigma (SVM). These methods were chosen to evaluate the prediction based on algorithms with different learning approaches (Table 3). The implementation of these algorithms was carried out in an R environment (R CORE TEAM, 2023), using the packages *base* (R CORE TEAM, 2023), *kkn* (SCHLIEP; HECHENBICHLER; LIZEE, 2022), *randomForest* (BREIMAN; CUTLER, 2022), and *kernelab* (KARATZOGLOU; SMOLA; HORNIK, 2023).

Table 3. Machine learning models used, their learning approaches, and R packages for application.

Model		Learning approach	R package
LM	Linear model	-	<i>base</i>
KNN	Weighted k -Nearest Neighbors	Similarity-based learning	<i>kkn</i>
RF	Random Forest	Information-based learning	<i>randomForest</i>
SVM	Support Vector Machine Radial Sigma	Error-based learning	<i>kernelab</i>

The k -Nearest Neighbors (KNN) model is a non-parametric learning algorithm that predicts values based on k neighboring records. As hyperparameters, there is basically the definition of the number k of nearest neighbors that will compose the neighborhood of the new observation and the choice of the distance measure responsible for identifying the k observations of the training set closest to the new observation (KUBAT, 2021; LANTZ, 2019).

As recommended by Kuhn and Johnson (2013), all predictors were centered and scaled prior to performing KNN to avoid potential bias. For the present study, the Weighted KNN (KKNN) (HECHENBICHLER; SCHLIEP, 2004) was used. The idea is to weigh the contribution of each of the k neighbors according to their distance to the point of prediction. So, the closer the neighbor, the more important it is. The KKNN algorithm differs from other KNN algorithms because it uses kernel functions to weigh the neighbors according to their Minkowski distances.

Methods based on decision trees, such as RF (BREIMAN, 2001), constitute an alternative for building predictive models when the relationship between predictors and the response of interest is non-linear and complex. In RF, there is a combination of prediction made by multiple decision trees, in which each tree depends on the values of a vector of predictors in the randomly sampled training set. As hyperparameters, three have to be defined: the number of trees in the forest (*ntree*), the minimum number of data points in each terminal node (*nodesize*), and the number of features tried at each node (*mtry*). As an estimator of the final response in each tree RF uses the average of the tree's predictions, which reduces its variance (TYRALIS; PAPACHARALAMPOUS; LANGOUSIS, 2019).

The SVM comprises a set of supervised learning techniques proposed by Cortes and Vapnik (1995). In the SVM, a hyperplane that allows the greatest possible margin between classes is used, leading to a higher probability of generalization. The estimation by SVM allows infinite results, and good performance depends on the adequate configuration of the hyperparameters. Among the main optimization parameters of the model, the penalty (cost) that controls the trade-off between margin and training errors, and the kernel width (sigma) that controls the degree of non-linearity of the model, have a great influence on the prediction results.

2.5 Selection of the prediction covariates

To select the best set of covariates for each model aiming to predict the annual RE map, three steps were carried out: removing variables with little variance, excluding correlated variables, and eliminating the less important ones. As explain Kuhn and Johnson (2013), this selection is advantageous since the excessive number of variables is time consuming and computationally expensive. For these three procedures functions from *caret* package (KUHN et al., 2020) were performed.

First, covariates with variance near zero were removed from the total dataset. This procedure was applied to eliminate the variables with little or no spatial variability across the

Brazilian territory. For this, the function *nearZeroVar* was employed. According to Kuhn et al. (2020), this function identifies covariates that have one unique value or that have both of the following characteristics: they have very few unique values relative to the number of samples and the ratio of the frequency of the most common value to the frequency of the second most common value is large. This procedure was also applied by Mello et al. (2022).

Second, the *findCorrelation* function was performed. This function quantifies and analyzes the correlation between all possible pairs of prediction covariates to avoid redundancy (KUHNS et al., 2020). In this study, the pair of covariates with more than 95% of Spearman's correlation was identified, and the covariate that presented the highest correlation with the other ones was removed. If two covariates are highly correlated, this implies that they are measuring the same information. Thus, removing one should not compromise the performance of the model and might lead to a more parsimonious and interpretable model (KUHNS; JOHNSON, 2013). This procedure was also applied by Ferreira et al. (2021) and Reis et al. (2021).

Third, the *recursive feature elimination (RFE)* also from *caret* package (KUHNS et al., 2020) was performed. This function implements a backward selection of covariates based on covariate importance ranking. The covariates are ranked and the less important ones are sequentially eliminated. The goal is to find the best subset of covariates to produce an accurate model. This procedure was also applied by Dias et al. (2021) and Gomes et al. (2019b).

The RFE was performed on the total set of covariates that were not eliminated by the second procedure (*findCorrelation* function) so that 21 subsets of covariates were tested for selection. These tested subsets consisted of 5, 6, 7, ..., 20, 25, 30, 40, and 50 covariates, as well as the total dataset left. The RFE was run using the cross-validation method with 10 folds, considering the internal optimization parameters for each tested algorithm. The internal parameters (hyperparameters) are internal characteristics intrinsic to each algorithm, which can be modified to improve the training results. The parameters of the models used are described by Kuhn (2019) in the *caret* package manual available at <https://topepo.github.io/caret/available-models.html>. As metric criteria for this selection, we used the Lin's concordance correlation coefficient (*LCCC*), as shown in Equation (2), which evaluates the agreement between the observed and predicted values (LIN, 1989).

The *LCCC* varies between -1 and 1 and, as explained by Khaledian and Miller (2020), it assesses both the precision and accuracy of the prediction. Precision refers to the spread of points around the regression line (correlation coefficient) for the estimated versus observed values. Accuracy refers to the correspondence between that regression line and the perfect line (1:1 line). *LCCC* can also be considered a more appropriate evaluator compared to the

coefficient of determination (R^2 , Equation (3)) since it captures bias in the model prediction (KHALEDIAN; MILLER, 2020).

$$LCCC = \frac{2p O_i E_i}{\sigma^2 O_i + \sigma^2 E_i + (\bar{O}_i - \bar{E}_i)^2} \quad (2)$$

$$R^2 = \frac{[\sum(E_i - \bar{E}_i) \times (O_i - \bar{O}_i)]^2}{[\sum(E_i - \bar{E}_i)^2] \times [\sum(O_i - \bar{O}_i)^2]} \quad (3)$$

in which $LCCC$ is the Lin's concordance correlation coefficient; R^2 is the coefficient of determination; O_i and E_i are, respectively, the observed and estimated RE values ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$); p is the Pearson correlation coefficient between the observed and estimated values; the $\sigma^2 O_i$ and $\sigma^2 E_i$ are, respectively, the variances of the observed and estimated values; \bar{O}_i and \bar{E}_i are, respectively, the means of the observed and estimated values.

2.6 Training, test, and selection of the models

To select the most accurate model, we predicted annual RE values applying each model separately using their own optimal subset of covariates. For this, data were randomly split into 75% for training and 25% for the test.

The training of the models was performed with the covariates selection followed by the optimization of the hyperparameters. For this, we also used repeated cross-validation (10-folds and 3 repetitions) to optimize the model considering the training dataset. The statistical indicator used as a criteria for this optimization was the $LCCC$. It is worth to mention that, among the adopted models, only the KKNN, RF, and SVM models present hyperparameters. Besides $LCCC$, the performance of the models was also assessed by the R^2 , the mean absolute error (MAE , Equation (4)), and the root mean square error ($RMSE$, Equation (5)).

$$MAE = \frac{\sum_{i=1}^n (|O_i - E_i|)}{n} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (E_i - O_i)^2}{n}} \quad (5)$$

in which MAE is the mean absolute error ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$); $RMSE$ is the root mean square error ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$); O_i is the observed annual RE value ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$); E_i is the estimated annual RE value ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ year}^{-1}$); n is the number of observations.

The test process was performed for each model separately. This procedure was based on the insertion of the selected covariates into the adjusted model, considering only the test dataset. This step allowed the obtaining of the estimated and observed pairs of values from which the statistical metrics were calculated. Despite the detailed description of each methodological step being explained separately, the training and test processes were performed subsequently, iteratively, and simultaneously. For each model, this methodology (covariate selection, training, and test process) was repeated 100 times. As shown by Kuhn and Johnson (2013), the process of several repetitions is important since the different groups of training and test datasets can lead to different accuracy results.

Taking all models tested into account (LM, RF, KNN, and SVM), the one considered with the best performance was chosen, that is, the model with higher $LCCC$ and R^2 values, as well as lower MAE and $RMSE$ values, for the test subset. Finally, Kruskal-Wallis's test (KRUSKAL; WALLIS, 1952) followed by Dunn's post-hoc test (DUNN, 1961) was applied to identify statistical differences between the models' performance. While Kruskal-Wallis' test evidence the occurrence of statistical differences between groups (machine learning models), Dunn's test compares pair-to-pair to identify which groups are similar and which are different from each other. For these tests, the level of significance was 5%. This procedure was also done by Souza et al. (2022).

2.7 Final map prediction and uncertainty analysis

The final annual RE map was created as a mean of the 100 maps (runs) generated by the model that presented the best prediction performance. In addition, to evidence the prediction uncertainties in the map created, maps with the standard deviation (SD) and coefficient of variation ($CV \% = \text{standard deviation}/\text{mean}$) considering the 100 runs were also created (GOMES et al., 2019b; SIQUEIRA et al., 2023). Thus, a qualitative evaluation of the spatial patterns of uncertainty regarding the annual RE prediction across Brazil was made.

3. Results and discussion

3.1 Elimination of covariates

The preprocessing steps for selecting the best set of covariates for the prediction of annual RE maps eliminated seven covariates that presented variance near zero (Table 4), all of them related to terrain aspects. In addition, 35 covariates (Table 5) presented more than 95% of Spearman's correlation with other covariates, and so, were also eliminated from the prediction process. The covariates left from the total set were considered for the RFE process during the models' training.

Table 4. Covariates that presented variance near zero, and that were eliminated from the prediction process.

Covariates eliminated						
1. FLC	2. TC	3. H	4. HI	5. S	6. TPI	7. VI

Table 5. Covariates that presented more than 95% of Spearman's correlation with other covariates, and that were eliminated from the prediction process.

Covariates eliminated				
1. Bio1	8. TmaxAcc	15. Tmax10	22. Tmax12	29. TRI
2. Bio11	9. Tmin8	16. Tmin12	23. Tmin1	30. RSA
3. Tmin10	10. Tmin5	17. Tmax6	24. Tmax1	31. Prec7
4. Tmin9	11. Bio10	18. Tmax5	25. Tmax3	32. Bio16
5. TminAcc	12. Bio6	19. Tmax7	26. MDE	33. MBI
6. Tmin11	13. Tmin7	20. Tmin3	27. Bio14	34. PROC
7. Tmin4	14. Tmax11	21. Tmax9	28. Bio17	35. TANC

3.2 Performance of the models

According to the statistical metrics obtained for the models' training and test (Table 6), the RF algorithm was considered the model with the best performance (higher LCCC and R^2 values, as well as lower MAE and RMSE values for the test subset) for predicting the annual RE map for Brazil.

