Separation and Purification Technology 364 (2025) 132565

Contents lists available at ScienceDirect

Separation and Purification Technology

journal homepage: www.elsevier.com/locate/seppur

A metadata survey of photothermal membranes for solar-driven membrane distillation

Ersin Aytaç [a,b], Farah Ejaz Ahmed [c], Faissal Aziz [d,e], Mohamed Khayet [a,f,*], Nidal Hilal [c,*]

[a] Department of Structure of Matter, Thermal Physics and Electronics, Faculty of Physics, University Complutense of Madrid, Avda. Complutense s/n, 28040 Madrid, Spain

[b] Department of Environmental Engineering, Zonguldak Bülent Ecevit University, 67100 Zonguldak, Turkey

[c] NYUAD Water Research Center, New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, United Arab Emirates

[d] National Center for Research and Studies on Water and Energy (CNEREE), Cadi Ayyad University, B. 511, 40000 Marrakech, Morocco

[e] Cadi Ayyad University, Faculty of Science Semlalia of Marrakech, Laboratory of Water Sciences, Microbial Biotechnologies and Natural Resources Sustainability, BP 2390, Marrakech 40000, Morocco

[f] Madrid Institute for Advanced Studies of Water (IMDEA Water Institute), Avda. Punto Com No 2, 28805 Alcalá de Henares, Madrid, Spain

ARTICLE INFO

Editor: Cintia Marangoni

Keywords:
Photothermal material
Solar energy
Bibliometric analysis
Machine learning
Natural language processing
Zero-shot classification

ABSTRACT

With the rapid advancement of effective photothermal materials, photothermal membrane distillation (P-MD) has emerged as a new technology for water treatment using solar energy. This technology captures large amounts of solar energy and localizes heating on the membrane-feed water interface using photothermal effects. This research analyses P-MD membrane engineering articles from multiple standpoints using machine learning (ML), data mining (DM), and bibliometrics. The results show the outlines (number of authors, annual growth rate, average citations, etc.) and details (influential authors, countries, articles, social networks, metrics, sentiment, emotion, objectivity, similarity, etc.) of 102 articles published between 2017–2024. We find that 385 authors contribute to the P-MD membrane engineering field, 44 journals are identified, and the average global citation per article is 41.66. Most articles (23) were published in 2022. Based on different metrics, Dr. Li Q. is the leader in the field. 5 core clusters were identified in the country analysis. Authors have a positive sentiment about P-MD membrane engineering, and their writing language is objective and emotion-free (neutral).

1. Introduction

Membrane distillation (MD) is a non-isothermal separation process that involves the separation of volatile compounds from aqueous solutions through a vapor pressure gradient across the membrane. While often associated with desalination, MD is a versatile technique that also finds application in various other areas, including wastewater treatment for removal of contaminants and treatment of landfill leachate, the concentration of fruit juices and dairy products in the food and beverage industry, purification, and concentration for biotechnological and chemical processes [1].

In MD, a temperature difference across a porous and typically hydrophobic membrane causes a vapor pressure difference and induces more volatile compounds to vaporize on the hot feed side, be transported through the membrane pores and condense on the cooler permeate side. There are four main MD configurations, differing in how the permeate is condensed and collected: direct contact membrane distillation (DCMD), air gap membrane distillation (AGMD), vacuum membrane distillation (VMD) and sweeping gas membrane distillation (SGMD) [2]. MD offers high rejection rates and operates at moderately low operating temperatures of 40 – 80 °C [3,4]. In general, MD is a low-pressure process and, as such, is less prone to fouling than pressure-driven membrane processes such as reverse osmosis (RO) and is also suitable for brine treatment as it can handle hypersaline feed solutions. In contrast to other thermal desalination methods, MD functions with a compact, modular system design with a smaller land footprint [5].

Although it was first patented in the 1960 s, MD still shies from full-scale commercial use for desalination due to high energy consumption,

* Corresponding authors at: Department of Structure of Matter, Thermal Physics and Electronics, Faculty of Physics, University Complutense of Madrid, Avda. Complutense s/n, 28040 Madrid, Spain (M. Khayet).

E-mail addresses: khayetm@fis.ucm.es (M. Khayet), nhh2@nyu.edu (N. Hilal).

https://doi.org/10.1016/j.seppur.2025.132565
Received 29 November 2024; Received in revised form 7 March 2025; Accepted 15 March 2025
Available online 17 March 2025

low energy efficiency and lack of suitable membranes with stable long-term performance [6]. One of the main contributors to energy consumption in MD is the thermal energy needed to heat the circulating feed. The bulk feed entering the module is heated, but due to the latent heat of vaporization, the membrane surface temperature on the feed side becomes lower than that of the bulk feed. Similarly, the membrane surface temperature on the permeate side is higher than that of the bulk permeate. Termed 'temperature polarization', this phenomenon reduces the effective driving force across the membranes and consequently lowers vapor transport [7]. While MD research in the 21st century has focused heavily on developing non-wetting, high-performing membranes, and improving the energy efficiency of MD systems through novel configurations, attention to alternative heating techniques, specifically direct heating of the membrane surface, has only recently garnered interest [8]. Such localized heating can reduce temperature polarization, improve mass transport, and enable greater recovery rates [9]. Direct surface heating techniques investigated in the past decade include photothermal, microwave [10], induction [11] and electrothermal [12–14] mechanisms. Of these, the incorporation of photothermal materials to induce localized photothermal surface heating of MD membranes has received considerable attention, necessitating a thorough survey of existing scientific literature that integrates bibliometrics, data mining and machine learning.

Photons interacting with bulk materials cause the excitation of electrons, and the subsequent release of energy absorbed by these electrons is dissipated through conversion processes using various activation modes, which include photoconductive, photoelectrochemical, photovoltaic, fluorescence and photothermal [15]. In this sense, photothermal materials primarily used in MD (i.e. photothermal membrane distillation, P-MD) include solar-absorbing inorganic semiconductors, plasmonic metallic nanomaterials and carbon-based materials [8]. Metallic and semiconductor materials operate under surface plasmon resonance (SPR), which indicates the wavelength at which collective electron oscillations are excited on the material's surface, transforming incoming light to heat [16]. Carbon-based materials such as carbon black nanoparticles [17], carbon nanotubes (CNTs) and activated carbon [18] demonstrate photothermal properties across a broad spectrum of light wavelengths with high conversion efficiencies.

The challenge of incorporating photothermal materials in MD membranes is to optimize light absorption and photothermal conversion efficiency without compromising membrane characteristics. In the case of MD, the membrane should be porous with a mean pore size between 0.1 and 1 μm [19]. MD membranes should either be hydrophobic or, if made of as a composite material, at least one layer should be hydrophobic to prevent wetting of the membrane pores and subsequent deterioration of permeate quality. Photothermal membranes should not only be cost-effective and easily scalable, but they should also absorb light at a broad range of wavelengths (250–2500 μm) [20]. Typically, photothermal materials are simply coated onto MD membranes through spraying, coating, and filtration techniques. However, considerable work still needs to be put into the long-term stability of these coatings because of the delamination effect. Additionally, the wettability and porosity of the coating layer may impact MD performance. Photothermal materials can also be incorporated into the bulk membrane by blending prior to phase separation for example, although the limited irradiation received might decrease energy efficiency. More work should be directed to develop compatible materials and incorporate photothermal materials to maximize the effects of integrating these materials and retain high salt rejection. An analysis of current literature will help determine research trends and directions of future research studies.

The fundamentals of P-MD technology are unique. P-MD is a relatively new technique within MD in which light-absorbing photothermal materials are incorporated within the membrane to convert light energy into heat in order to induce evaporation at the feed/membrane pore interface and reverse the temperature polarization effect at the feed membrane side [21]. The temperature profile is distinct from that of conventional MD as seen in Fig. 1. In P-MD, the feed temperature at the surface of the membrane $T_{f,m}$ is greater than that of the bulk feed temperature $T_{f,b}$ due to localized membrane surface heating, whereas in conventional MD heat losses through the membrane and latent heat cause $T_{f,m}$ to be lower than $T_{f,b}$. The profile of the permeate temperature (i.e., $T_{p,m}$ and $T_{p,b}$, which are surface and bulk temperatures on the permeate side, respectively) is maintained the same in both P-MD and conventional MD.

Since Otlet introduced in 1934 in his book "Trait'e de Documentation" [22], which was later employed by Alan Pritchard in 1969 in his article entitled "Statistical bibliography or bibliometrics" [23], bibliometric method (or the science of measuring and assessing academic literature) has grown considerably [24]. This method is used to study scientific publications on a certain topic and evaluates patterns, relationships, frequencies, and trends via statistical and mathematical approaches [25]. Bibliometrics has been applied in a variety of fields, such as patent analysis [26], pedagogy [27], European resuscitation congresses [28], olfactory marker protein [29], etc.

Data Mining (DM) is the technique of retrieving meaningful patterns, information, and knowledge from big data sets [30]. Because of its unique capabilities, DM has gained substantial popularity in the discovery of the inherent correlation criteria inside datasets [31]. Using data mining to process data and present the results to the audience in a visual format is a unique way to provide audiences with useful information that can be fully utilized to improve the domain of interest [32]. DM technique often comprises data processing, data conversion, data analysis, and other tasks, with a particular correlation between each to produce a full evaluation [33].

Text Mining (TM), also known as Knowledge Discovery from Text (KDT), Intelligent Text Analysis (ITA) and Text Data Mining (TDM), aims to extract and analyze important insights or patterns from irregular and unstructured textual data automatically and effectively. TM intersects various fields, including Statistics, Data Mining, Information Retrieval, Knowledge Discovery, and Natural Language Processing (NLP) [34–36]. Text mining approaches may successfully capture semantic characteristics, gather necessary feature information, and improve unstructured data extraction capabilities [37].

The sub-discipline of Artificial Intelligence (AI), Natural Language Processing, began in the 1940s by creating software models of language recognition phrases. NLP algorithms enable computers to understand, modify, and comprehend human language and sentences and to analyze textual data using sentence modeling. NLP evaluates words extracted from qualitative records to create quantitative data. In short, NLP algorithms convert text data into small fundamental parts known as tokens, which include words and punctuation, then attempt to extract meaning by analyzing the connections between tokens [38–40].

Machine learning (ML) is the NextGen of computer science and programming that depends on training algorithms to learn from data [41,42]. ML has found a place on a broad scale in numerous application areas, including healthcare, speech analytics, transportation, export forecasting, market analysis, land cover modeling, life sciences, and others [43–45]. Basically, ML depends on an algorithm that autonomously extracts general patterns inside the data, which can further be used to make predictions about new inputs (data) [46,47]. It has three major categories: supervised, unsupervised, and reinforcement learning. The most common kind of ML is supervised learning, in which the model is built using labeled data that match to predict the label of the new inputs. Using unlabeled data, unsupervised learning trains models to identify patterns and clusters in an input dataset. The reinforcement learning algorithm learns behavior through trial and error, starting with simply input data and the goal of optimizing the cumulative reward [48]. ML developments have led to the addition of new or hybrid approaches to these three main categories. One of them is weakly supervised learning. Instead of gathering expensive clean annotations, weakly supervised learning employs weak labels from multiple weak labeling sources such as heuristic rules, knowledge bases, or quality lower-

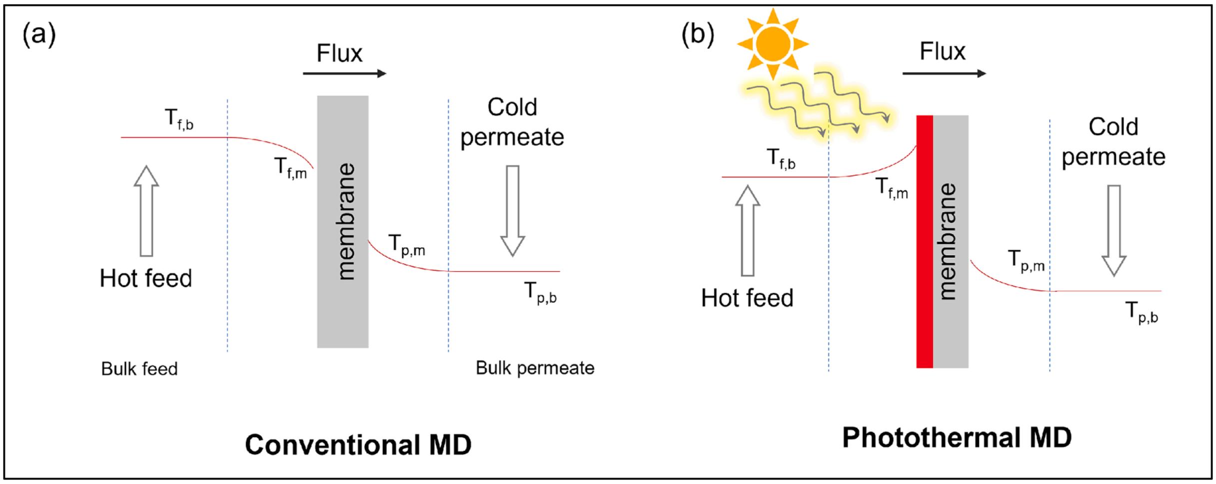

Fig. 1. Temperature profiles in **(a)** conventional and **(b)** P-MD.

quality crowdsourcing. These fragile labels are cheap to acquire but are frequently noisy and carry biases from their origins [49]. Semi-supervised learning is another ML model. The semi-supervised learning goal is to train an algorithm on unlabeled data by creating a pretext goal that trains an unsupervised learning model to get a general or task-oriented data model and then fine-tune it for subsequent actions [50]. Another innovative approach is self-supervised learning. The self-supervised learning approach seeks to detect supervised indicators in data automatically, eliminating the need for manual labeling. Contrastive and non-contrastive self-supervised learning methods are the two classes into which it is divided [51,52]. The combination of NLP and ML can expand text-based data processing analysis. Sentiment analysis, summarization, text classification, keyword extraction, named entity recognition, topic modeling, text clustering, and lemmatization are the processes that NLP and/or ML can perform on textual data [40]. Zero-shot models of unsupervised learning method, such as zero-shot classification of text/image data, may perform a variety of actions without the requirement for labeled data. Zero-shot classification is a text analysis technique that labels a textual data point without any prior training [53–55].

This work applies bibliometric methods, data analysis tools, and machine learning approaches to provide an overview of research in P-MD. As such, we aim to assess literature in this area, identify sub-areas of research and assess the output of research groups, universities, and countries globally. We believe this study will facilitate recognition of emerging research areas within P-MD, including membrane materials, fabrication techniques and their effect on both photothermal and MD characteristics, as well as contribute to future research plans. This study strives to bridge research gaps by finding and assessing current research on elements impacting P-MD membrane engineering to give a comprehensive knowledge of the fundamental driving factors. To this end, from a wide range of analysis methods (i.e. data mining, machine learning, bibliometric and manual) using a combination of different tools (i.e. R-Biblioshiny (in Bibliometrix), VOSviewer, Python and Orange Data Mining Tool), we try to respond to some questions frequently raised by researchers:

- Is P-MD membrane engineering a trending topic?
- What are the basic statistics of this domain?
- Which journals, countries, authors, or articles have a great impact?
- What is the status of collaboration in P-MD membrane engineering?
- What are the latest trending topics in P-MD?
- What are the dominant words used by researchers in the P-MD field?
- Are researchers optimistic about P-MD membranes?

2. Data and methods

The dataset used in the study was collected on 24th January 2025. The database chosen for collecting the dataset was Scopus because this database has a broader coverage of articles than other databases, has a sophisticated search feature, a list of indexed publishers, and premier resources for scholarly research in the sciences [56,57]. As of January 2025, Scopus contains more than 7 K publishers, more than 28.3 K active serial titles, more than 368 K books, 24.6 M open access items, more than 97.3 M records, more than 19 M authors, more than 2.4B cited references dating back to 1970, more than 2.33 M preprints and more than 94 thousand affiliation profiles [58]. The keywords used to gather the raw dataset were “membrane distillation – MD – direct contact membrane distillation – DCMD – air gap membrane distillation – AGMD – sweeping gas membrane distillation – sweep gas membrane distillation – SGMD – vacuum membrane distillation – vacuum enhanced membrane distillation – VMD – thermostatic sweeping gas membrane distillation – TSGMD – conductive gap membrane distillation – CGMD – material gap membrane distillation – MGMD – permeate gap membrane distillation – PGMD – liquid gap membrane distillation – LGMD and water gap membrane distillation – WGMD”. For yearly comparison and ML/NLP analysis, other criteria to clear the dataset can be seen in Table 1.

