

Sustainability evaluation of watershed indicator: A bibliometric analysis using Latent Dirichlet Allocation

Rebeca Moreno-Santoyo¹, Andrés García-Medina^{2*}, Juana Claudia Leyva Aguilera³, María Concepción Arredondo⁴

¹ Instituto de Investigaciones Oceanológicas, Autonomous University of Baja California, Transpeninsular Highway Ensenada - Tijuana 3917, Ensenada, 22860, Baja California, Mexico.

² Tijuana Institute of Technology, Tecnológico Nacional de México, Calz. del Tecnológico 12950, Tomás Aquino, Tijuana, 22414, Baja California, Mexico.

³ Faculty of Sciences, Autonomous University of Baja California, Transpeninsular Highway Ensenada - Tijuana 3917, Ensenada, 22860, Baja California, Mexico.

⁴ Faculty of Marine Sciences, Autonomous University of Baja California, Transpeninsular Highway Ensenada - Tijuana 3917, Ensenada, 22860, Baja California, Mexico.

* e-mail: andgarm.n@gmail.com

Abstract: A watershed is a geographic space with ecosystem functions shaped and maintained by ecological and socioeconomic processes that directly or indirectly influence human well-being. This study examines the evolution of 2,675 scientific publications through a systematic literature review (SLR) that integrates bibliometric and content analysis. Given the vast amount of information, the use of Big Data tools from machine learning—specifically the Latent Dirichlet Allocation (LDA) model—was necessary to construct a TopicScore index for selecting the articles with the highest co-occurrence. This approach enables us to address the following research questions: How is the watershed conceptualized as an object of evaluation? What does watershed sustainability entail? Finally, the study examines existing publication trends in relation to the United Nations' global agendas. The findings identify four main thematic axes: Groundwater, Hydrological and Land-use Indicators, Climate Change, and Socioecosystem, as well as the diverse conceptual and methodological approaches applied to assessing watershed sustainability over time.

Key words: Systematic Literature Review (SLR); Latent Dirichlet Allocation (LDA); watershed; sustainability.

1. INTRODUCTION

1.1 Watershed approach and the global agenda

In 1992, a breakthrough was made in the global agenda regarding water resources and sustainable development, as agreements were reached that remain in force today (Koudstaal et al., 1992). First, the principles of Integrated Water Resources Management (IWRM) were presented at the Conference on Water and the Environment (Dublin), which highlighted the importance of adopting a holistic, territorial framework for global water management (Moriarty et al., 2000). This conference highlighted the importance of a participatory approach and the fundamental role of women in decision-making. Water was recognized as an economic good, as was the promotion of decentralization processes in its management (Rahaman et al., 2004).

These principles were complemented by those established months later at the United Nations Conference on Environment and Development (Rio de Janeiro), known as the Earth Summit. At this summit, the watershed was recognized as the ideal spatial unit for the sustainable planning and management of natural resources, integrating ecological, social, and economic dimensions in a single territory. This is due to its importance in water conservation, biodiversity, and adaptation to climate change. Furthermore, the need to monitor territorial elements to evaluate the effectiveness of transitional actions toward sustainable development and to follow up on these actions was emphasized (UN, 1992).

These conferences had a major impact on the global development agenda, with the consolidation of the concept of sustainability, initially promoted at the Earth Summit in Rio de Janeiro in 1992 and taken up again at the Rio+20 conference (Elder and Grünwald, 2025; Linnér and Selin, 2013), which led to a reassessment of the relevance of the global development agenda, that is, the Millennium Development Goals (MDGs), focused on poverty reduction in developing countries from a multidimensional perspective (Loewe, 2012). As a result, two independent United Nations working groups were created: one to debate whether to adopt a new global development agenda; the other to compile a list of possible sustainability goals that more broadly integrate the economic, social, and environmental dimensions to complement the MDGs. This led to the Millennium Development Goals being phased out in 2015, paving the way for a new global agenda, called the Sustainable Development Goals (SDGs), known as the 2030 Agenda (Gain et al., 2016).

The 2030 Agenda for Sustainable Development was adopted by the United Nations General Assembly in September 2015. The Agenda comprises 17 SDGs and 169 targets, measured by more than 200 indicators (UN, 2015). These objectives aim to guide global development toward sustainability (Glaser, 2012), which goes beyond the fight against poverty.

While the indicators and targets are connected by the social, environmental, and economic dimensions, the agenda has a goal dedicated to water and sanitation (SDG6) with eight global targets and indicators, as shown in Table 1 (UN, 2025). These are evaluated annually by each member country of the United Nations. Compliance with the SDG 6 indicators depends largely on the protection and restoration of natural ecosystems, such as watersheds (Mulligan et al., 2020).

Table 1. Targets and Indicators of SDG 6: Clean Water and Sanitation

Target Name	Indicator Name
6.1. Achieve universal and equitable access to safe and affordable drinking water for all	6.1.1. Proportion of population use safely managed drinking water services.
6.2. Achieve access to adequate and equitable sanitation and hygiene for all and end open defecation	6.2.1. Proportion of population use (a) safely managed sanitation services and (b) a hand-washing facility with soap and water. 6.2.2. Proportion of population using (a) safely managed sanitation services and (b) facilities that are not shared with other households.
6.3. Improve water quality by reducing pollution, eliminating dumping and minimizing release of hazardous chemical and materials	6.3.1. Proportion of domestic and industrial wastewater flows safely treated. 6.3.2. Proportion of bodies of water with good ambient water quality.
Substantially increase water-use efficiency across all sectors and ensure sustainable withdrawals and supply of freshwater	6.4.1. Change over time (2015-2020) in water-use efficiency – total water-use efficiency (US\$/m ³). 6.4.2. Level of water stress: freshwater withdrawal as a proportion of available freshwater resources.
6.5. Implement integrated water resources management at all levels, including through transboundary cooperation	6.5.1. Degree of integrated water resources management implementation (0-100). 6.5.2. Proportion of transboundary basin area with an operational arrangement for water cooperation.
6.6. Protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes	6.6.1. Change in the extent of water-related ecosystems over time.
6.a. Expand international cooperation and capacity-building support to developing countries	6.a.1. Amount of water- and sanitation-related official development assistance is part of a government-coordinated spending plan.
6.b. Support and strengthen the participation of local communities for improving water and sanitation management	6.b.1. Proportion of local administrative units with established and operational policies and procedures for participation of local communities in water and sanitation management.

Currently, more than 30 years after the recognition of watersheds on the global agenda, it is essential to understand the complexity of the different approaches to long-term water resource management through synthesis, precision, and high-level detail in watershed sustainability assessments (Vollmer et al., 2016). Therefore, this paper explores the various approaches of the disciplines that study watersheds by analyzing existing scientific evidence within the context of the big data era. The aim is to strengthen understanding of scientific findings assessing watershed

sustainability by identifying gaps, progress, and their potential contributions to achieving global goals by 2030.

2. SUSTAINABILITY IN WATERSHEDS

2.1 Assessing sustainability in watersheds

After more than 30 years of promoting watersheds as an ideal space for territorial planning and water resource management, numerous studies now evaluate development that seeks to meet the needs of the present without compromising the capacity of future generations, with special attention to water resources (Muñoz, 2021). This is due to ecosystem degradation resulting from population growth, industrial activities, urbanization, and climate change, which together have exacerbated water pollution and availability (Ighalo et al., 2020), posing a serious threat to livelihoods and global sustainability (Lawal et al., 2023).

The watershed is conceived as a strategic territorial unit for the sustainable management and administration of water, due to the intrinsically linked and interdependent relationships between the physical and biotic systems (Jarzebski et al., 2024). It represents the space where the natural regeneration cycle of ecosystems occurs, as well as where complex interactions among social actors, economic activities, and unique ecological characteristics unfold (Perales, 2016; Perevochtchikova, 2008; Pérez and Le Blas, 2004).

Therefore, monitoring and evaluating the watershed are essential to understand the spatiotemporal dynamics of water resources, determine ecosystem functionality, and the services that support human, social, cultural, and economic well-being (Folke et al., 2016; Naeem et al., 2012).