Table 6. Models' training and test statistical metrics obtained for the prediction of the annual RE map for Brazil. The best metrics are shown in bold.

Model	Training				Test			
	LCCC	R ²	MAE	RMSE	LCCC	R ²	MAE	RMSE
LM	0.893	0.808	1164.3	1619.2	0.892	0.802	1209.9	1667.1
KKNN	0.931	0.872	815.9	1319.4	0.931	0.869	823.1	1351.1
RF	0.938	0.886	774.5	1248.7	0.942	0.890	767.9	1238.8
SVM	0.932	0.873	838.2	1314.9	0.933	0.874	847.5	1328.5

Despite the better results of the RF algorithm, the other models also presented satisfactory metrics. This can be evidenced in Figure 3, which compares the prediction errors (MAE and RMSE) obtained by using the analyzed models instead of simply using the mean value of all observations for the rainfall gauges selected in the test subset (Null Model). Thus, the use of the tested models leads to accuracy gains while predicting the annual RE map for Brazil.

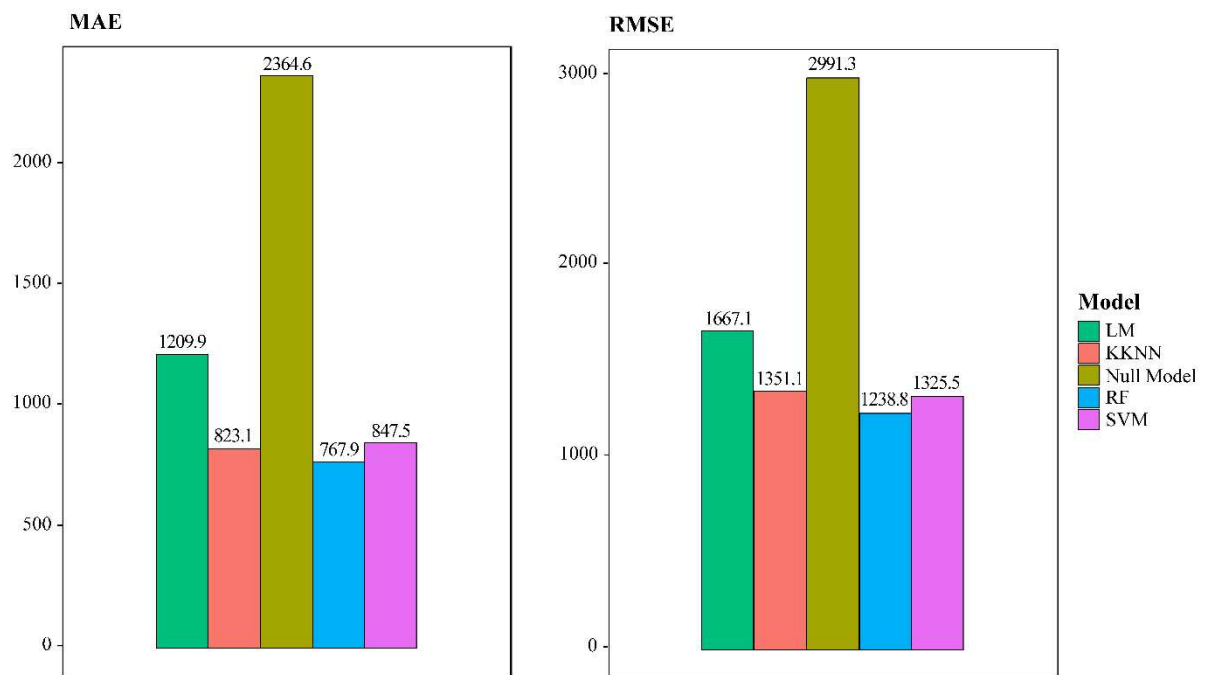


Figure 3. Comparison of the MAE and RMSE values obtained by the analyzed models and those by the Null Model (values obtained considering the mean value of all observations of the rainfall gauges selected). The values refer to the test subset.

Even with the satisfactory results for all tested models, significant differences regarding the statistical metrics obtained was found (Figure 4). The Kruskal-Wallis' and Dunn's tests evidenced that all algorithms statistically differed from each other and that using the RF model is the most suitable option for obtaining a better annual RE map for Brazil (better statistical performance). Figure 4 also evidences that the use of the RF promoted not only higher accuracy (better metrics) but also higher precision in the prediction processes since the estimated values vary less from each other over the multiple runs. Therefore, the RF model was used to obtain the annual RE map for Brazil.

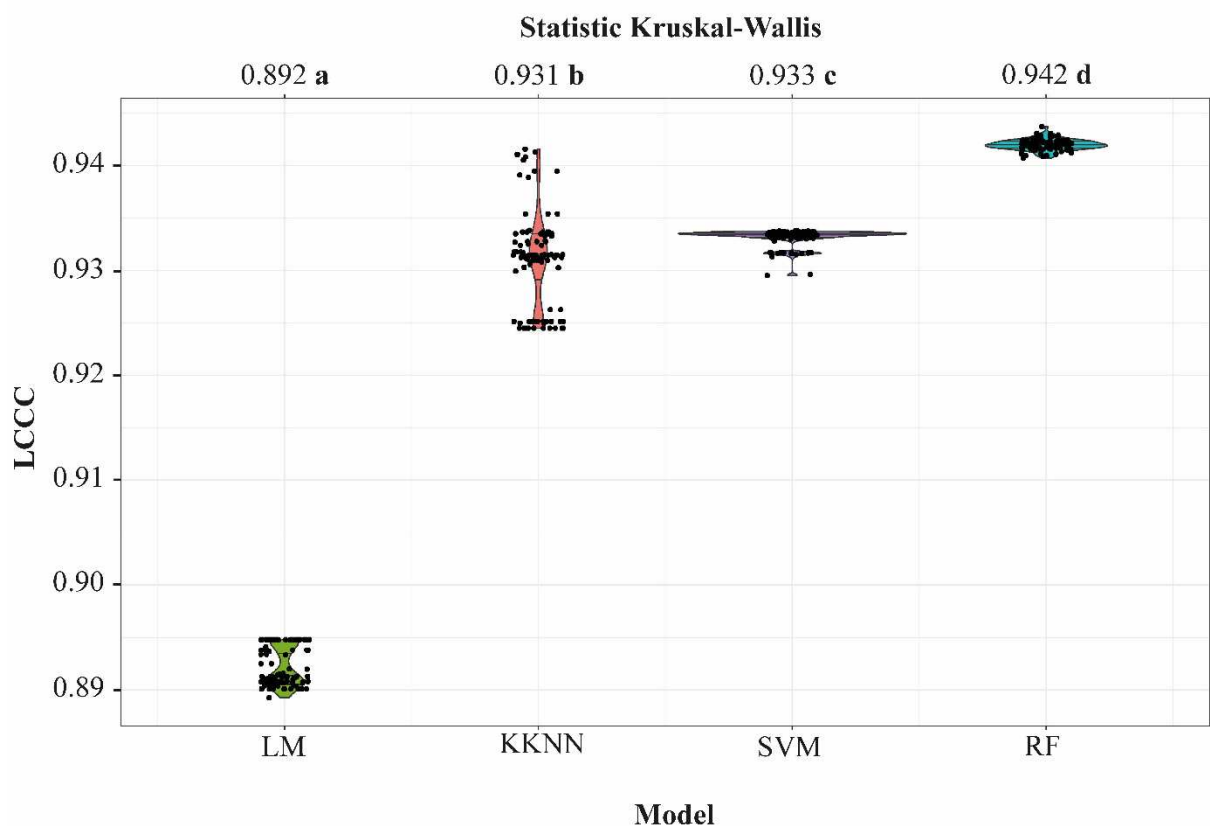


Figure 4. LCCC values for the predictions by the analyzed models, considering the test subset. The different letters following the mean values correspond to statistical difference occurrence.

3.3 Rainfall erosivity map

The annual RE map predicted by the RF model is shown in Figure 5a. According to this map, annual RE values range from 633 to 14,288 MJ mm ha⁻¹ h⁻¹ year⁻¹ over the Brazilian territory. These values differ from those for punctual annual RE values shown in the Chapter 2 of this thesis (252 to 23,916 MJ mm ha⁻¹ h⁻¹ year⁻¹), which evidences difficulty in predicting

the most extreme values in the series by the selected model. This may be related to the RF's inability to extrapolate predictions beyond the range of data considered for its training (ESTEBAN et al., 2019; HASHIMOTO et al., 2019). In addition, the decrease in the range of RE values mapped compared to the punctual estimates was also observed by Trindade et al. (2016).

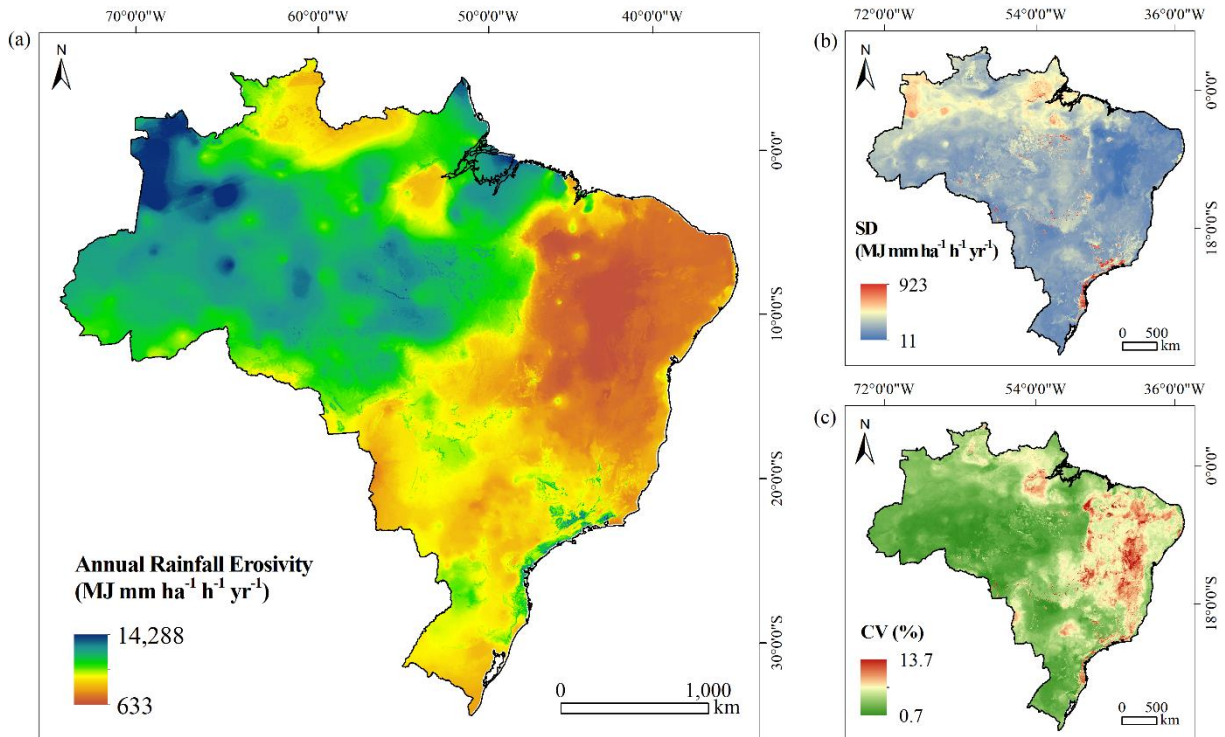


Figure 5. (a) Mean annual RE map predicted by the RF model for Brazil. Maps of (b) standard deviation (SD) and (c) coefficient of variation (CV) of the annual RE maps. These maps were created considering the 100 runs of the RF model's test.