This raw dataset was screened with the following keywords to find the membrane-engineered MD articles “engineering – engineered – fabrication – fabricated – preparation – prepared – synthesis – tailoring – tailored – modifying – modified”. A further search was conducted within the membrane-engineered MD articles to find out the photothermal membrane engineering MD articles with the following keywords: “photothermal – self-heating – P-MD – solar – localized heating – thermoplasmonic – plasmonic – nanophotonic – photonic”. Finally, a manual screening process was conducted to remove irrelevant papers. The final dataset covers 102 articles concerning P-MD membrane engineering.

To depict bibliometric associations, two scientific mapping applications were used: VOSviewer (version 1.6.20) and Biblioshiny (version 4.1.4). Biblioshiny (part of the Bibliometrix package) is an innovative

Table 1
Search filters.

Criteria	Description
Year	***Exclude:*** 2025
Document Type	***Limit to:*** Article
Publication Stage	***Limit to:*** Final
Source Type	***Limit to:*** Journal
Language	***Limit to:*** English

open-source software for undertaking extensive scientific mapping analysis [59]. VOSviewer is a publicly accessible tool for creating and viewing bibliometric maps that focuses on their graphical depiction [60]. The following are the fundamental calculations made in bibliometric analyses:

Compound Annual Growth Rate ($CAnGR_C(\%)$) value of the collection is expressed as in Eq. (1) [61],

$$CAnGR_C(\%) = \left(\left(\frac{nd_{C_l}}{nd_{C_f}}\right)^{\frac{1}{\left(y_{C_l} - y_{C_f}\right)}} - 1\right) \times 100 \tag{1}$$

where, n_i is the total number of articles in the first year, n_f is the total number of articles in the last year, Y_i is the collection's last year and Y_f is the collection's last year.

In Eq. (1), the numbers of articles from the first year of the collection are given by nd_{C_f}, the number of articles from the final year of the collection by nd_{C_l}, the first year of the dataset by y_{C_f}, and the last year of the dataset by y_{C_l}.

The collection's co-authors per document ($cApD_C$) value is expressed as in Eq. (2) [62],

$$cApD_C = \frac{na_C}{nd_C} \tag{2}$$

where nd_C is the total number of documents in the collection and na_C is the total number of authors (author's appearances) in the collection (recurring author names are taken into account).

Furthermore, the documents average age value of the dataset ($DAvA_C$) or an item (author, country, or journal) ($DAvA_{It}$) (indicated as average publication year in VOSviewer software) is defined as [63];

$$DAA_C = \frac{1}{nd_C}\sum_{C_i=1}^{C_n} a_{C_i} \text{or } DAA_{It} = \frac{1}{nd_{It}}\sum_{It_i=1}^{It_n} a_{It_i} \tag{3}$$

where x_i indicates the age of the document i, which can be determined via $(Y_f - Y_x)$ and Y_x is the publication year of document i.

In Eq. (3), a_{C_i} is the age of document i in the data stack that can be calculated by $(y_C - y_i)$ being y_C is the year that the dataset downloaded and y_i is the document i's publication year. a_{It_i} is the age of a document of the item, which can be found $(y_C - y_{It_i})$, being y_{It_i} is the publication year of the document of the item, and nd_{It} is the total number of documents of the item.

Global citations (GC) (sometimes referred to Total Citations (TC) in Biblioshiny) indicates the total number of times an item (author, paper, or journal) has been cited in a bibliographic database [64,65]. Local citations (LC) measure the number of citations of other authors for a certain article inside the gathered dataset [66].

The average global citations per document value of an item (author, country, or journal) ($AvGCpD_{It}$) (depicted as average citations per doc in Biblioshiny package and average citations in VOSviewer) or the dataset ($AvGCpD_C$) can be expressed as in Eq. (4) [67]:

$$AvGCpD_C = \frac{1}{nd_C}\sum_{i=1}^{nd_C} GC_i \text{or } AvGCpD_{It} = \frac{1}{nd_{It}}\sum_{It_i=1}^{It_n} GC_{It_i} \tag{4}$$

being GC_i the global citations of document i in the collection and GC_{It_i} is the global citations of document i of an item.

The average global citations per document published in a year y (in the corresponding year) metric ($AvGCpD_{C_y}$) (depicted as mean total citations per article, MeanTCperArt, in Biblioshiny) of the data stack is calculated as in Eq. (5) [63]:

$$AvGCpD_{C_y} = \frac{1}{nd_{C_y}}\sum_{i_y=1}^{nd_{C_y}} GC_{i_y} \tag{5}$$

where nd_{C_y} is the number of articles in the dataset in a year y and GC_{i_y} is the global citation of a document i in the dataset published in the same year.

The $AvNGCpD_{C_y}$ value (depicted as mean TC per year, MeanTCperYear, in Biblioshiny), which is the average normalized global citations per document published in the corresponding year is calculated as is calculated as in Eq. (6) [68]:

$$AvNGCpD_{C_y} = \frac{AvGCpD_{C_y}}{a_{C_y}} \tag{6}$$

where a_{C_y} represents the age of any document utilized in the $AvGCpD_{C_y}$ computation.

The global citations per year value of a document ($GCpY_i$) (depicted as times cited per year in Biblioshiny tool) value can be calculated as in Eq. (7) [69,70]:

$$GCpY_i = \frac{GC_i}{a_{C_i}} \tag{7}$$

where GC_i indicates the global citations of the document i.

RC_{i_y}, the relative global citations (depicted as normalized TC in Biblioshiny tool) of document i published in corresponding year (in year y) can be computed as in Eq. (8) [71]:

$$RGC_{i_y} = \frac{GC_{i_y}}{\sum_{i_y=1}^{nd_{C_y}} GC_{i_y}/nd_{C_y}} \tag{8}$$

being, GC_{i_y} is the total number of global citations of document i published in year y.

$AvRGC_{It}$ is the average relative global citations value (depicted as average normalized citations in VOSviewer) of an item can be calculated as [72]:

$$AvRGC_{It} = \frac{1}{nd_{It}}\sum_{i=1}^{nd_{It}} RGC_{It_{i_y}} \tag{9}$$

where $RGC_{It_{i_y}}$ denoted the relative global citations of document i published in the corresponding year of an item.

The international co-authorship ($IcoA$) ratio of the gathered dataset is determined as in Eq. (9) [73],

$$IcoA_C(\%) = \frac{MuCP}{nd_C} x100 \tag{10}$$

where $MuCP$ is the multiple country publications in the collection.

The articles fractionalized (AFr) value of an author can be calculated as in Eq. (11) [64],

$$AFr = \sum_{i=1}^{m} \frac{1}{na_{a_i}} \tag{11}$$

where na_{a_i} indicates the number of co-authors in the corresponding documents and m indicates the number of documents of an author.

Lotka's law, a reliable indicator of scientific productivity, explains the correlation between an author's number of published publications. According to this rule, the fraction of researchers having z publications correlates to the quotient $1/z^2$. Lotka's Law of scientific productivity (frequency) of authors (Y) can be calculated as in Eq. (12) [64],

$$Y = \frac{c}{n^z} \tag{12}$$

where c and z are the constants calculated for the collection.

The Hirsch index (h-index) was created to quantify both the quality and quantity of research and is described as the number of documents that have been referenced at least h times [74]. The h-index tends to rise with the career duration of an author/journal/country. Comparing items' within the same area but with very different career lengths creates a problem that can be corrected by *m*-index (also called *m*-quotient). The *m*-index is determined by dividing an items' *h*-index by the total number of years from the items' first publication, as shown in Eq. (13) [75].

$$m - index(quotient) = \frac{h - index}{y_{It_l} - y_{It_f}} \tag{13}$$

where y_{It_f} is the initial year of publication and y_{It_l} is the final year of publication of the author. But the Biblioshiny package orients y_C rather than y_{It_l} when calculating an item's *m*-index.

The *g*-index is described as the maximum number of publications that have achieved at least g^2 citations combined or equivalently at least *g* citations on average [76].

For the DM, TM, ML and NLP approaches, Orange Data Mining Tool and Python were employed. Word count, word cloud and similarity analyses (word similarity, not semantic similarity) of titles/abstracts were employed using Orange Data Mining tool. Orange is an open-source ML and DM package allowing data analysis using Python scripting and visual programming Janez Demšar and Blaž Zupan [77,78]. Before applying word cloud and similarity analyzes a preprocessing step was conducted. This preprocessing step includes transforming (converting to lowercase, removing URLs and accents and parsing html), tokenizing (using regular expressions), normalizing (using lemmatization) and filtering (removing stop words, lexicon, numbers, and some regular expressions) the corpus. In the similarity analysis, after preprocessing the corpus, bag-of-words (BoW) (counting the words of data instances to create a word frequency vector) and distances (calculating the distances between data instances) steps were employed, respectively. BoW method was based on term frequency. In this case, a text is represented as a vector, where a word's frequency is 1 (or equal to the number of the word frequency) if it appears in all words in the corpus and 0 otherwise [79]. Distances were determined first by calculating cosine similarity (CS). CS involves calculating the cosine value of an angle between two vectors to determine their similarity. Cosine similarity ranges between −1 and 1. The CS value equals −1 when two vectors point in opposing directions and 1 when they point in the same direction. The CS value between vector $A^T = \{x_1, x_2, \cdots\cdots, x_n\}$ and vector $B^T = \{y_1, y_2, \cdots\cdots, y_n\}$ can be expressed as in Eq. (14) [80].

$$CS(A,B) = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} {x_i}^2 \sum_{i=1}^{n} {y_i}^2} \tag{14}$$

Afterwards, the distances among each data point were determined using the cosine distance (CD), which is equal to 1 – *CS* [80]. This returns a number between 0 and 2, representing similarity and dissimilarity, respectively.

The sentiment, objectivity and emotion analyzes of abstracts were conducted using Python with Facebook's "bart-large-mnli" (zero-shot classification task), Moritz Laurer's "DeBERTa-v3-base-mnli-fever-anli" (zero-shot classification task) and J. Hartmann's "emotion-english-distilroberta-base" (text classification task) models respectively, from Hugging Face Community (https://huggingface.co/). bart-large-mnli is an NLP model from Facebook that is the checkpoint for bart-large after training on the Multi-Genre Natural Language Inference (MultiNLI) dataset. The MultiNLI collection is a crowd-sourced collection of 433 K sentence pairs tagged with textual entailment information [81,82]. This model classifies a text into 3 sentiments (positive, negative, and neutral). DeBERTa-v3-base-mnli-fever-anli model was also trained using the MultiNLI, Fever-NLI, and Adversarial-NLI (ANLI) datasets, which included 763–913 NLI hypothesis-premise pairings. DeBERTa-v3-base-mnli-fever-anli approach segments a text into 2 categories (objective and subjective) [83]. Emotion-english-distilroberta-base allows users to categorize emotions in English text data. The model represents a fine-tuned DistilRoBERTa-based checkpoint and is trained on six different datasets. It predicts Ekman's six primary emotions (surprise, sadness, joy, fear, disgust, and anger), plus a neutral class [84]. In these analyses, the sum of all scores is equal to one, and the higher the score, the better it matches the mood of the expression in a given text. The flowchart summarizing the methodology can be seen in Fig. 2.

3. Results and Discussions

3.1. Main information on the P-MD membrane engineering collection

The outputs from our analyses provide a multidimensional outcome of P-MD membrane engineering. However, before diving deeper into these complex results, a grasp of the basic statistics provides a quick overview of the overall state of this domain. Fig. 3 indicates the key features of the collection.

In the collection, a total of 102 articles from 44 different journals were published between 2017 and 2024. This represents a compound annual growth rate of 26.73 % and shows that research activities in this area are continuously increasing. The first publications about P-MD membrane engineering appeared in 2017; "*Thermoplasmonic membrane distillation*" by Politano et al. [85], "*Photothermal nanocomposite membranes for direct solar membrane distillation*" by Wu et al. [17], "*Solar desalination of seawater using double-dye-modified PTFE membrane*" by Fujimara and Kikuchi [86] and "*Photothermal* membrane *distillation for seawater desalination*" by Politano et al. [87]. On average, there were 41.66 global citations per document. This shows that other studies widely cited the studies in the domain. In total there are 385 different authors, with an average of 5.93 authors per document showing that research in P-MD membrane engineering is often conducted in collaboration. In addition, international cooperation is 20.59 %, which shows that research is carried out in a global framework. Fig. 4 indicates the time series of article publishing, the average citations per article published in the corresponding year metrics and the normalized average citations per article published in the corresponding year.

As can be seen in Fig. 4, the number of articles published annually has generally increased. In 2023 there is a slight decline, but in 2024 the situation seems to be recovered. From 2017 to 2024, the $AvGCpD_{C_y}$ and $AvNGCpD_{C_y}$ values have decreased. Obviously, the articles published in 2017 have more impact on the domain, but the articles in recent years are less effective. This situation arouses the idea that scientists producing P-MD membranes should do more groundbreaking and sound studies.

Fig. 5 shows the subject of the dataset based on Scopus segmentation. Although the main subjects of P-MD membrane engineering seem to be chemistry, chemical engineering, and materials science in Fig. 5, the fact that the field also touches scientific branches such as biochemistry, genetics, molecular biology, energy, physics, and astronomy is an important indicator of how wide a framework it can reach. Fig. 6 depicts the scatter matrix (distributions and correlations) of the number of pages, number of references, number of global citations and number of authors values of the articles.

In Fig. 6 the bottom triangle displays the scatter plot distribution of element pairs, the upper triangle shows the correlation coefficients between element pairs, and the diagonal cells between these two display the histogram distribution of each element. First, we can consider the distribution of each variable with histograms. On average, a P-MD membrane engineering article contains 11 pages. The shortest article has 5, while the longest article has 21 pages. About 53 references were used on average in the P-MD articles (maximum 98, minimum 12, mode 42). While the number of global citations ranges widely between 1 and 419 (~31 on average). This indicates that most articles have low citations. A given P-MD membrane engineering article was written by at least 2 and no more than 16 authors. We can gain a better understanding of the relationship between pairs of variables by interpreting scatter plots (i.e., lower triangle) and correlation coefficient values (i.e., upper triangle) together. The correlation coefficient ranges from −1 to 1 (i.e., 1

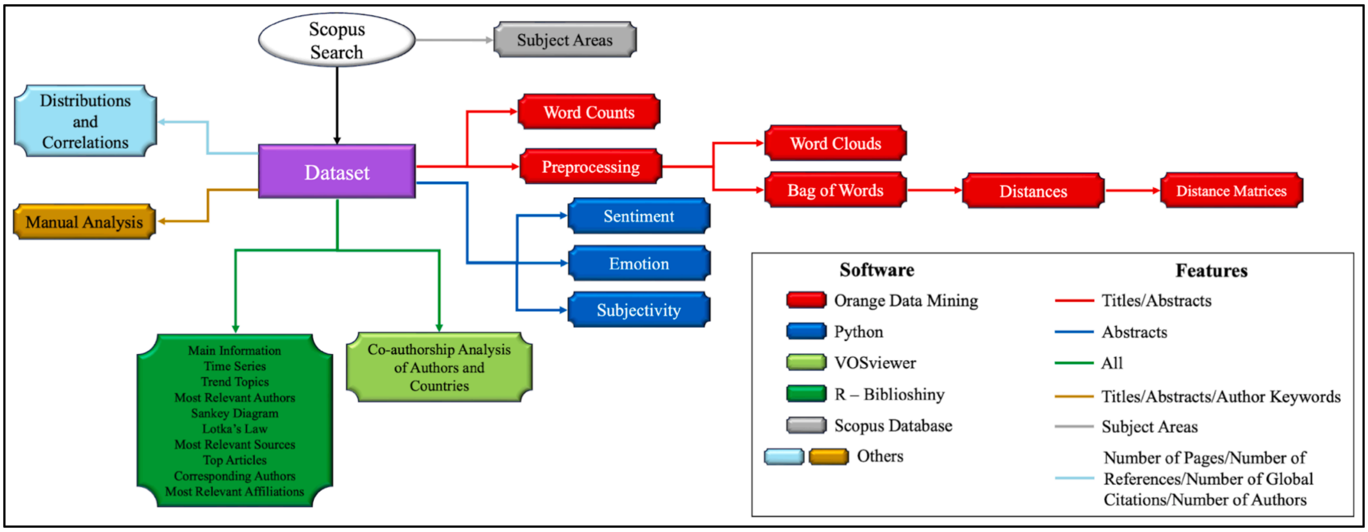

Fig. 2. Flowchart of the method followed.