In this sense, indices and indicators function as an evaluation tool that facilitates understanding of the status of natural resources: water, soil, and vegetation (Vollmer et al., 2016), through the measurement, estimation, and comparison of situations, activities, or results. This allows decision-makers to have synthesized, precise, and detailed information on the reality of watersheds in the face of the negative or positive impacts they face (Mosaffaie et al., 2021; Senent-Aparicio et al., 2016; Turner et al., 2016), which allows the implementation of management programs aimed at sustainable management (Pereira et al., 2020). However, achieving watershed sustainability is not easy, as it requires a holistic understanding of the variables that affect water resource management and their interrelationships (Karamouz et al., 2017).

Lawal et al. (2023), identified the watershed as a geological transition zone (GTZ), composed of aquifers with complex hydrogeology and vulnerable to contamination. Therefore, the study focuses on analysing the differences in water quality across different soils. It considers groundwater quality monitoring and management strategies and affirms that long-term territorial planning is necessary to achieve the watershed's sustainability.

Zhang et al. (2020) mentioned that the watershed is a hydraulic system sensitive to changes in land use, where the interaction between vegetation, precipitation, and surface and groundwater dynamics is essential for the availability of water resources. The study presented different land-use change scenarios (deforestation, afforestation, and urbanization) that alter hydrological processes and the water balance in some Australian watersheds. It is emphasized that achieving watershed sustainability involves appropriately managing land-use changes to minimize negative impacts on the hydrological cycle, and using an improved model that more accurately reflects vegetation dynamics, thereby enabling informed decisions for sustainable water management and ecological conservation.

Wabela et al. (2024), the authors considered the watershed in the sustainability evaluation as a functional unit of agricultural management, local production, and vulnerable to the decrease in water resources. Therefore, the sustainability of the watershed is related to the efficient

management of water use and the availability and equitable distribution of the resource, and seeks to establish measures to avoid deficits or excesses in management.

Based on the above studies, it can be observed that the adoption of models and methodologies varies according to the research objective, perspective, and disciplinary background, which in turn define the conceptualization of what constitutes a sustainable watershed in assessment terms. In many cases, the focus is on addressing a specific problem in the territory, to monitor the progress of the implemented actions (Barbier and Burgess, 2017), with the direct or indirect fulfilment of the current global agenda - SDGs -, especially those related to water, sanitation, environmental conservation, and sustainable socioeconomic development (Liu et al., 2023).

The above highlights the importance of recording and linking advances in natural resource management, given the large number of publications on the subject and the diverse approaches, especially in the face of persistent environmental problems. In this regard, it is pertinent to develop innovative approaches that improve the synthesis, precision, and level of detail in the analysis of the evolution of watershed sustainability assessments (Vollmer et al., 2016). These approaches would allow for identifying gaps and advances in the sustainable development of water resources, after three decades of promoting the watershed as the ideal unit for territorial planning and integrated natural resource management.

2.2 Latent Dirichlet Analysis (LDA)

To achieve a deeper understanding of the scientific literature, this study conducts a Systematic Literature Review (SLR) by collecting, examining, and synthesizing the characteristics of academic research within the domain (Card, 2015; Chien et al., 2024; Paul et al., 2021) through bibliometric and content analysis.

The relevance of this study lies in the application of the Latent Dirichlet Allocation (LDA) model for content analysis, a machine learning technique that, although it has been widely used in research related to topic mining in texts —particularly in areas such as climate change, energy storage, health cybersecurity during the COVID-19 pandemic, and distributed photovoltaic energy (Li et al., 2024; Lieneck et al., 2023; Wang et al., 2024; Wu et al., 2023)—, still has limited application in the analysis of studies on sustainability in watersheds.

This research, therefore, represents a novel contribution by integrating the LDA model into an SLR focused on watershed sustainability assessment. This approach will identify trends, gaps, and critical points in the evolution of scientific knowledge on the topic, which is especially relevant in the context of SDGs, whose global target for 2030 demands rigorous, evidence-based monitoring of the current state of natural resources and the approaches applied to their sustainable management.

LDA was originally proposed by Blei et al. (2003) as a generative probabilistic model for discrete data such as text corpora. LDA can be seen as an extension of Principal Component Analysis (PCA) for discrete data. Buntine and Jakulin (2005) introduced Discrete Component Analysis (DCA) as a generalization of PCA for discrete data, where LDA is a particular case of DCA. The review by Blei (2012) is an excellent guide to understand the intuition behind LDA, which gives an overview of these models from a practical perspective.

2.3 Organization of the paper

Section 2 addresses the methodology, systematically describing the steps of the bibliometric and content analysis using LDA. Section 3 presents the most relevant findings derived from the study, which are compared with the priorities established in the global agenda. Section 4 then discusses these results in depth, identifying key stages and turning points in the evolution of the study on watershed sustainability assessment. Finally, Section 5 presents the conclusions and summarizes the main contributions of the study.

3. METHODOLOGY

To evaluate the existing scientific evidence and identify gaps in the understanding of watershed sustainability, the different disciplinary approaches are examined and their influence on global agendas is analyzed through SLR. This method ensures that the bibliographic search and analytical process are transparent and replicable (Kitchenham et al., 2009; Pati and Lorusso, 2018).

The first step was to establish criteria for bibliographic search engines, such as Web of Science. The PICO methodology proposed by Haynes et al. (1997) was used. PICO methodology allows the inclusion and exclusion of specialized documents based on a research question consisting of four components: *Patient/Problem*, *Intervention*, *Comparison*, and *Outcome*, which constitute the PICO anagram, as recommended by Galvao and Ricarte (2019) (see Table 2).

Table 2. PICO methodology. Query used on Web of Science based on our Research question

Research question	RQ: What are the topics studied in the assessment of watershed sustainability?
Population or problem	"watershed" or "basin"
Intervention	("assessment" OR "evaluation" OR "estimate") AND ("Sustainability" OR "Sustainable")
Comparison	No comparison
Outcome	("index" OR "indicator")

The Web of Science data search engine was selected for its extensive coverage of the scientific literature (De Lemos et al., 2024). A global scope was accepted, and any specific study area was considered. The articles were retrieved through a search across the entire document using the following criterion: (("watershed*" OR "basin*") AND ("Sustainability" OR "Sustainable") AND ("assessment" OR "evaluation" OR "estimate") AND ("index" OR "indicator")). The search yielded 2,675 scientific articles published from 1992 to December 2024. The next step was to perform a bibliometric analysis, in which the production of scientific articles was identified by year, disciplinary area, and country. Once the documents were collected, the LDA technique was applied using the general-purpose programming language Python.

Let us define some elements to formalize the LDA method, following the works of Blei et al. (2003) and Blei (2012). First, the basic unit, namely a word, is defined as an element of the vocabulary. A document is a sequence of N words, denoted by $\mathbf{w} = (w_1, w_2, \dots, w_N)$, where w_n is the n -th word in the document. A corpus is a collection of D documents, denoted by $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_D)$. In this way, $w_{n,d}$ denotes the n -th word in document d , i.e., the (d, n) -th element of the corpus matrix \mathbf{W} . The topics are represented by $\boldsymbol{\beta}$, where each β_k , for $k = 1, \dots, K$, is a distribution over the vocabulary. The topic proportion for topic k in document d is denoted by $\theta_{d,k}$. Moreover, the topic assigned to the n -th word in document d is represented as $z_{d,n}$.

LDA can be defined as a generative probabilistic model of a corpus. In this model, documents are assumed to be represented as random mixtures over latent topics. Specifically, the generative process for LDA is given by the following joint distribution probability p of the hidden $(\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{z})$ and observed (\mathbf{W}) variables Eq. 1:

$$p(\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{z}, \mathbf{W}) = \left(\prod_{i=1}^k p(\beta_i)\right) \prod_{d=1}^D p(\theta_d) \left(\prod_{n=1}^N p(z_{d,n}|\theta_d)\right) p(w_{d,n}|\boldsymbol{\beta}, \mathbf{z}) \tag{1}$$

A graphical representation of this model can be seen in Figure 1. Here, each node represents a random variable. Hidden variables are unshaded, while observed variables are shaded. In this case, the only observed variable is the set of document words, $w_{d,n}$. On the other hand, the rectangles (plates) represent repeated structures, and the arrows represent the dependency between the variables.

