SEVEN publicações acadêmicas

ANALYSIS OF PRECIPITATION ESTIMATION MODELS USING TIME SERIES: A BIBLIOMETRIC ANALYSIS

ANÁLISE DOS MODELOS DE ESTIMATIVA DE PRECIPITAÇÃO UTILIZANDO SÉRIES TEMPORAIS: UMA ANÁLISE BIBLIOMÉTRICA

ANÁLISIS DE MODELOS DE ESTIMACIÓN DE PRECIPITACIÓN UTILIZANDO SERIES DE TIEMPO: UN ANÁLISIS BIBLIOMÉTRICO

https://doi.org/10.56238/sevened2025.036-019

Jessica Almeida Monteiro Arruda[1], Rodrigo Nogueira de Vasconcelos[2], Rosângela Leal Santos[3]

ABSTRACT
This study conducts a bibliometric analysis of precipitation estimation models using time series, aiming to identify trends, methods, and scientific contributions over the last four decades. The research is justified by the importance of precipitation as a primary variable in hydrological modeling and the scarcity of accurate rainfall data, especially in regions with high spatial and temporal variability. The work seeks to map the main precipitation estimation models, analyze their spatial and temporal resolution, and identify methodological trends and innovations in the field. The methodology adopted included a bibliometric approach, using the Scopus database to collect scientific articles, followed by statistical analyses and term co-occurrence networks using the VOSviewer software. Seventy-eight articles published between 1982 and 2023 were analyzed, focusing on the 40% most cited in each decade. The results reveal a significant evolution in precipitation modeling, with the transition from simple statistical methods to advanced techniques such as artificial neural networks (ANNs) and deep learning (LSTM). Daily temporal and regional spatial resolution predominated in the studies, and the integration of satellite data and artificial intelligence (AI) techniques has become a dominant trend in the last decade. The conclusion is that advances in precipitation modeling have significant practical impacts on water resource management and agriculture, contributing to the prediction of extreme events and adaptation to climate change.

Keywords: Precipitation. Time Series. Hydrological Modeling. Bibliometric Analysis. Artificial Intelligence. Climate Change.

RESUMO
Este estudo realiza uma análise bibliométrica sobre os modelos de estimativa de precipitação utilizando séries temporais, com o objetivo de identificar tendências, métodos e contribuições científicas ao longo das últimas quatro décadas. A pesquisa justifica-se pela importância da precipitação como variável primordial na modelagem hidrológica e pela escassez de dados pluviométricos precisos, especialmente em regiões com alta variabilidade espacial e temporal. O trabalho busca mapear os principais modelos de estimativa de precipitação,

[1] Master of Science in Environmental Sciences. Universidade Estadual de Feira de Santana.
Bahia, Brazil. E-mail: jamarruda@uefs.br
[2] Dr. in Ecology. Universidade Estadual de Feira de Santana. Bahia, Brazil. E-mail: rnvuefsppgm@gmail.com
[3] Dr. in Transportation Engineering Universidade Estadual de Feira de Santana.
Bahia, Brazil. E-mail: rosaleal@uefs.br

analisar sua resolução espacial e temporal, e identificar as tendências metodológicas e inovações na área. A metodologia adotada incluiu uma abordagem bibliométrica, utilizando a base de dados Scopus para coleta de artigos científicos, seguida de análises estatísticas e de redes de coocorrência de termos com o auxílio do software VOSviewer. Foram analisados 78 artigos publicados entre 1982 e 2023, com foco nos 40% mais citados de cada década. Os resultados revelam uma evolução significativa na modelagem da precipitação, com a transição de métodos estatísticos simples para técnicas avançadas, como redes neurais artificiais (ANN) e deep learning (LSTM). A resolução temporal diária e espacial regional predominou nos estudos, e a integração de dados de satélites e técnicas de inteligência artificial (IA) tornou-se uma tendência dominante na última década. Conclui-se que os avanços na modelagem da precipitação têm impactos práticos significativos na gestão de recursos hídricos e na agricultura, contribuindo para a previsão de eventos extremos e a adaptação às mudanças climáticas.

Palavras-chave: Precipitação. Séries Temporais. Modelagem Hidrológica. Análise Bibliométrica. Inteligência Artificial. Mudanças Climáticas.

RESUMEN
Este estudio realiza un análisis bibliométrico de modelos de estimación de precipitación utilizando series de tiempo, con el objetivo de identificar tendencias, métodos y contribuciones científicas en las últimas cuatro décadas. La investigación se justifica por la importancia de la precipitación como variable primaria en el modelado hidrológico y la escasez de datos precisos de lluvia, especialmente en regiones con alta variabilidad espacial y temporal. El trabajo busca mapear los principales modelos de estimación de precipitación, analizar su resolución espacial y temporal e identificar tendencias metodológicas e innovaciones en el campo. La metodología adoptada incluyó un enfoque bibliométrico, utilizando la base de datos Scopus para recopilar artículos científicos, seguido de análisis estadísticos y redes de co-ocurrencia de términos utilizando el software VOSviewer. Se analizaron setenta y ocho artículos publicados entre 1982 y 2023, centrándose en el 40% más citado en cada década. Los resultados revelan una evolución significativa en el modelado de precipitación, con la transición de métodos estadísticos simples a técnicas avanzadas como redes neuronales artificiales (ANN) y aprendizaje profundo (LSTM). La resolución espacial temporal y regional diaria predominó en los estudios, y la integración de datos satelitales y técnicas de inteligencia artificial (IA) se ha convertido en una tendencia dominante en la última década. La conclusión es que los avances en la modelización de precipitaciones tienen un impacto práctico significativo en la gestión de los recursos hídricos y la agricultura, contribuyendo a la predicción de eventos extremos y la adaptación al cambio climático.

Palabras clave: Precipitación. Series Temporales. Modelado Hidrológico. Análisis Bibliométrico. Inteligencia Artificial. Cambio Climático.

1 INTRODUCTION

The understanding of hydrological phenomena is essential for rainfall-flow modeling, in which precipitation stands out as a primary input variable. However, the difficulty in obtaining accurate rainfall data compromises the realistic representation of hydrological events, especially in basins where high spatial variability and flaws in records can result in significant discrepancies between the model and reality. Studies such as the one by Arai *et al.* (2009) demonstrate the relevance of precipitation in the characterization of the regional climate and its direct impact on agriculture, while Silva *et al.* (2011) show that prolonged periods of drought affect agriculture and reduce water levels in rivers, reservoirs and springs, even impacting the generation of hydroelectric power. Lucena *et al.* (2008) complements this view, emphasizing that rainfall intensity is influenced by large-scale atmospheric and oceanic phenomena.

In the Brazilian context, the network of stations of the National Institute of Meteorology (INMET) faces challenges due to the limited distribution of stations and inconsistencies in the historical series, arising from technical and data processing problems. To mitigate such challenges, techniques such as multiple linear regression have been shown to be effective in adjusting and filling gaps in historical data (Pedro Júnior *et al*., 1994), and methodologies based on multiple regression analysis correlate precipitation with physiographic and temporal variables (Antonini *et al*., 2019). Given the scarcity of rainfall data and the complexity inherent to precipitation modeling, it is essential to identify the main models used, as well as the spatial and temporal resolution (hourly, daily, monthly or other periodicity) of these models, in addition to understanding the trends and methodological innovations over the decades.

Bibliometric analysis emerges, in this scenario, as a robust tool that uses statistical and mathematical techniques to map scientific productivity and identify authorial, institutional, and thematic patterns (Wu, 2017; Guedes & Borshiver, 2005; Café & Brascher, 2008; Job, 2018). Thus, this study proposes to integrate precipitation modeling with a bibliometric and network approach, aiming to identify the main precipitation estimation models from time series, analyze the resolution and periodicity of these models, map the main researchers and institutions in the area, and evaluate the most used methods in the calibration and validation of the models.

2 THEORETICAL TABLEWORK

2.1 CLIMATE

Climate is the set of meteorological phenomena that characterize the average state of the atmosphere over a given place on the earth's surface. (HANN, 1882 apud MONTEIRO, 1991). Through this definition, it is possible to understand the central idea of what climatology would be.

It is extremely important to understand that climate and weather are not the same, and therefore, they have some differences. Time is understood as a set of studied elements that expresses the conditions of the atmosphere observed at a current moment, that is, at an instant in the atmosphere. Climate, on the other hand, will have its understanding associated with a set of elements studied through records, left over years and years.

According to Borsato (2000), the climatic elements somewhere on the earth's surface can be defined as: air temperature, relative humidity, atmospheric pressure, cloudiness, solar radiation, dust, wind direction and speed, electric field, water vapor voltage, positive and negative ions, horizontal visibility and vertical electric current, acting simultaneously and together, considered as a specific property of the action of each air mass.

2.1.1 Climate Components

According to Silva (2001), the study of climate and knowledge of its origin lead us to an understanding related to factors that can be defined as: generics and modifiers.