The highest annual RE magnitudes are found for the northern region as well as the São Paulo State coast. As discussed in Chapter 2, the lowest annual RE values are concentrated in northeastern Brazil. Compared to the other annual RE maps available for Brazil (Table 7), the minimum RE magnitude obtained in the present study was inferior to the observed in most of the previous maps. Oppositely, the maximum RE value presented by the map here proposed was inferior to all other available maps. Despite this, the map here presented was created using more than three times the number of rainfall gauges considered in the other studies, which represents the greater reliability in the predicted values on the map shown by the present study.

Table 7. Annual RE range shown in the present study and in the previous ones, as well as the mapping technique used and the number of rainfall gauges considered for mapping.

Reference	Annual RE range	Mapping technique	Number of gauges
Present study	633 to 14,288	Random Forest	5,166
Silva (2004)	3,116 to 20,035	Not mentioned	1,600
Oliveira et al. (2012)	1,672 to 22,452	Kriging	80
Mello et al. (2013)	2,216 to 23,187	Kriging	928
Oliveira et al. (2015)	468 to 20,000	Inverse Distance Weighted	142
Trindade et al. (2016)	1,782 to 16,583	Kriging	1,521

For obtaining the annual RE map, the RFE process considered 25 covariates as those with more importance in the prediction (Figure 6). Figure 6 evidences the number of times that the covariates are included in the five (Top5, red bars) and ten (Top10, blue bars) most important ones, considering the 100 models' runs. The three most important were, respectively, the rainfall depth for August (Prec8), the rainfall of the coldest quarter (Bio19), and the total annual rainfall (Bio12). Disregarding the obvious relationship between annual rainfall totals (Bio12) and annual RE mapping, the best climatic predictors are related to the absence of rainfalls. Prec8 and Bio19 occur in the winter season when the lowest rainfall depths are observed for Brazil. Therefore, the spatial differences between the rainfall patterns that occur in this season present more relevance for improving the annual RE map accuracy.

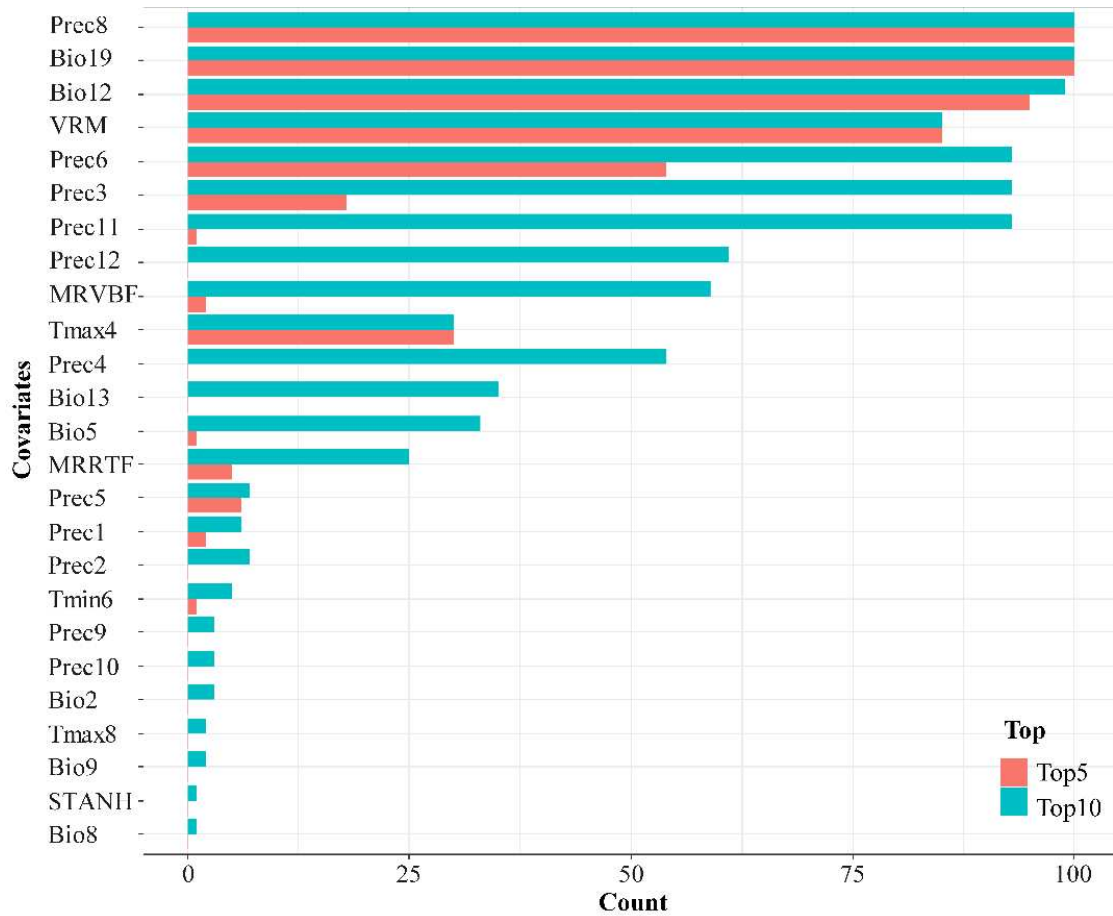


Figure 6. Number of times that the covariates selected by the RFE process are included in the five (Top5, red bars) and ten (Top10, blue bars) most important ones for predicting the annual RE map, considering the 100 RF model's runs.

In general, the terrain characteristics play a significant role in the rainfall occurrence in Brazil (FERREIRA; REBOITA, 2022; NADEEM et al., 2022). Despite this, for obtaining the annual RE map here presented only the multiresolution index of ridge top flatness (MRRTF), multiresolution index of valley bottom flatness (MRVBF), standardized height (STANH), and vector ruggedness measure (VRM) were considered as important covariates in the prediction process.

3.4 Uncertainty maps

In Figures 5b and Figure 5c are shown, respectively, the maps of SD and CV regarding the annual RE values for Brazil. As expected, higher SD values were observed for areas with the highest annual RE values. Despite this, the maximum SD values are found for the coast of the southeastern region, while the maximum annual RE values were observed for the extreme

northwestern region of the country (Figure 5a). These high SD values may be related to the orographic effect of the Serra do Mar mountains on the coast of southeastern Brazil. The terrain characteristics in this area are complex and highly influence rainfall occurrence (LUIZ-SILVA; OSCAR-JÚNIOR, 2022).

The SD analysis only shows the variability in the magnitudes of the annual RE throughout the model's runs. To overcome this limitation, the CV expresses this variation as a percentage and is not influenced by the RE magnitudes. Thus, the CV is more adequate for evidencing the uncertainties in the predictions.

In general, the CV values for the annual RE map are considered low (maximum of 13.7%) which evidence the satisfactory precision of the RF to map this variable for Brazil. The highest uncertainties are located in the northeastern of the country where also are found the lowest RE magnitudes. In addition, high CV values were found for the Serra do Mar mountains which characterizes this region not only as one of the most susceptible to high RE magnitudes but also with the lowest precision in the values mapped. This evidence the need for deeper investigations regarding the RE patterns in this area.

4. Conclusion

In this study, a gridded annual RE map was created for Brazil using machine learning techniques. For this, the best subset of covariates was selected for training and testing the models, and the uncertainty of the final map was assessed.

The LM, KNN, RF, and SVM models deferred statistically from each other in predicting annual RE values. Among these models, RF showed the best predictive performance, generating a map with RE magnitudes ranging from 633 to 14,288 MJ mm ha⁻¹ h⁻¹ year⁻¹. In addition to the total annual rainfall, the other covariates with higher importance for the predictions were the rainfall depth for August and rainfall of the coldest quarter, which evidence the relevance of the months with lower rainfall amounts in the annual RE mapping.

Finally, analyses revealed that the areas with the highest uncertainties in the values mapped are the northeastern of the country as well as the Serra do Mar mountains region in the southeastern part of the country. Despite this, the RF model presented satisfactory precision for the obtention of a gridded annual RE map for Brazil. The created map is considered an advancement regarding the availability of RE values for the entire Brazilian territory and can be used as R-factor data in the USLE and RUSLE models for predicting soil loss in Brazil.

GENERAL CONCLUSION

The assessment of RE in Brazil increased over the last decade, which demonstrates a demand from the scientific community and society in general to the obtaining reliable soil loss estimates. For calculating RE values, pluviographic rainfall data and regression equations are the main used methods, while the use of synthetic series of rainfall is a promising alternative for this purpose. In Brazil, the erosivity of rainfalls is mainly represented by the EI₃₀ index. In addition, a disparity in the spatial distribution of erosivity studies is observed, with a concentration of these studies in the southeast region.

The use of a large national database made possible the obtention of long-term mean RE values never described in the literature before for Brazil, ranging from 252 to 23,916 MJ mm ha⁻¹ h⁻¹ year⁻¹. Due to the continental scale of the Brazilian territory, a great spatial variation of the RE values was observed, with a mean annual value of 5,620 MJ mm ha⁻¹ h⁻¹ year⁻¹. Complementing the RE assessment, the ED magnitudes were introduced to the entire country and allowed the identification of high-intensity rainfall spots in Brazil all year long, which can contribute to the understanding of the RE impacts on soil loss.

Another novelty to the RE assessment in the country, the RECI and RESI indexes evidenced that in western Ceará and central Rio Grande do Norte states the RE occurs mostly concentrated in specific months while the southern region of the country has the most regular distribution of the RE values throughout the year. In addition, the RE gravity center analyses showed a latitudinal path variation (north-south migration) of the erosivity throughout the months and that the annual RE gravity center of Brazil is within the Goiás state. These results constitute an advancement in the RE studies of the country and have the potential to improve the understanding of the spatial RE patterns.

A division of the Brazilian territory into homogeneous regions regarding RE was here presented. The applied methodology defined eleven regions with different patterns concerning the monthly RE magnitudes. In addition, regionalized regression models with good estimation accuracy were adjusted and validated for each region. These established models solve the common problem of the unavailability of RE estimation models for the entire country which used to lead to a generalization of some models to many regions and to incorrect RE estimates. Another advantage of using the proposed models is that they only require monthly rainfall depths for the predictions, so they are important tools for improving the availability of RE values in Brazil.