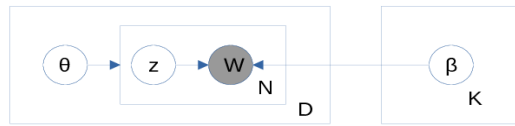


Figure 1. Graphical representation of LDA.

The statistical dependencies appearing in both Eq. 2 and its graphical representation (Figure 1) define the structure of an LDA model. In theory, the parameters (β, θ, z) of the LDA model are estimated using Bayesian statistics with the posterior probability distribution p :

$$p(\beta, \theta, z | W) = \frac{p(\beta, \theta, z, W)}{p(W)} \quad (2)$$

However, estimating these probabilities is computationally intractable since the number of possible topic structures is exponentially large. In practice, topic modeling algorithms fall into two main categories: sampling-based and variational algorithms.

The first approach requires collecting information to approximate the probability distribution of topics using computational statistics techniques such as Markov chains and Gibbs sampling. The second strategy is a deterministic, optimization-based methodology, which is much better suited to big data. A proposal in this regard is online stochastic optimization (Hoffman et al., 2010), which can be applied to massive and streaming document collections. Moreover, it has been shown that LDA parameters can be estimated using stochastic variational inference (Hoffman et al., 2013).

In practice, it is necessary to determine the number of topics K a priori. In Roder and Hinneburg (2015), the authors propose the Coherence Measure as a metric to guide this choice. They define a set of statements as coherent if the statements support each other. In this sense, a coherent set of facts should cover most of the relevant information. As an illustrative example, consider the following statements: (1) The forest is a place with abundant vegetation, (2) The vegetation in the forest is mostly composed of pine trees, (3) The pine trees are covered with snow, (4) It snows in the forest during the winter season. From an intuitive perspective, these statements may be regarded as coherent, given their internal consistency.

To calculate coherence without human intervention, joint and marginal probability functions associated with such statements are constructed. The procedure for estimating coherence can be divided into four main stages: (1) Segmentation, (2) Probability Estimation, (3) Confirmation Measure, and (4) Aggregation. In the first stage, clustering techniques are applied to segment the text. The second stage involves computing probabilistic co-occurrence measures for the words within each segment or group. In the third stage, the confirmation measure is evaluated using, for example, log-likelihood metrics. Finally, in the fourth stage, these evaluations are aggregated to produce a coherence score for the topics.

The three publications with the highest TopicScore for each topic identified using the LDA method are selected and analyzed to answer the following questions: How is the watershed viewed as an object of assessment? What does the watershed's sustainability entail? Finally, the relationship between global agendas and the identified topics and keywords is analyzed.

4. RESULTS

4.1 Exploratory data analysis

Fifty-five (55) disciplinary areas were identified, with Environmental Sciences and Ecology accounting for the largest proportion at 36%, followed by Water Resources at 15%. Figure 2 shows the top 10 disciplines by article count on watershed sustainability.

The fields of Environmental Science and Ecology have shown uninterrupted productivity since 2004 and 2006, respectively; Engineering and Geology have also demonstrated continuous productivity since 2007 and 2008. Most of these fields are related to the exact and biological sciences.

However, academic publications were recorded in the area of Public, Environmental, and Occupational Health, which accounted for 0.9% of the total published in 2011, ranking eleventh in terms of output. Public Administration was also recently incorporated into these studies in 2021. The first study was published in 1992, and research output has been continuous since 1998. Production peaked in 2024 at 18.3%, followed by 2023 at 16.8%.

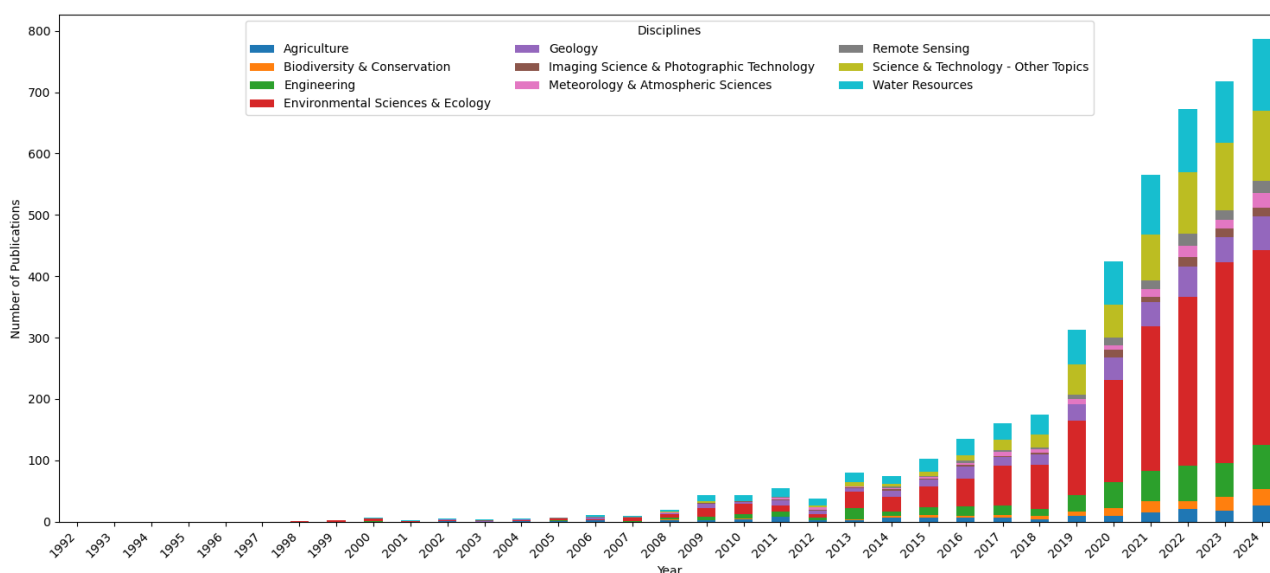


Figure 2. Annual production of scientific articles by major disciplines.

China leads in the number of publications, with 28.6%, followed by the United States at 9.5%. The United States was a pioneer in publishing the first watershed sustainability assessments in 1999. China joined this field in 2006 and has since maintained average production rates higher than those of other countries.

4.2 LDA analysis

For the purposes of content analysis, the variable corresponding to the article abstract was considered. Since this field contains unstructured textual information, the LDA technique - described in the methodology section - was applied to identify the main topics discussed across the dataset's abstracts. To ensure the quality of the input for the LDA model, the abstracts were preprocessed following a standard text-cleaning pipeline. Specifically, the following steps were performed:

1. Converted all text to lowercase.
2. Removed punctuation marks.
3. Tokenized the text into individual words.
4. Removed common *stopwords*.
5. Lemmatized the tokens to their base forms.

To represent the textual data in a format suitable for topic modeling, the preprocessed text was converted to a numerical representation using the Bag-of-Words (BoW) model (Harris,1954). This approach transforms each document into a vector that records the frequency of each word in the vocabulary across the corpus. LDA was applied to the corpus, which had been previously

preprocessed and transformed into a Bag-of-Words (BoW) representation, using the Python implementation of Blei et al. (2017).

The LDA model was trained across a range of topic numbers, from 2 to 5. For each model, the coherence score was computed based on the top 10 words in each topic. The optimal number of topics was determined by identifying the first local maximum of the coherence score. The implementation for computing topic coherence follows the approach of the Python library *gensim* (Řehůřek, 2024) and requires several inputs. The first input is the set of topics estimated by LDA. The second input consists of the preprocessed texts, prepared in the same way as for the LDA training. The third input involves assigning a unique identifier to each distinct word in the corpus. The fourth input is the confirmation measure used to evaluate topic coherence.

The coherence measure as a function of the number of topics is shown in Figure 3. The first local maximum was identified at $K = 4$, and this value was subsequently selected as the number of topics for the LDA model.