According to Vecchia (2001) the elements of climate are: air temperature, air humidity, atmospheric pressure, wind direction and speed, cloudiness and atmospheric precipitation, that is, it includes not only meteorological phenomena, but also phenomena that are influenced by the conditions of the place

2.1.1.1 Generic factors

These are dynamic factors that condition the climate, these meteorological phenomena originate from solar radiation and atmospheric and maritime movements, they are defined:

- Solar radiation: it is the factor that promotes global thermal exchange processes due to the warming of the Earth's surface. (CARLETTO, 2005)

- General circulation of the atmosphere: it combines the thermal forces, whose circulation is meridian, that is, in the north-south direction, and the dynamic forces that are caused by the rotation of the earth and that promotes a modification in the air flow that tries to create a zonal circulation, in an East-West direction, according to Vecchia (2001).

- Sea currents: transfer heat from regions with temperature differences. They act in two ways mainly: by increasing the pressure of water vapor, that is, the humidity of the air, and by modifying the wind regime, according to Vecchia (2001).

2.1.1.2 Modifying factors

They are the static factors that act on the climate and create local variations, they are: geographical, topological, biological and cultural. Most of the time, the variations caused by these factors take on large magnitudes. Vecchia (2001) summarizes the main modifying factors in descending order of magnitude:

- Geographic: latitude, altitude, and the relationship between land and water masses.
- Topographic: elevation, sun orientation and soil structure.
- Biological: fauna and flora.
- Cultural: changes in the soil, pollution of the atmosphere and hydrological modifications.

The elements and especially the factors of climate are sometimes so complex and variable, highly dynamic, sometimes with chaotic behavior, that it is impossible to define their precise functioning.

2.2 PRECIPITATION

Precipitation, as defined by Bertoni and Tucci (2009), consists of all water from the atmosphere that reaches the earth's surface, regardless of its physical state, such as drizzle, rain, snow, hail, hail, frost or dew. In Brazil, approximately 99% of precipitation occurs in the form of rain, while less than 1% manifests itself in the other forms (Santos *et al*., 2001). Rainfall is a widely studied climatic element, especially due to meteorological and hydrological records, which in some regions date back to the beginning of the twentieth century (Cavalcanti *et al*., 2009).

Clouds, which allow precipitation to occur under specific conditions, are formed by the mixture of air, water vapor, and liquid or solid droplets, whose diameters vary between 0.001 and 0.03 mm (Bertoni and Tucci, 2013). The main characteristics of precipitation include its duration, total precipitated volume, temporal and spatial distribution. In terms of hydrometeorological processes, rainfall is the element that presents the greatest monthly and annual variability when compared between different regions.

According to Bertoni and Tucci (2013), the formation of precipitation is associated with the upward movement of humid air masses, which causes a dynamic or adiabatic cooling, leading the water vapor to reach its saturation point. From the condensation level, under favorable conditions and in the presence of hygroscopic nuclei, water vapor condenses, forming tiny droplets around these nuclei. For the water droplets to precipitate, it is necessary that they reach such a volume that their weight exceeds the forces that keep them in suspension, thus acquiring a speed of fall superior to the ascending vertical components of atmospheric movements. Raindrops have diameters ranging from 0.5 to 2.0 mm, and can reach maximum values of 5.0 to 5.5 mm. When a droplet reaches a diameter of 7.00 mm, its fall velocity reaches approximately 9 m/s, at which time its deformation occurs and subsequent fragmentation into smaller droplets due to air resistance (Bertoni and Tucci, 2009).

3 METHODOLOGY

For the development of the research, a unified methodological approach of bibliometric analysis was implemented, providing greater robustness to the study. Table 1 presents the relationship between the research questions, the analyses proposed in our methodological approach and the data used for their execution. The table is structured in three main fields: the questions of interest, the analysis used to answer them, and the type of data used.

Table 1

Description of the relationship between work issues, analyses, and data sources

Questions	Analysis	Data
What are the trends in publications on precipitation models using time series?	General statistics/Word co-occurrence network/Co-authoring spatial network	All papers
Which countries contributed the most to research on precipitation models using time series?	General statistics/Word co-occurrence network/Co-authoring spatial network	All papers
Do the articles specify the geographical area covered by the precipitation model and the spatial and/or temporal resolution of the model?	Metrics	40% more cited

Do the surveys identify the relevant climate and weather variables that affect precipitation and the data sources used to provide these variables?	Metrics	40% more cited
Do studies consider different types of precipitation and the physical and meteorological processes in the precipitation model?	Metrics	40% more cited
The studies present the most used methods for calibration and validation of precipitation models and what are these methods?	Metrics	40% more cited
How can advances in precipitation modeling impact water resources management and agriculture?	General statistics/general characteristics	40% more cited
What are the emerging trends in precipitation modeling, and how are artificial intelligence and machine learning being integrated into precipitation models?	General characteristics	40% more cited
What are the most influential publications on precipitation models using time series?	General statistics/Word co-occurrence network/Co-authoring spatial network	40% more cited

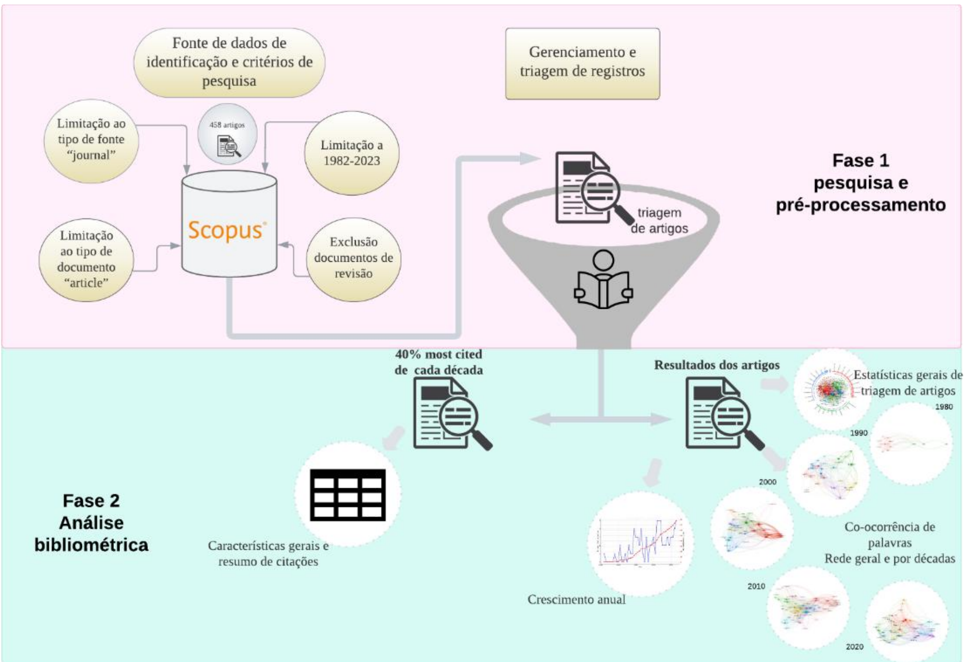
Source: Prepared by the authors.

The approach and methodological steps used are illustrated in Figure 1, which involves the choice of the search database, the identification of terms relevant to the theme and the application of filters in the search engine. Then, a manual screening was carried out, with the reading of the titles and abstracts of the retrieved articles, in order to identify and exclude those that did not align with the research objectives, characterizing Phase 1 of the research. Subsequently, analyses were carried out for two sets of data: all selected articles, a total equivalent to 78 articles, and for the 40% most cited in each decade.

F

Figure 1

Logical scheme, methodological approach and data analysis



Source: Prepared by the authors themselves.

3.1 DATABASE

The selection of the Scopus database for the collection of bibliometric data is justified by its comprehensiveness and quality, being recognized as one of the largest global multidisciplinary databases. Scopus provides access to more than 27,800 peer-reviewed journals and more than 330,000 books, as well as conference proceedings, ensuring extensive and reliable coverage of scientific information (ELSEVIER, 2024). Its advanced data analysis and retrieval tools allow the identification of trends, citation patterns, and collaboration networks, essential aspects for the mapping of scientific knowledge.

In the present study, Scopus is particularly relevant for the analysis of time-series-based precipitation estimation models, as it facilitates the retrieval of articles that address the validation of specific models and methodologies. The restriction of the search to articles published in journals aims to ensure the originality and robustness of the data, avoiding common duplications in reviews, book chapters or conference proceedings. This rigor in the selection of sources reflects the need for a consolidated and reliable basis for detailed bibliometric analyses, reinforcing the choice of Scopus as the primary source.

The use of Scopus, with its robust tools for exporting and analyzing data in multiple formats, contributes significantly to bibliometric studies. Such studies allow us to understand scientific developments in specific areas, identify gaps in the literature, and guide future research and the development of new methodologies, standing out as an essential resource for academic research.

