Journal of Informetrics 18 (2024) 101543

Contents lists available at ScienceDirect

Journal of Informetrics

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

Research Paper

Are the bibliometric growth patterns of excellent scholars similar? From the analysis of ACM Fellows

Xianzhe Peng, Huixin Xu, Jin Shi *

School of Information Management, Nanjing University, Nanjing 210023, China

ARTICLE INFO

Keywords:
Excellent scholars
Growth pattern
Scientific productivity
Academic impact
Academic comprehensive ability
Dynamic time wrap

ABSTRACT

The growth of excellent scholars provides paradigmatic career paths leading to research success, as their research capabilities ultimately manifest as fluctuations in bibliometric indexes. Examining the commonalities in the trajectories of these bibliometric indexes displays the universal characteristics of their growth process, and furtherly shows exemplary routes to scientific success. In this study, we examine 287 excellent scholars elected as ACM Fellows in the field of computer science from 2016s to 2020s. Based on their changes in productivity, impact, and comprehensive abilities, we categorize them into three categories, four categories, and six categories, respectively. Most of these scholars experience continuous growth in productivity during the early development stages, maintaining a prolonged period of high productivity in the mid-later maturity stages. Their impact rises smoothly and consistently, while the growth of their comprehensive abilities is relatively gradual, remaining at above-average levels in the mid-later maturity stages. Furthermore, the level of recognition within the scientific research community varies for different categories of scholars, and there are also differences in the growth patterns between scholars from Asia and those from Western regions.

1. Introduction

In the 21st century characterized by the prominence of human resource, the cultivation and identification of excellent scholars has become a focal point in the field of scientific research. It is imperative to investigate methods for effectively identifying potential researchers based on their scholarly attributes and development models (Liu et al., 2021; Sinatra et al., 2016). Consequently, extensive research has been conducted on identifying scholarly characteristics and career development trajectories (Feichtinger et al., 2019; Ram et al., 2022; Sun et al., 2023; Way et al., 2017; Way et al., 2019), which provides valuable insights for evaluating existing researchers and allocating resources. Due to the gradual and intricate nature of scholars' development (Wang et al., 2019), it is disputed to ensure scientific validity by identifying elite characteristics from a static perspective. This has led recent studies to adopt a developmental perspective in elite identification, and these studies has gained insights into the growth trajectories of different types of scholars, such as the randomness in the productivity changes of most scholars (Zhang, LaBerge, Way, Larremore, & Clauset, 2023) and several dynamic indicators of evaluating scholars' research abilities in previous studies (Fiala, 2014; Pan & Fortunato, 2014).

To explore and nurture these promising scholars, there were numerical investigations about the exclusive characteristics of successful scientists. Investigation findings indicate that successful scholars possess diverse characteristics, including outstanding collaborators (Li et al., 2019; Zeng et al., 2022), affiliations with prestigious institutions (Zhang et al., 2022), and a consistent pattern of

* Corresponding author.
E-mail address: shijin@nju.edu.cn (J. Shi).

https://doi.org/10.1016/j.joi.2024.101543
Received 2 November 2023; Received in revised form 4 May 2024; Accepted 7 May 2024
Available online 18 May 2024
1751-1577/

producing high-impact works (Liu et al., 2018). Unfortunately, from the analysis of the outstanding scholars' academic development patterns based on the above features, it is found that there is significant variability (Li et al., 2019; Sinatra et al., 2016), which means that these features cannot be considered as sufficient conditions to identify scholars with potential. However, the common characteristic among outstanding scholars with these features is high productivity or high impact. It is currently unknown whether outstanding scholars' growth patterns have commonalities when analyzed from the growth of academic ability. Existing research has only shown the general trend of academic productivity and impact of a general group of scholars (Győrffy et al., 2020), and has not analyzed whether outstanding scholars have similar patterns of changes in academic capabilities from this perspective. If outstanding scholars have similar academic growth patterns, it suggests that their developmental processes possess universal stage-specific characteristics. Based on these features, timely academic support and assistance can be provided to them, serving as a reference basis for adjusting the funding system.