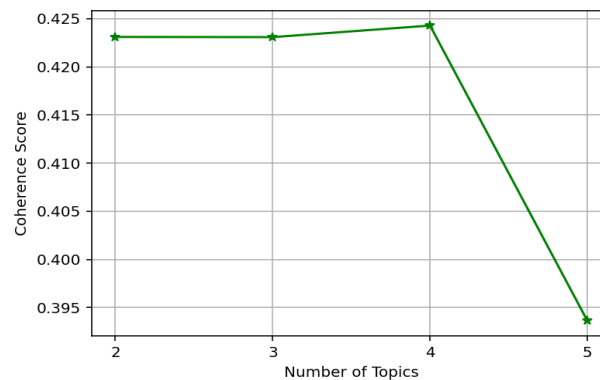


Figure 3. Coherence measure as a function of the number of topics.

In Figure 4, the T-distributed Stochastic Neighbor Embedding (t-SNE) is computed in two dimensions using the scikit-learn implementation (Maaten and Hinton, 2008; Pedregosa et al., 2011). t-SNE is a technique for visualizing high-dimensional data. The idea behind the method is to convert distances between points into joint probabilities and to minimize the Kullback-Leibler divergence between the joint probabilities of the low-dimensional embedding and the high-dimensional data. In this plot, each point represents a document from the corpus. In this two-dimensional space, the distance between the documents reflects their relationship in terms of the topics with the highest probability of occurrence. In this figure, the four topics are clearly differentiated geometrically, allowing us to interpret each one individually. The analysis suggests that the scientific literature on watershed sustainability assessment has evolved into distinct research clusters, reflecting differentiated methodological traditions and disciplinary perspectives across time and countries.

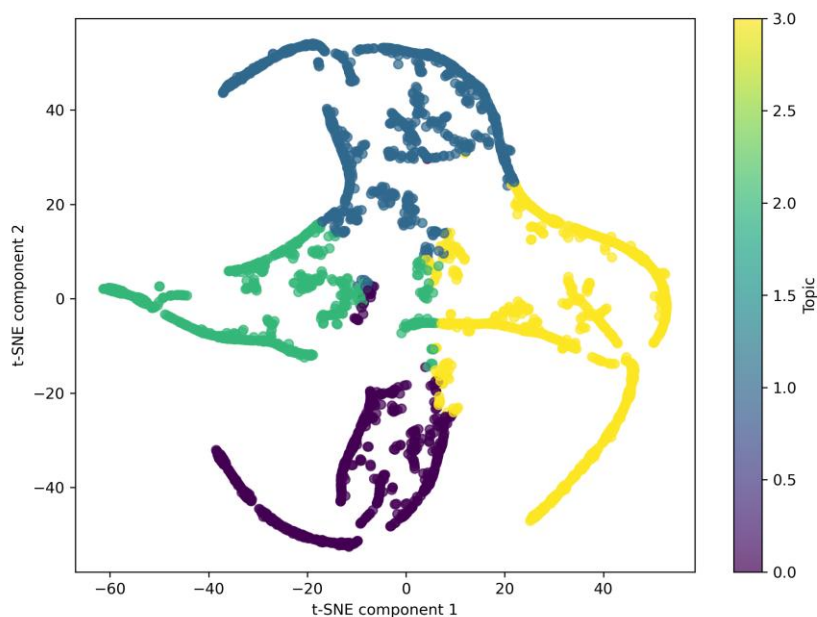


Figure 4. t-SNE representation in a 2D space.

4.3 Topic identification and interpretation

Table 3 presents the top 10 words with the highest probability of belonging to each of the topics identified by the LDA model. Topic 1 primarily concerns the words water, groundwater, and quality, and can therefore be associated with *groundwater*. Topic 2 mainly focuses on the words water, soil, and indicator, which are associated with *hydrological and land-use indicators*. Topic 3 is represented by the word water, followed by drought, vegetation, and climate change, which is associated with topics derived from *climate change*. Finally, Topic 4 is represented by the words area, ecological, territory, space, and development, which are associated with *socioecosystem* topics. Table 4 summarizes the thematic axes along with the topics identified using the LDA technique.

Table 3. Top ten words for each identified topic

Topic	Keywords
1	water, groundwater, quality, index, study, sample, area, irrigation, river, risk
2	water, soil, indicator, resource, study, management, index, system, environmental, basin
3	water, drought, model, index, change, study, climate, data, area, vegetation
4	area, ecological, study, model, land, development, index, factor, spatial, result

Table 4. Summary of thematic axes and corresponding topics extracted with LDA

Topic	Thematic Axes
1	Groundwater
2	Hydrological and Land-use Indicators
3	Climate change
4	Socioecosystem

Figure 5 shows the distribution of the 10 words that define each thematic axis, based on their probability of occurrence in the text. The axes of *groundwater*, *hydrological and land-use indicators*, and *climate change* are primarily associated with the word *water*, whereas the *socioecosystem* axis is more closely linked to the word *area*.

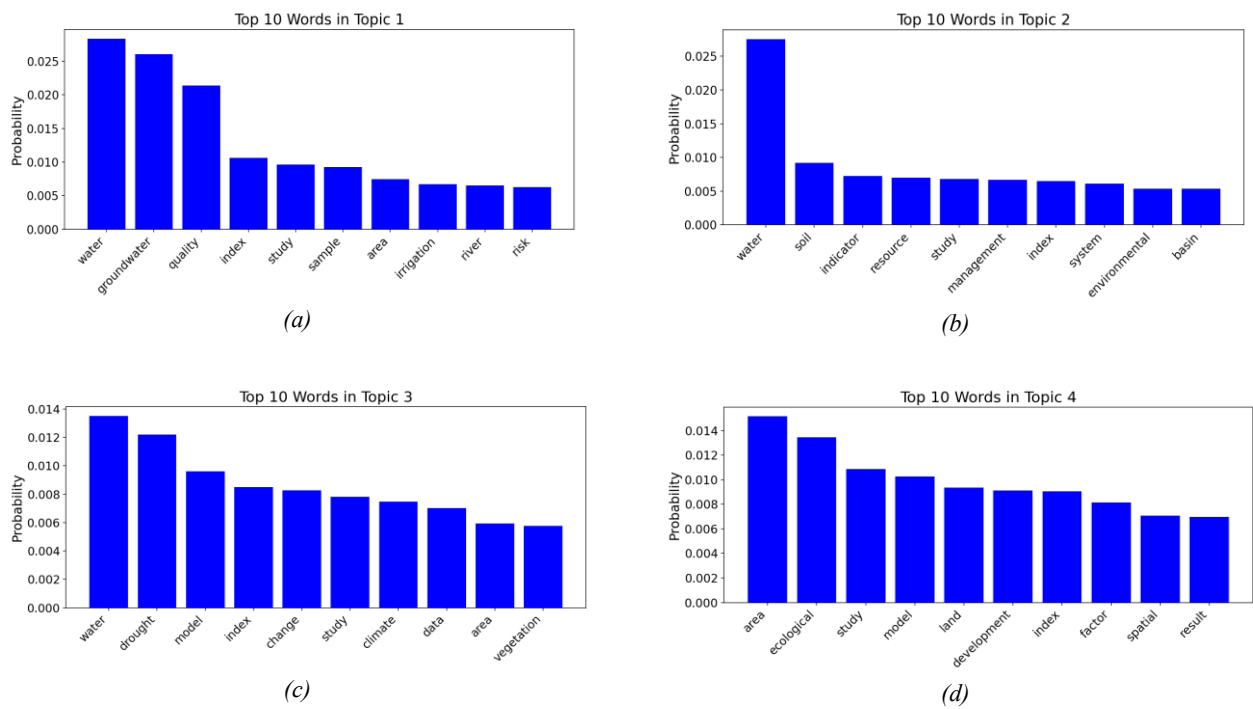


Figure 5. Probability of occurrence for the top ten words within each topic: (a) Topic 1, (b) Topic 2, (c) Topic 3, (d) Topic 4.

As shown in Figure 6, the *climate change* topic accounts for 30% of the analyzed documents, followed by *hydrological* and *land-use indicators* at 28%, and finally *groundwater* and *socioecosystem*, each with over 20%.