3.2 SCREENING SEARCH AND REGISTRATION STRATEGY

The research strategy used in the Scopus database focused on the analysis of precipitation estimation models and time series analysis, taking into account their validation methods. For the construction of the dataset, the search was structured using Boolean operators, such as OR and AND, which allowed combining different terms and synonyms to refine the results and include only the articles relevant to the scope of the study. The terms chosen, such as "Rainfall Estimation Statistic Model*", "Rainfall Data Analysis Model*", "Rainfall Model*", among others, aimed to identify the most significant publications in the area of precipitation modeling. The use of these Boolean operators helped to select articles that addressed modeled time series and the analyses associated with mathematical modeling, essential aspects for scientific development on the subject.

To structure the dataset, Boolean operators OR and AND were used. The keywords used in the research were selected with the objective of identifying essential aspects that contribute to the understanding of the scientific development related to precipitation modeling, as well as to identify gaps in the literature on the subject. Several preliminary tests were conducted with different terms and synonyms until the final set of keywords pertinent to the study was reached.

The terms of the final equation: TITLE-ABS-KEY (("Rainfall Model*" OR "Precipitation Model*" OR "Rainfall Estimation Model*" OR "Precipitation Estimation Model*" OR "Rainfall Estimation Statistic Model*" OR "Precipitation Estimation Statistic Model*" OR "Rainfall Data Analysis Model*") AND ("Time Series Analysis*" OR "Time Series*" OR "Temporal Data Analysis" OR "Temporal Data" OR "Rainfall Data Model")) AND PUBYEAR > 1981 AND PUBYEAR < 2024 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (DOCTYPE , "ar")).

The final search equation further refined the results by limiting the year of publication between 1982 and 2023 and the type of document to finished research articles (peer-reviewed articles), ensuring that the data collected was of high quality and relevance.

After this initial stage, the titles and abstracts of the retrieved articles were carefully reviewed, excluding those that were not aligned with the objectives of the study. This approach allowed the bibliometric analysis to be based on a robust and well-defined dataset, which was later exported in CSV format for qualitative and quantitative analyses. These analyses included the geographic distribution of the studies, allowing a comprehensive view of the evolution and impact of research on precipitation models over time (SANTANA *et al*., 2021).

This initial process made it possible to identify emerging patterns and evolution in research topics over time. Subsequently, a detailed systematic review was carried out on 40% of the most cited articles in each period.

This study, by integrating a quantitative and qualitative approach, not only maps the historical development of the area of environmental sciences, but also highlights the gaps and future opportunities, reinforcing the importance of methodological innovation and international collaboration in scientific research.

3.3 SEMANTIC AND GENERAL NETWORK ANALYSIS

The construction of the term networks was carried out using the VOSviewer program, a software developed for bibliometric analysis of scientific literature (VAN ECK; WALTMAN, 2010) that uses the VOS mapping technique to visualize similarities and allow the development of bibliometric maps. The software was used to make maps based on data of co-occurrences between the terms. The use of this software was based on the possibility of running on many hardware platforms, operating systems, being free and also the possibility of being used directly from the internet (VAN ECK; WALTMAN, 2010; VAN NUNEN *et al*., 2018). The minimum number of occurrences of a keyword is five times for the decades and general network, for titles, abstracts, and keywords of all publications to build a network.

VOSviewer can be used to build networks of scientific publications, scientific journals, researchers, research organizations, countries, and keywords. Items in these networks can be connected by co-authorship, co-occurrence, citation, bibliographic coupling, or co-citation links (VAN ECK; WALTMAN, 2010). The bibliometric maps produced in VOSviewer use principles related to text mining (VAN ECK; WALTMAN, 2010). The analysis of the keywords between titles and abstracts shows the connection between related terms, providing the division into groups called clusters, which groups all terms considered similar.

The counting method for this study was full counting, corresponding to the total count, where each link of co-occurrences has the same weight, which means that all occurrences of the term in the document will be counted, but there is binary counting only observing the presence or not of the term existing in the document (VAN ECK; WALTMAN, 2010).

The graphical representations expressed in the maps relate the proportionality of similarities and co-occurrences of terms in the publications: the co-occurrence networks are composed of nodes and edges. In which nodes are objects such as the co-occurrence of words and countries. Between any pair of knots there can be an edge. An edge is a connection or a relationship between two nodes. The distance between two nodes in the visualization roughly indicates the relationship between the searched terms in terms of co-occurrence. Node size is determined by the weight of an item in a given network (VAN ECK; WALTMAN, 2010).

For the creation of the clusters, a thesaurus file was prepared, which represents a thesaurus dictionary in txt format, where the word is as the label is and replace by, which is the way the term is to be (WALTMAN; VAN ECK; NOYONS, 2010). The thesaurus is used to clean up terms, and can ignore some, exclude, change the way they are written and/or unify terms that have similar meanings.

3.4 STATISTICAL ANALYSIS OF THE NETWORK

The bibliometric analysis, essential to understand the panorama of scientific production in certain areas of study, was conducted through the Bibliometrix statistical package, an open-source tool integrated with the R language, widely recognized for its versatility in scientific data analysis (ARIA; CUCCURULLO, 2017).

For the development of this analysis, the necessary commands for the installation and activation of the Bibliometrix package were implemented in the R Studio environment. First, install.packages ("bibliometrix") was used to ensure the installation of the package, followed by library(bibliometrix) and biblioshiny(), which activated the interactive graphical interface, allowing a more intuitive visualization of the data. From there, a database with scientific publications was entered into the system, and variables such as the contribution of different countries and the productivity of the authors were analyzed, these data were processed and organized manually, using adjustments in the frequency histograms to refine the presentation of the results.

In short, the use of the Bibliometrix package in R Studio, with its wide range of resources for analysis and visualization, made it possible to elaborate an in-depth mapping of scientific production, allowing a better understanding of the dynamics and trends of the field under study.

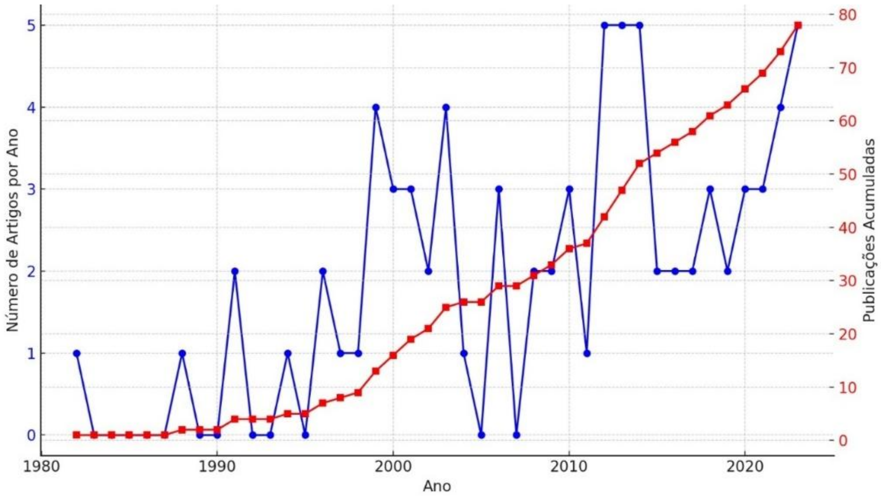
4 RESULTS AND DISCUSSIONS

4.1 PRECIPITATION MODELING TRENDS PUBLICATIONS

After processing the bibliographic data and manual filtering, the final sample resulted in 78 articles published between 1982 and 2023, with an average of 1.86 articles per year and a standard deviation of 1.59, i.e., 1.86 ± 1.59. This indicates that, on average, there were approximately 1.86 annual publications, with a variation of 1.59 articles around this average. There was also an annual growth of 4.0%, as shown in Figure 2.

Figure 2

Articles published by year and accumulated in the period 1982-2023



Source: Prepared by the authors themselves.

When annual growth is evaluated, the data indicate a modest and irregular growth in the number of articles published between 1982 and 2023. The analysis reveals years with production peaks, such as 1999, 2003, 2012 and 2023, as well as periods without publications or with few productions (for example, between 1983 and 1987). This behavior reflects the growing interest in precipitation modeling as new technologies and data have become available, as well as the impact of global and regional events.

Precipitation modeling, especially utilizing techniques such as kriging and Fourier analysis, has benefited from advances in computing and remote sensing, such as the use of satellites, which have become more accessible in recent decades. Factors such as the climate crisis and the need to predict extreme events (floods and droughts) also drove the growth in scientific production from the 2000s onwards. The increase in production can be explained by the growing need for more accurate predictive tools, especially in regions with scarce rainfall data, such as Bahia, for example.

Table 2 shows a clear trend of increase in the number of articles published over time.

Table 2

Evolution of the number of publications over the decades, with percentage, mean and standard deviation

Decade	Total Publications	Percentage of Total (%)	Average Articles/year	Standard deviation
1982-1989	2	2,56%	0,25	0,43
1990-1999	10	12,82%	1,1	1,22
2000-2009	26	33,33%	2	1,26
2010-2019	36	46,15%	3	1,41
2020-2023	12	15,38%	3,75	0,83

Source: Prepared by the authors themselves.