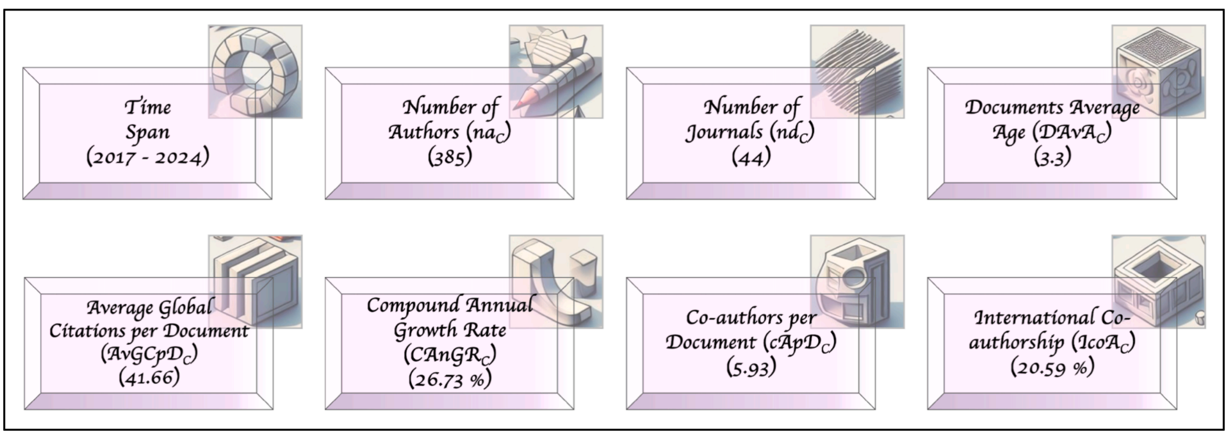

Fig. 3. Main elements of the P-MD membrane engineering dataset.

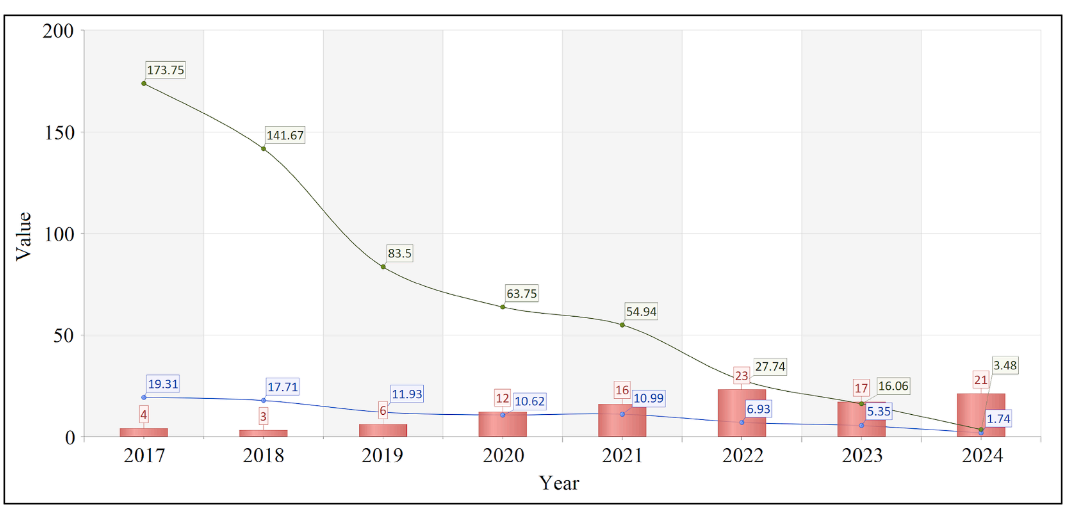

Fig. 4. Time series of publishing and citation values.

means a perfect positive correlation, 0 means no correlation between variables, and −1 means a perfect negative correlation). The number of pages *vs.* the number of global citations is found to be −0.381, indicating a weak negative relationship between them (i.e., as the number of pages increases, the number of citations tends to decrease slightly). The obtained value of the number of pages *vs.* the number of references is 0.546. This is evidence of a moderate positive relationship between them. As the article length increases, the number of references tends to

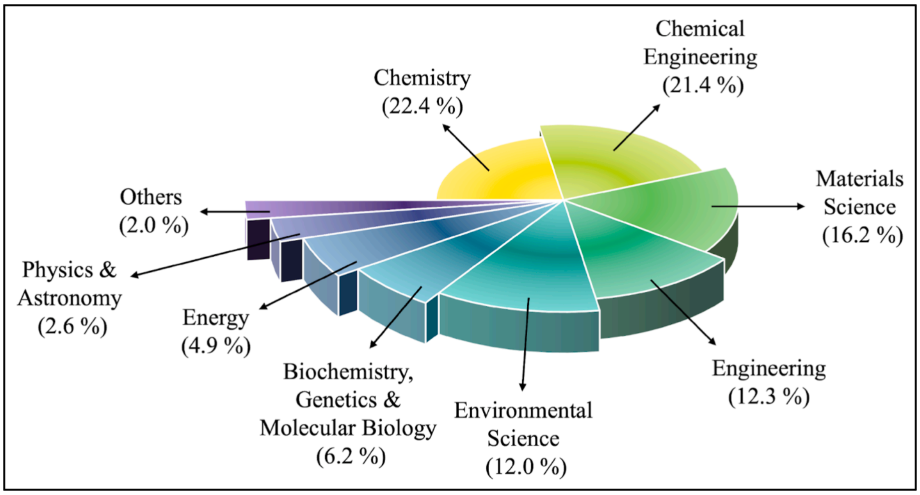

Fig. 5. Subject areas of the publications.

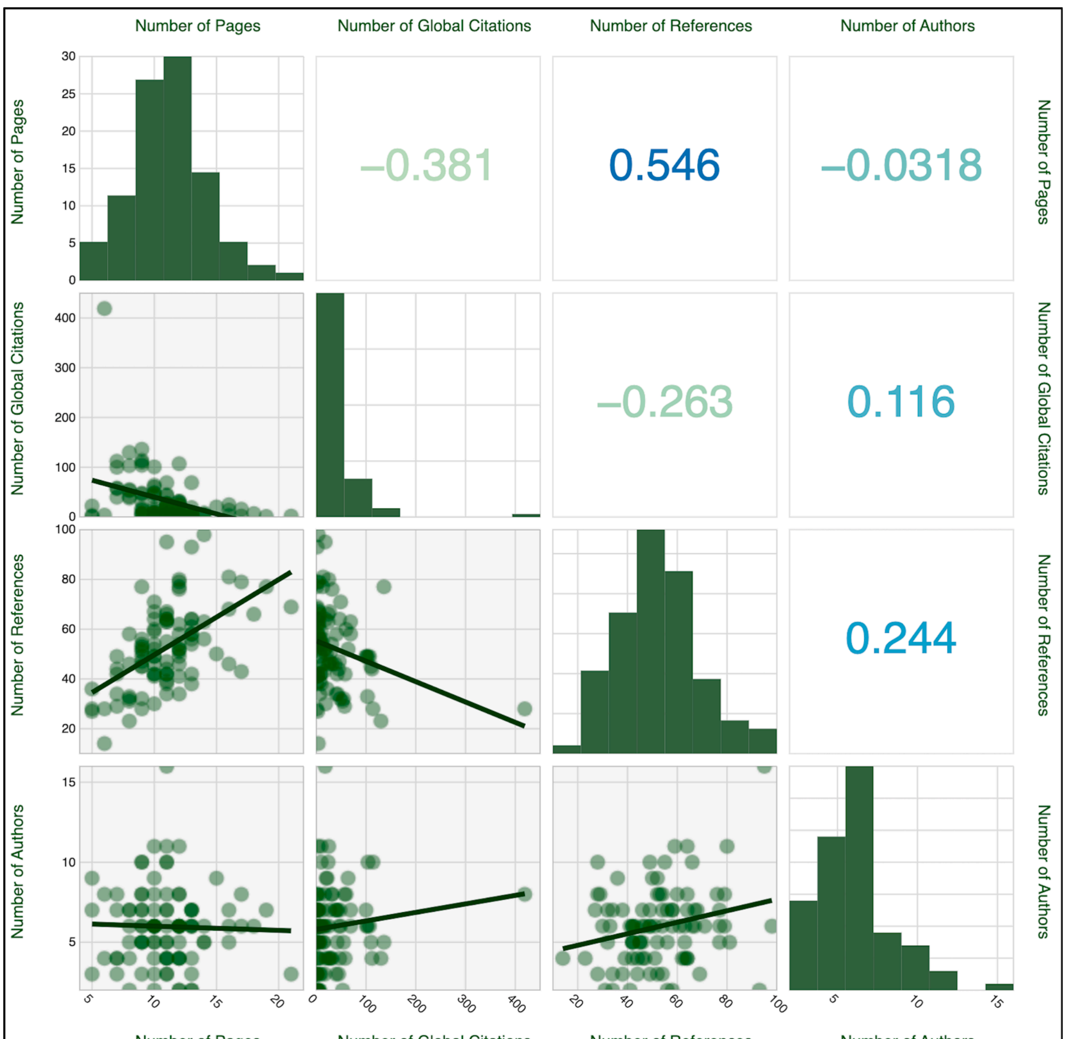

Fig. 6. Scatter matrix (distributions and correlation) of the number of pages, number of references, number of global citations and number of authors values of the articles.

increase, too. This is a logical conclusion since the longer the article, the more references it includes. The correlation coefficient between the number of pages and the number of authors is −0.0318, indicating a very weak negative relationship between them. In other words, there is almost no significance between the article length and the number of authors. The number of global citations *vs.* the number of references is found to be −0.263, indicating a weak negative relationship between them. This proves that increasing the number of references in a P-MD membrane engineering article results in a slight decrease in the number of citations. The correlation coefficient between the number of global citations and the number of authors is 0.116, showing a weak positive relationship. The number of citations increases slightly with the increase in the number of authors. This result makes sense, given that more authors can provide more visibility and a wider network. The correlation coefficient between the number of references and the number of authors is 0.244. This weak positive relationship between them means that the number of authors tends to increase slightly as the number of references increases. It may make sense that more complex or comprehensive studies require more authors and references.

Trend topics analysis is an effective way to find out which direction P-MD membrane engineering research is heading. From Fig. 7, it is noticeable that dispersing blue 14 was the trending topic in 2017. Disperse blue 14 is a type of dye that can be used to modify polytetrafluoroethylene (PTFE) membranes [86]. In 2018, PTFE membrane was the trending topic. The choice of using PTFE polymer in membrane engineering is evident because the fluorine atom's high electronegativity, low polarizability, tiny van der Waals radius (1.32 Å), and strong C-F bond (485 kJ mol^{-1}) make PTFE a chemically inert, durable, and extremely hydrophobic material. PTFE's exceptional characteristics make it appropriate for a multitude of common uses and industrial operations [88]. Anti-oil fouling, which is the trending topic of 2019, is a property that membrane engineers desire and research has been conducted to prepare a host membrane matrix to repel oil-based impurities and to eliminate membrane fouling [89]. Localized heating is the trending topic for 2020. It refers to the use of solar energy to induce water vapor by applying local heating on the photothermal membrane interface [90]. This localized heating eliminates the use of feed recirculation power outside the membrane module and reduces the temperature polarization effect by using less energy to attain the same temperature close to the membrane than by heating the bulk feed water. This technique can be applied to various applications beyond desalination thanks to the localized heating at the feed/membrane interface, allowing for use in small, modular designs [91]. The use of the keywords solar membrane distillation, membrane distillation and photothermal membrane distillation is naturally the main theme of the study and is therefore discussed extensively throughout this article. Based on electrohydrodynamics, electrospinning is a technique applied to generate films from different polymer solutions at the micro and nanoscale [92]. Being the trending topic of 2024, electrospinning is a simple, innovative, adaptable, and cost-effective technology to produce fibers with high and customizable void volume fraction, thickness, and surface area. The shape of electrospun nanofibers is greatly influenced by several factors, including polymer concentration and its molecular weight, solvent, viscosity, electrical conductivity, and surface tension of the polymer solution, applied voltage and tip-to-collector distance, among others. Controlling these factors allows for easy production of electrospun nanofiber scaffolds for suitable purposes, making it an ideal membrane preparation technique for P-MD membranes [93].

Word counts of the titles and abstracts of P-MD membrane engineering articles were also analyzed (Fig. 8). Note that 1 article does not contain an abstract, therefore, abstract analyses are performed on 101 entries.

Fig. 8(a) represents the histogram of titles. As can be seen, the authors use a wide range of number of words for the titles, ranging from 7 to 18. The shortest title consists of 3 words, while the longest title consists of 24 words. Besides, the distribution of the title word counts follows a normal distribution. When Fig. 8(b) – word count of abstracts – is analyzed, it is seen that the length of abstracts of the authors is gathered around 200 words. This is mainly since journals tend to limit abstracts to 200 words. The shortest abstract consists of 111 words, and the longest of 343 words. Abstract word counts also fits a normal distribution. The word cloud approach visually represents word frequency in the constructed text. The more frequently the term appears in the analyzed article, the larger it appears in the produced figure [94]. A word cloud method was applied to the titles and abstracts of the articles in the collection. The results are visualized in Fig. 9.

As can be seen in Fig. 9(a), word cloud of titles, the top 5 most frequently used words in the titles of the articles are membrane (148) (or membranes before lemmatization process), distillation (80), photothermal (59), solar (55), and drive (23) (or driven/driving before lemmatization process). The frequent use of these words reflects the purpose and target audience of the researchers. This frequent use of the words "membrane" and distillation point toward the core focus on the process of MD, while the frequent occurrence of "solar" and "photothermal" highlights the integration of solar energy as a critical component in advancing photothermal MD. The consistent use of these key terms reflects the purpose of the studies and their relevance to the broader community. It also indicates that the authors wanted to preserve the significance or essence of their articles.

Fig. 9(b) depicts the word cloud for article abstracts, and the top 5 words that appear are membrane (632) (or membranes before lemmatization process), photothermal (264), solar (252), water (251), and energy (203). Again, the frequent mention of "membrane" and "water" suggests a strong emphasis on fundamental components and applications of MD, while the prevalence of "solar" and "photothermal" demonstrates how the availability of solar energy justifies advancements in photothermal MD. The word "energy" further emphasizes the focus on efficient energy use within this context. In addition, words such as

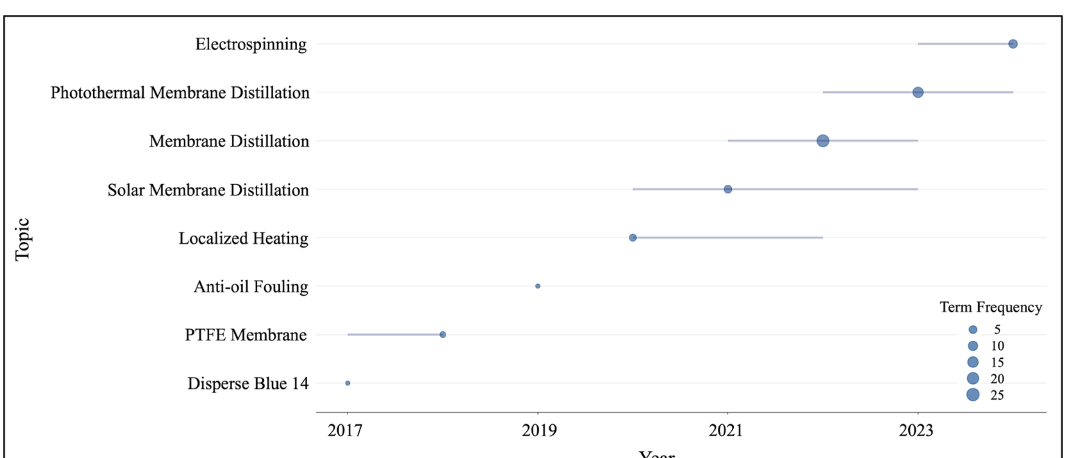

Fig. 7. Trend topics based on the author's keywords (word minimum frequency = 1, number of words per year = 1).