Among the machine learning models tested, RF showed the best predictive performance

for the obtention of a gridded annual RE map, and constitutes an alternative approach to the use of traditional interpolation methods for mapping RE in Brazil. For this prediction, the covariates with higher importance were the total annual rainfall, as well as the rainfall depth for August and rainfall of the coldest quarter, which evidence the relevance of the months with lower rainfall amounts in the annual RE mapping.

Further analyses revealed that northeastern Brazil as well as the Serra do Mar mountains region in the southeastern part of the country are the areas with the highest uncertainties in the map here presented. Despite this, the RF model presented satisfactory precision for the obtention of a gridded annual RE map for Brazil. This map improves the understanding of the RE spatial patterns and can be used as R-factor data in the USLE and RUSLE models, which may enhance planning soil and water management and conservation in the country.

For future studies, a reconstruction of the past RE values using the regionalized models here proposed may contribute to the identification of different RE patterns over the years as well as to the assessment of non-stationarity and increasing or decreasing trends. Another recommendation is the assessment of the machine learning algorithms tested here to the obtention of monthly RE maps, as well as for creating the first map of ED for the entire Brazilian territory. Also, the use of covariate datasets different from those here presented is suggested, which can help map RE with greater accuracy. Finally, the results presented in this thesis highlight the need for further studies in areas that remain with lack of RE information, such as the north and center-west regions of the country, as well as for areas in which the RE has notably high magnitudes, such as the Serra do Mar region in the southeast of Brazil.

REFERENCES

- AGUIAR, L. F.; CATALDI, M. Social and environmental vulnerability in Southeast Brazil associated with the South Atlantic Convergence Zone. **Natural Hazards**, v. 109, n. 3, p. 2423–2437, 14 dez. 2021.
- ALBUQUERQUE, A. W. et al. Parâmetros erosividade da chuva e da enxurrada correlacionados com as perdas de solo de um solo Bruno Não-Cálcico Vértico em Sumé (PB). **Revista Brasileira de Ciência do Solo**, v. 22, n. 4, p. 743–749, 1998.
- ALBUQUERQUE, A. W. et al. Parâmetros erosividade da chuva, da enxurrada e da chuva-enxurrada correlacionados com as perdas de solo de um luvisolo. **Revista Brasileira de Ciência do Solo**, v. 26, p. 695–703, 2002.
- ALBUQUERQUE, A. W. DE; CHAVES, I. B. DE; VASQUES FILHO, J. Características físicas da chuva correlacionadas com as perdas de solo num regossolo eutrófico de Caruaru (PE). **Revista Brasileira de Ciência do Solo**, v. 18, n. 2, p. 279–283, 1994.
- ALEWELL, C. et al. Using the USLE: Chances, challenges and limitations of soil erosion modelling. **International Soil and Water Conservation Research**, v. 7, n. 3, p. 203–225, 2019.
- ALEWELL, C. et al. Global phosphorus shortage will be aggravated by soil erosion. **Nature Communications**, v. 11, n. 1, p. 4546, 11 set. 2020.
- ALMAGRO, A. et al. Projected climate change impacts in rainfall erosivity over Brazil. **Scientific Reports**, v. 7, n. 1, p. 1–12, 2017.
- ALMEIDA, A. Q. et al. Modelling the spatial dependence of the rainfall erosivity index in the Brazilian semiarid region. **Pesquisa Agropecuária Brasileira**, v. 52, n. 6, p. 371–379, 2017.
- ALMEIDA, C. O. S. et al. Erosividade da chuva em municípios do Mato Grosso: distribuição sazonal e correlações com dados pluviométricos Rainfall erosivity in municipal districts of Mato Grosso: seasonal distribution and correlation with rainfall data. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 16, n. 2, p. 142–152, 2012.
- ALVARENGA, L. A. Precipitação no sudeste brasileiro e sua relação com a Zona de Convergência do Atlântico Sul. **Revista Agrogeoambiental**, v. 4, n. 2, p. 1–7, 2012.
- ALVARES, C. A. et al. Köppen’s climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.
- ALVES, G. J. et al. Natural disaster in the mountainous region of Rio de Janeiro state, Brazil: Assessment of the daily rainfall erosivity as an early warning index. **International Soil and Water Conservation Research**, abr. 2022.
- ALVES, K. M. A. DA S.; CAVALVANTI, L. C. DE S.; NÓBREGA, R. S. Eventos extremos e risco de inundação: uma análise do comportamento evolutivo dos Distúrbios Ondulatórios de Leste em junho de 2010 sobre a bacia do rio Una - Pernambuco. **GeoTextos**, v. 9, n. 2, p. 173–189, 2013.

- ANDRADE, A. F. et al. Erosividade e padrões de precipitação pluvial para Aragarças – GO. **Brazilian Journal of Development**, v. 6, n. 4, p. 21931–21950, 2020a.
- ANDRADE, A. F. et al. Precipitation Patterns and Rainfall Erosivity Return Period Under Savanna Conditions in Formosa, Goiás, Brazil. **Journal of Agricultural Studies**, v. 8, n. 4, p. 554, 2020b.
- ANJOS, J. C. R. et al. Intensity and distribution in the space-time of the rain erosivity in Goiás and Federal District states. **Scientific Electronic Archives**, v. 13, n. 10, p. 1, 2020.
- AQUINO, R. F. et al. Spatial variability of the rainfall erosivity in southern region of Minas Gerais state, Brazil. **Ciência e Agrotecnologia**, v. 36, n. 5, p. 533–542, 2012.
- ARAÚJO, M. L. S. et al. Spatiotemporal dynamics of soybean crop in the Matopiba region, Brazil (1990–2015). **Land Use Policy**, v. 80, p. 57–67, 2019.
- ATTORRE, F. et al. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. **International Journal of Climatology**, v. 27, n. 13, p. 1825–1843, 2007.
- AVANZI, J. C. et al. Modeling of the rainfall and R-factor for Tocantins state, Brazil. **Revista Brasileira de Ciência do Solo**, v. 43, p. 1–14, 2019.
- BACK, Á. J. et al. Erosive rainfall in rio do peixe valley in Santa Catarina, Brazil: Part I - determination of the erosivity index. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 21, n. 11, p. 774–779, 2017.
- BACK, Á. J. Erosividade da chuva para a região do Planalto Serrano de Santa Catarina, Brasil. **Revista de Ciências Agrárias**, v. 41, n. 2, p. 298–308, 2018.
- BACK, Á. J.; ALBERTON, J. V.; POLETO, C. Erosivity Index and Characteristics of Erosive Rainfall from the Far Western Region of Santa Catarina, Brazil. **Journal of Environmental Engineering**, v. 144, n. 7, p. 04018049, 2018.
- BACK, Á. J.; GONÇALVES, F. N.; FAN, F. M. Spatial, seasonal, and temporal variations in rainfall aggressiveness in the South of Brazil. **Engenharia Agrícola**, v. 39, n. 4, p. 466–475, 2019.
- BACK, A. J.; POLETO, C. Distribuição espacial e temporal da erosividade das chuvas no estado de Santa Catarina, Brasil. **Revista Brasileira de Climatologia**, v. 14, n. 22, p. 381–403, 2018.
- BAENA, L. G. N. et al. Programa computacional para geração de séries sintéticas de dados climáticos. **Engenharia na Agricultura**, v. 13, p. 210–220, 2005.
- BAGWAN, W. A. An assessment of rainfall-induced land degradation condition using Erosivity Density (ED) and heatmap method for Urmodi River watershed of Maharashtra, India. **Journal of Sedimentary Environments**, v. 5, n. 3, p. 279–292, 8 set. 2020.
- BALLABIO, C. et al. Mapping monthly rainfall erosivity in Europe. **Science of the Total Environment**, v. 579, p. 1298–1315, 2017.

BAZZANO, M. G. P.; ELTZ, F. L. F.; CASSOL, E. A. Erosividade, coeficiente de chuva, padrões e período de retorno das chuvas de Quaraí, RS. **Revista Brasileira de Ciência do Solo**, v. 31, n. 5, p. 1205–1217, 2007.

BAZZANO, M. G. P.; ELTZ, F. L. F.; CASSOL, E. A. Erosividade e características hidrológicas das chuvas de Rio Grande (RS). **Revista Brasileira de Ciência do Solo**, v. 34, n. 1, p. 235–244, 2010.

BERTOL, I. et al. Erodibilidade de um nitossolo háplico alumínico determinada em condições de campo. **Revista Brasileira de Ciência do Solo**, v. 31, n. 3, p. 541–549, 2007.

BERTOSSI, A. P. A. et al. Seleção e agrupamento de indicadores da qualidade de águas utilizando Estatística Multivariada. **Semina: Ciências Agrárias**, v. 34, n. 5, p. 2025, 17 out. 2013.

BIVAND, R. **rgrass7: interface between grass 7 geographical information system and R. R package version 0.1-10.** , 2017.

BOLAR, K. **STAT: Interactive Document for Working with Basic Statistical Analysis. R package version 2.3.** , 2019.

BORRELLI, P. et al. Land use and climate change impacts on global soil erosion by water (2015-2070). **Proceedings of the National Academy of Sciences**, v. 117, n. 36, p. 21994–22001, 8 set. 2020.

BORRELLI, P. et al. Soil erosion modelling: A global review and statistical analysis. **Science of the Total Environment**, v. 780, p. 146494, 2021.

BRASIL. **Gross Value of Agricultural Production. Ministry of Agriculture, Livestock and Supply.** Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/valor-bruto-da-producao-agropecuaria-vbp>>. Acesso em: 22 mar. 2022.

BREIMAN, LEO; CUTLER, A. **Package ‘randomForest’. Breiman and Cutler’s Random Forests for Classification and Regression.** , 2022.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

BRENNING, A. Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. . **Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie**, v. 19, n. 23–32, p. 410, 2008.

BRITO, C. S. DE et al. Long-term basin-scale comparison of two high-resolution satellite-based remote sensing datasets for assessing rainfall and erosivity in a basin in the Brazilian semiarid region. **Theoretical and Applied Climatology**, n. 0123456789, 2021.

CABEDA, M. S. V. **Computation of storm EI value.** [s.l.] West Lafayette: Purdue University, 1976.

CARDOSO, D. P. et al. RainfallErosivityFactor: An R package for rainfall erosivity (R-factor) determination. **Catena**, v. 189, n. February, p. 104509, 2020.

CARUSO, C.; QUARTA, F. Interpolation methods comparison. **Computers & Mathematics with Applications**, v. 35, n. 12, p. 109–126, 1998.

CARVALHO, D. F. DE et al. Distribuição, probabilidade de ocorrência e período de retorno dos índices de erosividade EI30 e KE>25 em Seropédica - RJ. **Engenharia Agrícola**, v. 30, n. 2, p. 244–252, 2010.

CARVALHO, D. F. DE et al. Rainfall erosivity for the State of Rio de Janeiro estimated by artificial neural network. **Engenharia Agrícola**, v. 32, n. 1, p. 197–207, 2012.