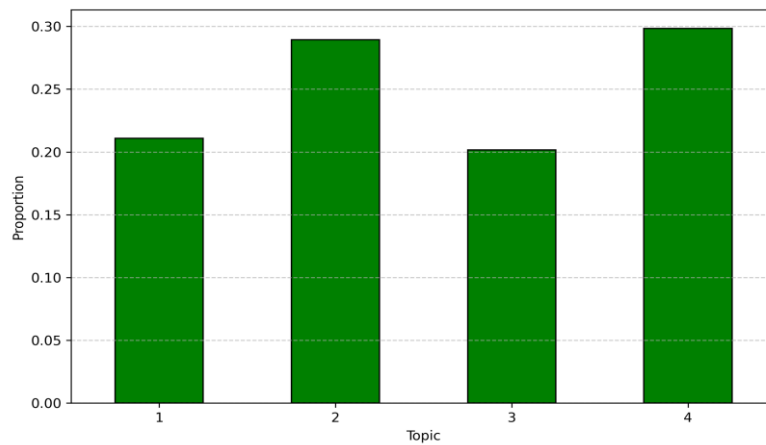


Figure 6. Proportion of documents based on the most representative topic.

Based on the above and as part of the content analysis, the three publications with the highest TopicScore were selected for each topic identified using the LDA technique, resulting in a total of 12 articles, as shown in Tables 4-7.

Table 5. Top articles assigned to Topic 1

Article	Publication Year	Topic	Topic Score
Lawal et al. (2023)	2023	1	0.997118
Geburu et al. (2024)	2024	1	0.996921
Elsayed et al. (2020)	2020	1	0.996914

Table 6. Top articles assigned to Topic 2

Article	Publication Year	Topic	Topic Score
Wabela et al. (2024)	2024	2	0.998012
Vera et al. (2019)	2019	2	0.997311
Valenti et al., (2017)	2017	2	0.997269

Table 7. Top articles assigned to Topic 3

Article	Publication Year	Topic	Topic Score
Kirchner et al. (2020)	2020	3	0.997959
Zhang et al. (2020)	2020	3	0.997574
Liu et al. (2015)	2015	3	0.997482

Table 8. Top articles assigned to Topic 4

Article	Publication Year	Topic	Topic Score
Zhu et al. (2023)	2023	4	0.997323
Wei et al. (2021)	2021	4	0.996978
Yang et al. (2022)	2022	4	0.996889

Table 5 presents the main articles in the *groundwater* thematic axis, where Lawal et al. (2023), from the disciplinary area *Environmental Sciences and Ecology*, has the highest Score Topic. The study focuses on evaluating groundwater quality in northeastern Nigeria, derived from geogenic factors (mineral dissolution, silicate weathering, etc.) and anthropogenic factors (solid waste leachate, wastewater, agricultural runoff). Therefore, they state that to achieve watershed sustainability, groundwater quality monitoring and management strategies with a territorial approach must be considered, along with the integration of hydrochemical and geological data for long-term planning.

Regarding Table 6, which is associated with *hydrological and land-use indicators*, the main study identified is Wabela et al. (2024), corresponding to the disciplinary area of *agriculture*. This study views the watershed as a functional unit for agricultural management and as a productive space vulnerable to water efficiency, and thus evaluates irrigation performance at the local scale.

Table 7 shows the main articles on *climate change*, where Kirchner et al. (2020) corresponds to the disciplinary area *Geology* published in 2020. The study, which focuses on understanding water cycles in aquifers in mountain ecosystems, considers the watershed as a complex hydro-ecological system where physical processes (snowmelt, aquifer recharge) and biological processes (plant transpiration) interact, resulting in daily and seasonal variations. It highlights that the sustainability of the watershed is determined by the dynamic balance between water supply from snow and vegetation demand, considering seasonal and daily climatic effects to preserve the watershed's hydrological and ecological functions.

According to Table 8, the article corresponding to the *socioecosystem* topic is Zhu et al. (2023), published in 2023 in the disciplinary area of *Science and Technology*. The study examines ecological pressures on the lake watersheds of the central Yunnan Plateau, driven by both natural and anthropogenic factors, including urban expansion, recurrent droughts, soil erosion, biodiversity loss, and water pollution. It emphasizes that watershed sustainability requires maintaining ecological carrying capacity—particularly in areas most affected by human activity, such as the plains—through measures such as ecological restoration, land-use regulation, and forest protection. The authors stress the importance of long-term planning strategies that treat the watershed as a key socio-ecological unit, with its spatial evaluation informing territorial planning and safeguarding its ecological functions.

4.4 Content analysis

The three articles from each identified thematic axis are then analyzed to answer the following questions related to the assessment of watershed sustainability: How is the watershed visualized? What does watershed sustainability entail?

Regarding *groundwater*, the watershed is viewed as a complex, agricultural, and vulnerable hydrogeological and productive system, dependent on water, conditioned by natural and human factors, where sustainability involves ensuring the quality of irrigation water through groundwater quality monitoring and management strategies, as well as the responsible use of fertilizers and the selection of water quality-tolerant crops.

According to *hydrological and land-use indicators*, the watershed is viewed as a diverse, productive agricultural and forestry territory vulnerable to water efficiency. This space is used to manage and implement sustainable plantations. Sustainability implies efficiency, the availability and equitable distribution of water resources, the restoration of degraded areas, the utilization of biomass, and territorial planning.

The *climate change* thematic axis views the watershed as a hydro-ecological and productive system, where natural and human processes determine the availability, quality, and balanced use of water. Therefore, watershed sustainability involves balancing water supply and demand, adapting to climatic variability, reducing soil impacts, optimizing agricultural irrigation, and conserving limited resources to maintain productivity.

The *socioecosystem* axis identifies the watershed as a socioecological and environmental unit at the regional scale, where biophysical and social factors interact, and is key to territorial planning and management. Therefore, to be sustainable, it must maintain ecological carrying capacity and reduce vulnerabilities through the system's ecological development and adaptability.

4.5. Thematic axes and global agenda

On the other hand, Figure 7 shows the number of documents, the year of publication of the articles associated with a topic, and their relationship with events on the global agenda related to sustainable watershed management. The year 1992 corresponds to the Earth Summit and the Dublin Conference on Water and Sustainable Development. The year 2000 marked the beginning of the Millennium Development Goals. The RIO+20 Summit was held in 2012, and 2015 marked the publication of the SDGs.

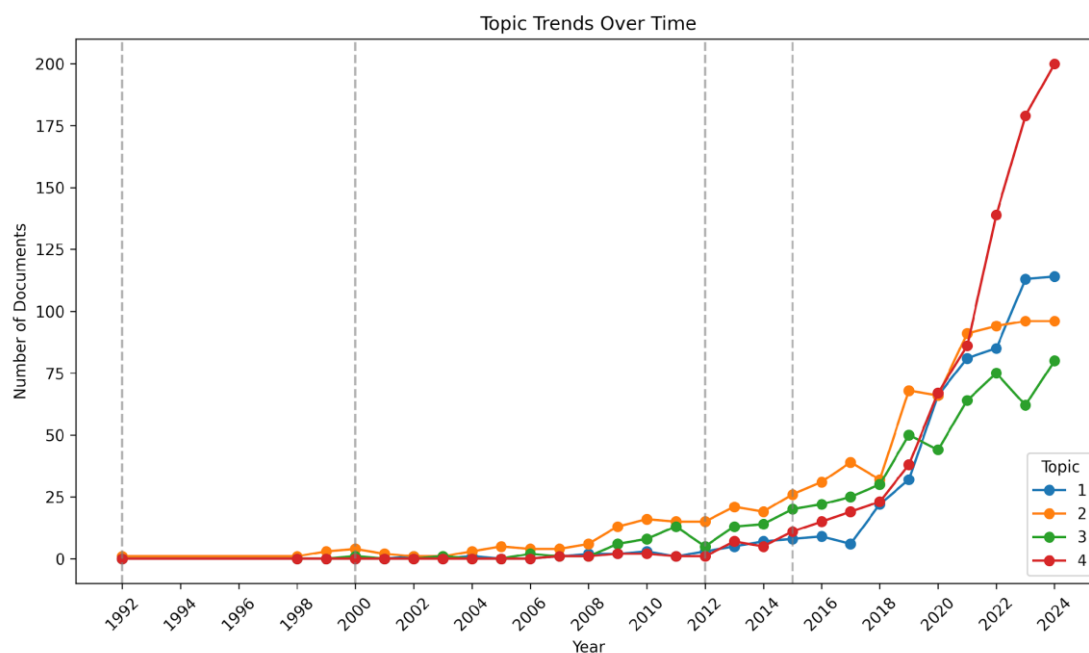


Figure 7. Count of documents as a function of the publication year.