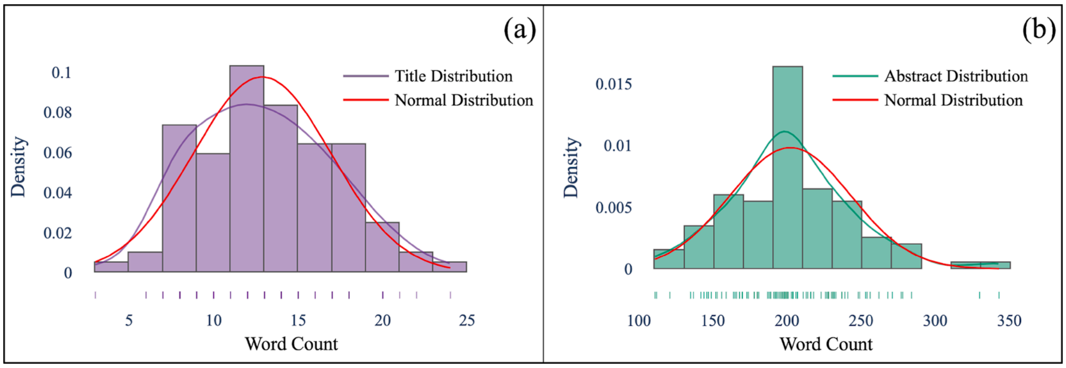

Fig. 8. Word count densities of **(a)** titles and **(b)** abstracts.

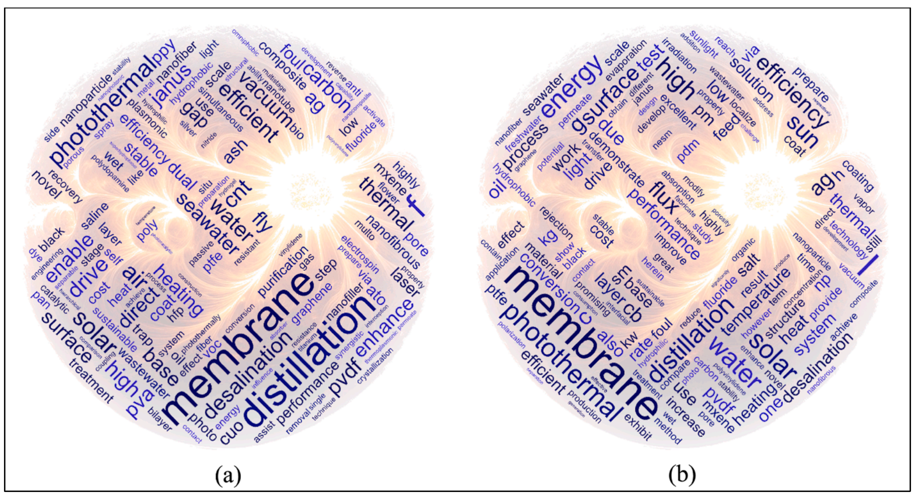

Fig. 9. Word cloud of **(a)** titles and **(b)** abstracts of the articles.

"high" (150), and "efficiency" (115) indicate the desire to create high-efficiency systems. The use of "flux" (108) indicates a focus on enhancing flux, while "surface" (108) suggests efforts in modifying membrane surfaces to enhance performance or induce photothermal behavior potentially. Interestingly, PVDF (polyvinylidene fluoride) was mentioned 94 times, indicating the dominantly used polymer to fabricate MD membranes, although polyethylene (PE), polypropylene (PP) and polytetrafluoroethylene (PTFE) are also commonly applied [95]. Manual searching was conducted to identify MD configurations, photothermal materials, field of applications and host matrix materials preferred by scientists interested in photothermal membrane engineering. Manual searching was performed on the titles, abstracts, and author keywords of the articles to determine the number of articles pointing to the use of above-mentioned applications. Note that some articles refer to more than one MD configuration/photothermal material/field of application/host matrix material. VMD is the most mentioned configuration in the dataset (13 times), followed by DCMD configuration (8 times). AGMD was mentioned 5 times, while SGMD was only mentioned in 1 article. Photothermal materials such as gold (Au), silver (Ag), aluminum (Al), carbon black, graphite (graphitic carbon), graphene, (reduced) graphene oxide (RGO/GO), multi-walled carbon nanotubes (WCNT), polypyrrole (PPy), polyaniline (PANI) nanofiber, Fe_3O_4 (iron (II, III) oxide/black iron oxide), TiN (titanium nitride), MXene, MoS_2 (molybdenum disulfide), nickel chalcogenide, activated carbon, CuO (copper (II) oxide/cupric oxide), SnO_2 (Tin (IV) oxide/stannic oxide) and Ti/MgF_2 (Titanium/Magnesium Fluoride) are encountered 5, 12, 2, 16, 3, 7, 4, 4, 4, 2, 2, 3, 10, 3, 1, 3, 2, 1 and 1 times in the collected articles, respectively. Desalination, volatile organic compounds (VOCs) removal, oily wastewater treatment, dye removal (textile), heavy metal rejection, pharmaceutical wastewater treatment and food industry wastewater are the fields of of applications discussed in the collection 73, 12, 8, 4, 2, 1 and 1 times, respectively. Among the host matrix materials, Polytetrafluoroethylene (PTFE) is mentioned in 13 articles, Polyvinylidene fluoride/ polyvinylidene fluoride-co-hexafluoropropylene (PVDF/PVDF-HFP) is mentioned in 49 articles, Polyacrylonitrile (PAN) is mentioned in 3 articles, Polypropylene (PP) is mentioned in 3 articles, Polydimethylsiloxane (PMDS) is mentioned in 9 articles and Bacteria Nanocellulose (BNC) is mentioned in 1 article. For photothermal membranes, it was found that the most mentioned techniques in the abstracts of the collected dataset are coating (21 articles), electrospinning (20 articles) and phase inversion (14 articles). Coating the photothermal membrane layer is necessary to carry out P-MD. The dope solution is used to prepare the host matrix and the photothermal material of the membranes via electrospinning or phase inversion. This last technique remains common for the preparation of membranes both at laboratory and industrial scales. Vacuum filtration and chemical modification have also been considered in 6 and 7 articles, respectively. Other techniques used in photothermal membrane engineering include electrospraying (3 articles), laser one-step direct writing (2 articles) and chemical vapor deposition (CVD) (3 articles).

3.2. Contributing researchers on P-MD membrane engineering

The most relevant authors publishing P-MD membrane engineering with their specific metrics (based on their number of publications) and the journals in which these authors published are illustrated in Figs. 10

and 11, respectively.

As Fig. 10 indicates, Li Q. is the dominant author in terms of most articles (9), articles fractionalized value (1.74), which measures an individual author's contribution to a published collection of articles, *h*-index (8) and *g*-index (9). As can be seen in Fig. 11, Li Q. publishes in various journals, including Separation and Purification Technology, Desalination, and the Journal of Membrane Science, among others. The foremost scientist with global citations is Politano A. with 713 citations in his 8 articles (Fig. 10). When *m*-index is considered, Wang Y. is the first author with a value of 1.4, as indicated in Fig. 10.

Social networks are very important for the scientific community for reasons such as scientific communication, visibility, effectiveness, research, collaboration, education, and lifelong learning. We used VOSviewer to reveal the social networks of authors publishing in the field of photothermal membrane engineering of MD, and the results are shown in Fig. 12. Note that for the readability of the figure, the minimum number of documents of an author was limited to 4.

In Fig. 12 the vertex (node) size is proportional to the number of articles of the author. A more prominent node denotes more articles, and vice versa. The thickness of the edge "link" indicates the relationship "degree of co-authorship" between two nodes. As can be seen in Fig. 12, P-MD membrane engineering domain includes 4 social networks "clusters". Wang Y. gears the largest cluster with 7 members (cluster 3). Cluster 4 has 6 members. Wang Y. has the highest number of links (8 connections). Fig. 12(a) indicates the average publication year of authors. Cluster 2 appears to have the publications with the oldest average age (2020,21), whereas cluster 4 has more recent publications with a high $DAvA_{It}$ value (2022.25). Fig. 12(b) shows the authors' average global citations per article value. Despite having recently published on P-MD membrane engineering, the authors in cluster 4 have a relatively low citation number, which is an expected result. The researchers in cluster 2 seem to have high $AvGCpD_{It}$ values (90.3). Cupolillo A. in cluster 1 has the highest $AvGCpD_{It}$ value (159.50). Furthermore, the highest link strength (i.e., the highest number of co-publications) is found between Politano A. and Curcio E. among the authors in cluster 1 with a value of 7. To remove the issue of disparities in citation counts owing to a document's age and to better comprehend a document's influence, the average relative global citations metric, which considers the age of the publications, is more descriptive (Fig. 12(c)) [72,96,97]. When Fig. 12(c) is examined, it is noticed that cluster 4 reaches an $AvRGC_{It}$ value of 1.48. The top 3 authors with the highest $AvRGC_{It}$ metric are Kang W. (1.96) from the 4th cluster, An A. K. (1.90) from the 3rd cluster, and Cupolillo A. (1.53) from the 1st cluster. To foster knowledge transfer and bridging different areas of expertise, researchers should establish more international partnerships and engage in interdisciplinary projects. In addition, dominant groups and authors in the field should collaborate more extensively to bring together different perspectives and expertise in favor of P-MD to develop it into a world-

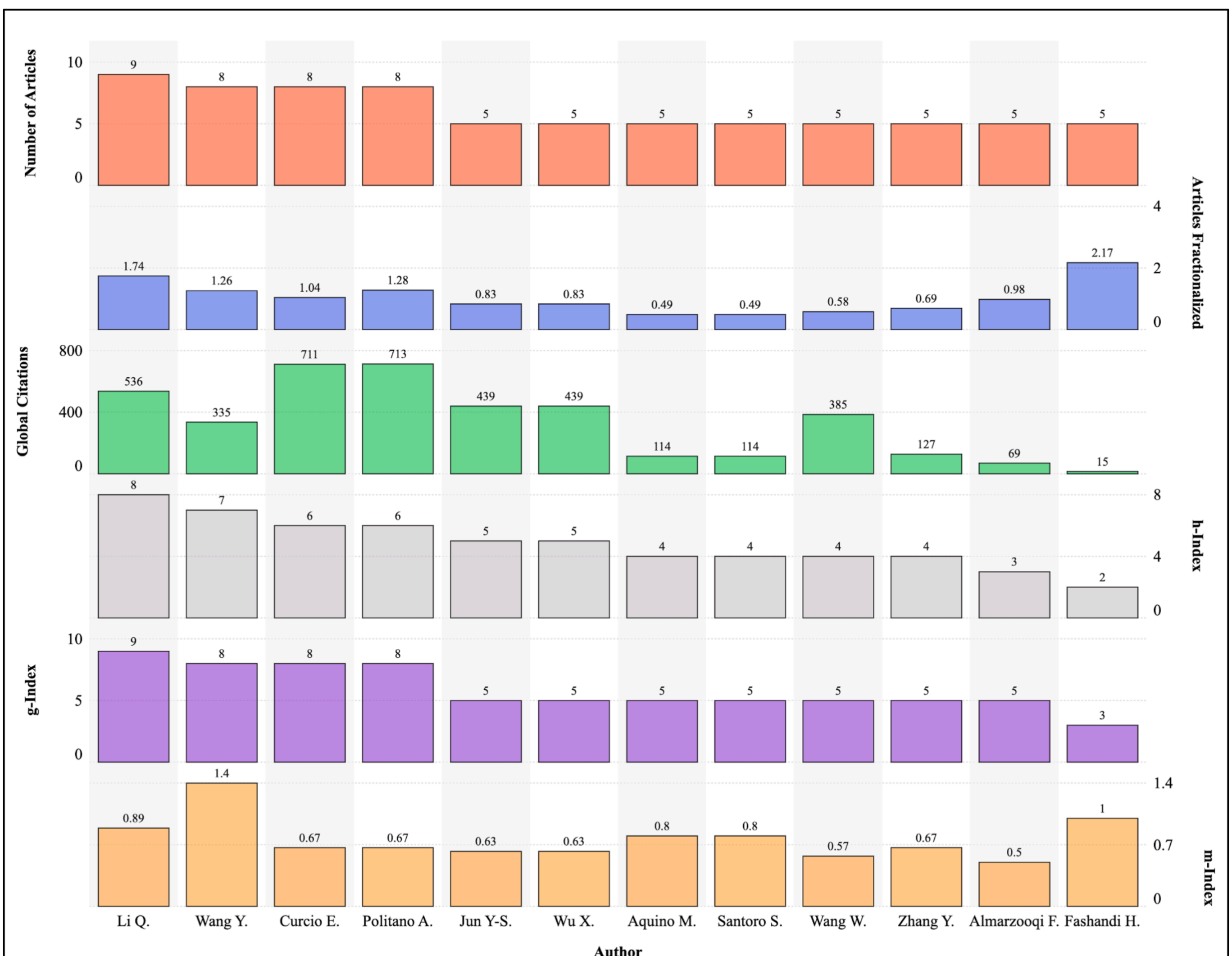

Fig. 10. Metrics of most relevant authors based on the number of published articles.

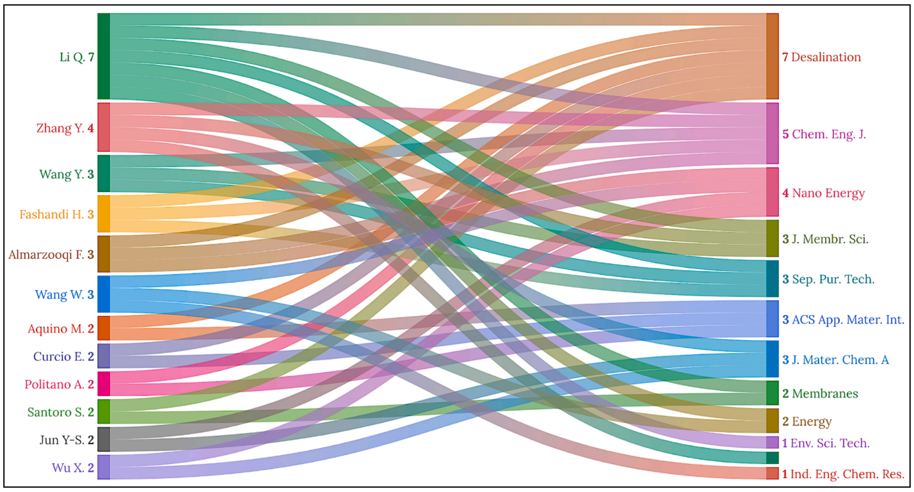

Fig. 11. Most relevant journals where P-MD articles are published.

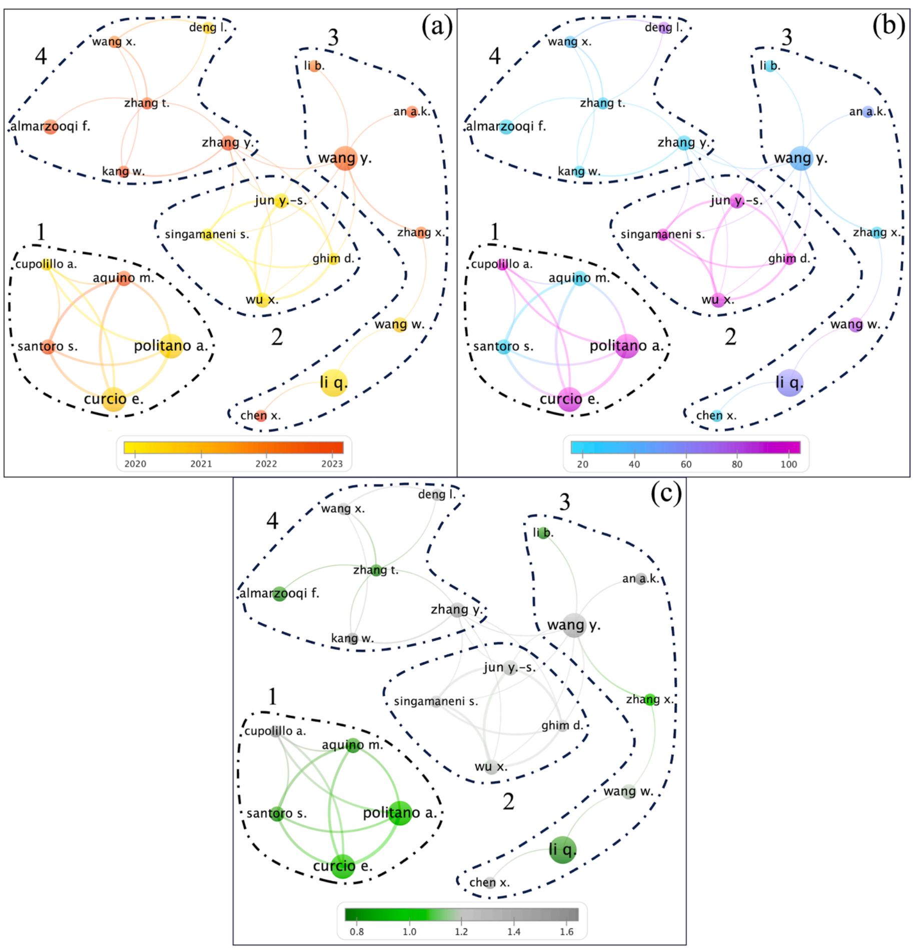

Fig. 12. Co-authorship analysis of authors **(a)** documents average age (average publication year) ($DAvA_{It}$), **(b)** average global citations per document ($AvGCpD_{It}$), and **(c)** average relative global citations ($AvRGC_{It}$) (weights = documents, min. number of documents of an author = 4, clusters with single item are removed).

class desalination technology.