CARVALHO, L. M. V.; CAVALCANTI, I. F. A. The South American Monsoon System (SAMS). Em: [s.l.] Springer Climate, 2016. p. 121–148.

CARVALHO, M. P.; HERNANI, L. C. Parâmetros de erosividade da chuva e da enxurrada correlacionados com perdas de solo e erodibilidade de um Latossolo Roxo de Dourados (MS). **Revista Brasileira de Ciência do Solo**, v. 25, n. 1, p. 137–146, 2001.

CARVALHO, N. O. **Hidrossedimentologia Prática**. Rio de Janeiro: CPRM–Companhia de Pesquisa em Recursos Minerais, 1994.

CARVALHO, N. O. **Hidrossedimentologia Prática**. Rio de Janeiro: Interciência, 2008.

CASSOL, E. A. et al. Erosividade, padrões hidrológicos, período de retorno e probabilidade de ocorrência das chuvas em São Borja, RS. **Revista Brasileira de Ciência do Solo**, v. 32, p. 1239–1251, 2008.

CASTAGNA, D. et al. Rainfall erosivity in municipalities of the Brazilian Cerrado biome. **Nativa**, v. 10, n. 3, p. 373–386, 20 ago. 2022.

CASTRO, R. M. et al. Soil losses related to land use and rainfall seasonality in a watershed in the Brazilian Cerrado. **Journal of South American Earth Sciences**, v. 119, p. 104020, nov. 2022.

CECÍLIO, R. A. et al. Assessing rainfall erosivity indices through synthetic precipitation series and artificial neural networks. **Anais da Academia Brasileira de Ciências**, v. 85, n. 4, p. 1523–1535, 2013.

CECÍLIO, R. A. et al. Database of rainfall erosivity factor for 141 locations in Brazil. **Latin American Data in Science**, v. 1, n. 3, p. 95–101, 2021.

CHALISE, D. et al. Assessing the Impacts of Tillage and Mulch on Soil Erosion and Corn Yield. **Agronomy**, v. 10, n. 1, p. 63, 2 jan. 2020.

CHALISE, D.; KUMAR, L.; KRISTIANSEN, P. Land Degradation by Soil Erosion in Nepal: A Review. **Soil Systems**, v. 3, n. 1, p. 12, 8 fev. 2019.

CHEN, Y. et al. Applicability of two satellite-based precipitation products for assessing rainfall erosivity in China. **Science of The Total Environment**, v. 757, p. 143975, fev. 2021.

COGO, C. M.; ELTZ, F. L. F.; CASSOL, E. A. Erosividade das chuvas em Santa Maria, RS, determinada pelo índice EI30. **Revista Brasileira de Agrometeorologia**, v. 14, n. 3, p. 1–11, 2006.

- COLMAN, C. B. et al. Effects of climate and land-cover changes on soil erosion in Brazilian Pantanal. **Sustainability**, v. 11, n. 24, p. 7053, 2019.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, set. 1995.
- CUNHA, E. R. DA et al. Assessment of current and future land use/cover changes in soil erosion in the Rio da Prata basin (Brazil). **Science of The Total Environment**, v. 818, p. 151811, abr. 2022.
- DA SILVA JÚNIOR, J. C. et al. Random forest techniques for spatial interpolation of evapotranspiration data from Brazilian's Northeast. **Computers and Electronics in Agriculture**, v. 166, p. 105017, nov. 2019.
- DAS, S.; JAIN, M. K.; GUPTA, V. A step towards mapping rainfall erosivity for India using high-resolution GPM satellite rainfall products. **CATENA**, v. 212, p. 106067, maio 2022.
- DASH, C. J.; DAS, N. K.; ADHIKARY, P. P. Rainfall erosivity and erosivity density in Eastern Ghats Highland of east India. **Natural Hazards**, v. 97, n. 2, p. 727–746, 2019.
- DEHGHAN, Z.; ESLAMIAN, S. S.; MODARRES, R. Spatial clustering of maximum 24-h rainfall over Urmia Lake Basin by new weighting approaches. **International Journal of Climatology**, v. 38, n. 5, p. 2298–2313, abr. 2018.
- DEMŠAR, U. et al. Principal Component Analysis on Spatial Data: An Overview. **Annals of the Association of American Geographers**, v. 103, n. 1, p. 106–128, jan. 2013.
- DI LENA, B.; CURCI, G.; VERGNI, L. Analysis of rainfall erosivity trends 1980–2018 in a complex terrain region (Abruzzo, central italy) from rain gauges and gridded datasets. **Atmosphere**, v. 12, n. 6, p. 657, 2021.
- DIAS, L. C. et al. Patterns of land use, extensification, and intensification of Brazilian agriculture. **Global Change Biology**, v. 22, n. 8, p. 2887–2903, 2016.
- DIAS, R. L. S. et al. Machine learning models applied to TSS estimation in a reservoir using multispectral sensor onboard to RPA. **Ecological Informatics**, v. 65, p. 101414, nov. 2021.
- DIODATO, N. et al. Reconstruction of erosivity density in northwest Italy since 1701. **Hydrological Sciences Journal**, v. 66, n. 7, p. 1185–1196, 19 maio 2021.
- DISSANAYAKE, D.; MORIMOTO, T.; RANAGALAGE, M. Accessing the soil erosion rate based on RUSLE model for sustainable land use management: a case study of the Kotmale watershed, Sri Lanka. **Modeling Earth Systems and Environment**, v. 5, n. 1, p. 291–306, 22 mar. 2019.
- DUAN, X. et al. Quantifying soil erosion effects on soil productivity in the dry-hot valley, southwestern China. **Environmental Earth Sciences**, v. 75, n. 16, p. 1164, 2016.
- DUNN, O. J. Multiple Comparisons among Means. **Journal of the American Statistical Association**, v. 56, n. 293, p. 52–64, mar. 1961.

DURÃES, M. F.; MELLO, C. R. DE. Distribuição espacial da erosão potencial e atual do solo na Bacia Hidrográfica do Rio Sapucaí, MG. **Engenharia Sanitária e Ambiental**, v. 21, n. 4, p. 677–685, 2016.

DURKEE, J. D.; MOTE, T. L.; SHEPHERD, J. M. The Contribution of Mesoscale Convective Complexes to Rainfall across Subtropical South America. **Journal of Climate**, v. 22, n. 17, p. 4590–4605, 1 set. 2009.

ELTZ, F. L. F.; CASSOL, E. A.; PASCOTINI, P. B. Potencial erosivo e características das chuvas de Encruzilhada do Sul, RS. **Revista Brasileira de Engenharia Agrícola e Ambiental**, n. 55, p. 331–337, 2011.

ESTEBAN, J. et al. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. **Remote Sensing**, v. 11, n. 16, p. 1944, 20 ago. 2019.

FALCÃO, C. J. L. M.; DUARTE, S. M. DE A.; DA SILVA VELOSO, A. Estimating potential soil sheet Erosion in a Brazilian semiarid county using USLE, GIS, and remote sensing data. **Environmental Monitoring and Assessment**, v. 192, n. 1, 2020.

FERREIRA, G. W. S.; REBOITA, M. S. A New Look into the South America Precipitation Regimes: Observation and Forecast. **Atmosphere**, v. 13, n. 6, p. 873, 26 maio 2022.

FERREIRA, R. G. et al. Machine learning models for streamflow regionalization in a tropical watershed. **Journal of Environmental Management**, v. 280, p. 111713, fev. 2021.

FICK, S. E.; HIJMANS, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. **International Journal of Climatology**, v. 37, n. 12, p. 4302–4315, 2017.

FOSTER, G. R. et al. Conversion of the universal soil loss equation to SI metric units. **Journal of Soil and Water Conservation**, v. 36, n. 6, p. 335–359, 1981.

FOSTER, G. R. **User's Reference Guide: Revised Universal Soil Loss Equation (RUSLE2)**. Washington, DC: [s.n.].

GALILI, T. **dendextend: Extending “dendrogram” Functionality in R. R package version 1.14.**, 2020.

GARRETT, R. D. et al. Intensification in agriculture-forest frontiers: Land use responses to development and conservation policies in Brazil. **Global Environmental Change**, v. 53, p. 233–243, 2018.

GOLLNOW, F.; LAKES, T. Policy change, land use, and agriculture: The case of soy production and cattle ranching in Brazil, 2001–2012. **Applied Geography**, v. 55, p. 203–211, 2014.

GOMES, L. et al. Agricultural Expansion in the Brazilian Cerrado: Increased Soil and Nutrient Losses and Decreased Agricultural Productivity. **Land**, v. 8, n. 1, p. 12, 8 jan. 2019a.

GOMES, L. C. et al. Modelling and mapping soil organic carbon stocks in Brazil. **Geoderma**, v. 340, p. 337–350, abr. 2019b.

GONÇALVES, F. A. et al. Índices e espacialização da erosividade das chuvas para o Estado do Rio de Janeiro. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 10, n. 2, p. 269–276, 2006.

GUDURU, J. U.; JILO, N. B. Assessment of rainfall-induced soil erosion rate and severity analysis for prioritization of conservation measures using RUSLE and Multi-Criteria Evaluations Technique at Gidabo watershed, Rift Valley Basin, Ethiopia. **Ecohydrology & Hydrobiology**, v. 23, n. 1, p. 30–47, jan. 2023.

GUIMARÃES, D. V. et al. Modeling of soil losses on a yellow argisol under planted forest. **Floresta e Ambiente**, v. 26, n. 1, p. 1–11, 2019.

GUO, B. et al. Multiple spatial–temporal scale change patterns of rainfall erosivity in China over past 58 years based on gravity centre model. **Geomatics, Natural Hazards and Risk**, v. 10, n. 1, p. 2200–2219, 1 jan. 2019.

HAIR, JR. J. F.; BLACK, W. C.; SANT'ANNA, A. SCHLUP. **Análise multivariada de dados (6a. ed.)**. [s.l.: s.n.].

HASHIMOTO, H. et al. High-resolution mapping of daily climate variables by aggregating multiple spatial data sets with the random forest algorithm over the conterminous United States. **International Journal of Climatology**, v. 39, n. 6, p. 2964–2983, 10 maio 2019.

HECHENBICHLER, K.; SCHLIEP, K. **Weighted k-nearest-neighbor techniques and ordinal classification, Discussion Paper 399**. [s.l.: s.n.].

HICKMANN, C. et al. Erosividade das chuvas em Uruguaiana, RS, determinada pelo índice EI30, com base no período de 1963 a 1991. **Revista Brasileira de Ciência do Solo**, v. 32, p. 825–831, 2008.

HIJMANS, R. J. **The raster package**. <https://rspatial.org/raster/pkg/RasterPackage.pdf>, 2023.

HUANG, J. et al. Climatology of rainfall erosivity during 1961–2012 in Jiangsu Province, southeast China. **Natural Hazards**, v. 98, n. 3, p. 1155–1168, 16 set. 2019.

HUDSON, N. **Soil conservation**. New York: Ithaca, 1971.