Figure 2 and Table 2 show that with regard to the growth of accumulated publications, it is necessary to reflect the consolidation of precipitation modeling as an important research area. The cumulative graph shows a consistent increase from 1999 onwards, which coincides with increased access to technological tools such as models based on geospatial data and the use of advanced computational techniques. From 2010 onwards, growth accelerated, highlighting the fundamental role of modeling to face contemporary climate challenges, such as extreme variations in precipitation, changes in rainfall patterns and the sustainable management of water resources.

When the analysis is carried out by decades, it is found that, in the 80s (1982-1989), 2 publications were made, corresponding to 2.56% of the total. This period presents a low number of articles, reflecting the initial stage of research in precipitation modeling. In the 90s (1990-1999) we had 10 publications, representing 12.82% of the total. The 1990s marks the beginning of a more significant increase in interest in climate modeling, possibly driven by growing concern about the impact of climate change. In the 2000s (2000-2009) there were 26 publications, representing 33.33% of the total. This decade sees a notable increase in scientific output, which can be associated with the emergence of remote sensing technologies, more sophisticated climate models, and the widespread use of meteorological data from satellites. From 2010-2019, 2010s, 36 publications, corresponding to 46.15% of the total. This period consolidates precipitation modeling as a mature research area, with the intensive use of big data, artificial intelligence, and the development of more advanced predictive modeling techniques, such as kriging and spectral analysis. And finally, the decade still in formation, 2020-2023, there are already 12 publications so far, representing 15.38% of the total. Even with fewer years, production continues to grow, reflecting the increasing relevance of precipitation modeling, especially in response to extreme weather events and the need to more accurately predict the impacts of climate change.

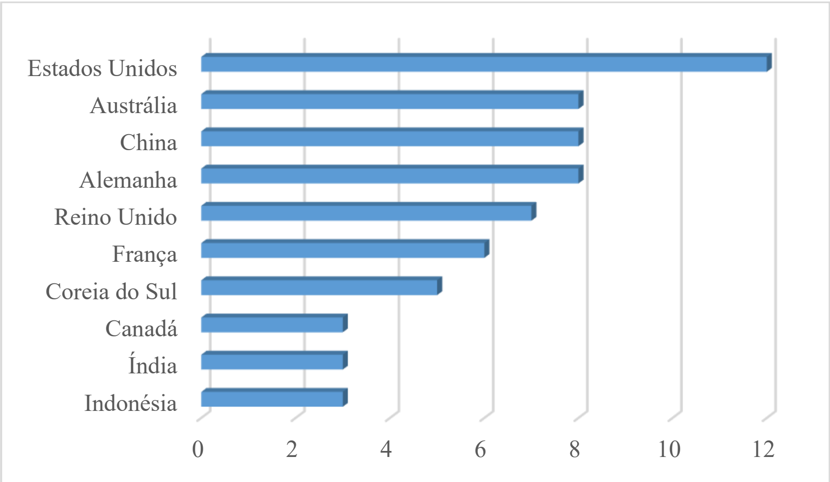
Still analyzing Figure 2, it is possible to observe that the peaks and falls in publications can be justified by technological factors and global events. The 1999 peak may be related to increased interest in precipitation modeling due to the Kyoto Protocol and growing concern about climate change. In the 2000s, the use of satellites and techniques such as kriging boosted publications. Between 2012 and 2013, the use of big data and artificial intelligence further expanded interest in the field. The relative drop in 2020-2023 may have been influenced by the COVID-19 pandemic, which impacted global scientific production, although interest remained high due to the relevance of the topic. These data justify the mean and standard deviation presented in Table 2.

4.2 COUNTRIES CONTRIBUTED TO RESEARCH ON PRECIPITATION MODELING

A bibliometric network was built to visualize the collaboration networks and the distribution of publications around the world related to the application on precipitation modeling. Figure 3 shows how the quantities of publications were distributed among the 10 countries with the highest number of publications in the area in question.

Figure 3

Number of publications by country



Source: Prepared by the authors themselves.

The United States occupies the first place, with a total of 30 publications, and Brazil occupies the ninth position in the global ranking of the geographical distribution of scientific productions relative to the number of citations by countries, and Table 3 shows the percentage of each country.

Table 3

Number of articles published and percentages

Countries	Frequency	Percentage (%)
United States	30	12,66
United Kingdom	26	10,97
Australia	22	9,28
France	19	8,02
China	18	7,59
Germany	14	5,91
Indonesia	11	4,64
South Korea	10	4,22
Brazil	6	2,53
Ecuador	5	2,11

Source: Prepared by the authors themselves.

4.3 GEOGRAPHICAL AREA AND RESOLUTION

The analysis of the 32 articles reveals that all of them (100%) specify the geographical area covered by the precipitation model and the spatial and/or temporal resolution of the model. This indicates that the scientific community has a consensus on the importance of clearly defining the scope and resolution of precipitation models.

Regarding temporal resolution, the most common is the daily resolution, present in 90.6% of the articles. Hourly resolution appears in only 3.1% of cases, while the combination of daily and monthly resolution is observed in 6.3% of articles. This suggests that everyday resolution is preferred in most studies, likely due to its balance between detail and computational feasibility.

With regard to spatial resolution, the most frequent is regional, present in 93.8% of the articles. Local resolution appears in only 6.2% of cases, indicating that most studies focus on larger geographic areas, such as regions or watersheds, rather than specific locations.

When analyzing the trends over the decades, a consolidation of daily and regional resolution is observed. In the 1980s, only daily and regional resolution was observed. In the

1990s, there was a small diversification, with the introduction of time and local resolution. However, from the 2000s onwards, daily and regional resolution became dominant, with 100% prevalence in the 2010s and 2020s. This reflects a standardization in modeling approaches, possibly driven by technological advances and the need for comparability between studies.

Therefore, the data show that the scientific community prioritizes daily and regional resolution in precipitation modeling, with a clear trend of consolidation of these characteristics over time. Time and local resolution, although present in some studies, is less common, suggesting that they are applied in specific contexts or in early stages of research.

4.4 CLIMATE VARIABLES AND DATA SOURCE

The analysis of the 32 articles reveals that 59.4% of them identify the relevant climatic and meteorological variables that affect precipitation, while 40.6% do not identify them, but mention the data sources used. This indicates that, although most studies seek to understand the factors that influence precipitation, a significant portion focus more on the collection and use of data than on the exploration of specific climate variables.

Over the decades, a clear evolution has been observed both in the climate variables studied and in the data sources used. In the 1980s, articles focused on precipitation-related variables, such as depth and duration of events, using primarily rain gauge data. In the 1990s, there was an introduction of more complex variables, such as atmospheric circulation, temperature, and pressure, with the use of atmospheric pressure data and meteorological records.

In the 2000s, precipitation modeling expanded to include variables such as topography, terrain gradient, and relative humidity, using reanalyses, interpolated data, and radiosonde The 2010s marked an even greater diversification, with the introduction of variables such as altitude, latitude, longitude, and oceanic-atmospheric indices (IPO-PDO), in addition to the expansion of the use of satellite data, regional climate models, and paleoclimate series.

In the 2020s, there was a reduction in the number of articles identifying climate variables, with only 20% of studies highlighting specific variables. However, there is a consolidation of the use of advanced data sources, such as satellites and global climate models (GCMs), reflecting a trend to prioritize the collection and processing of high-resolution data.

It is thus concluded that, over time, there has been a transition from studies focused on simple variables and local data (such as rain gauges) to a more complex and integrated approach, which combines detailed climate variables with advanced data sources, such as satellites and global models. This evolution reflects technological advances and the need for more accurate models to deal with the challenges of climate change and water resources management.

4.5 TYPES OF PRECIPITATION AND PHYSICAL PROCESSES

The analysis of the 32 articles reveals that 59.4% of them consider physical and meteorological processes in their precipitation models, while 40.6% do not address these processes. This indicates that, although most studies seek to understand the mechanisms that influence precipitation, a significant portion focus more on the collection and use of data than on the exploration of the underlying physical and meteorological processes.

Over the decades, there has been a clear evolution in the approach to physical and meteorological processes. In the 1980s, studies focused on specific processes, such as convective storms. In the 1990s, there was an introduction of broader concepts, such as the relationship between atmospheric circulation and the occurrence of rainfall, in addition to the influence of temperature and humidity on the formation of precipitation. In the 2000s, studies expanded to include spatial gradients and orographic effects, reflecting greater complexity in modeling.

From the 2010s onwards, there was a significant diversification of the processes considered, with the introduction of concepts such as advection, diffusion, convection, monsoon circulation and interannual variability. This trend continued in the 2020s, with a consolidation of the use of statistical relationships between temperature and precipitation, indicating a prioritization of quantitative analysis and modeling.

It is important to highlight that, in all articles, the only form of precipitation considered was rain. There was no mention of other types of precipitation, such as snow or hail, which reflects the focus of the studies on regions and contexts in which rain is the predominant form of precipitation.

Thus, over time, there has been a transition from studies focused on specific and local processes to a more complex and integrated approach, which combines multiple physical and meteorological processes with advanced statistical analysis techniques. This evolution reflects technological advances and the need for more accurate models to deal with the challenges of climate change and water resources management.