Furthermore, to understand the distribution of productivity of authors in P-MD membrane engineering domain, Lotka's law analysis has been conducted, and the results can be seen in Table 2.

As can be seen in Table 2, most of the authors (68.6 %) published on photothermal MD membrane engineering contribute to only 1 article. This situation shows no continuity in researchers working on photothermal membrane production. The number of authors who published 5 or more publications is 2.1 %. This rate is very low. We hope that the number of researchers interested in photothermal MD membrane engineering will increase for the sake of desalination and membrane science.

Table 2
Lotka's Law results of the dataset.

Documents written	Number of Authors	Proportion of Authors
1	264	0.686
2	70	0.182
3	28	0.073
4	11	0.029
5	8	0.021
7	3	0.008
9	1	0.003

3.3. P-MD membrane engineering research journals

Scientific journals are one of the cornerstones of science and academia. These are gates of scientific communication ensuring the protection of scientific quality, they are the first places to be consulted as scientific resources, and they are very important in distributing and disseminating information [98]. For these reasons, in this section, we reveal certain metrics of the journals to the readers together with the resources to be consulted on P-MD membrane engineering. Fig. 13 illustrates the academic performance of the top journals of the collection.

As Fig. 13 indicates, the Journal of Membrane Science is the source with the highest metrics in a number of publications (15), *h*-index (12), *m*-index (1.5) and *g*-index (15). Desalination has the highest number of global citations, with 597. Separation and Purification Technology is the journal that occupies the third place in the number of articles published (10), *h*-index (7), and *g*-index (10); and the second in *m*-index (1.17). This indicates that articles published in this journal have a high impact relative to the career span of the authors publishing in this journal.

3.4. Influential articles in P-MD membrane engineering

The top 10 articles of the dataset, depending on the global citations and their corresponding other metrics, can be seen in Fig. 14.

According to global citations in Fig. 14, the leading article in terms of the number of global citations (477), global citations per year (53) and relative global citations (2.75) about P-MD membrane engineering is "*Photothermal Membrane Distillation for Seawater Desalination*" written by Politano et al. (2016) [87], highlighting its impact on the early development of the field. Its influence arises from the introduction of thermo-plasmonic effect, a groundbreaking mechanism in which Ag nanoparticles embedded in microporous PVDF membranes enhance localized heating when exposed to UV irradiation. This breakthrough demonstrated the use of the thermos-plasmonic effect to mitigate the temperature polarization phenomenon, a key challenge in MD, enabling high permeate flux and thermal efficiency. The experimental validation of this effect laid the foundation for further studies in nanomaterial-enhanced photothermal membranes. Considering the global citations, there are no obvious differences between the second and tenth articles, as can be observed in Fig. 14. Ranked second with 165 times globally cited is the study of Wu X. et al. (2018) with the title "*Localized heating with a photothermal polydopamine coating facilitates a novel membrane distillation process*" [99]. The importance of this paper lies in its demonstration of a polydopamine (PDA)-coated PVDF membrane as a scalable and efficient approach to solar-driven MD. The paper introduced an easily implementable, flexible, and robust coating method that

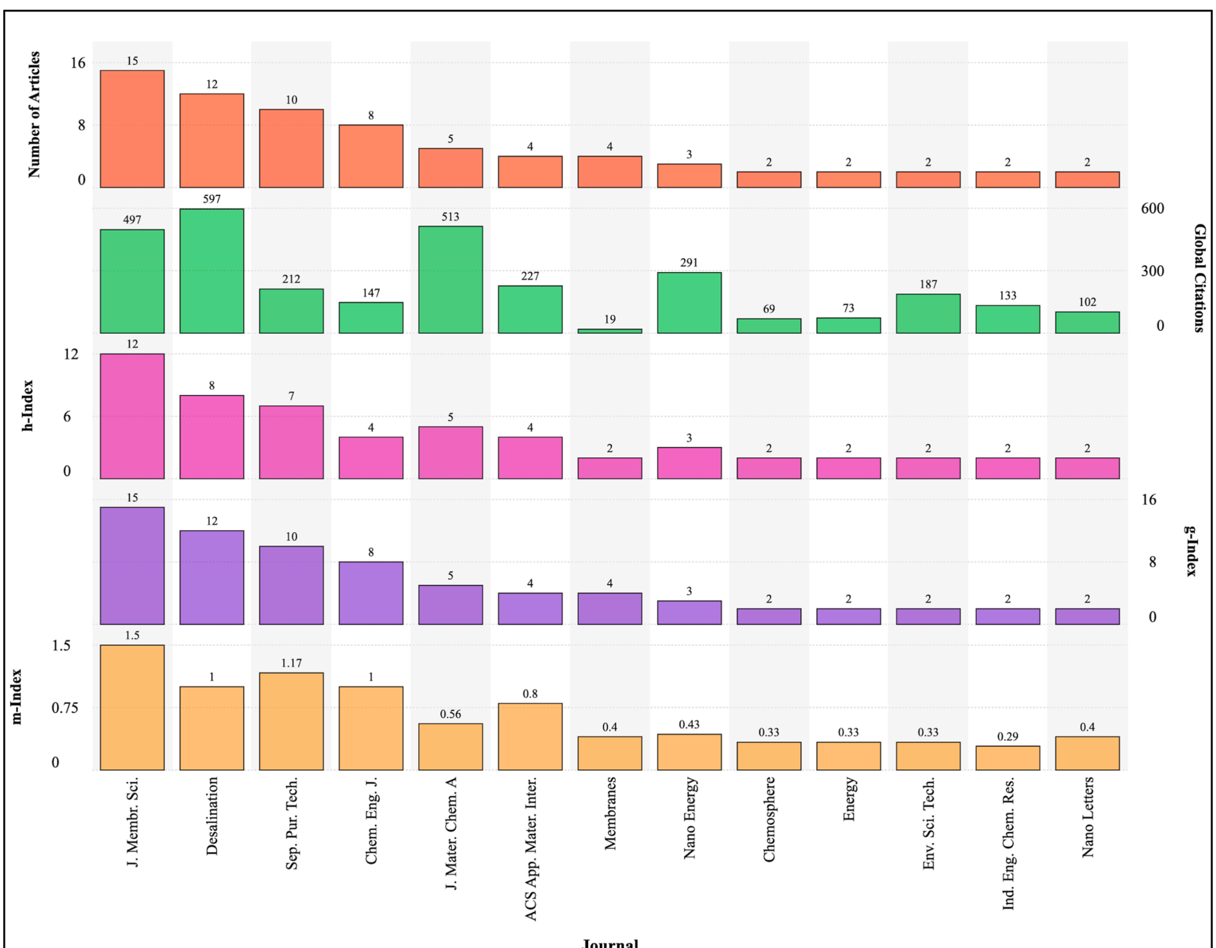

Fig. 13. Metrics of most relevant sources based on the number of published articles.

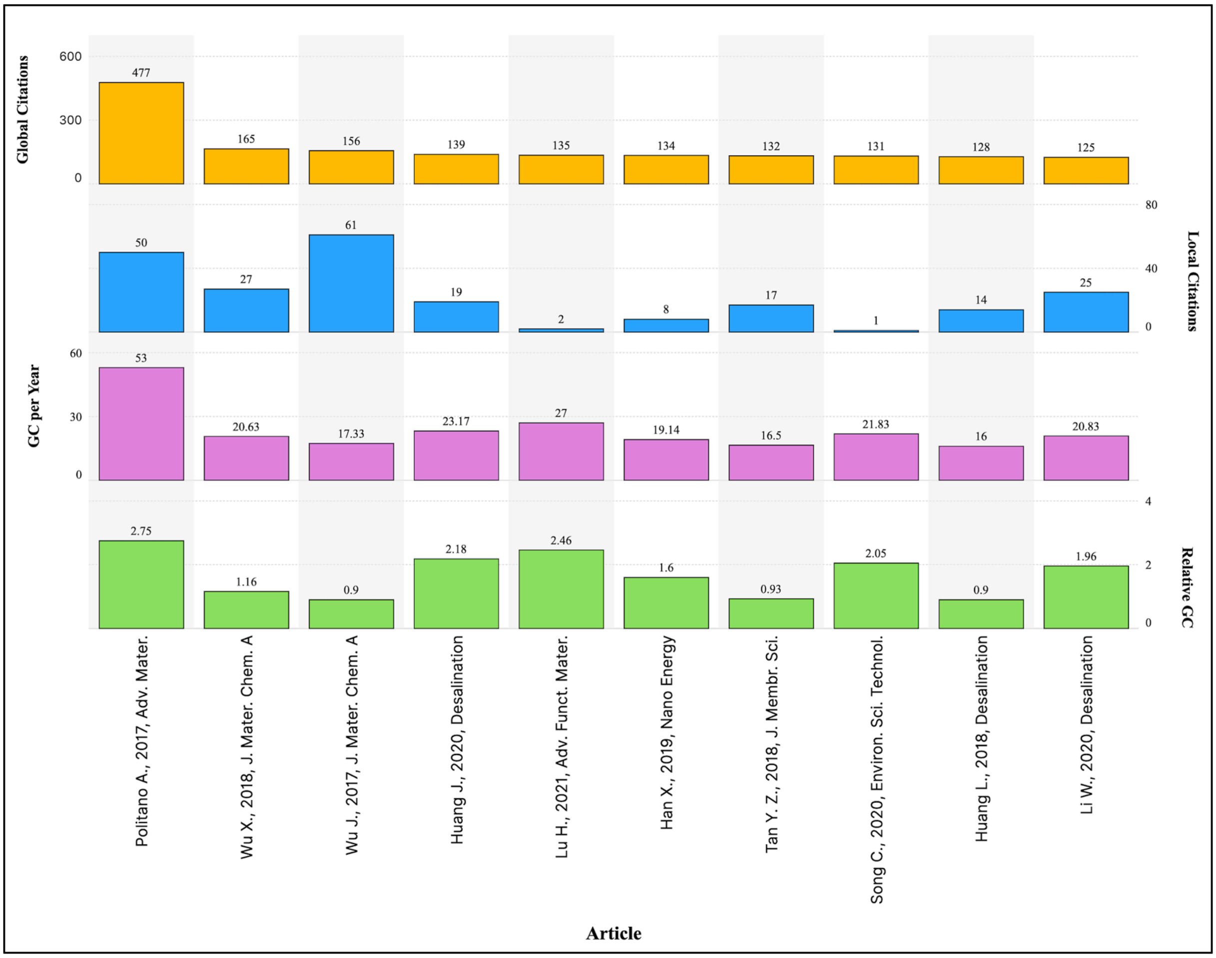

Fig. 14. Top 10 articles based on their global citations.

exploited PDA's broad-spectrum light absorption and high photothermal conversion to yield superior water flux (0.49 kg m^{-2} h^{-1}) and high energy efficiency (45 %) among known P-MD membranes when exposed to 0.75 kW m^{-2} solar irradiation. The authors applied fluorosilanization to enhance hydrophobicity, which enhanced wetting resistance and salt rejection. When local citations are counted, the article entitled "*Photothermal Nanocomposite Membranes for Direct Solar Membrane Distillation*" written by Wu et al. (2017) [17] has a value of 61. This study pioneered a direct solar membrane distillation approach using photothermal nanoparticle coatings. This work also addressed the temperature polarization phenomenon by embedding carbon black nanoparticles and SiO_2/Au nano-shells. Their findings demonstrated that solar energy could be directly utilized to drive the MD process without needing external heating, increasing therefore the permeate flux by up to 33 % under 1 sun irradiation. This innovation established a new paradigm in self-sustained solar desalination, influencing subsequent studies that sought to enhance MD's energy efficiency.

Beyond these individual contributions, a common theme among the most cited papers is their focus on addressing efficiency bottlenecks in P-MD through novel material engineering. These studies shifted the field from conventional polymeric membranes to multifunctional, nanomaterial-integrated membranes that optimize heat localization, transmembrane temperature as well as permeate flux and salt rejection. As a result, their influence extends beyond citations, serving as key references that guide the design principles of next-generation P-MD membranes.

3.5. *Country involvement in P-MD membrane engineering research*

A country's interest in scientific development in a particular field reflects its willingness to seek answers to questions about the nature and function of that field and its importance to scientific and technological development. In addition, a country's privileging of a particular branch of science often indicates its desire to progress and develop in that field. To reveal the P-MD membrane engineering prioritized countries, analyses were conducted, and the results can be seen in Figs. 15 and 16, respectively.

China, the USA, and Italy appear to be the leading countries in P-MD membrane engineering research with 50, 20 and 12 articles, respectively (Fig. 15). Furthermore, China and the USA have the highest number of co-publications (4). There are 5 social networks in Fig. 15 and the highest number of links belongs to Italy, with 8 countries. The country with the most recent articles is Iran, which has a value of 2023.6. Australia has the oldest publications, with an average of 2020.50 years. Based on $DAvA_{It}$ metric, cluster 4 comes first with an average document age of 2022.66. In Fig. 15(b), the country with the highest average global citations per document value is the UAE, with an average of 91.17 citations for the published 6 articles. Italy achieved an $AvGCpD_{It}$ value of 63.92 for 12 published articles. Egypt, on the other hand, has been the lagging country in this field, with an average of 2 citations for its 2 articles on P-MD membrane engineering. Cluster 4 occupies the first place with an average of 49.23 average citations per document value. Fig. 15 (c) indicates the average relative global citations of countries. In this metric, Hong Kong ranks first with a value of 1.84; Saudi Arabia ranks second with an $AvRGC_{It}$ value of 1.56 and Tunisia the last country with

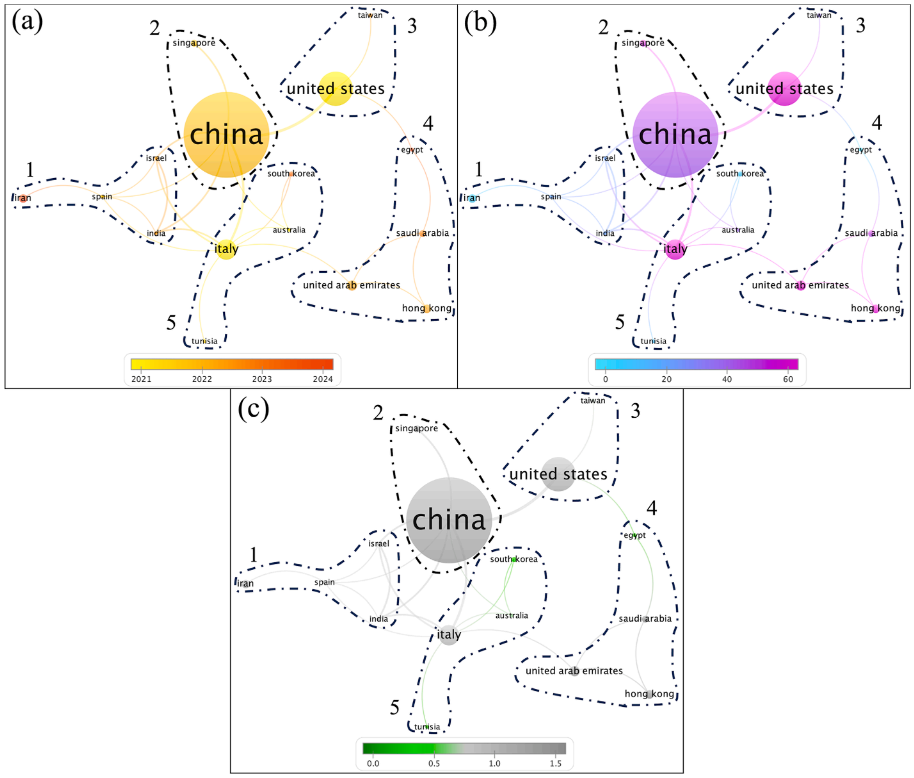

Fig. 15. Co-authorship analysis of countries **(a)** documents average age (average publication year, $DAvA_{It}$), **(b)** average global citations per document ($AvGCpD_{It}$), and **(c)** average relative global citations ($AvRGC_{It}$) (weights = documents, min. number of documents of a country = 2, clusters with single item are removed).