IBGE. **Biomass e sistema costeiro-marinho do Brasil: compatível com a escala 1:250000. Série Relatórios Metodológicos, vol. 45**. Rio de Janeiro: [s.n.]. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv101676.pdf>>.

IBGE. **Bacias e Divisões Hidrográficas do Brasil. Série Relatórios Metodológicos, vol. 48**. Rio de Janeiro: [s.n.]. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv101854.pdf>>. Acesso em: 20 set. 2022.

INMET. **Monthly and Annual Normal Rainfall (1991-2020). Brazilian Climatological Normals**. Disponível em: <<https://portal.inmet.gov.br/normais>>. Acesso em: 22 mar. 2022.

JIA, L. et al. Temporal and spatial variation of rainfall erosivity in the Loess Plateau of China and its impact on sediment load. **CATENA**, v. 210, p. 105931, mar. 2022.

JOHANNSEN, L. L. et al. An update of the spatial and temporal variability of rainfall erosivity (R-factor) for the main agricultural production zones of Austria. **CATENA**, v. 215, p. 106305, ago. 2022.

KARATZOGLU, A.; SMOLA, A.; HORNIK, K. Package ‘kernlab’. Kernel-Based Machine Learning Lab. 2023.

KARIMI, S. M. et al. Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. **ISH Journal of Hydraulic Engineering**, v. 26, n. 4, p. 376–386, 1 out. 2020.

KASSAMBARA, A. Multivariate Analysis II: Practical Guide to Principal Component Methods in R. **Sthda**, p. 170, 2017.

KASSAMBARA, A.; MUNDT, F. Package ‘factoextra’. **Extract and visualize the results of multivariate data analyses**. [s.l.: s.n.]. Disponível em: <<https://cran.microsoft.com/snapshot/2016-11-30/web/packages/factoextra/factoextra.pdf>>.

KHALEDIAN, Y.; MILLER, B. A. Selecting appropriate machine learning methods for digital soil mapping. **Applied Mathematical Modelling**, v. 81, p. 401–418, maio 2020.

KIM, J. et al. Use of a high-resolution-satellite-based precipitation product in mapping continental-scale rainfall erosivity: A case study of the United States. **CATENA**, v. 193, p. 104602, out. 2020.

KÖPPEN, W. Das geographische System der Klimate. **Gebruder Borntrager**, v. 1, p. 1–44, 1936.

KRUSKAL, W. H.; WALLIS, W. A. Use of Ranks in One-Criterion Variance Analysis. **Journal of the American Statistical Association**, v. 47, n. 260, p. 583, dez. 1952.

KUBAT, M. **An Introduction to Machine Learning**. Cham: Springer International Publishing, 2021.

KUHN, M. **The caret Package**.

KUHN, M. et al. Package ‘caret’. **The R Journal**, v. 223, n. 7, 2020.

KUHN, M.; JOHNSON, K. **Applied Predictive Modeling**. New York, NY: Springer New York, 2013.

LANTZ, B. **Machine learning with R: expert techniques for predictive modeling**. Third edition ed. [s.l.] Packt publishing ltd., 2019.

LÊ, S.; JOSSE, J.; HUSSON, F. FactoMineR: An R package for multivariate analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 1–18, 2008.

LEE, J. et al. Evaluation of Rainfall Erosivity Factor Estimation Using Machine and Deep Learning Models. **Water**, v. 13, n. 3, p. 382, 1 fev. 2021.

LEE, S. et al. Estimation of rainfall erosivity factor in Italy and Switzerland using Bayesian optimization based machine learning models. **Catena**, v. 211, p. 105957, abr. 2022.

LENSE, G. H. E. et al. Soil losses in the State of Rondônia, Brazil. **Ciência Rural**, v. 51, n. 5, 2021.

LI, X.; YE, X. Variability of rainfall erosivity and erosivity density in the Ganjiang River Catchment, China: Characteristics and influences of climate change. **Atmosphere**, v. 9, n. 2, p. 48, 2018.

LIMA, C. A. DE et al. Characteristics of rainfall and erosion under natural conditions of land use in semiarid regions. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 17, n. 11, p. 1222–1229, 2013.

LIN, L. I.-K. A Concordance Correlation Coefficient to Evaluate Reproducibility. **Biometrics**, v. 45, n. 1, p. 255, mar. 1989.

LIU, B. et al. The assessment of soil loss by water erosion in China. **International Soil and Water Conservation Research**, v. 8, n. 4, p. 430–439, dez. 2020a.

LIU, Y. et al. Global rainfall erosivity changes between 1980 and 2017 based on an erosivity model using daily precipitation data. **Catena**, v. 194, p. 104768, 2020b.

LOBO, G. P. et al. Evaluation and improvement of the CLIGEN model for storm and rainfall erosivity generation in Central Chile. **Catena**, v. 127, p. 206–213, 2015.

LOMBARDI NETO, F. **Rainfall erosivity-its distribution and relationship with soil loss at Campinas, Brazil**. [s.l.] Purdue University, West Lafayette, IN, 1977.

LOMBARDI NETO, F.; MOLDENHAUER, W. C. Rainfall erosivity: Its distribution and relationship with soil loss at Campinas, state of São Paulo, Brazil. **Bragantia**, n. 51, p. 189–196, 1992.

LUIZ-SILVA, W.; OSCAR-JÚNIOR, A. C. Climate extremes related with rainfall in the State of Rio de Janeiro, Brazil: a review of climatological characteristics and recorded trends. **Natural Hazards**, v. 114, n. 1, p. 713–732, 6 out. 2022.

MACHADO, C. C. C. et al. Distúrbio Ondulatório de Leste como condicionante a eventos extremos de precipitação em Pernambuco. **Revista Brasileira de Climatologia**, v. 11, p. 146–188, 2012.

MACHADO, D. O. et al. Rainfall erosivity for Pantanal biome. **Engenharia Sanitária e Ambiental**, v. 19, n. 2, p. 195–201, 2014.

MACHADO, R. L. et al. Erosividade das chuvas associada a períodos de retorno e probabilidade de ocorrência no Estado do Rio de Janeiro. **Revista Brasileira de Ciência do Solo**, v. 37, n. 2, p. 529–547, 2013.

MACHADO, R. L. et al. Multivariate analysis of erosivity indices and rainfall physical characteristics associated with rainfall patterns in Rio de Janeiro. **Revista Brasileira de Ciência do Solo**, v. 41, p. 1–15, 2017.

MAPBIOMAS PROJECT. **Collection 6.0 of Annual Land Use Land Cover Maps of Brazil**. , 2021. Disponível em: <<http://mapbiomas.org>>

- MARQUES, J. J. G. S. M. et al. Índices de erosividade da chuva, perdas de solo e fator erodibilidade para dois solos da região dos cerrados - primeira aproximação. **Revista Brasileira de Ciência do Solo**, v. 21, n. 3, p. 427–434, 1997.
- MAZURANA, J. et al. Erosividade, padrões hidrológicos e período de retorno das chuvas erosivas de Santa Rosa (RS). **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 13, n. suppl, p. 975–983, 2009.
- MEDEIROS, G. DE O. R. et al. Estimates of Annual Soil Loss Rates in the State of São Paulo, Brazil. **Revista Brasileira de Ciência do Solo**, v. 40, n. 0, 2016.
- MELLO, C. R. et al. Multivariate models for annual rainfall erosivity in Brazil. **Geoderma**, v. 202–203, p. 88–102, 2013.
- MELLO, C. R. et al. Assessing the climate change impacts on the rainfall erosivity throughout the twenty-first century in the Grande River Basin (GRB) headwaters, Southeastern Brazil. **Environmental Earth Sciences**, v. 73, p. 8683–8698, 2015a.
- MELLO, C. R. D. et al. Distribuição espacial da precipitação e da erosividade da chuva mensal e anual no Estado do Espírito Santo. **Revista Brasileira de Ciência do Solo**, v. 36, p. 1878–1891, 2012.
- MELLO, C. R. DE et al. Erosividade mensal e anual da chuva no Estado de Minas Gerais. **Pesquisa Agropecuária Brasileira**, v. 42, n. 4, p. 537–545, 2007.
- MELLO, C. R. DE et al. Interpolation methods for improving the RUSLE R-factor mapping in Brazil. **Journal of Soil and Water Conservation**, v. 70, n. 3, p. 182–197, 2015b.
- MELLO, C. R. DE et al. Daily rainfall erosivity as an indicator for natural disasters: assessment in mountainous regions of southeastern Brazil. **Natural Hazards**, v. 103, n. 1, p. 947–966, 2020.
- MELLO, D. C. DE et al. A new methodological framework for geophysical sensor combinations associated with machine learning algorithms to understand soil attributes. **Geoscientific Model Development**, v. 15, n. 3, p. 1219–1246, 10 fev. 2022.
- MELVILLE, T.; WUDDIVIRA, M.; SUTHERLAND, M. Geospatial modelling of rainfall erosivity in the humid tropics using remotely sensed data. **Earth Science Informatics**, v. 15, n. 2, p. 891–904, 8 jun. 2022.
- MONTEBELLER, C. A. et al. Variabilidade espacial do potencial erosivo das chuvas no Estado do Rio de Janeiro. **Engenharia Agrícola**, v. 27, n. 2, p. 426–435, 2007.
- MONTENEGRO, A. A. DE A. et al. Temporal dynamics of soil moisture and rainfall erosivity in a tropical volcanic archipelago. **Journal of Hydrology**, v. 563, n. October 2017, p. 737–749, 2018.
- MORAES, F. D. S. et al. Atmospheric characteristics favorable for the development of mesoscale convective complexes in southern Brazil. **Climate Research**, v. 80, n. 1, p. 43–58, 9 abr. 2020.