4.6 CALIBRATION AND VALIDATION METHODS

The analysis of the methods of calibration and validation of precipitation models over the decades reveals a significant evolution, driven by the advancement of technology and the need for more accurate and reliable models. Table 4 shows the most used methods by decade.

Table 4

Most used methods over the decades

Decade	Most Used Methods
1980	Correlation between depth and duration; observational data.
1990	Conditional probabilities; stochastic models, maximum verisimilitude.
2000	Markov Chains; simulated annealing; nonlinear regression; adjustment of distributions.
2010	Neural Networks (ANN); MCMC; cross-validation; adjustment of distributions; satellite data.
2020	Deep learning (LSTM); copulations; heuristic optimization (ISPSO); integration of satellites.

Source: Prepared by the authors themselves.

In the 1980s, calibration and validation methods were relatively simple and based on descriptive statistics. The article analyzed from this period used correlations between depth and duration of precipitation events and reliable observational data in semi-arid zones. These methods reflected the technological limitations of the time, focusing on empirical adjustments and basic correlations. The lack of diversification of techniques can be attributed to the scarcity of data and limited computational capacity.

In the 1990s, there was a significant advance with the introduction of more sophisticated statistical methods. The articles of this decade used conditional probabilities based on time series of atmospheric circulation patterns and stochastic models applied on a regional scale. In addition, methods such as maximum likelihood and comparison of observed and simulated statistics began to be employed. These techniques allowed for a better representation of climate variability and a more robust validation of the models.

The 2000s marked a period of diversification and optimization of calibration and validation methods. The articles analyzed showed the use of techniques such as Markov chains, simulated annealing and nonlinear regression. These methods allowed for more precise adjustments of the model parameters and a better representation of precipitation at different spatial and temporal scales. In addition, adjustment of distributions (Weibull, GEV, Kappa) has become common, reflecting a growing concern with the statistical representation of data.

The 2010s were marked by the use of computationally intensive methods and the concern with the quantification of uncertainties. With 12 articles analyzed, techniques such as artificial neural networks (ANN), Bayesian methods (MCMC) and cross-validation stand out. These methods allowed for more accurate calibration and the incorporation of uncertainties into the models. In addition, adjustment of distributions (Poisson, geometric, exponential) and integration of satellite data have become commonplace, reflecting the increasing availability of data and the advancement of technology.

In the 2020s, calibration and validation methods reached a new level of sophistication. The analyzed articles observed the use of techniques such as deep learning (LSTM), copulas and heuristic optimization (ISPSO). These methods allowed for a more accurate representation of precipitation, especially at thin temporal and spatial scales. In addition, the integration of satellite data and distributed hydrological models (such as W-flow) has become a trend, reflecting the need for more realistic models that are applicable in different contexts.

Over the decades, calibration and validation methods have evolved from simple, empirical techniques to computationally intensive and sophisticated approaches. This evolution has been driven by the advancement of technology and the increasing availability of data. From the 2010s onwards, there was a growing concern with the quantification of uncertainties, using methods such as MCMC and cross-validation. This reflects the need for more reliable and robust models. In the 2020s, the integration of satellite data and hydrological models became a trend, allowing for a more realistic representation of precipitation and its variability. The use of artificial intelligence techniques such as neural networks (ANN) and deep learning (LSTM) became commonplace in the 2010s and 2020s, reflecting the potential of these techniques to improve the accuracy of models.

4.7 IMPACTS ON WATER RESOURCES MANAGEMENT AND AGRICULTURE

Advances in precipitation modeling have positively impacted water resource management and agriculture in a number of ways. Firstly, improved forecasting of precipitation events allows for more efficient management of reservoirs and irrigation systems. More accurate models help predict water availability, which is crucial for agricultural planning, especially in semi-arid regions where water is a scarce resource. Additionally, the ability to predict extreme events such as droughts and floods has been essential for risk mitigation and natural disaster preparedness.

The integration of satellite data and ground-based observations has brought significant advances in extreme event prediction and water resource management. Models that incorporate this data allow early warnings for floods and droughts, helping in the operation of reservoirs and in the planning of water infrastructure. This is particularly important for agriculture, where predicting water availability and mitigating agricultural risks are critical to ensuring food security.

Another important impact of advances in precipitation modeling is adaptation to climate change. Models that incorporate future climate projections allow farmers and water resource managers to prepare for climate change scenarios, such as changes in precipitation patterns and an increase in the frequency of extreme events. This is crucial for the sustainable management of water resources and for the resilience of agriculture in the face of climate change.

Exhibit 5 below summarizes the main impacts of advances in precipitation modeling in water resources management and agriculture by decade.

Table 5

Major advances in water resources management and agriculture

Decade	Impacts on Water Resources Management and Agriculture
1980	Improved prediction of precipitation events; support for agriculture in semi-arid areas.
1990	Forecast of droughts and floods; water resources planning and agriculture based on atmospheric patterns.
2000	Generation of realistic spatial patterns; improvement in the allocation of water resources; support for agriculture.
2010	Early warnings for extreme events; forecast of water availability for agriculture; water infrastructure planning.
2020	Flow forecasting and reservoir management; mitigation of agricultural risks; adaptation to

	climate change.

Source: Prepared by the authors themselves.

Therefore, advances in precipitation modeling have had a significant impact on water resources management and agriculture. Improved precipitation event forecasting, satellite data integration, and climate change adaptation are some of the key benefits of these advancements.

4.8 EMERGING TRENDS AND INTEGRATION OF NEW TECHNOLOGIES

In the 1980s and 1990s, precipitation modeling was dominated by traditional statistical methods, such as event-based and stochastic models, focusing on semi-arid zones and atmospheric patterns. These approaches, while effective for the context of the time, were limited by a lack of data and computational power. From the 2000s onwards, there was a significant advance with the introduction of techniques such as random cascade models and statistical downscaling, which allowed for a more accurate representation of precipitation at regional and local scales.

The 2010s marked the beginning of the integration of AI/ML techniques in precipitation modeling, with the use of artificial neural networks (ANN) for downscaling and precipitation prediction. In addition, the integration of satellite data and the use of Bayesian methods for quantification of uncertainties have brought significant advances in the accuracy of forecasts. In the 2020s, the consolidation of deep learning techniques, such as LSTM (Long Short-Term Memory), and the use of copulas and heuristic optimization (ISPSO) stood out as emerging trends. These techniques allowed for more accurate and adaptable predictions to climate change scenarios.

The integration of AI/ML in precipitation modeling began to gain prominence in the 2010s, with the use of neural networks (ANN) for downscaling and precipitation forecasting. However, it was in the 2020s that these techniques took hold, with the widespread use of deep learning (LSTM) and heuristic optimization techniques. These advances have made it possible not only to improve the accuracy of forecasts, but also to quantify uncertainties and adapt models to climate change scenarios.

Table 6 presents a summary of the evolution of trends and modeling over the decades according to the articles analyzed.

Table 6

Trend analysis and integration with AI/ML for decades

Decade	Emerging Trends	AI/ML integration
1980	Event-based models; focus on semi-arid areas.	Not mentioned.
1990	Stochastic models; atmospheric classification.	Not mentioned.
2000	Random cascade models; statistical downscaling; adjustment of distributions.	Not mentioned.
2010	Integration of satellite data; Bayesian models; neural networks (ANN).	Neural networks (ANN) for downscaling and forecasting.
2020	Deep learning (LSTM); copulations; integration of satellite data; Heuristic optimization.	Deep learning (LSTM); ML techniques for parameter tuning and optimization.

Source: Prepared by the authors themselves.

4.9 CHARACTERISTICS OF THE MOST INFLUENTIAL PUBLICATIONS ON PRECIPITATION MODELS USING TIME SERIES

The four percent of the most cited articles in each decade of analysis, between 1982-2023, were evaluated. Table 7 shows the number of articles per decade and the corresponding total read to 40% of each decade.

Table 7

Data on the number of articles after final screening divided by decades and quantity read

Decade	Total Articles	Amount read
1980	2	1
1990	11	5
2000	20	8
2010	30	12
2020	15	6

Source: Prepared by the authors themselves.

The final number of articles after all exclusion analysis was 78, in which when extracting the 40% of each decade that were read, the sum corresponds to the equivalent of 32 articles, which corresponds to about 41% of the final total of selected articles.

The most cited published articles are shown in Table 8.