To address the deficiencies of existing research, this study evaluates the changes in academic productivity, impact, and overall capability of outstanding scholars from the perspective of bibliometrics. The goal is to determine if they have similar academic growth patterns, analyze the combined forms and distinctive characteristics of these patterns, and provide a micro and comprehensive perspective on understanding the academic growth process of excellent scholars. Specifically, this study focuses on 287 excellent scholars who were elected as ACM Fellows in the field of computer science between 2016s and 2020s, and these scholars have a minimum scientific research experience of 20 years. It examines their trajectory of changes in publication volume, h-index, and p-index from a retrospective perspective, aiming to determine whether these outstanding scholars share similar patterns and experiences in terms of their productivity, impact, and overall abilities. Furthermore, clustering analysis is employed to categorize scholars based on the similarity of their trajectories to gain insights into the diverse paths leading to academic success.

2. Related works

Scientific elites deserve considerable attention, not just due to their outstanding contributions to research, which have earned them reputation and influence, but also because they offer established paradigms for attaining scientific success (Li et al., 2020). Gaining insights into the career development patterns of these excellent researchers across diverse scientific domains proves invaluable for scientists and decision-makers aspiring to identify and cultivate their individual professional paths and institutions (Fortunato et al., 2018). Consequently, numerous quantitative studies have extensively relied on expansive datasets to scrutinize patterns associated with the academic productivity and impact progression of these elites (Fortunato, 2014; Jones & Weinberg, 2011), the evolutionary dynamics of scientific collaborative networks (Milojević, 2014; Petersen, 2015), as well as the transitions in job positions within their respective organizations (Petersen, 2018; Sun et al., 2023). These patterns ultimately forge the career trajectories of scientific elites.

In recent years, there has been significant research focused on analyzing the career trajectories of scientists. This research is motivated by the necessity to formulate policies that enhance scholars' career advancements and identify key characteristics that contribute to successful scientific careers (Bornmann & Williams, 2017; Vinkenburg et al., 2020). While indications of research success may encompass prestigious awards and high-ranking positions, they ultimately rely on a sufficient level of research productivity and impact (Nielsen & Andersen, 2021). As a result, previous studies have employed bibliometric methods to examine potential characteristics associated with the academic progression of exceptional scientists, with their citation counts serving as an evaluation criterion. These characteristics include early academic education and publication (Laurance et al., 2013), prolific first-author publications in high impact factor journals (van Dijk et al., 2014), and a high publication count particularly in top journals during their early careers (Lee, 2019; Lindahl, 2018; Schilling & Green, 2011). However, whether these characteristics can be considered exclusive features of outstanding scholars, such as sustained high research productivity (Li et al., 2020; Sinatra et al., 2016), still requires further investigation and validation.

Distinguishing unique characteristics of outstanding scholars provides an effective method for identifying promising young scholars (Haunschild & Bornmann, 2023). Existing research in this regard is primarily based on comprehensive scholar datasets, comparing aspects such as scientific collaborative networks, bibliometric indicators, and personal social attributes to uncover the characteristics of scholars who have had successful careers (Jin et al., 2021; Li et al., 2020). Among these approaches, comparative analysis through bibliometric methods has gained wide recognition in academia when examining the differences between ordinary scholars and outstanding scholars (Bornmann & Williams, 2017; Lindahl, 2018). However, previous research has predominantly focused on extensive datasets encompassing ordinary scholars, utilizing primarily comprehensive analysis methods encompassing multiple aspects (Li et al., 2019; Liu et al., 2018; Ram et al., 2022). Even studies specifically focusing on the career development trajectories of exceptional scientists lack a systematic analysis of scholars' growth paths from a bibliometric perspective (Jin et al., 2021; Li et al., 2020).