- MOREIRA, L. L. et al. Spatial–temporal dynamics of rainfall erosivity in the state of Espírito Santo (Brazil) from remote sensing data. **World Journal of Science, Technology and Sustainable Development**, v. 17, n. 3, p. 297–309, 2020.
- MOREIRA, M. C. et al. Desenvolvimento e análise de uma rede neural artificial para estimativa da erosividade da chuva para o Estado de São Paulo. **Revista Brasileira de Ciência do Solo**, v. 30, p. 1069–1076, 2006a.
- MOREIRA, M. C. et al. Programa computacional para estimativa da erosividade da chuva no estado de São Paulo utilizando redes neurais artificiais. **Engenharia na Agricultura**, v. 14, n. 2, p. 88–92, 2006b.
- MOREIRA, M. C. et al. NetErosividade MG: erosividade da chuva em Minas Gerais. **Revista Brasileira de Ciência do Solo**, v. 32, n. 3, p. 1349–1353, 2008.
- MOREIRA, M. C. et al. Redes neurais artificiais para estimativa mensal da erosividade da chuva no estado de Minas Gerais. **Engenharia na Agricultura**, v. 17, n. 1, p. 75–83, 2009.
- MOREIRA, M. C. et al. Programa computacional para estimativa da erosividade da chuva no Espírito Santo. **Engenharia na Agricultura**, v. 20, n. 4, p. 350–356, 2012.
- MOREIRA, M. C. et al. Spatial Interpolation of Rainfall Erosivity Using Artificial Neural Networks for Southern Brazil Conditions. **Revista Brasileira de Ciência do Solo**, v. 40, p. 1–11, 2016.
- MOSAVI, A. et al. Susceptibility mapping of soil water erosion using machine learning models. **Water**, v. 12, n. 7, p. 1995, 2020.
- NADEEM, M. U. et al. Multiscale Ground Validation of Satellite and Reanalysis Precipitation Products over Diverse Climatic and Topographic Conditions. **Remote Sensing**, v. 14, n. 18, p. 4680, 19 set. 2022.
- NAZUHAN, M.; RAHAMAN, Z. A.; OTHMAN, Z. A Review on Rainfall Erosivity (R factor) in Universal Soil Loss Equation. **International Journal of Academic Research in Business and Social Sciences**, v. 8, n. 2, p. 816–822, 2018.
- NEARING, M. A. et al. Rainfall erosivity: An historical review. **Catena**, v. 157, n. May, p. 357–362, 2017.
- NERY, J. T.; CARFAN, A. C. Re-analysis of pluvial precipitation in southern Brazil. **Atmosfera**, v. 27, n. 2, p. 103–115, abr. 2014.
- NEVES, D. J. D.; ALCÂNTARA, C. R.; SOUZA, E. P. DE. Estudo de Caso de um Distúrbio Ondulatório de Leste sobre o Estado do Rio Grande do Norte - Brasil. **Revista Brasileira de Meteorologia**, v. 31, n. 4, p. 490–505, 2016.
- NEVES, M.; DI LOLLO, J. A. Erosividade da Chuva no Município de São Pedro - SP: Análise entre 1960-2020. **Sociedade & Natureza**, v. 34, n. 1, 21 jun. 2022.
- NIELSEN, D. M. et al. Local indices for the South American monsoon system and its impacts on Southeast Brazilian precipitation patterns. **Natural Hazards**, v. 83, n. 2, p. 909–928, 19 set. 2016.

- NOVARA, A. et al. The impact of soil erosion on soil fertility and vine vigor. A multidisciplinary approach based on field, laboratory and remote sensing approaches. **Science of The Total Environment**, v. 622, p. 474–480, 2018.
- OLIVA, F. G. Climatologia e variabilidade dos principais sistemas meteorológicos atuantes no Brasil, relação com chuvas intensas e impactos associados. **GeoPUC – Revista da Pós-Graduação em Geografia da PUC-Rio**, v. 12, n. 23, p. 74–99, 2019.
- OLIVEIRA, F. P. DE et al. Potencial erosivo da chuva no Vale do Rio Doce, região centro leste do estado de Minas Gerais: primeira aproximação. **Ciência e Agrotecnologia**, v. 33, n. 6, p. 1569–1577, 2009.
- OLIVEIRA, J. P. B. DE et al. Espacialização da erosividade das chuvas no Brasil a partir de séries sintéticas de precipitação. **Revista Brasileira de Ciências Agrárias**, v. 10, n. 4, p. 558–563, 2015.
- OLIVEIRA, J. P. B. DE et al. Assessing the use of rainfall synthetic series to estimate rainfall erosivity in Brazil. **Catena**, v. 171, n. October 2017, p. 327–336, 2018.
- OLIVEIRA, V. A. DE et al. Soil erosion vulnerability in the verde river basin, southern minas gerais. **Ciência e Agrotecnologia**, v. 38, n. 3, p. 262–269, 2014.
- OLIVEIRA, P. T. S. et al. Spatial variability of the rainfall erosive potential in the State of Mato Grosso do Sul, Brazil. **Engenharia Agrícola**, v. 32, n. 1, p. 69–79, 2012.
- OLIVEIRA, P. T. S.; WENDLAND, E.; NEARING, M. A. Rainfall erosivity in Brazil : A review. **Catena**, v. 100, p. 139–147, 2012.
- OLIVEIRA, V. P. S.; ZANETTI, S. S.; PRUSKI, F. F. CLIMABR Parte I: Modelo para a geração de séries sintéticas de precipitação. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 9, n. 3, p. 348–355, 2005a.
- OLIVEIRA, V. P. S.; ZANETTI, S. S.; PRUSKI, F. F. CLIMABR parte II: geração do perfil de precipitação. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 9, n. 3, p. 356–363, 2005b.
- OLIVER, J. E. Monthly precipitation distribution: a comparative index. The Professional Geographer, 32(3), 300-309. **The Professional Geographer**, v. 32, n. 3, p. 300–309, 1980.
- PADULANO, R.; RIANNA, G.; SANTINI, M. Datasets and approaches for the estimation of rainfall erosivity over Italy: A comprehensive comparison study and a new method. **Journal of Hydrology: Regional Studies**, v. 34, p. 100788, 2021.
- PANAGOS, P. et al. Rainfall erosivity in Europe. **Science of the Total Environment**, v. 511, p. 801–814, 2015.
- PANAGOS, P. et al. Spatio-temporal analysis of rainfall erosivity and erosivity density in Greece. **Catena**, v. 137, p. 161–172, 2016.
- PANAGOS, P. et al. Global rainfall erosivity assessment based on high-temporal resolution rainfall records. **Scientific Reports**, v. 7, n. 1, p. 1–12, 2017a.

- PANAGOS, P. et al. Towards estimates of future rainfall erosivity in Europe based on REDES and WorldClim datasets. **Journal of Hydrology**, v. 548, p. 251–262, maio 2017b.
- PANAGOS, P. et al. Global rainfall erosivity projections for 2050 and 2070. **Journal of Hydrology**, v. 610, p. 127865, jul. 2022.
- PANAGOS, P.; BORRELLI, P.; ROBINSON, D. FAO calls for actions to reduce global soil erosion. **Mitigation and Adaptation Strategies for Global Change**, v. 25, n. 5, p. 789–790, 2020.
- PANDEY, S. et al. Recent advances in assessment of soil erosion vulnerability in a watershed. **International Soil and Water Conservation Research**, v. 9, n. 3, p. 305–318, 2021.
- PAULA, G. M. DE et al. Influência do fenômeno El Niño na erosividade das chuvas na região de Santa Maria (RS). **Revista Brasileira de Ciência do Solo**, v. 34, p. 1315–1323, 2010.
- PINHEIRO, A. G. et al. Rainfall pattern and erosion potential in the physiographic regions of the state of Pernambuco, Brazil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 22, n. 12, p. 849–853, 2018.
- PONTES, L. M. et al. Spatial distribution of annual and monthly rainfall erosivity in the jaguari river basin. **Revista Brasileira de Ciência do Solo**, v. 41, p. 1–13, 2017.
- R CORE TEAM. **R: A language and environment for statistical computing**. Disponível em: <<https://www.r-project.org/>>.
- RAIMO, L. A. D. L. DI et al. Spatio-temporal variability of erosivity in Mato Grosso, Brazil. **Revista Ambiente e Água**, v. 13, n. 6, p. 1–14, 2018.
- REIS, G. B. et al. Effect of environmental covariable selection in the hydrological modeling using machine learning models to predict daily streamflow. **Journal of Environmental Management**, v. 290, p. 112625, jul. 2021.
- RENARD, K. G. et al. **Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE)**: Agriculture Handbook, 703. Washington, DC: [s.n.].
- RIQUETTI, N. B. et al. Rainfall erosivity in South America: Current patterns and future perspectives. **Science of the Total Environment**, v. 724, p. 138315, 2020.
- RODRIGUES, J. A. M. et al. Estimating soil erosion vulnerability in the Cervo River basin – MG. **Geociências**, v. 36, n. 3, p. 531–542, 2017.
- RODRIGUEZ, R. D. G. et al. Using entropy theory to improve the definition of homogeneous regions in the semi-arid region of Brazil. **Hydrological Sciences Journal**, v. 61, n. 11, p. 2096–2109, 17 ago. 2016.
- ROSA, A. G. et al. Erosividade da chuva em Rondon do Pará, PA, Brasil de 1999 a 2015 e projetada para 2035. **Revista Ambiente e Agua**, v. 11, n. 4, p. 1006–1021, 2016.

- SAITO, N. S. et al. Uso da geotecnologia na estimativa da erosividade das chuvas e sua relação com o uso e ocupação do solo para o Espírito Santo. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 4, n. 2, p. 51–63, 2009.
- SANTOS, A. H. M. et al. Distúrbio Ondulatório de Leste e seus impactos na cidade de Salvador. **Revista Brasileira de Meteorologia**, v. 27, n. 3, p. 355–364, 2012.
- SANTOS, C. N. **El Niño, La Niña e a erosividade das chuvas no Estado do Rio Grande do Sul**. [s.l.] Federal University of Pelotas, 2008.
- SANTOS, T. E. M. DOS; MONTENEGRO, A. A. A. Erosividade e padrões hidrológicos de precipitação no Agreste Central pernambucano Erosivity and rainfall hydrological patterns in the Pernambuco Central “Agreste”. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 16, n. 8, p. 871–881, 2012.
- SANTOS, W. P. et al. Projections of rainfall erosivity in climate change scenarios for the largest watershed within Brazilian territory. **CATENA**, v. 213, p. 106225, jun. 2022.
- ŠARAPATKA, B.; BEDNÁŘ, M. Rainfall Erosivity Impact on Sustainable Management of Agricultural Land in Changing Climate Conditions. **Land**, v. 11, n. 4, p. 467, 24 mar. 2022.
- SARTORI, M. et al. A linkage between the biophysical and the economic: Assessing the global market impacts of soil erosion. **Land Use Policy**, v. 86, p. 299–312, 2019.
- SCHICK, J. et al. Erosividade Das Chuvas De Lages, Santa Catarina. **Revista Brasileira de Ciência do Solo**, v. 38, n. 6, p. 1890–1905, 2014.
- SCHLIEP, K.; HECHENBICHLER, K.; LIZEE, A. **Package ‘kkn**. **Weighted k-Nearest Neighbors**. , 2022.
- SEKULIĆ, A. et al. Random Forest Spatial Interpolation. **Remote Sensing**, v. 12, n. 10, p. 1687, 2020.
- SENA, N. C. et al. Soil sampling strategy in areas of difficult access using the cLHS method. **Geoderma Regional**, v. 24, p. e00354, mar. 2021.
- SENANAYAKE, S.; PRADHAN, B. Predicting soil erosion susceptibility associated with climate change scenarios in the Central Highlands of Sri Lanka. **Journal of Environmental Management**, v. 308, p. 114589, 2022.
- SERIO, M. A.; CAROLLO, F. G.; FERRO, V. Raindrop size distribution and terminal velocity for rainfall erosivity studies. **Journal of Hydrology**, v. 576, p. 210–228, 2019.
- SHAHANA SHIRIN, A. H.; THOMAS, R. Regionalization of Rainfall in Kerala State. **Procedia Technology**, v. 24, p. 15–22, 2016.
- SHI, D. et al. Relationship between the periodicity of soil and water loss and erosion-sensitive periods based on temporal distributions of rainfall erosivity in the Three Gorges Reservoir Region, China. **CATENA**, v. 202, p. 105268, jul. 2021.
- SHIN, J. Y. et al. Spatial and temporal variations in rainfall erosivity and erosivity density in South Korea. **Catena**, v. 176, p. 125–144, 2019.