Table 8

Published articles

No.	Title	Authors	Year
1	Practical generation of synthetic rainfall event time series in a semi-arid climatic zone	Bogárdi, János J.; Duckstein, Lucien; Rumambo, Omar H.	1988
2	Modeling daily rainfall using a semi-Markov representation of circulation pattern occurrence	Bardossy, Andras; Plate, Erich J.	1991
3	Regression model for generating time series of daily precipitation amounts for climate change impact studies	Buishand T.A.; Klein Tank A.M.G.	1996
4	Further developments of the neyman-scott clustered point process for modeling rainfall	Cowpertwait, P.S.P.	1991
5	Precipitation and air flow indices over the British Isles	Conway, D., Wilby, R.L., Jones, P.D.	1996
6	Generating precipitation time series using simulated annealing	Bárdossy, A.	1998
7	The collaborative historical African rainfall model: Description and evaluation	Funk, Chris; Michaelsen, Joel; Verdin, Jim; Artan, Guleid; Husak, Greg; Senay, Gabriel; Gadain, Hussein; Magadazire, Tamuka	2003
8	Time series analysis model for rainfall data in jordan: Case study for using time series analysis	King Abdul Aziz University, Jeddah, Saudi Arabia	2009
9	Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades	Jothityangkoon, Chatchai; Sivapalan, Murugesu; Viney, Neil R.	2000
10	Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks	Schoof, J.T., Pryor, S.C.	2001
11	Multivariate rainfall disaggregation at a fine timescale	Koutsoyiannis, D., Onof, C., Wheater, H.S.	2003
12	A spatial rainfall generator for small spatial scales	Willems, P.	2001
13	A space-time hybrid hourly rainfall model for derived flood frequency analysis	Haberlandt, U., Ebner Von Eschenbach, A.-D., Buchwald, I.	2008

14	Coupled rainfall model and discharge model for flood frequency estimation	Arnaud, P., Lavabre, J.	2002
15	A rainfall model based on a Geographically Weighted Regression algorithm for rainfall estimations over the arid Qaidam Basin in China	Lv, Aifeng; Zhou, Lei	2016
16	Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in Nordic watersheds	Li, Zhi; Brissette, François; Chen, Jie	2013
17	Downscaling transient climate change using a Neyman-Scott Rectangular Pulses stochastic rainfall model	Burton, A., Fowler, H.J.,Blenkinsop, S., Kilsby, C.G.	2010
18	A dynamic nonstationary spatio-temporal model for short term prediction of precipitation	Sigrist, F., Künsch, H.R., Stahel, W.A.	2012
19	Application of a stochastic weather generator to assess climate change impacts in a semi-arid climate: The Upper Indus Basin	Forsythe, N., Fowler, H.J., Blenkinsop, S., ... Harpham, C., Hashmi, M.Z.	2014
20	Climate-informed stochastic hydrological modeling: Incorporating decadal-scale variability using paleo data	Henley, B.J., Thyer, M.A., Kuczera, G., Franks, S.W.	2011
21	Precipitation Modeling by Polyhedral RCMARS and Comparison with MARS and CMARS	Özmen, A., Batmaz, İ., Weber, G.-W.	2014
22	Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns - an evaluation of assumptions	Haberlandt, U., Belli, A., Bárdossy, A.	2015
23	Regionalization of the modified bartlett-lewis rectangular pulse stochastic rainfall model	Kim, D., Olivera, F., Cho, H., Socolofsky, S.A.	2013
24	Short time step continuous rainfall modeling and simulation of extreme events	Callau Poduje, A.C., Haberlandt, U.	2017
25	Models of daily rainfall cross-correlation for the United Kingdom		2013

		Burton, A., Glenis, V., Jones, M.R., Kilsby, C.G.	
26	Regionalization of the Modified Bartlett–Lewis rectangular pulse stochastic rainfall model across the Korean Peninsula	Kim, D., Kwon, H.-H., Lee, S.-O., Kim, S.	2016
27	Forecast of rainfall distribution based on fixed sliding window long short-term memory	Chen, C., Zhang, Q., Kashani, M.H., ... Dash, S.S., Chau, K.-W.	2022
28	Rainfall Generation Revisited: Introducing CoSMoS-2s and Advancing Copula-Based Intermittent Time Series Modeling	Papalexiou, S.M.	2022
29	Trends of rainfall and temperature in Northeast Brazil | Rainfall and temperature trends in Northeast Brazil	de Carvalho, A.A., A. Montenegro, A.A., da Silva, H.P., ... de Morais, J.E.F., da Silva, T.G.F.	2020
30	Development and hydrometeorological evaluation of a new stochastic daily rainfall model: Coupling Markov chain with rainfall event model	Gao, C., Booij, M.J., Xu, Y.-P.	2020
31	A robust method to develop future rainfall IDF curves under climate change condition in two major basins of Iran	Khazaei, M.R.	2021
32	The Reliability of W-flow Run-off-Rainfall Model in Predicting Rainfall to the Discharge	Tama, D.R., Limantara, L.M., Suhartanto, E., Devia, Y.P.	2023

Source: Prepared by the authors themselves.

4.10 TOPOLOGY OF THE DECADES AS RESEARCH FOCUSES AND SEMANTIC NETWORKS

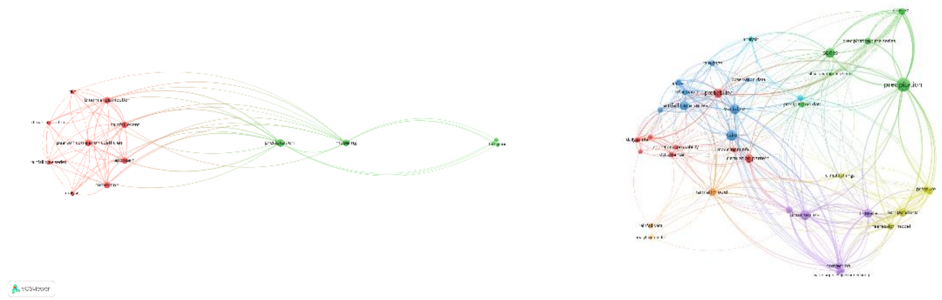
Figure 2 shows an analysis of the network of co-word occurrence, which can be used to identify the state of the art. The research topic that is Mathematical Modeling of precipitation was categorized into four colored clusters, following the composition over time: 1980s (2 clusters), 1990s (7 clusters), 2000s (9 clusters), 2010s (10 clusters), 2020s (7 clusters).

The 10 most cited terms across the clusters include: "precipitation," "rainfall," "probability," "time series," "performance," "climate change," "precipitation intensity," "distribution," "accuracy," and "forecast." These terms reflect the evolution of precipitation modeling, moving from simple rainfall and probability analyses to complex predictive models that integrate time-series data, extreme event forecasts, and the impact of climate change.

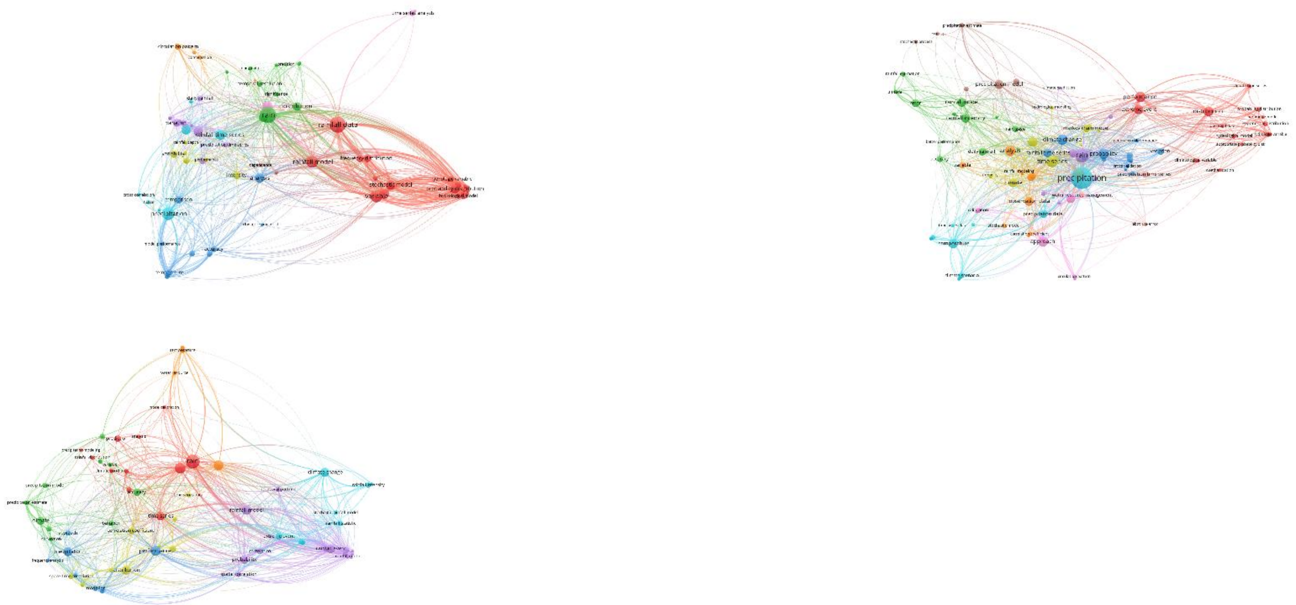
Figure 1

Co-occurrence network of words presented in the titles, abstracts and keywords for each decade (1980, 1990, 2000, 2010 and 2020)

1980s 1990s



2000s 2010s 2020s



Source: Prepared by the authors themselves.