Analyzing the similarities in the career trajectories of excellent scholars can help to uncover the universal characteristics of their growth process, showcasing the typical paths that researches take towards scientific success. Bibliometric indicators, which quantitatively measure scholars' academic abilities, can reveal their developmental trajectories over time. This offers an opportunity to investigate whether excellent scholars share common patterns of growth. Consequently, this study focuses on outstanding scholars who were elected as ACM Fellows in the field of computer science between 2016 and 2020. By applying the Dynamic Time Warping (DTW) method to calculate the temporal similarity of their publication count, h-index, and p-index, and utilizing the K-means clustering algorithm to analyze the similarity of these curves, this research uncovers diverse growth patterns exhibited by excellent scholars from various regions and backgrounds. These findings provide a detailed and comprehensive insights into the academic growth process of excellent scholars in the field of computer science.

3. Data and methodology

3.1. Dataset

The Association for Computing Machinery (ACM), established in 1947, is the largest scientific and educational computing society worldwide (https://www.acm.org) (Fernandes et al., 2022). The ACM Fellow title, established in 1994, is the most prestigious member grade and recognizes the top 1 % of its members for their outstanding accomplishments in the computer science field. The selection criteria for this title have always been that scholars need to have a lasting impact on the computer science field in terms of technical and leadership contributions, and this impact can be evidenced by publications, products, awards, or other artifacts that are publicly recognized as worthy of merit (https://awards.acm.org/fellows/nominations). Previous research has defined excellent scholars as those ranking in the top 1 % of scholars in their research field and age group (Tol, 2013; Van Leeuwen et al., 2003). Therefore, scholars who have been awarded the title of ACM Fellow can be considered excellent scholars in the field of computer science. As of October 2023, 1444 scholars have been awarded the title of ACM Fellow. This study aims to understand the complete academic growth patterns of contemporary excellent scholars. In selecting the sample, it is important to include scholars whose attainment of the ACM Fellow title is as close to the present as possible, without being too late, as scholars who receive the title too late may have incomplete academic experiences at the present time. Therefore, this study focused on 316 scholars who were awarded the ACM Fellow title between 2016s and 2020s. Measuring the length of a scholar's research experience by the time span between the publication of the scholar's first and last papers, the selection criteria for the scholars include having at least 20 years of research experience (Fernandes et al., 2022). The filtering process yielded a list of 287 scholars, including their names, regions, and year of election. These selected scholars are predominantly from North America, Europe, and Asia, consistent with the geographical distribution of the overall sample (See Fig. 1A), showing a certain level of representativeness.

Based on above list of 287 scholars, the next step is to obtain their list of publications. DBLP, a treasure trove of open bibliographic data, offers an abundance of information regarding significant publications in journals and conference proceedings in the field of computer science (Ley, 2009; Rosenfeld, 2023). As of now, it indexes over 7 million publications from over 3.4 million authors, published in more than 6000 conferences and 1800 journals. For this fact, we utilized the DBLP website which was widely adopted in previous studies (Fernandes & Monteiro, 2017; Kim, 2018). In addition, it performs better in scholar disambiguation tasks compared to the Scopus and Web of Science databases (Kim, 2018), resulting in more comprehensive and accurate publication data collected from scholars. This study obtained the scholar homepages of 287 ACM Fellows on the DBLP platform (https://dblp.org/search/author/api)