- SILVA, B. K. DA N. et al. Avaliação de Extremos de Erosividade Causados pela Precipitação na Bacia do Rio Apodi/Mossoró-RN. **Revista Brasileira de Meteorologia**, v. 35, n. Special Issue, p. 871–879, 2020a.
- SILVA, B. P. C. et al. Soil and water losses in eucalyptus plantation and natural forest and determination of the USLE factors at a pilot sub-basin in Rio Grande do Sul, Brazil. **Ciência e Agrotecnologia**, v. 40, n. 4, p. 432–442, 2016.
- SILVA, D. S. DOS S. et al. Modeling of the spatial and temporal dynamics of erosivity in the Amazon. **Modeling Earth Systems and Environment**, v. 6, n. 1, p. 513–523, 5 mar. 2020b.
- SILVA, A. M. DA. Rainfall erosivity map for Brazil. **Catena**, v. 57, n. 3, p. 251–259, 2004.
- SILVA, A. M. DA et al. Erosividade da chuva e erodibilidade de Cambissolo e Latossolo na região de Lavras, sul de Minas Gerais. **Revista Brasileira de Ciência do Solo**, v. 33, n. 6, p. 1811–1820, 2009.
- SILVA, L. R. S. DA. Hydrometeorological monitoring in Brazil from the perspective of governance. **Revista de Gestão de Água da América Latina**, v. 18, p. e3, 2021.
- SILVA, M. A. DA et al. Avaliação e espacialização da erosividade da chuva no Vale do Rio Doce, região centro-leste do Estado de Minas Gerais. **Revista Brasileira de Ciência do Solo**, v. 34, p. 1029–1039, 2010a.
- SILVA, R. M. DA et al. Erosivity, surface runoff, and soil erosion estimation using GIS-coupled runoff-erosion model in the Mamuaba catchment, Brazil. **Environmental Monitoring and Assessment**, v. 185, n. 11, p. 8977–8990, 2013.
- SILVA, R. M. DA et al. Spatial distribution and estimation of rainfall trends and erosivity in the Epitácio Pessoa reservoir catchment, Paraíba, Brazil. **Natural Hazards**, v. 102, n. 3, p. 829–849, 2020c.
- SILVA, M. J.; GALVÍNCIO, J. D.; COSTA, V. S. DE O. Abordagem interdisciplinar sobre a influência da Zona de Convergência Intertropical - ZCIT no nordeste brasileiro. **Revista Movimentos Sociais e Dinâmicas Espaciais**, v. 6, n. 1, p. 107–117, 2017.
- SILVA, R. B. et al. Assessing Rainfall Erosivity with Artificial Neural Networks for the Ribeira Valley, Brazil. **International Journal of Agronomy**, v. 2010, p. 1–7, 2010b.
- SILVA, S. DE A. et al. Variabilidade espacial do potencial erosivo das chuvas para o estado do Espírito Santo, Brasil. **Irriga**, v. 15, n. 3, p. 312–323, 2010c.
- SIMPLÍCIO, A. A. F. et al. Erosion at hillslope and micro-basin scales in the Gilbués desertification region, Northeastern Brazil. **Land Degradation & Development**, v. 32, p. 1487–1499, 2021.
- SINGH, J.; SINGH, O. Assessing rainfall erosivity and erosivity density over a western Himalayan catchment, India. **Journal of Earth System Science**, v. 129, n. 1, p. 1–22, 2020.
- SIQUEIRA, R. G. et al. Machine learning applied for Antarctic soil mapping: Spatial prediction of soil texture for Maritime Antarctica and Northern Antarctic Peninsula. **Geoderma**, v. 432, p. 116405, abr. 2023.

- SOBRINHO, T. A. et al. Estimativa da erosividade local das chuvas, utilizando redes neurais artificiais. **Ambiente & Água**, v. 6, n. 2, p. 246–254, 2011.
- SOUSA, F. R. C. DE; PAULA, D. P. DE. Analysis of soil loss by erosion in Coreaú river basin (Ceará-Brazil). **Revista Brasileira de Geomorfologia**, v. 20, n. 3, p. 491–507, 2019.
- SOUZA, C. M. et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. **Remote Sensing**, v. 12, n. 17, p. 2735, 2020a.
- SOUZA, C. M. P. et al. Multicriteria analysis and machine learning algorithm for definition of areas for micro-dam, Southeastern Brazil. **Caminhos de Geografia**, v. 22, n. 84, p. 01–13, 15 dez. 2021.
- SOUZA, C. M. P. et al. Spatiotemporal prediction of rainfall erosivity by machine learning in southeastern Brazil. **Geocarto International**, v. 37, n. 26, p. 11652–11670, 13 dez. 2022.
- SOUZA, E. DE O. et al. Erosivity estimation and spatialization in climatic mesoregions in the state of Alagoas. **Revista Brasileira de Meteorologia**, v. 35, n. Special Issue, p. 769–783, 2020b.
- SPERA, S. Agricultural Intensification Can Preserve the Brazilian Cerrado: Applying Lessons From Mato Grosso and Goiás to Brazil's Last Agricultural Frontier. **Tropical Conservation Science**, v. 10, p. 194008291772066, 30 jan. 2017.
- STEFANIDIS, S. et al. Assessing Soil Loss by Water Erosion in a Typical Mediterranean Ecosystem of Northern Greece under Current and Future Rainfall Erosivity. **Water**, v. 13, n. 15, p. 2002, 21 jul. 2021.
- SUPRIYONO, S. et al. Spatial-Temporal Trend Analysis of Rainfall Erosivity and Erosivity Density of Tropical Area in Air Bengkulu Watershed, Indonesia. **Quaestiones Geographicae**, v. 40, n. 3, p. 125–142, 1 set. 2021.
- TEIXEIRA, D. B. S. et al. Rainfall erosivity and erosivity density through rainfall synthetic series for São Paulo State, Brazil: Assessment, regionalization and modeling. **International Soil and Water Conservation Research**, v. 10, n. 3, p. 355–370, set. 2022a.
- TEIXEIRA, D. B. S. et al. Recent advancements in rainfall erosivity assessment in Brazil: A review. **CATENA**, v. 219, p. 106572, dez. 2022b.
- TERASSI, P. M. DE B. et al. Rainfall and erosivity in the municipality of Rio de Janeiro - Brazil. **Urban Climate**, v. 33, n. April, 2020.
- TERASSI, P. M. DE B.; GALVANI, E. Identification of homogeneous rainfall regions in the Eastern watersheds of the state of Paraná, Brazil. **Climate**, v. 5, n. 3, 2017.
- TRINDADE, A. L. F. et al. Variabilidade espacial da erosividade das chuvas no Brasil. **Pesquisa Agropecuária Brasileira**, v. 51, n. 12, p. 1918–1928, 2016.
- TYRALIS, H.; PAPACHARALAMPOUS, G.; LANGOUSIS, A. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. **Water**, v. 11, n. 5, p. 910, 30 abr. 2019.

- VALVASSORI, M. L.; BACK, Á. J. Avaliação do potencial erosivo das chuvas em urussanga, sc, no período de 1980 a 2012. **Revista Brasileira de Ciência do Solo**, v. 38, n. 3, p. 1011–1019, 2014.
- VANTAS, K.; SIDIROPOULOS, E.; LOUKAS, A. Robustness Spatiotemporal Clustering and Trend Detection of Rainfall Erosivity Density in Greece. **Water**, v. 11, n. 5, p. 1050, 2019.
- VAROUCHAKIS, E. A. Geostatistics: Mathematical and Statistical Basis. **Spatiotemporal Analysis of Extreme Hydrological Events**, p. 1–38, 1 jan. 2019.
- VIEIRA, S. R.; LOMBARDI NETO, F. Variabilidade espacial do potencial de erosão das chuvas do Estado de São Paulo. **Bragantia**, v. 54, n. 2, p. 405–412, 1995.
- VIOLA, M. R. et al. Distribuição e potencial erosivo das chuvas no Estado do Tocantins. **Pesquisa Agropecuária Brasileira**, v. 49, n. 2, p. 125–135, 2014.
- WALSH, R. P. D.; LAWLER, D. M. Rainfall seasonality: description, spatial patterns and change through time. **Weather**, v. 36, n. 7, p. 201–208, 1981.
- WALTRICK, P. C. et al. Estimativa da erosividade de chuvas no estado do Paraná pelo método da pluviometria: atualização com dados de 1986 a 2008. **Revista Brasileira de Ciência do Solo**, v. 39, n. 1, p. 256–267, 2015.
- WARD JR, J. H. Hierarchical grouping to optimize an objective function. **Journal of the American statistical association**, v. 58, n. 301, p. 236–244, 1963.
- WILLMOTT, C. J. On the validation of models. **Physical Geography**, v. 2, n. 2, p. 184–194, 1981.
- WISCHMEIER, W. H. A rainfall erosion index for a universal soil loss equations. **Soil Science Society of America Proceedings**, v. 23, p. 246–249, 1959.
- WISCHMEIER, W. H.; SMITH, D. D. Rainfall energy and its relationship to soil loss. **Transactions, American Geophysical Union**, v. 39, n. 2, p. 285, 1958.
- WISCHMEIER, W. H.; SMITH, D. D. **Predicting Rainfall Erosion Losses. A guide to conservation planning**: Agriculture Handbook, 537. Washington, DC: [s.n.].
- WUEPPER, D.; BORRELLI, P.; FINGER, R. Countries and the global rate of soil erosion. **Nature Sustainability**, v. 3, n. 1, p. 51–55, 2 dez. 2019.
- WUEPPER, D.; BORRELLI, P.; FINGER, R. Countries and the global rate of soil erosion. **Nature Sustainability**, v. 3, n. 1, p. 51–55, 2020.
- XU, Z. et al. Spatial–temporal distribution of rainfall erosivity, erosivity density and correlation with El Niño–Southern Oscillation in the Huaihe River Basin, China. **Ecological Informatics**, v. 52, p. 14–25, 2019.
- YU, B. Using CLIGEN to generate RUSLE climate inputs. **Transactions of the ASAE**, v. 45, n. 4, p. 993, 2002.

ZANETTI, S. S. et al. Programa computacional para geração de séries sintéticas de precipitação. **Engenharia Agrícola**, v. 25, p. 96–104, 2005.

ZHANG, Y. et al. Evaluation of CLIGEN for storm generation on the semiarid Loess Plateau in China. **Catena**, v. 73, n. 1, p. 1–9, 2008.

ZHU, D.; XIONG, K.; XIAO, H. Multi-time scale variability of rainfall erosivity and erosivity density in the karst region of southern China, 1960–2017. **Catena**, v. 197, p. 104977, 2021.