Between 1982 and 1989, studies on precipitation modeling, with only two publications, focused on two main clusters. Cluster 1 includes terms such as "rainfall", "binomial distribution" and "precipitation event", reflecting an approach based on statistical and probabilistic methods, such as the binomial distribution and Pearson's correlation coefficient, to analyze rainfall events. The presence of "time series" and "convection" indicates the use of temporal data and atmospheric processes in the study of these phenomena. Cluster 2, on the other hand, highlights terms such as "precipitation", "modeling" and "behavior", suggesting a focus on modeling precipitation processes and analyzing the behavior of these phenomena under different climatic conditions. These clusters mark the beginning of investigations into rainfall patterns and their variations.

Between 1990 and 1999, studies on precipitation modeling expanded, with 11 publications distributed in seven clusters. Cluster 1 highlighted terms such as "probability", "daily precipitation" and "conditional probability", indicating the use of advanced statistical methods to model rainfall events and their probabilities. Cluster 2 focused on "precipitation", "time series" and "precision", reflecting temporal analysis and the search for greater accuracy in the prediction of precipitation data. Cluster 3 combined "rainfall", "probability", "observation data" and "performance", pointing to the use of observational data and performance metrics in climate models. Cluster 4 introduced variables such as "temperature", "pressure", "climate change" and "regression model", relating climate change to precipitation patterns. Cluster 5 highlighted "time series", "climate" and "convection", with a focus on climate analysis and atmospheric circulation models. Cluster 6 emphasized "variation" and "analysis," exploring climatic variations and analytical techniques. Finally, Cluster 7 highlighted "precipitation model" and "precipitation data", reflecting quantitative modeling based on observational data. These clusters show an evolution in methods, with greater integration of statistics, time series, atmospheric variables, and climate change in the study of precipitation.

Between 2000 and 2009, studies on precipitation modeling became more complex, with 20 publications distributed in nine clusters. Cluster 1 highlighted terms such as "precipitation data", "precipitation model", "stochastic model" and "probability distribution", indicating the use of statistical and hydrological models to analyze precipitation variability. Cluster 2 focused on "rainfall", "distribution", "temporal resolution" and "daily data", with emphasis on the spatial and temporal distribution of precipitation. Cluster 3 combined "analysis", "accuracy" and "temperature", reflecting the evaluation of climate models and the analysis of atmospheric data. Cluster 4 highlighted "variability", "intensity", "climate scenario" and "depth of precipitation", exploring variations and intensities of rainfall in different scenarios. Cluster 5 focused on "time series" and "variation", with modeling of precipitation in space-time. Cluster 6 included "precipitation," "precipitation time series," "rain gauge," and "precipitation estimation," using observational data to feed models. Cluster 7 highlighted "circulation pattern", analyzing the interaction between atmospheric patterns and precipitation. Cluster 8 focused on "dependence" and "amplitude", studying interrelationships between climate variables and precipitation events. Cluster 9 emphasized "approach" and "time series analysis", with a methodological focus on time series and detailed data analysis. These clusters show a significant advance in precipitation modeling, with greater complexity, integration of time series, hydrological variables, and stochastic models, reflecting the evolution of climate research techniques.

Between 2010 and 2019, studies on precipitation modeling presented a diversified and highly technical approach, with 30 publications distributed in ten clusters. Cluster 1 highlighted terms such as "performance", "distribution", "extreme event" and "probability distribution", with a focus on the evaluation of precipitation models and prediction of extreme weather events. Cluster 2 emphasized "precipitation model", "precipitation intensity", "error" and "rain gauge", indicating interest in measuring rainfall intensity and the accuracy of models based on observational data. Cluster 3 addressed "probability", "climate change" and "precipitation time series", relating climate change to precipitation modeling with an emphasis on temporal data. Cluster 4 focused on "time series", "variability", "intensity" and "Markov chain model", using advanced methods to predict variations in precipitation. Cluster 5 highlighted "rainfall" and "forecast", with terms such as "statistical model" and "time series analysis", showing the application of statistical models for rainfall forecasting. Cluster 6 explored "precipitation", "temperature" and "climate scenario", analyzing interactions between climate variables and future scenarios. Cluster 7 included "analysis", "validation" and "observation data", with a focus on model validation and precision analysis. Cluster 8 highlighted "precipitation model", "kriging" and "stochastic process", using advanced spatial and stochastic techniques. Cluster 9 emphasized "approach" and "calibration," reflecting the importance of model calibration to ensure accuracy. Cluster 10 focused on "precipitation time series" and "water resources management", highlighting the relevance of temporal analysis for the efficient management of water resources. These clusters demonstrate a high degree of sophistication, with advances in statistical techniques, prediction of extreme events, spatial modeling, and integration of climate variables, reflecting the complexity and interdisciplinarity of climate research in this period.

Between 2020 and 2023, studies on precipitation modeling delved into topics such as climate predictions, variability, and rainfall events, with 15 publications distributed in seven clusters. Cluster 1 highlighted "rainfall", "forecast", "analysis" and "precipitation data", with a focus on rain event forecasts based on detailed analyses of observational data. Cluster 2 emphasized "accuracy," "climate," "meteorological data," "precipitation model," and "precipitation estimation," reflecting the search for greater reliability in climate models. Cluster 3 focused on "performance", "approach" and "modeling", with studies aimed at evaluating and improving modeling techniques. Cluster 4 highlighted "distribution", "variability", "correlation coefficient" and "time series data", indicating detailed analyses of precipitation variability over time. Cluster 5 focused on "precipitation model", "temporal pattern", "probability" and "precipitation event", with a focus on stochastic modeling and analysis of rainfall temporal patterns. Cluster 6 highlighted "climate change", "extreme event", "hydrological modelling" and "precipitation intensity", exploring the impacts of extreme events and climate change on hydrology. Cluster 7 emphasized "temperature," "water resources," and "precipitation time series," analyzing the effects of temperature on water availability and precipitation variability. These clusters show a strong emphasis on improving the accuracy of models, analyzing extreme events and climate change, and using time series to understand precipitation patterns and variations.

5 CONCLUSION

This study performed a comprehensive bibliometric analysis on precipitation estimation models using time series, identifying trends, methods, and significant contributions over the last four decades. The main findings reveal a marked evolution in precipitation modeling, with a progressive increase in the complexity and precision of the methods used. Initially, in the 1980s, models were based on simple statistical techniques, such as binomial distributions and basic correlations. Over time, there has been a transition to more sophisticated methods such as stochastic models, artificial neural networks (ANN), and more recently, deep learning (LSTM) and heuristic optimization (ISPSO) techniques.

The analysis highlighted the importance of temporal and spatial resolution, with most studies adopting daily and regional resolutions, reflecting a balance between detail and computational feasibility. Additionally, the integration of satellite data and the utilization of artificial intelligence (AI) and machine learning (ML) techniques have become dominant trends over the past decade, allowing for more accurate and adaptable predictions to climate change scenarios.

The practical impacts of these advances are significant, especially in water resources management and agriculture. More accurate models allow for more reliable prediction of extreme events, such as droughts and floods, contributing to risk mitigation and sustainable planning of water and agricultural infrastructures. Adaptation to climate change was also a central theme, with models that incorporate future projections helping to prepare farmers and managers for adverse climate scenarios.

In terms of theoretical contributions, the study mapped the main methodological trends, highlighting the importance of quantifying uncertainties, the integration of multiple climate variables, and the use of advanced statistical and spatial analysis techniques. The bibliometric analysis also identified the main contributing countries and institutions, with the United States leading in number of publications, followed by the United Kingdom and Australia.

In summary, this study reinforces the relevance of precipitation modeling as a crucial research area to address contemporary climate challenges. The findings highlight the continued need for methodological innovation, international collaboration, and the integration of new technologies, such as AI and ML, to enhance the accuracy and applicability of precipitation models. These advances not only contribute to scientific advancement, but also have significant practical implications for sustainable water resource management and food security in a context of global climate change.

REFERENCES

Antonini, J. C. A., Silva, E. M., Oliveira, L. F. C., & Sano, E. E. (2009). Modelo matemático para estimativa da temperatura média diária do ar no Estado de Goiás. Pesquisa Agropecuária Brasileira, 44(4), 331–338.

Arai, H., et al. (2009). Estudo sobre a significância da precipitação na caracterização do clima regional. Revista de Climatologia, X(Y), 123–145.

Arnaud, P., & Lavabre, J. (2002). Coupled rainfall model and discharge model for flood frequency estimation. Water Resources Research, 38(11), 1241.

Bárdossy, A. (1998). Generating precipitation time series using simulated annealing. Water Resources Research, 34(7), 1733–1741.

Bárdossy, A., & Plate, E. J. (1991). Modeling daily rainfall using a semi-Markov representation of circulation pattern occurrence. Journal of Hydrology, 122(1–4), 33–47.

Bertoni, J. C., & Tucci, C. E. M. (2009). Precipitação. In C. E. M. Tucci & A. L. L. Silveira (Eds.), Hidrologia: Ciência e aplicação. ABRH.

Bogárdi, J. J., Duckstein, L., & Rumambo, O. H. (1988). Practical generation of synthetic rainfall event time series in a semi-arid climatic zone. Journal of Hydrology, 102(1–4), 335–355.