According to the Figure 7, the first article related to the thematic axis of *hydrological and land use indicators* Ripl and Feibicke (1992), which is related to the proposal of a hydrological model. The first article on climate change was published in 2000; the first on *groundwater* in 2002; and the first on *socioecosystems* in 2007.

Regarding the trend in publications and events on the global agenda, it is observed that publications on watershed sustainability assessment began in 1992 at the United Nations Conference on Environment and Development. A significant turning point is noted starting in 2015, with the publication of the SDGs, followed by an increase in scientific studies mainly focused on *hydrological and land-use indicators*. However, since 2020, the trend has changed, with an increase in publications on watershed sustainability assessments associated with the *socioecosystem* axis.

5. DISCUSSION

There are numerous and diverse approaches to studying river basins, given their importance in the hydrological cycle and the sustainable management of a territory's natural resources. Therefore, given the era of Big Data, this study evaluated the existing scientific evidence and gaps in understanding the assessment of river basin sustainability through LDA. This allowed for the identification and analysis of different approaches, as well as their impact on global agendas, using an SLR in a repeatable and transparent manner (Kitchenham et al., 2009; Pati and Lorusso, 2018).

Using the PICO technique, 2,675 scientific articles published since 1992 were identified. This relates to the United Nations Conference on Environment and Development, where the watershed was recognized as the ideal geographic unit for the sustainable management and planning of natural resources, given its importance in the hydrological cycle. Furthermore, the need to monitor its components was highlighted to evaluate the effectiveness and follow up on transitional actions toward the Sustainable Development Declaration (1992).

Likewise, most research areas are related to *Ecology*. Monitoring and the techniques and/or technologies used to measure it are related to the areas of *Environmental Sciences and Ecology*, *Water Resources*, *Science and Technology - Other Topics*, and *Agriculture*. Starting in 2011, academic publications from disciplines other than those initially predominant were incorporated, such as *Public, Environmental and Occupational Health* and, more recently, *Public Administration*, which began to contribute studies in 2021. This evolution makes evident that there is a growing multidisciplinary in the study of the evaluation of the sustainability of watersheds, which responds to the recognition that the watershed is not only a geographical unit in which the natural cycle of ecosystem regeneration develops, but also a space of complex interrelationships between users, interdependent of the physical and biotic system (Perales et al., 2016; Perevochtchikova, 2008; Pérez and Le Blas, 2004).

On the other hand, given the large amount of data and the unstructured textual information they contain, it was decided to apply the Natural Language Processing (NLP) technique known as LDA to identify the main topics discussed in the abstracts of the entire set of articles. For this study, four thematic axes were identified that relate to the predominant terms in the articles. In the assessment of watershed sustainability from 1992 to 2004, the following topics were examined: *groundwater* (topic 1), *hydrological and land-use indicators* (topic 2), *climate change* (topic 3), and *socioecosystem* (topic 4). *Climate change* accounted for 30% of the documents analyzed, *hydrological and land-use indicators* for 28%, and *groundwater* and *socioecology* for just over 20%.

A fundamental part of this study's findings is the calculation of the *Topic Score*, which allows us to identify the most relevant articles by measuring the probability of co-occurrence of words identified with the PICO technique and to discern from a large amount of information using objective, replicable criteria. The latter was used to select the top three articles per topic.

The *hydrological and land-use indicators* axis corresponds to the highest topic score and is related to the words water, soil, indicator, resource, study, management, index, system, environmental, and watershed. From this dataset, only the words indicator, index, and watershed were used for the bibliographic search of the Web of Science. It was found that they are coupled with the words water, soil, resource, study, management, system, and environment, highlighting fundamental concepts in sustainability assessment, such as the management of water and soil resources. In 2015, an increase in publications was observed, coinciding with the launch of the 2030 Agenda.

The article with the highest Score Topic is Wabela et al. (2024), published in 2024, in the agriculture discipline. This article views the watershed as a functional unit for agricultural management and a productive space vulnerable to water efficiency and therefore evaluates irrigation performance at the local level using indicators.

The selection of articles with the highest topic scores for each thematic axis provides answers to the following questions: How is the watershed visualized? What does watershed sustainability entail? This is done to identify the diversity of approaches and gaps in understanding watershed sustainability assessment.

Topic *climate change* views the watershed as a hydro-ecological and productive system, where natural and human processes determine the availability, quality, and balanced use of water. Therefore, watershed sustainability involves balancing water supply and demand, adapting to climatic variability, reducing soil impacts, optimizing agricultural irrigation, and conserving limited resources to maintain productivity. *Socioecological* axis 4 identifies the watershed as a socioecological and environmental unit at the regional scale, where biophysical and social factors interact, and is key to territorial planning and management. Therefore, to be sustainable, it must maintain ecological carrying capacity and reduce vulnerabilities through the system's adaptability and ecological development.

Based on the above and in accordance with Jarzebski et al. (2024), the watershed is conceived as a strategic territorial unit for the sustainable management of water and soil, given the intrinsically linked and interdependent relationships between the physical and biotic systems. Likewise, global agendas have marked a change in the trend of publications on topics 2, 3, and 4, particularly in 2015, related to the start of the 2030 Agenda.

The 2030 Agenda marks a turning point in the assessment of watershed sustainability, as it establishes a global framework for sustainable development that, at its core, includes the SDGs. In particular, SDG 6, dedicated to water and sanitation, includes clearly defined indicators that are assessed annually by each United Nations member country. This framework promotes integrated water resources management at the national level and encourages periodic reporting on progress through monitoring and follow-up. It establishes concrete goals and requires annual reports from countries. This has motivated states to strengthen their actions in the face of the global environmental crisis.

The results obtained allowed us to identify the main themes regarding watershed sustainability assessment. Likewise, research gaps were initially observed, as no terms in the topics reflect the social sphere, despite the growing number of studies and the development of global frameworks for watershed sustainability assessment. However, with the help of the thematic integration resulting from this analysis, it was possible to visualize the publication trend. It was found that, starting in 2020, there has been a considerable increase in studies addressing watershed sustainability assessment from a socioecological conceptual framework. This framework considers the watershed as a socioecological and environmental unit at the regional scale, where biophysical and social factors interact, and is key to territorial planning and management. This definition reflects human action and social structures as an integral part of nature (Gunderson and Holling, 2002; Folke et al., 2010).

Consequently, the application of machine learning techniques, particularly the LDA model, is essential for conducting a rigorous analysis and achieving a more precise interpretation of

documentary content, especially in the context of big data, characterized by continuous growth, heterogeneity, and complexity of information and datasets being generated.

6. CONCLUSION

The Earth Summit in Rio de Janeiro and the Conference on Water and Environment in Dublin in 1992 reflected humanity's need to reach agreements that would allow for sustainable development. They also motivated the generation of scientific articles that study the evaluation of watershed sustainability within the 2030 Agenda. The first two summits defined the watershed as the ideal unit for the management of territorial resources, given that it is a natural geographic divider and container of the hydrological and social cycle. A global framework for sustainable development was subsequently established, including the SDGs at its core, particularly SDG 6, dedicated to water and sanitation, with clearly defined indicators evaluated annually by each member country of the United Nations. This led to an increase in the number of scientific articles evaluating watershed sustainability.

In the analysis of scientific production, the application of the LDA machine learning model enabled an exhaustive examination of a large dataset of 2,675 scientific articles. Key steps included calculating the number of topics, identifying the set of words associated with each topic, and computing the Topic Score. This approach identifies the most relevant articles by measuring the probability of word co-occurrence, as determined using the PICO technique, and provides objective, replicable criteria for discerning information from a large dataset. As such, it serves as a powerful tool for selecting articles before conducting detailed and exhaustive document analysis.

Four topics were identified in the text, which allow us to identify the main thematic axes, which in the case of the evaluation of watershed sustainability in the period from 1992 to 2004 are groundwater, hydrological and land use indicators, climate change and socioecosystem.