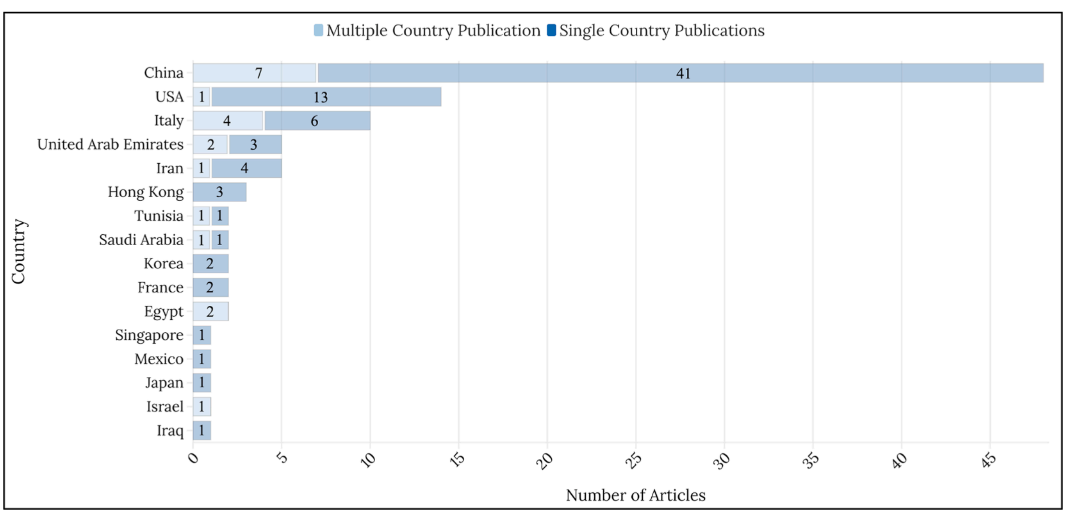

Fig. 16. Corresponding author's country.

0.23 average relative citations. When $AvRGC_{It}$ value is analyzed, cluster 4 ranks first with a value of 1.20, while cluster 2 ranks second with a value of 1.11. International cooperation in the field of P-MD is still limited. Researchers should establish more international partnerships and engage in interdisciplinary projects to promote knowledge exchange and convergence of different areas of expertise. In particular, projects that bridge the gap between laboratories in developed countries and practices in water-scarce regions should be supported.

Fig. 16 illustrates the corresponding author's country. Since China has the highest number of articles (50), it is unsurprising that it also ranks first in the corresponding authors' country value (48 correspondence). The interesting situation is that China has international partnerships in only 7 articles if the corresponding author is from China. 41 articles consist only of studies carried out by Chinese scientists of the correspondence from China. We believe that China's further international cooperation in photothermal MD membrane engineering will be instrumental in making membrane distillation one of the world's leading membrane processes. Again, countries such as Hong Kong, France and the Republic of Korea are far from international cooperation when the corresponding author is in their own country, as can be seen in Fig. 16. As both Fig. 15 and Fig. 16 indicate, China has the upper hand in P-MD membrane engineering. At the same time, Italy and UAE appear to be more willing to work internationally. In this context, it is important to examine in more detail the reasons for China's success in this area.

China's rise to the top in terms of number of published papers in 2018 [100]. Several factors contribute to China's continued progress in research and development (R&D): a large population and human capital base, a labor market that favors academic meritocracy, a significant diaspora of Chinese-origin scientists, and a centralized government ready to invest in science and development [101]. China, one of the world's lowest-income countries at the beginning of the 21st century, has transformed itself into a scientific knowledge superpower in less than two decades, a unique achievement in scientific history. The way China employs its freshly expanded scientific resources will shape the future of science and technology, as well as our increasingly knowledge-based economy [102]. We believe this will be a significant opportunity for the P-MD membrane engineering field.

3.6. Institutional contributions to P-MD membrane engineering research

Affiliations provide the necessary facilities to carry out particular scientific research. The institutions involved in P-MD membrane engineering are given in Fig. 17. Note that Biblioshiny package counts author-wise affiliation (i.e. if 3 authors in an article are from the same affiliation, the Biblioshiny package counts this affiliation as 3, not as 1).

The top 3 affiliations more active in the field of P-MD membrane engineering shown in Fig. 17 are from USA, Italy, and China. Washington University in St. Louis (USA) tops the list (34 times), followed by the University of Calabria (Italy) with 33 times and then Tiangong University (China) in third place (32 times). The strong presence of these affiliations in the field may be attributed to a combination of funding, specialized research facilities, and international collaborations. Washington University in St. Louis (USA) is likely benefiting from its well-funded research programs in advanced membrane technologies and nanoengineered materials. The university's partnerships with government agencies such as the U.S. Department of Energy (DOE) and the National Science Foundation (NSF) have supported the development of innovative materials for desalination and water purification. Additionally, their state-of-the-art laboratories provide a foundation for experimental studies in P-MD. The University of Calabria (Italy) has emerged as a leading European center for P-MD research. This prominence is partly due to its participation in large-scale European research initiatives, such as Horizon 2020, which funds sustainable water treatment technologies. The institution also benefits from strong industry-academic partnerships, particularly in developing plasma-enhanced and thermos-plasmonic membranes, which have contributed to the advancement of solar-driven MD. Tiangong University (TGU, China) is one of the first Chinese universities to conduct research on hollow fiber membrane science and technology. With the ongoing support of the National "Sixth Five-Year" to "Twelfth Five-Year" Science and Technology Plan, the Ministry of Science and Technology and Tianjin Municipal Government jointly approved the establishment of the provincial and ministerial key laboratories for membrane separation and membrane process in TGU in January 2015. With these advantages, these affiliations can help emerging P-MD membrane engineering researchers build their careers on a stronger foundation and succeed in the research field.

3.7. Machine learning and natural language processing on P-MD membrane engineering articles

Fig. 18 shows the analyses of sentiment, emotion, and objectivity based on the abstracts of the articles. Please note that because 1 article does not contain an abstract, sentiment and subjectivity analyses were carried out on 101 abstracts, whereas emotion analysis was done on 100 entries due to computational difficulties in processing data instances larger than 512 tokens. Sentiment analysis is a technique that assesses and determines the polarity (positive/optimistic, negative/pessimistic, or neutral) expressed in text data [103]. Emotion analysis is the process of extracting and assessing user's emotional states in the text [104]. Subjectivity analysis examines textual data to evaluate if it reflects a personal opinion (private state) or not [105].

Fig. 18(a) shows the sentiment scores of abstracts. As can be clearly seen from this figure, positive sentiment is mostly distributed between 0.6 and 1.0, which indicates that abstracts carry a general optimism about P-MD membrane engineering. Besides it was also observed that positive sentiment score was dominant in all 101 abstracts. The positive sentiment analysis results reflect the researchers' confidence in P-MD technology. To translate this into concrete outputs, future research should be more application-oriented, focusing on developing cost-effective and scalable solutions. Experimental studies and pilot-scale projects will be key in translating theoretical potential into real-world applications. Fig. 18(b) covers the emotion scores. This analysis aims to determine whether the authors include expressions of personal feelings in the abstracts. It can be seen from Fig. 18(b) that except for neutral emotion, the remaining emotions (anger, disgust, fear, joy, sadness, and surprise) have scores around 0, which specifies that authors considered not reflect personal feelings when writing the abstracts. In 99 abstracts, neutral is the dominant emotion, and only 1 abstract has a dominant disgust emotion score. Fig. 18(c) indicates the objectivity scores. This type of analysis was carried out to assess the objectivity of the abstracts by examining whether the language, wording, and arguments in the abstracts comply with the principle of impartiality. As can be seen in Fig. 18(c), scientists working on P-MD membrane engineering have paid attention to objectivity, one of the most important aspects of scientific language, when writing abstracts (objectivity scores are mostly around 0.8 to 1.0). In all 101 abstracts, the dominant value is in favor of objectivity.

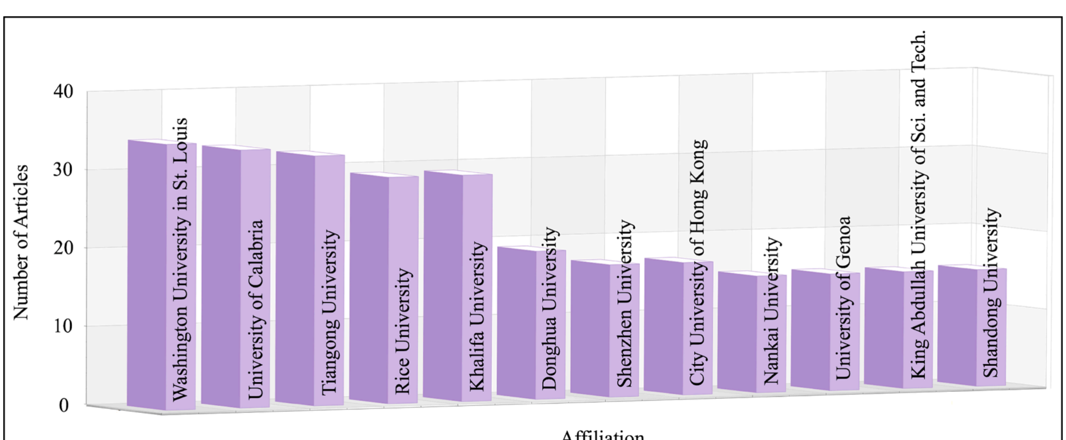

Fig. 17. Most relevant affiliations.

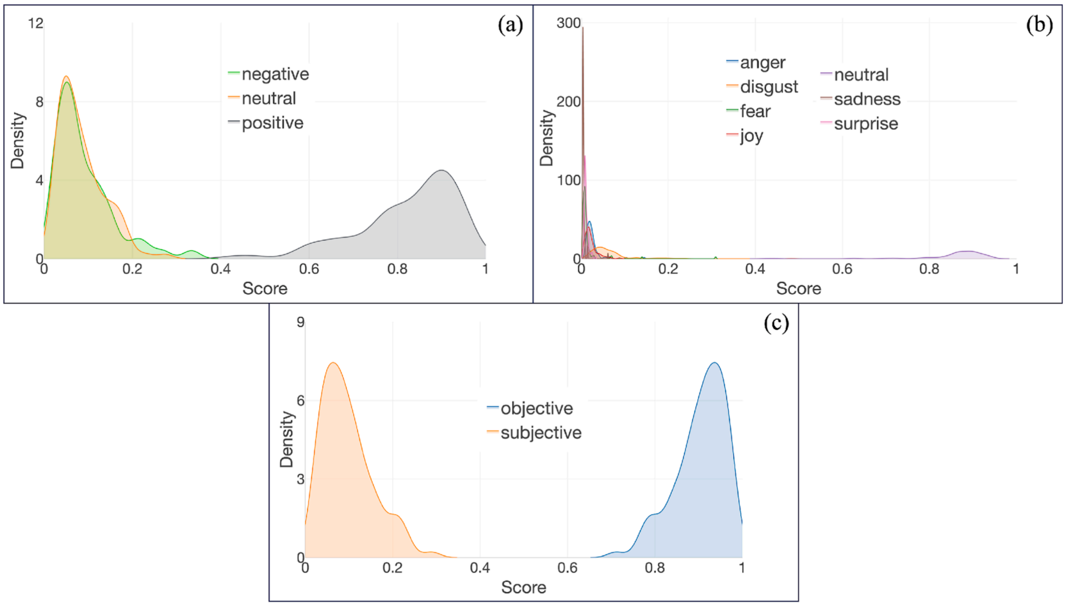

Fig. 18. **(a)** Sentiment, **(b)** emotion and **(c)** subjectivity analyze results on abstracts.

Similarity analysis is a part of TM and NLP used to identify similarities within a given set of text or documents [106,107]. A similarity analysis (word similarity, not semantic similarity) was conducted to find the similarity of titles and abstracts of the articles.

Fig. 19(a) shows the similarity of the titles of the studies in the P-MD domain. The title of an article is often the first thing readers notice before deciding whether to read further. Authors must select a title that not only piques researchers' interest but also accurately conveys the work's substance and encourages additional reading. Besides, reviewers and editors of journals may avoid papers with similar titles, even if the article's content differs. Using similar titles might mislead readers about the study's uniqueness [63]. For all these reasons, we conducted a similarity analysis on the titles of the articles in the dataset to see the outcomes. In the similarity matrix in Fig. 19(a), the closer the cell color is to light blue, the more similar the two items are, and the closer to dark blue, the more different the two items are. As can be seen from Fig. 19 (a), the color of the cells in the collection is close to dark blue, and the average cosine distance value is 0.75. This score reflects the authors' success in finding original names for papers published on P-MD membrane engineering. The cosine distance value of the articles with the most similar titles was calculated as 0.12. Fig. 19(b) depicts the similarity of abstracts. Performing similar analysis in the abstracts of the articles will help scientists who want to publish their work to write unique abstracts. Editors and reviewers of journals will evaluate the submitted manuscripts to examine the distinctive abstracts. In Fig. 19 (b), the two items are more similar, the closer the cell color is to light green and the more distinct the two items are when the cell color is close to dark green. The average cosine similarity score of Fig. 19(b) is 0.68, which can also be confirmed by the color of the figure (close to dark green). This result indicates how well the authors can develop distinctive abstracts for their publications. The minimum cosine similarity score was found out to be about 0.35.

4. Conclusions

P-MD technology is a promising solution to alleviate water scarcity on our planet. Without relying on conventional energy sources, localized heating at the membrane-feed interface has the potential to reduce water costs and improve the sustainability of water production processes. This technology utilizes available solar energy to localize heating on the membrane-feed water interface in MD systems using photothermal materials. The growing interest in MD membrane development, as reflected in P-MD research over the past decade, underscores the need for continued innovation in material engineering. This study examines P-MD membrane engineering articles from several perspectives, such as machine learning, data mining, bibliometrics, and manual analysis,

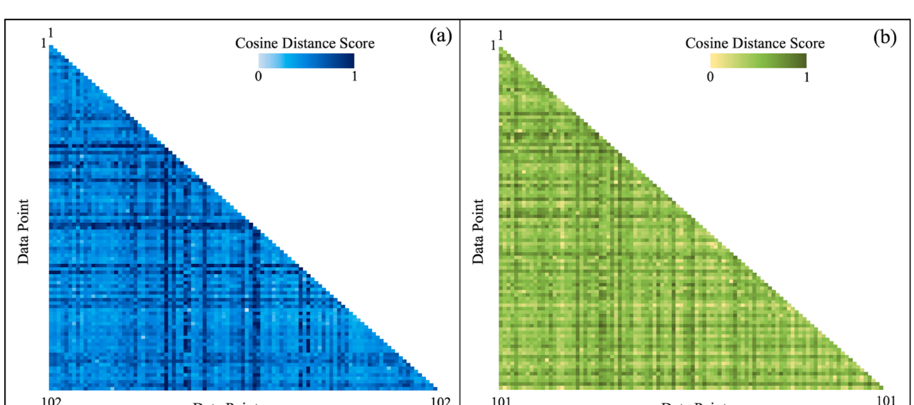

Fig. 19. Similarity matrices of **(a)** titles and **(b)** abstracts of the articles.