Borsato, V. A. (2000). A climatologia dinâmica e o ensino da geografia no segundo grau: Uma aproximação ao problema. Revista GeoNotas, 4(1). ISSN 1415-0646.

Buishand, T. A., & Klein Tank, A. M. G. (1996). Regression model for generating time series of daily precipitation amounts for climate change impact studies. Theoretical and Applied Climatology, 53(1–3), 1–19.

Burton, A., Fowler, H. J., Blenkinsop, S., & Kilsby, C. G. (2010). Downscaling transient climate change using a Neyman-Scott Rectangular Pulses stochastic rainfall model. Journal of Hydrology, 381(1–2), 18–32.

Burton, A., Glenis, V., Jones, M. R., & Kilsby, C. G. (2013). Models of daily rainfall cross-correlation for the United Kingdom. Environmental Modelling & Software, 49, 22–33.

Café, L., & Brascher, M. (2008). Organização da informação e bibliometria. Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, (esp.), 54–75.

Callau Poduje, A. C., & Haberlandt, U. (2017). Short time step continuous rainfall modeling and simulation of extreme events. Journal of Hydrology, 550, 324–336.

Carletoo, C. (2005). Subcoberturas: Aplicação de isolantes térmicos em sistemas de coberturas [Dissertação de Mestrado]. Escola de Engenharia de São Carlos, Universidade de São Paulo.

Cavalcanti, I. F. A., Ferreira, N. J., Silva, M. G. A. J., & Dias, M. A. F. S. (Eds.). (2009). Tempo e clima no Brasil. Oficina de Textos.

Chen, C., Zhang, Q., Kashani, M. H., Dash, S. S., & Chau, K.-W. (2022). Forecast of rainfall distribution based on fixed sliding window long short-term memory. Water Resources Research, 58(1), e2021WR031641.

Conway, D., Wilby, R. L., & Jones, P. D. (1996). Precipitation and air flow indices over the British Isles. Climate Research, 5(2), 169–185.

Cowpertwait, P. S. P. (1991). Further developments of the Neyman-Scott clustered point process for modeling rainfall. Water Resources Research, 27(7), 1431–1438.

De Carvalho, A. A., De A. Montenegro, A. A., Da Silva, H. P., De Morais, J. E. F., & Da Silva, T. G. F. (2020). Trends of rainfall and temperature in Northeast Brazil. Revista Brasileira de Engenharia Agrícola e Ambiental, 24(1), 15–23.

Forsythe, N., Fowler, H. J., Blenkinsop, S., Harpham, C., & Hashmi, M. Z. (2014). Application of a stochastic weather generator to assess climate change impacts in a semi-arid climate: The Upper Indus Basin. Journal of Hydrology, 517, 1019–1034.

Funk, C., Michaelsen, J., Verdin, J., Artan, G., Husak, G., Senay, G., Gadain, H., & Magadzire, T. (2003). The collaborative historical African rainfall model: Description and evaluation. International Journal of Climatology, 23(1), 47–66.

Gao, C., Booij, M. J., & Xu, Y.-P. (2020). Development and hydrometeorological evaluation of a new stochastic daily rainfall model: Coupling Markov chain with rainfall event model. Journal of Hydrology, 588, 125337.

Guedes, V. L. S., & Borshiver, S. (2005). Bibliometria: Uma ferramenta estatística para a gestão da informação e do conhecimento. In Encontro Nacional de Ciência da Informação, 6., Salvador. Anais. ENANCIB.

Haberlandt, U., Belli, A., & Bárdossy, A. (2015). Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns - An evaluation of assumptions. International Journal of Climatology, 35(2), 217–231.

Haberlandt, U., Ebner von Eschenbach, A.-D., & Buchwald, I. (2008). A space-time hybrid hourly rainfall model for derived flood frequency analysis. Hydrology and Earth System Sciences, 12(6), 1353–1367.

Henley, B. J., Thyer, M. A., Kuczera, G., & Franks, S. W. (2011). Climate-informed stochastic hydrological modeling: Incorporating decadal-scale variability using paleo data. Water Resources Research, 47(10), W00H10.

Job, I. (2018). Bibliometria aplicada aos estudos do campo da Educação Física: Confiabilidade, qualidade e relevância nas publicações. Motrivivência, 30(54), 18–34.

Jothityangkoon, C., Sivapalan, M., & Viney, N. R. (2000). Tests of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades. Water Resources Research, 36(1), 29–42.

Khazaei, M. R. (2021). A robust method to develop future rainfall IDF curves under climate change condition in two major basins of Iran. Theoretical and Applied Climatology, 144(1–2), 1–15.

Kim, D., Kwon, H.-H., Lee, S.-O., & Kim, S. (2014). Regionalization of the Modified Bartlett–Lewis rectangular pulse stochastic rainfall model across the Korean Peninsula. Journal of Hydro-environment Research, 8(4), 317–328.

Kim, D., Olivera, F., Cho, H., & Socolofsky, S. A. (2012). Regionalization of the modified Bartlett-Lewis rectangular pulse stochastic rainfall model. Terrestrial, Atmospheric and Oceanic Sciences, 23(6), 631–642.

King Abdul Aziz University, Jeddah, Saudi Arabia. (2009). Time series analysis model for rainfall data in Jordan: Case study for using time series analysis. American Journal of Environmental Sciences, 5(5), 599–604.

Koutsoyiannis, D., Onof, C., & Wheater, H. S. (2003). Multivariate rainfall disaggregation at a fine timescale. Water Resources Research, 39(7), 1173.

Li, Z., Brissette, F., & Chen, J. (2013). Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in Nordic watersheds. Hydrological Processes, 27(26), 3718–3729.

Lucena, D. B. (2008). Impacto dos Oceanos Pacífico e Atlântico no Clima do Nordeste do Brasil [Tese de Doutorado]. Universidade Federal de Campina Grande.

Lv, A., & Zhou, L. (2016). A rainfall model based on a Geographically Weighted Regression algorithm for rainfall estimations over the arid Qaidam Basin in China. Remote Sensing, 8(4), 311.

Monteiro, C. A. F. (1991). Clima e excepcionalismo: Conjecturas sobre o desempenho da atmosfera como fenômeno geográfico. UFSC.

Özmen, A., Batmaz, İ., & Weber, G.-W. (2014). Precipitation modeling by Polyhedral RCMARS and comparison with MARS and CMARS. Environmental Modeling & Assessment, 19(5), 425–435.

Papalexiou, S. M. (2022). Rainfall generation revisited: Introducing CoSMoS-2s and advancing copula-based intermittent time series modeling. Water Resources Research, 58(1), e2021WR031641.

Pedro Júnior, M. J., Sentelhas, P. C., Pommer, C. V., & Martins, F. P. (1994). Determinação da temperatura-base, graus-dia e índice biometeorológico para a videira 'Niagara Rosada'. Revista Brasileira de Agrometeorologia, 2, 51–56.

Santana, M. M. M., Mariano-Neto, E., de Vasconcelos, R., & Dodonov, P. (2021). Mapping the research history, collaborations and trends of remote sensing in fire ecology. Scientometrics, 126, 1–20. https://doi.org/10.1007/s11192-020-03805-x

Santos, A. F., Santos, C. S., Gomes, A. S. (2023). Análise bibliométrica das tendências e avanços nas pesquisas em sustentabilidade no período de 2019 a 2023. Universidade do Estado de Mato Grosso – UNEMAT, Campus Universitário de Sinop.

Santos, I., Fill, H. D., Sugai, M. R. V. B., Buba, H., Kishi, R. T., Marone, E., & Lautert, L. F. C. (2001). Hidrometria aplicada. Editora do Instituto de Tecnologia para o Desenvolvimento.

Schoof, J. T., & Pryor, S. C. (2001). Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. International Journal of Climatology, 21(7), 773–790.

Sigrist, F., Künsch, H. R., & Stahel, W. A. (2012). A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. Annals of Applied Statistics, 6(4), 1452–1477.

Silva, A. C. (2001). Estudos climáticos e ambiente construído no município de Descalvado - SP [Dissertação de Mestrado]. Escola de Engenharia de São Carlos, Universidade de São Paulo.

Tama, D. R., Limantara, L. M., Suhartanto, E., & Devia, Y. P. (2023). The reliability of W-flow Run-off-Rainfall Model in predicting rainfall to the discharge. Civil Engineering Journal, 9(7), 15.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2), 523–538.

Van Nunen, K., Li, J., Reniers, G., & Ponnet, K. (2018). Bibliometric analysis of safety culture research. Safety Science, 108, 248–258.

Vecchia, F. (2001). Projeto Conforto Térmico. Estudo comparativo do comportamento térmico de sistemas de cobertura. EESC/FIPAI.

Willems, P. (2001). A spatial rainfall generator for small spatial scales. Journal of Hydrology, 252(1–4), 126–144.

Wu, J. (2017). Thirty years of Landscape Ecology (1987–2017): Retrospects and prospects. Landscape Ecology, 32(12), 2225–2239.