This study found that watershed assessment has taken a different direction since 2020, where the approach is based on the conceptualization of the watershed as a socio-ecological unit, in which biophysical and social factors interact, key to territorial planning and management. Therefore, sustainability is associated with reducing the vulnerability of watershed components through the system's adaptability and ecological development. These findings are particularly relevant in the context of the SDGs, as they offer a comprehensive view of the state of scientific knowledge, the gaps in knowledge, and their potential contribution to achieving the global goals by 2030.

STATEMENTS AND DECLARATIONS

Author RMS acknowledges support from Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), Mexico, through scholarship number 1140258. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. Conceptualization, Methodology, Data Curation, Software, Validation, Writing – Original Draft, Writing – Review and Editing by Rebeca Moreno-Santoyo and Andrés García-Medina. Supervision by Concepción Arredondo García and Juana Claudia Leyva Aguilera. All authors have read and approved the final manuscript. The authors have no competing interests to declare that are relevant to the content of this article. The data that support the findings of this study are available from the corresponding author upon request.

REFERENCES

- Barbier, E.B., and Burgess, J.C. (2017). The sustainable development goals and the systems approach to sustainability. *Economics*, 11 (1), 20170028, <https://doi.org/10.5018/economics-ejournal.ja.2017-28>.
- Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). Latent Dirichlet allocation. *Journal of machine Learning research*, 3 (Jan), 993–1022.

- Blei, D.M. (2012). Probabilistic topic models. *Communications of the ACM*, 55 (4), 77–84, <https://doi.org/10.1145/2133806.2133826>.
- Blei, D.M., Kucukelbir, A., and McAuliffe, J.D. (2017). Online LDA vb. Retrieved from <https://github.com/blei-lab/onlineLDAvb> (Accessed: 2025-03-31).
- Buntine, W. and Jakulin, A. (2005). Discrete component analysis. *International statistical and optimization perspectives workshop subspace, latent structure and feature selection* (pp. 1–33).
- Card, N.A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.
- Chien, H., Phan, L.T., and Jou, S.C. (2024). Social science approaches to critical zone studies: a systematic review. *Environment, Development and Sustainability* 1–57, <https://doi.org/10.1007/s10668-024-05673-4>.
- De Lemos, G.L., de Aquino Eslabão, A., dos Santos, J.F., Rodrigues, A.T., da Costa, L.F., de Medeiros Costa, H.K., Fagá, M.T.W. and dos Santos, E.M. (2024). Nature-based solutions experiences: a systematic literature review for public policies. *Nature-Based Solutions*, 5, p.100121. <https://doi.org/10.1016/j.nbsj.2024.10012120>.
- Elder, M., and Grünewald, S. (2025). Insights from Agenda 21 for enhancing the implementation of the SDGs and shaping the post-2030 sustainable development agenda. *Earth System Governance*, 25, 100262. <https://doi.org/10.1016/j.esg.2025.100262>.
- Elsayed, S., Hussein, H., Moghanm, F. S., Khedher, K. M., Eid, E. M., and Gad, M. (2020). Application of Irrigation Water Quality Indices and Multivariate Statistical Techniques for Surface Water Quality Assessments in the Northern Nile Delta, Egypt. *Water*, 12(12), 3300. <https://doi.org/10.3390/w12123300>.
- Folke, C., Carpenter, S.R., Walker, B., Scheffer, M., Chapin, T., and Rockström, J. (2010). Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and society*, 15 (4).
- Folke, C., Biggs, R., Norström, A.V., Reyers, B., and Rockström, J. (2016). Socioecological resilience and biosphere-based sustainability science. *Ecology and Society*, 21 (3).
- Gain, A.K., Giupponi, C., and Wada, Y. (2016). Measuring global water security towards sustainable development goals. *Environmental Research Letters*, 11 (12), 124015, <https://doi.org/10.1088/1748-9326/11/12/124015>.
- Galvao, M. C. B., and Ricarte, I. L. M. (2019). Revisão sistemática da literatura: conceituação, produção e publicação. *Logeion: Filosofia da informação*, 6(1), 57-73.
- Gebru, H., Gebreyohannes, T., and Hagos, E. (2024). Evaluation of groundwater quality for irrigation purposes and impact of irrigation on water in golina river basin, northern Ethiopia. *Momona Ethiopian Journal of Science*, 16 (1), 144–166.
- Glaser, G. (2012). Base sustainable development goals on science. *Nature*, 491 (7422), 35–35, <https://doi.org/10.1038/491035a>.
- Gunderson, L.H., and Holling, C.S. (2002). *Panarchy: Understanding transformations in human and natural systems*. Island Press.
- Harris, Z.S. (1954). Distributional structure. *Word*, 10 (2-3), 146–162, <https://doi.org/10.1080/00437956.1954.11659520>.
- Haynes, R., Sackett, D., Richardson, W., Rosenberg, W., and Langley, G. (1997). Evidence based medicine: How to practice and teach. *Canadian Medical Association Journal*, 157 (6), 788, <https://doi.org/10.1503/cmaj.1070057>.
- Hoffman, M., Bach, F., and Blei, D. (2010). Online learning for Latent Dirichlet Allocation. *Advances in neural information processing systems* (Vol. 23).
- Hoffman, M.D., Blei, D.M., Wang, C., and Paisley, J. (2013). Stochastic variational inference. *The Journal of machine Learning research*, 14 (1), 1303–1347.
- Ighalo, J.O., Adeniyi, A.G., Adeniran, J.A., and Ogunniyi, S. (2021). A systematic literature analysis of the nature and regional distribution of water pollution sources in Nigeria. *Journal of Cleaner Production*, 283, 124566.
- Jarzebski, M.P., Karthe, D., Chapagain, S.K., Setiawati, M.D., Wadumestrige Dona, C.G., Pu, J., and Fukushima, K. (2024). Comparative analysis of water sustainability indices: A systematic review. *Water*, 16 (7), 961, <https://doi.org/10.3390/w16070961>.
- Karamouz, M., Mohammadpour, P., and Mahmoodzadeh, D. (2017). Assessment of sustainability in water supply-demand considering uncertainties. *Water Resources Management*, 31, 3761–3778, <https://doi.org/10.1007/s11269-017-1703-7>.
- Kirchner, J.W., Godsey, S.E., Solomon, M., Osterhuber, R., McConnell, J.R., and Penna, D. (2020). The pulse of a montane ecosystem: coupling between daily cycles in solar flux, snowmelt, transpiration, groundwater, and streamflow at Sagehen creek and independence creek, Sierra Nevada, Usa. *Hydrology and Earth System Sciences*, 24 (11), 5095–5123.
- Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51 (1), 7–15, <https://doi.org/10.1016/j.infsof.2008.09.009>.
- Koudstaal, R., Rijsberman, F. R., and Savenije, H. (1992). Water and sustainable development. In *Natural Resources Forum* (Vol. 16, No. 4, pp. 277-290). Oxford, UK: Blackwell Publishing Ltd. <https://doi.org/10.1111/j.1477-8947.1992.tb00859>.
- Lawal, A., Tijani, M.N., Snow, D., and D'Alessio, M. (2023). Quality and hydrochemical assessment of groundwater in geological transition zones: a case study from NE Nigeria. *Environmental Science and Pollution Research*, 30 (4), 10643–10663.
- Li, N., Lv, T., Wang, X., Meng, X., Xu, J., and Guo, Y. (2024). Research progress and hot topics of distributed photovoltaic: Bibliometric analysis and latent Dirichlet allocation model. *Energy and Buildings*, 115056, <https://doi.org/10.1016/j.enbuild.2024.11505622>.
- Lieneck, C., McLauchlan, M., and Phillips, S. (2023). Healthcare cybersecurity ethical concerns during the covid-19 global pandemic: a rapid review. *Healthcare*, 11 (22), 2983, <https://doi.org/10.3390/healthcare11222983>.
- Linnér, B.-O., and Selin, H. (2013). The United Nations conference on sustainable development: forty years in the making. *Environment and Planning C: Government and Policy*, 31 (6), 971–987.
- Liu, C., Sun, G., McNulty, S.G., and Kang, S. (2015). An improved evapotranspiration model for an apple orchard in northwestern China. *Transactions of the ASABE*, 58 (5), 1253–1264.
- Liu, G., Chen, L., Wang, W., Wang, M., Zhang, Y., Li, J., Lin, C., Xiong, J., Zhu, Q., Liu, Y., Zhu, H., and Shen, Z. (2023). Balancing water quality impacts and cost-effectiveness for sustainable watershed management. *Journal of Hydrology*, 621, Article 129645. <https://doi.org/10.1016/j.jhydrol.2023.129645>.