shedding light on the domain's hidden information. The dataset used in the research was downloaded on 24th January 2025 from the Scopus database and covers 102 articles between 2017–2024. 385 authors contributed to the field and articles were published in 44 different journals. The average age of the documents is as young as 3.3 years, indicating that the field is still in its infancy. Although 2022 reached the highest number of articles (23), according to $AvGCpD_{C_y}$ and $AvNGCpD_{C_y}$ values, articles published in 2017 are the most influential articles in P-MD membrane engineering (173.75 and 19.31, respectively). An article contains 11 pages on average, includes ~53 references on average and is cited by ~31 times on average. Electrospinning is a trending topic of recent years. The words used by the authors in the titles and abstracts are membrane, solar, and photothermal. The most considered configurations in P-MD are VMD (13) and DCMD (8), the most mentioned photothermal materials are carbon black (16) and silver (12), the most mentioned application is desalination (73), and the most mentioned host matrix is the polymer PVDF and the copolymer PVDF-HFP (49). Li Q. is the most prolific author, with 9 articles, a fractionalized value of 1.74, an *h*-index of 8, and a *g*-index of 9. P-MD membrane engineering domain includes 4 core working groups of authors. Journal of Membrane Science is the source dominating the field based on metrics like the number of publications (15), *h*-index (12), *g*-index (15) and *m*-index (1.5). "*Photothermal Membrane Distillation for Seawater Desalination*" written by Politano et al. (2016) [87] is the most influential article in terms of number of global citations (477), times cited per year (53) and relative GC (2.75) values. China is the leading country in research about P-MD membrane engineering, with 50 published articles. Sentiment analysis conducted on abstracts indicates that authors are very optimistic about the membranes they prepared. The abstracts were written emotion-free (neutral). The similarity analysis executed on titles and abstracts has an average of 0.75 and 0.68 cosine distance scores, respectively, which indicates that authors paid attention to writing uniquely. The insights gained from this study are important to enable researchers working in P-MD membrane engineering to assess the gaps in the field and guide future studies. P-MD technology can significantly contribute to the global water problem in areas such as water treatment, desalinization, efficient use of water resources, energy efficiency and sustainability. Therefore, the increasing number of articles in the field shows the growing interest of scientists aware of the P-MD potential and the intensive research and development efforts towards this technology. The high number of authors show the importance of interdisciplinary collaborations in the P-MD field. Future research should focus on conducting more comprehensive and innovative studies by bringing together researchers from different specialties. The relevance of certain journals (i.e. Journal of Membrane Science, Desalination and Separation and Purification Technology) will help researchers to have information about the recent publications. Future studies should reach a wider audience by publishing in journals with high-impact factors. The increasing demand for high-performance membranes, growing interest in nanotechnology, sustainability concerns, increasing research and publication activities, and technological advancements have played an important role in making electrospinning a trending topic in recent years. This technique has become an increasingly popular method for P-MD membrane formation. China's leadership in the number of publications in the P-MD field demonstrates the importance and investments made by this country. This is instructive for other scientists conducting research in the P-MD field in many ways, such as examining China's research focus, establishing international collaborations, assessing strategic priorities, and following emerging technologies. This study provides researchers working on P-MD membrane engineering with important insights into the field's current state. However, P-MD researchers can still create new research avenues and innovation opportunities for themselves by examining the results of this study in detail to gain deeper insights.

5. Key gaps and future direction

Despite significant advancements in P-MD, several challenges must be addressed to enhance technological readiness. The scalability of P-MD remains a critical issue, as most studies to date have been conducted at the laboratory scale, with limited data on long-term performance, as well as fouling resistance, and operational stability under real-world conditions. Further research should focus on upscaling effects and, if possible, operation of pilot plants to assess feasibility for industrial applications. Additionally, while photothermal materials improve localized heating, they are one of many contributors to the energy efficiency of MD systems. More work is needed to enhance solar absorption efficiency and optimize membrane design to minimize energy losses. Moreover, PVDF remains the dominant material for the formation of P-MD membranes. It is under scrutiny as part of per- and polyfluoroalkyl substances (PFAS). It may have limited future application, as environmental agencies such as the European Chemical Agency have increasingly expressed their resolve to eliminate the use of these substances. This highlights the need for novel materials, including biodegradable polymers and hybrid composites, combined with continuing to study nanostructured coatings. These materials must be optimized for photothermal efficiency, wetting resistance, and chemical stability when applied in P-MD membrane engineering. Furthermore, there is a lack of comprehensive techno-economic assessments that evaluate P-MD technology's financial and manufacturing feasibility, making collaboration with industry stakeholders essential to facilitate commercialization. Addressing these challenges through collaborative efforts will be crucial in advancing P-MD research from laboratory studies to real-world applications. Additionally, we observed that institutional contributions in P-MD membrane engineering remain geographically concentrated, with a significant number of publications emerging from a few countries. This calls for expanding funding opportunities and collaborative networks to involve more institutions to encourage a more globally inclusive research environment. By assessing the current research landscape on this topic, this study identifies key gaps and areas of improvement, serving as a guide to drive this technology towards efficient, cost-effective, and sustainable MD systems.

CRediT authorship contribution statement

Ersin Aytaç: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Farah Ejaz Ahmed:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Faissal Aziz:** Data curation, Formal analysis, Methodology, Validation, Writing – review & editing. **Mohamed Khayet:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Nidal Hilal:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This publication is part of the project PID2022-138389OB-C31, funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. This work was supported by the NYUAD Water Research Center, funded by Tamkeen under the NYUAD Research Institute Award (Project CG007). M. Khayet and F. Aziz thank FINCOME 2024 Long-Term Stay (CNRST,

Centre National pour la Recherche Scientifique et Technique, Morocco).

Data availability

Data for this article, (a csv file) is available at Google Drive at https://drive.google.com/file/d/18hsKT5u6pQUJB6P_RTa-ntX0ujyown3c/view?usp=drive_link

References

[1] O. Pelin, Membrane Distillation: Principle, Advances, Limitations and Future Prospects in Food Industry, in: Z. Sina (Ed.), Distillation, IntechOpen, Rijeka, 2012.
[2] L. Francis, F.E. Ahmed, N. Hilal, Advances in membrane distillation module configurations, Membranes 12 (2022) 81, https://doi.org/10.3390/membranes12010081.
[3] R.B. Saffarini, B. Mansoor, R. Thomas, H.A. Arafat, Effect of temperature-dependent microstructure evolution on pore wetting in PTFE membranes under membrane distillation conditions, J. Membr. Sci. 429 (2013) 282–294, https://doi.org/10.1016/j.memsci.2012.11.049.
[4] A. Criscuoli, A. Capuano, M. Andreucci, E. Drioli, Low-temperature direct contact membrane distillation for the treatment of aqueous solutions containing urea, Membranes 10 (2020) 176, https://doi.org/10.3390/membranes10080176.
[5] F. Han, S. Liu, K. Wang, X. Zhang, Enhanced performance of membrane distillation using surface heating process, Membranes 11 (2021) 866, https://doi.org/10.3390/membranes11110866.
[6] N. Thomas, M.O. Mavukkandy, S. Loutatidou, H.A. Arafat, Membrane distillation research & implementation: lessons from the past five decades, Sep. Purif. Technol. 189 (2017) 108–127, https://doi.org/10.1016/j.seppur.2017.07.069.
[7] K.P. Lopez, R. Wang, E.A. Hjelvik, S. Lin, A.P. Straub, Toward a universal framework for evaluating transport resistances and driving forces in membrane-based desalination processes, Sci. Adv. 9 (2023) eade0413, https://doi.org/10.1126/sciadv.ade0413.
[8] F.E. Ahmed, B.S. Lalia, R. Hashaikeh, N. Hilal, Alternative heating techniques in membrane distillation: a review, Desalination 496 (2020) 114713, https://doi.org/10.1016/j.desal.2020.114713.
[9] A. Deshmukh, C. Boo, V. Karanikola, S. Lin, A.P. Straub, T. Tong, D.M. Warsinger, M. Elimelech, Membrane distillation at the water-energy nexus: limits, opportunities, and challenges, Energy Environ. Sci. 11 (2018) 1177–1196, https://doi.org/10.1039/C8EE00291F.
[10] S. Roy, M.S. Humoud, W. Intrchom, S. Mitra, Microwave-induced desalination via direct contact membrane distillation, ACS Sustain. Chem. Eng. 6 (2018) 626–632, https://doi.org/10.1021/acssuschemeng.7b02950.
[11] A. Anvari, K.M. Kekre, A. Azimi Yancheshme, Y. Yao, A. Ronen, Membrane distillation of high salinity water by induction heated thermally conducting membranes, J. Membr. Sci. 589 (2019) 117253, https://doi.org/10.1016/j.memsci.2019.117253.
[12] F.E. Ahmed, B.S. Lalia, R. Hashaikeh, N. Hilal, Enhanced performance of direct contact membrane distillation via selected electrothermal heating of membrane surface, J. Membr. Sci. 610 (2020) 118224, https://doi.org/10.1016/j.memsci.2020.118224.
[13] K. Li, Y. Zhang, Z. Wang, L. Liu, H. Liu, J. Wang, Electrothermally driven membrane distillation for low-energy consumption and wetting mitigation, Environ. Sci. Technol. 53 (2019) 13506–13513, https://doi.org/10.1021/acs.est.9b04861.
[14] F.E. Ahmed, Y. Ibrahim, N. Hilal, A spacer-based approach for localized Joule heating in membrane distillation, npj Clean Water 7 (2024) 42, https://doi.org/10.1038/s41545-024-00337-2.
[15] N.S. Fuzil, N.H. Othman, N.H. Alias, F. Marpani, M.H.D. Othman, A.F. Ismail, W. J. Lau, K. Li, T.D. Kusworo, I. Ichinose, M.M.A. Shirazi, A review on photothermal material and its usage in the development of photothermal membrane for sustainable clean water production, Desalination 517 (2021) 115259, https://doi.org/10.1016/j.desal.2021.115259.
[16] J. Qiu, W.D. Wei, Surface plasmon-mediated photothermal chemistry, J. Phys. Chem. C 118 (2014) 20735–20749, https://doi.org/10.1021/jp5042553.
[17] J. Wu, K.R. Zodrow, P.B. Szemraj, Q. Li, Photothermal nanocomposite membranes for direct solar membrane distillation, J. Mater. Chem. A 5 (2017) 23712–23719, https://doi.org/10.1039/C7TA04555G.
[18] M. Alqaydi, M.O. Mavukkandy, I. Mustafa, A. Alnuaimi, H.A. Arafat, F. Almarzooqi, Activated carbon as a photothermal absorber in PVDF membranes for solar driven air-gap membrane distillation, Desalination 541 (2022) 116031, https://doi.org/10.1016/j.desal.2022.116031.
[19] L. Eykens, K. De Sitter, C. Dotremont, L. Pinoy, B. Van der Bruggen, How to optimize the membrane properties for membrane distillation: a review, Ind. Eng. Chem. Res. 55 (2016) 9333–9343, https://doi.org/10.1021/acs.iecr.6b02226.
[20] G.-R. Xu, M. Wang, K. Xu, H.-L. Zhao, Q. Liu, Membrane fabrication and configuration design development of photothermal membrane distillation (PMD), Desalination 565 (2023) 116833, https://doi.org/10.1016/j.desal.2023.116833.
[21] S. Tian, X. Li, J. Ren, Z. Zhou, F. Wang, K. Shi, J. Xu, T. Gu, H. Shon, Emerging heat-localized solar distillation systems: solar interfacial distillation VS photothermal membrane distillation, Desalination 572 (2024) 117147, https://doi.org/10.1016/j.desal.2023.117147.
[22] P. Otlet, Traiťe de Documentation, Mundaneum Bruxelles-Belgium, 1939.
[23] A. Pritchard, Statistical bibliography or bibliometrics, J. Doc. 25 (1969) 348–349.
[24] M.B. Kalıpçı, E.K. Şimşek, R. Eren, Decoding the trends and the emerging research directions of E-commerce and tourism in the light of resource dependence theory: a bibliometric analysis, Heliyon 10 (2024) e28076, https://doi.org/10.1016/j.heliyon.2024.e28076.
[25] L. Zhang, W. Wider, M.A. Fauzi, L. Jiang, J.C.M. Tanucan, L.N. Udang, Psychological capital research in HEIs: bibliometric analysis of current and future trends, Heliyon 10 (2024) e26607, https://doi.org/10.1016/j.heliyon.2024.e26607.
[26] A.R. Karataş, H. Kazak, A.T. Akcan, E. Akkaş, M. Arık, A bibliometric mapping analysis of the literature on patent analysis, World Patent Inf. 77 (2024) 102266, https://doi.org/10.1016/j.wpi.2024.102266.
[27] S. O'Connor, Over twenty years of pedagogical research from Nurse education in practice: a bibliometric analysis from 2001 to 2023, Nurse Educ. Pract. 76 (2024) 103912, https://doi.org/10.1016/j.nepr.2024.103912.
[28] N. Fijačko, R.M. Creber, B.S. Abella, P. Kocbek, Š. Metličar, R. Greif, G. Štiglic, Using generative artificial intelligence in bibliometric analysis: 10 years of research trends from the European Resuscitation Congresses, Resuscitation plus 18 (2024) 100584, https://doi.org/10.1016/j.resplu.2024.100584.
[29] S. Ab Aziz, M.H. Mohd Nasir, A.R. Jusoh, K.F. Azman, C.A.N. Ismail, A.H. Ahmad, Z. Othman, R. Zakaria, Global research Activity on olfactory marker protein (OMP): A bibliometric and visualized analysis, Heliyon 10 (2024) e26106, https://doi.org/10.1016/j.heliyon.2024.e26106.
[30] Shankar, D.D. Azhakath, A.S. Khalil, N.S. J, M. T, S. K, Data mining for cyber biosecurity risk management – a comprehensive review, Comput. Secur. 137 (2024) 103627, https://doi.org/10.1016/j.cose.2023.103627.
[31] B. Yang, Z. Liu, S. Lacasse, X. Liang, Spatiotemporal deformation characteristics of Outang landslide and identification of triggering factors using data mining, J. Rock Mech. Geotech. Eng. (2024), https://doi.org/10.1016/j.jrmge.2023.09.030.
[32] G. Feng, M. Fan, Research on learning behavior patterns from the perspective of educational data mining: evaluation, prediction and visualization, Expert Syst. Appl. 237 (2024) 121555, https://doi.org/10.1016/j.eswa.2023.121555.
[33] F. Lu, Z. Jin, X. Zhang, An analysis of data mining techniques in software engineering database design, Procedia Comput. Sci. 228 (2023) 47–56, https://doi.org/10.1016/j.procs.2023.11.007.
[34] Z. Wang, F. Zhang, M. Ren, D. Gao, A new multifractal-based deep learning model for text mining, Inf. Process. Manage. 61 (2024) 103561, https://doi.org/10.1016/j.ipm.2023.103561.
[35] E. Senave, M.J. Jans, R.P. Srivastava, The application of text mining in accounting, Int. J. Account. Inf. Syst. 50 (2023) 100624, https://doi.org/10.1016/j.accinf.2023.100624.
[36] E. Aytaç, M. Khayet, Visual footprint of separation through membrane distillation on youtube, Data 10 (2025) 24, https://doi.org/10.3390/data10020024.
[37] M. Xiong, H. Wang, C. Che, M. Sun, Application of text mining and coupling theory to depth cognition of aviation safety risk, Reliab. Eng. Syst. Saf. 245 (2024) 110032, https://doi.org/10.1016/j.ress.2024.110032.
[38] C. Maciejewski, K. Ozierański, A. Barwiołek, M. Basza, A. Bożym, M. Ciurla, M. Janusz Krajsman, M. Maciejewska, P. Lodziński, G. Opolski, M. Grabowski, A. Cacko, P. Balsam, AssistMED project: Transforming cardiology cohort characterisation from electronic health records through natural language processing – Algorithm design, preliminary results, and field prospects, Int. J. Med. Inf. 185 (2024) 105380, https://doi.org/10.1016/j.ijmedinf.2024.105380.
[39] D.Y. Lee, N. Kim, C. Park, S. Gan, S.J. Son, R.W. Park, B. Park, Explainable multimodal prediction of treatment-resistance in patients with depression leveraging brain morphometry and natural language processing, Psychiatry Res. 334 (2024) 115817, https://doi.org/10.1016/j.psychres.2024.115817.
[40] P. Carroll, B. Singh, E. Mangina, Uncovering gender dimensions in energy policy using Natural Language Processing, Renew. Sustain. Energy Rev. 193 (2024) 114281, https://doi.org/10.1016/j.rser.2024.114281.
[41] Y. Gao, J. Wang, X. Xu, Machine learning in construction and demolition waste management: progress, challenges, and future directions, Autom. Constr. 162 (2024) 105380, https://doi.org/10.1016/j.autcon.2024.105380.
[42] E. Aytaç, Object detection and regression based visible spectrophotometric analysis: a demonstration using methylene blue solution, ADCAIJ: Adv Distrib. Comput. Artif. Intell. J. 12 (2023) e29120, https://doi.org/10.14201/adcaij.29120.
[43] E. Aytaç, Havzaların Benzerliklerini Tanımlamada Alternatif Bir Yaklaşım: Hiyerarşik Kümeleme Yöntemi Uygulaması, Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi 21 (2021) 958–970, https://doi.org/10.35414/akufemubid.870649.
[44] E. Aytaç, Forecasting Turkey's hazelnut export quantities with facebook's prophet algorithm and box-cox transformation, ADCAIJ: Adv Distrib. Comput. Artif. Intell. J. 10 (2021) 33–47, https://doi.org/10.14201/ADCAIJ20211013347.
[45] E. Aytaç, Modeling future impacts on land cover of rapid expansion of hazelnut orchards: a case study on Samsun, Turkey, Eur. J. Sustain. Develop. Res. 6 (2022) em0193, https://doi.org/10.21601/ejosdr/12167.
[46] S. Eisbach, O. Mai, G. Hertel, Combining theoretical modelling and machine learning approaches: the case of teamwork effects on individual effort expenditure, New Ideas Psychol. 73 (2024) 101077, https://doi.org/10.1016/j.newideapsych.2024.101077.
[47] E. Aytaç, Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based k-means clustering, a