- Loewe, M. (2012). Post 2015: How to reconcile the millennium development goals (MDGS) and the sustainable development goals (SDGS)? (Briefing Paper No. 18/2012). Bonn: Deutsches Institut für Entwicklungspolitik (DIE).
- Maaten, L. V. D., and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov), 2579-2605.
- Moriarty, P.B., Visscher, J.T., Bury, P., and Postma, L. (2000). The Dublin principles revisited for wss. *Proceedings of the international symposium on water, sanitation and hygiene* (pp. 392–395). Dhaka, Bangladesh: IRC International Water and Sanitation Centre.
- Mosaffaie, J., Jam, A. S., Tabatabaei, M. R., and Kousari, M. R. (2021). Trend assessment of the watershed health based on DPSIR framework. *Land use policy*. <https://doi.org/10.1016/j.landusepol.2020.104911>.
- Mulligan, M., van Soesbergen, A., Hole, D.G., Brooks, T.M., Burke, S., and Hutton, J. (2020). Mapping nature's contribution to SDG 6 and implications for other SDGS at policy relevant scales. *Remote Sensing of Environment*, 239, 111671, <https://doi.org/10.1016/j.rse.2020.111671>.
- Muñoz, R.T. (2021). ¿Qué falta para entender la sustentabilidad? *Estudios sobre las culturas contemporáneas*, 27 (54), 77–117.
- Naeem, S., Duffy, J.E., Zavaleta, E. (2012). The functions of biological diversity in an age of extinction. *science*, 336 (6087), 1401–1406,23.
- Pati, D., and Lorusso, L.N. (2018). How to write a systematic review of literature. *HERD: Health Environments Research & Design Journal*, 11 (1), 15–30, <https://doi.org/10.1177/1937586717747384>.
- Paul, J., Lim, W.M., O' Cass, A., Hao, A.W., and Bresciani, S. (2021). Scientific procedures and rationales for systematic literature reviews (spar-4-slr). *International Journal of Consumer Studies*, 45 (4), O1–O16, <https://doi.org/10.1111/ijcs.12695>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine learning in Python (versión 1.5.1) [Software]. <https://scikit-learn.org/>.
- Perales Miranda, V.H. (2016). La cuenca social como aproximación sociológica a las intervenciones en cuencas hidrográficas. *Temas sociales*(39), 221–240.
- Pereira, M.A.F., Barbieiro, B.L., and Quevedo, D.M.d. (2020). Importance of river basin monitoring and hydrological data availability for the integrated management of water resources. *Sociedade & Natureza*, 32, 292–303.
- Perevochtchikova, M. (2008). Gestión de cuencas hidrográficas: experiencias y desafíos en México y Rusia. *Revista Latinoamericana de Recursos Naturales*, 4 (3), 313–325.
- Pérez, A.A., and Le Blas, F.N. (2004). Lineamientos para la aplicación del enfoque ecosistémico a la gestión integral del recurso hídrico. Programa de las Naciones Unidas para el Medio Ambiente, Oficina Regional.
- Rahaman, M.M., Varis, O., and Kajander, T. (2004). Eu water framework directive vs. integrated water resources management: The seven mismatches. *International Journal of Water Resources Development*, 20 (4), 565–575, <https://doi.org/10.1080/07900620412331319199>.
- Ripl, W., and Feibicke, M. (1992). Nitrogen metabolism in ecosystems—a new approach. *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, 77 (1), 5–27, <https://doi.org/10.1002/iroh.19920770103>.
- Roder, M., Both, A., and Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining* (pp. 399-408).
- Řehůřek, R. (2024). Gensim (Version 4.3.3) [Software]. <https://radimrehurek.com/gensim/>.
- Senent-Aparicio, J., Pérez-Sánchez, J., and Bielsa-Artero, A.M. (2016). Evaluación de la sostenibilidad de cuencas mediterráneas semiáridas. caso de estudio: cuenca Segura, España. *Tecnología y ciencias del agua*, 7 (2), 67–84.
- Turner, K.G., Anderson, S., Gonzales-Chang, M., Costanza, R., Courville, S., Dalgaard, T., Dominati, E., Kubiszewski, I., Ogilvy, S., Porfirio, L. and Ratna, N. (2016). A review of methods, data, and models to assess changes in the value of ecosystem services from land degradation and restoration. *Ecological Modelling*, 319, pp.190-207.
- UN (1992). United Nations. Rio Declaration on Environment and Development. United Nations Conference on Environment and Development, Rio de Janeiro, Brazil. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf.
- UN (2015). United Nations, Transforming our world: The 2030 agenda for sustainable development. New York, NY: Author. Retrieved from <https://sustainabledevelopment.un.org> (Accessed April 17, 2025).
- UN (2025). United Nations Statistics Division, SDG indicators: Metadata repository. <https://unstats.un.org/sdgs/metadata/?Goal=6> (Accessed February 10, 2026).
- Valenti, F., Arcidiacono, C., Chinnici, G., Cascone, G., and Porto, S.M. (2017). Quantification of olive pomace availability for biogas production by using a GIS-based model. *Biofuels, Bioproducts and Biorefining*, 11 (5), 784–797.
- Vera V, R.R., Cota-Sánchez, J.H., and Grijalva Olmedo, J.E. (2019). Biodiversity, dynamics, and impact of chakras on the Ecuadorian amazon. *Journal of plant ecology*, 12 (1), 34–44.
- Vollmer, D., Regan, H.M., and Andelman, S.J. (2016). Assessing the sustainability of freshwater systems: A critical review of composite indicators. *Ambio*, 45, 765–780, <https://doi.org/10.1007/s13280-016-0792-7>.
- Wabela, K., Hammani, A., Taky, A., and Tekleab, S. (2024). On-farm performance evaluation of small-scale irrigation schemes in the Ethiopian rift valley: Internal and external performance process approach. *Irrigation and Drainage*, 73 (4), 1579–1591.
- Wang, D., Liu, N., Chen, F., Wang, Y., and Mao, J. (2024). Progress and prospects of energy storage technology research: Based on multidimensional comparison. *Journal of Energy Storage*, 75, 109710, <https://doi.org/10.1016/j.est.2023.109710>.
- Wei, W., Guo, Z., Shi, P., Zhou, L., Wang, X., Li, Z., Pang, S. and Xie, B. (2021). Spatiotemporal changes of land desertification sensitivity in northwest China from 2000 to 2017. *Journal of Geographical Sciences*, 31(1), pp.46-68.
- Wu, M., Long, R., Chen, F., Chen, H., Bai, Y., Cheng, K., and Huang, H. (2023). Spatiotemporal difference analysis in climate change topics and sentiment orientation: Based on LDA and BILSTM model. *Resources, Conservation and Recycling*, 188, 106697, <https://doi.org/10.1016/j.resconrec.2022.106697>.

- Yang, X., Dai, X., Li, W., Lu, H., Liu, C., Li, N., Yang, Z., He, Y., Li, W., Fu, X. and Ma, L. (2022). Socio-ecological vulnerability in Aba prefecture, Western Sichuan Plateau: Evaluation, driving forces and scenario simulation. *ISPRS International Journal of Geo-Information*, 11(10), p.524.
- Zhang, H., Wang, B., Li Liu, D., Zhang, M., Leslie, L.M., and Yu, Q. (2020). Using an improved swat model to simulate hydrological responses to land use change: A case study of a catchment in tropical Australia. *Journal of Hydrology*, 585,124822.
- Zhu, H., Cheng, F., Wang, J., Jiao, Y., Zhou, J., Sha, J., Liu, F. and Nong, L., (2023). Variation in the Ecological Carrying Capacity and Its Driving Factors of the Five Lake Basins in Central Yunnan Plateau, China. *Sustainability*, 15(19), 14442.