demonstration on Western Black Sea Region of Turkey, Int. Soil Water Conserv. Res. 8 (2020) 321–331, https://doi.org/10.1016/j.iswcr.2020.05.002.
[48] E. Aytaç, M. Khayet, A topic modeling approach to discover the global and local subjects in membrane distillation separation process, Separations 10 (2023) 482, https://doi.org/10.3390/separations10090482.
[49] D. Zhu, X. Shen, M. Mosbach, A. Stephan, D. Klakow, Weaker than you think: a critical look at weakly supervised learning, arXiv (2023).
[50] Z. Tan, Y. Yu, J. Meng, S. Liu, W. Li, Self-supervised learning with self-distillation on COVID-19 medical image classification, Comput. Methods Programs Biomed. 243 (2024) 107876, https://doi.org/10.1016/j.cmpb.2023.107876.
[51] B. Voigt, O. Fischer, B. Schilling, C. Krumnow, C. Herta, Investigation of semi- and self-supervised learning methods in the histopathological domain, J. Pathol. Inf. 14 (2023) 100305, https://doi.org/10.1016/j.jpi.2023.100305.
[52] Q. Zhang, Z. Zhao, H. Zhou, X. Li, C. Li, Self-supervised contrastive learning on heterogeneous graphs with mutual constraints of structure and feature, Inf. Sci. 640 (2023) 119026, https://doi.org/10.1016/j.ins.2023.119026.
[53] C.F. Moreno-Garcia, C. Jayne, E. Elyan, M. Aceves-Martins, A novel application of machine learning and zero-shot classification methods for automated abstract screening in systematic reviews, Decision Anal. J. 6 (2023) 100162, https://doi.org/10.1016/j.dajour.2023.100162.
[54] E. Çelik, T. Dalyan, Unified benchmark for zero-shot Turkish text classification, Inf. Process. Manage. 60 (2023) 103298, https://doi.org/10.1016/j.ipm.2023.103298.
[55] W. Yin, J. Hay, D. Roth, Benchmarking zero-shot text classification: datasets evaluation and entailment approach, arXiv (2019), https://doi.org/10.48550/arXiv.1909.00161.
[56] H.S. Kusuma, R.A. Nida, V. Listiawati, D.E. Rahayu, L.A. Febryola, H. Darmokoesoemo, A.N. Amenaghawon, Trends on adsorption of lead (Pb) using water hyacinth: bibliometric evaluation of Scopus database, Environ. Res. 244 (2024) 117917, https://doi.org/10.1016/j.envres.2023.117917.
[57] N.Z.M. Afandi, R. Umar, N.H. Sabri, S. Safei, C.C. Lau, Revealing trends: A bibliometric analysis of 28 years of space weather event research publications using the Scopus databases (1994–2022), Adv. Space Res. 72 (2023) 5753–5766, https://doi.org/10.1016/j.asr.2023.09.029.
[58] Scopus, Scopus_Numbers. https://www.elsevier.com/products/scopus/content?dgcid=RN_AGCM_Sourced_300005030#4-titles-on-scopus, 2025 (accessed 31 January, 2025).
[59] M. Aria, C. Cuccurullo, bibliometrix: An R-tool for comprehensive science mapping analysis, J. Informet. 11 (2017) 959–975, https://doi.org/10.1016/j.joi.2017.08.007.
[60] N.J. van Eck, L. Waltman, Software survey: VOSviewer, a computer program for bibliometric mapping, Scientometrics 84 (2010) 523–538, https://doi.org/10.1007/s11192-009-0146-3.
[61] T.C. Hydrick, C. Zhang, B. Ruch, J. Wagler, K. Kumm, J.W. Harbell, W.R. Hewitt, C.C. Jadlowiec, N.N. Katariya, A.A. Moss, M.C. Nguyen, K.S. Reddy, A.L. Singer, A.K. Mathur, Declining Medicare reimbursement in abdominal transplantation from 2000 to 2021, Surgery 173 (2023) 1484–1490, https://doi.org/10.1016/j.surg.2023.01.015.
[62] M. Khayet, E. Aytaç, T. Matsuura, Bibliometric and sentiment analysis with machine learning on the scientific contribution of Professor Srinivasa Sourirajan, Desalination 543 (2022) 116095, https://doi.org/10.1016/j.desal.2022.116095.
[63] E. Aytaç, A. Fombona-Pascual, J.J. Lado, E.G. Quismondo, J. Palma, M. Khayet, Faradaic deionization technology: insights from bibliometric, data mining and machine learning approaches, Desalination (2023) 116715, https://doi.org/10.1016/j.desal.2023.116715.
[64] E. Aytaç, M. Khayet, A deep dive into membrane distillation literature with data analysis, bibliometric methods, and machine learning, Desalination 553 (2023) 116482, https://doi.org/10.1016/j.desal.2023.116482.
[65] Bibliometrix, Frequently Asked Questions. https://www.bibliometrix.org/home/index.php/faq, 2024 (accessed May 12th, 2024).
[66] E. Aytaç, N.K. Khanzada, Y. Ibrahim, M. Khayet, N. Hilal, Reverse osmosis membrane engineering: multidirectional analysis using bibliometric machine learning, data, and text mining approaches, Membranes 14 (2024) 259, https://doi.org/10.3390/membranes14120259.
[67] M. Khayet, E. Aytaç, A Glimpse into Dr Nidal Hilal's Scientific Achievements, J. Memb. Sci. Res. 10 (2024) 1999042, https://doi.org/10.22079/jmsr.2023.1999042.1598.
[68] S. Ngassa, K. Kilulya, R. Masalu, Scientific landscape on phthalates biodegradation research: a bibliometric and scientometric study Tanzania, J. Sci. 49 (2023) 842–858, https://doi.org/10.4314/tjs.v49i4.6.
[69] I. Hanafi, Z. Zahir, J. El Bahaoui, F.M. Cabrera, K. Haboubi, The evolution of net zero emissions research: a bibliometric investigation, Sci. Afr., 24 (2024) e02228, https://doi.org/10.1016/j.sciaf.2024.e02228.
[70] X. Meng, W. Cui, Q. Liang, B. Zhang, Y. Wei, Trends and hotspots in tea and Alzheimer's disease research from 2014 to 2023: a bibliometric and visual analysis, Heliyon 10 (2024) e30063, https://doi.org/10.1016/j.heliyon.2024.e30063.
[71] P. Chadha, P. Yadav, M. Sharma, Y. Sharma, R. Nangia, K. Sharma, V. K. Maksudovna, Global trends in waste materials: a bibliometric analysis, Mater. Today Proc. (2023), https://doi.org/10.1016/j.matpr.2023.10.059.
[72] C.Z. Li, L. Zhang, X. Liang, B. Xiao, V.W.Y. Tam, X. Lai, Z. Chen, Advances in the research of building energy saving, Energy Build., 254 (2022) 111556, https://doi.org/10.1016/j.enbuild.2021.111556.
[73] E. Aytaç, J. Contreras-Martínez, M. Khayet, Mathematical and computational modeling of membrane distillation technology: a data-driven review, Int. J. Thermofluids 21 (2024) 100567, https://doi.org/10.1016/j.ijft.2024.100567.
[74] R. Wang, M. Lewis, R. Zheng-Pywell, J. Julson, M. Smithson, H. Chen, Using the H-index as a factor in the promotion of surgical faculty, Heliyon 8 (2022) e09319, https://doi.org/10.1016/j.heliyon.2022.e09319.
[75] E. Roldan-Valadez, S.Y. Salazar-Ruiz, R. Ibarra-Contreras, C. Rios, Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics, Ir. J. Med. Sci., 188 (2019) 939–951, https://doi.org/10.1007/s11845-018-1936-5.
[76] M. Schreiber, How to derive an advantage from the arbitrariness of the g-index, J. Informet. 7 (2013) 555–561, https://doi.org/10.1016/j.joi.2013.02.003.
[77] J. Demšar, T. Curk, A. Erjavec, Č. Gorup, T. Hočevar, M. Milutinovič, M. Možina, M. Polajnar, M. Toplak, A. Starič, M. Štajdohar, L. Umek, L. Žagar, J. Žbontar, M. Žitnik, B. Zupan, Orange: data mining toolbox in python, J. Mach. Learn. Res. 14 (2013) 2349–2353.
[78] E. Aytaç, Exploring electrocoagulation through data analysis and text mining perspectives, Environ. Eng. Manage. J. 21 (2022) 671–685.
[79] A.P. Castro, G.A. Wainer, W.P. Calixto, Weighting construction by bag-of-words with similarity-learning and supervised training for classification models in court text documents, Appl. Soft Comput. 124 (2022) 108987, https://doi.org/10.1016/j.asoc.2022.108987.
[80] M. Islam, H. Zunair, N. Mohammed, CosSIF: Cosine similarity-based image filtering to overcome low inter-class variation in synthetic medical image datasets, Comput. Biol. Med. 172 (2024) 108317, https://doi.org/10.1016/j.compbiomed.2024.108317.
[81] Facebook, bart-large-mnli. https://huggingface.co/facebook/bart-large-mnli, 2019 (accessed March 28, 2024).
[82] New_York_University, Machine Learning for Language. https://huggingface.co/datasets/nyu-mll/multi_nli, 2024 (accessed March 28, 2024).
[83] M. Laurer, DeBERTa-v3-base-mnli-fever-anli. https://huggingface.co/MoritzLaurer/DeBERTa-v3-base-mnli-fever-anli, 2024 (accessed March 29, 2024).
[84] J. Hartmann, Emotion English DistilRoBERTa-base. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base, 2022 (accessed March 30, 2024).
[85] A. Politano, G. Di Profio, V. Sanna, E. Curcio, Thermoplasmonic membrane distillation, Chem. Eng. Trans. 60 (2017) 301–306.
[86] M. Fujiwara, M. Kikuchi, Solar desalination of seawater using double-dye-modified PTFE membrane, Water Res. 127 (2017) 96–103, https://doi.org/10.1016/j.watres.2017.10.015.
[87] A. Politano, P. Argurio, G. Di Profio, V. Sanna, A. Cupolillo, S. Chakraborty, H. A. Arafat, E. Curcio, Photothermal membrane distillation for seawater desalination, Adv. Mater. 29 (2017) 1603504, https://doi.org/10.1002/adma.201603504.
[88] Q. Guo, Y. Huang, M. Xu, Q. Huang, J. Cheng, S. Yu, Y. Zhang, C. Xiao, PTFE porous membrane technology: a comprehensive review, J. Membr. Sci. 664 (2022) 121115, https://doi.org/10.1016/j.memsci.2022.121115.
[89] B. Gong, H. Yang, S. Wu, J. Yan, K. Cen, Z. Bo, K.K. Ostrikov, Superstructure-enabled anti-fouling membrane for efficient photothermal distillation, ACS Sustain. Chem. Eng. 7 (2019) 20151–20158, https://doi.org/10.1021/acssuschemeng.9b06160.
[90] Y. Wan, H. Yu, R. Li, H. Chang, D. Qu, F. Qu, Membrane distillation localized heating in high-salinity wastewater treatment: mechanisms, performance and challenges, Resour. Conserv. Recycl. 215 (2025) 108086, https://doi.org/10.1016/j.resconrec.2024.108086.
[91] A. Alsaati, A.M. Marconnet, Energy efficient membrane distillation through localized heating, Desalination 442 (2018) 99–107, https://doi.org/10.1016/j.desal.2018.05.009.
[92] D. Ji, Y. Lin, X. Guo, B. Ramasubramanian, R. Wang, N. Radacsi, R. Jose, X. Qin, S. Ramakrishna, Electrospinning of nanofibres, Nat. Rev. Methods Primers 4 (2024) 1, https://doi.org/10.1038/s43586-023-00278-z.
[93] A. Haider, S. Haider, I.-K. Kang, A comprehensive review summarizing the effect of electrospinning parameters and potential applications of nanofibers in biomedical and biotechnology, Arabian J. Chem. 11 (2018) 1165–1188, https://doi.org/10.1016/j.arabjc.2015.11.015.
[94] R.L. Atenstaedt, Word cloud analysis of historical changes in the subject matter of public health practice in the United Kingdom, Public Health 197 (2021) 39–41, https://doi.org/10.1016/j.puhe.2021.06.010.
[95] J. Ravi, M.H.D. Othman, T. Matsuura, M. Ro'il Bilad, T.H. El-badawy, F. Aziz, A. F. Ismail, M.A. Rahman, J. Jaafar, Polymeric membranes for desalination using membrane distillation: A review, Desalination 490 (2020) 114530, https://doi.org/10.1016/j.desal.2020.114530.
[96] S. Dash, A.S. Kalamdhad, Science mapping approach to critical reviewing of published literature on water quality indexing, Ecol. Indic. 128 (2021) 107862, https://doi.org/10.1016/j.ecolind.2021.107862.
[97] L. Liao, L. Quan, C. Yang, L. Li, Knowledge synthesis of intelligent decision techniques applications in the AECO industry, Autom. Constr. 140 (2022) 104304, https://doi.org/10.1016/j.autcon.2022.104304.
[98] Y. Ibrahim, E. Aytaç, N.K. Khanzada, M. Khayet, N. Hilal, The role of feed spacers in membrane technology: 45 years of research, Sep. Purif. Technol. 357 (2025) 130109, https://doi.org/10.1016/j.seppur.2024.130109.
[99] X. Wu, Q. Jiang, D. Ghim, S. Singamaneni, Y.-S. Jun, Localized heating with a photothermal polydopamine coating facilitates a novel membrane distillation process, J. Mater. Chem. A 6 (2018) 18799–18807, https://doi.org/10.1039/C8TA05738A.
[100] J. Tollefson, China declared world's largest producer of scientific articles, Nature 553 (2018) 390, https://doi.org/10.1038/d41586-018-00927-4.

[101] Y. Xie, C. Zhang, Q. Lai, China's rise as a major contributor to science and technology, Proc. Natl. Acad. Sci. 111 (2014) 9437–9442, https://doi.org/10.1073/pnas.1407709111.

[102] Q. Xie, R.B. Freeman, China's Overwhelming Contribution to Scientific Publications, VoxChina (2018).

[103] J.R. Jim, M.A.R. Talukder, P. Malakar, M.M. Kabir, K. Nur, M.F. Mridha, Recent advancements and challenges of NLP-based sentiment analysis: a state-of-the-art review, Nat. Lang. Process. J., 6 (2024) 100059, https://doi.org/10.1016/j.nlp.2024.100059.

[104] S. Peng, L. Cao, Y. Zhou, Z. Ouyang, A. Yang, X. Li, W. Jia, S. Yu, A survey on deep learning for textual emotion analysis in social networks, Digital Commun. Networks 8 (2022) 745–762, https://doi.org/10.1016/j.dcan.2021.10.003.

[105] A. Montoyo, P. Martínez-Barco, A. Balahur, Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments, Decis. Support Syst. 53 (2012) 675–679, https://doi.org/10.1016/j.dss.2012.05.022.

[106] L. Yuan, S. Gao, P. Pan, CTSARF: A Chinese text similarity analysis model based on residual fusion, Neurocomputing 559 (2023) 126801, https://doi.org/10.1016/j.neucom.2023.126801.

[107] M. Wang, H. Sakaji, H. Higashitani, M. Iwadare, K. Izumi, Discovering new applications: Cross-domain exploration of patent documents using causal extraction and similarity analysis, World Patent Inf. 75 (2023) 102238, https://doi.org/10.1016/j.wpi.2023.102238.