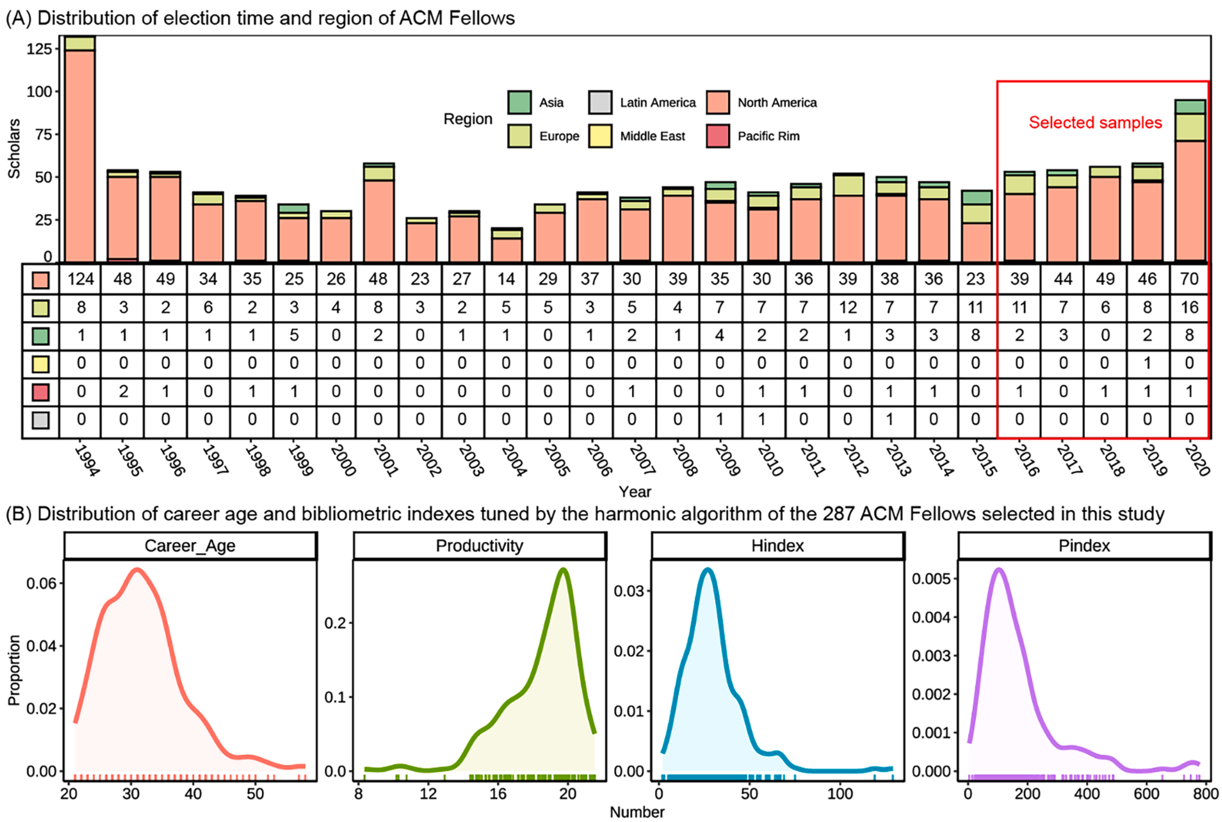

Fig. 1. Descriptive statistics of all ACM Fellows and ACM Fellows filtered in this study.

through manual search. Subsequently, by using the public Application Programming Interface (API) provided by DBLP, a total of 52, 429 publication entries for them were accessed. By constructing search queries in Web of Science database (Garfield, 1972) based on the titles, publication years, and DOIs included in the bibliographic records, we can obtain citation information of these publications from DBLP. Specifically, for 29,489 publications with DOIs, we can directly obtain citation statistics for these publications based on their DOIs. For other publications without DOIs, we primarily used the titles and publication years to construct search queries in the Web of Science platform to obtain their citation information on that platform. Through this search method and combined with manual verification, we obtained a total of 52,429 publications' citation statistics on the Web of Science platform. Although the accuracy of DBLP in disambiguating author names is high, with and recall both above 90 %, there may still be a small number of errors, especially in dealing with the disambiguation of Chinese scholars' names, which are prone to homonyms (Kim, 2018). Therefore, we have individually checked the bibliographic and citation statistics of Chinese scholars' publications, and randomly sampled and verified the bibliographic and corresponding citation statistics of 30 scholars, all of which showed no errors or omissions. Thus, we can consider the collected data to have a certain level of accuracy.

In general, the productivity of a scholar is often measured by the volume of their publications. The scholar's h-index is defined as the number of his or her papers with citation number $\geq h$, giving particular emphasis to the high-impact papers (Hirsch, 2005). This metric serves to quantify the influence and number of a scholar's high-quality academic publications and is widely utilized to gauge scholarly impact (Saad, 2010). To provide a more comprehensive assessment of a scholar's academic productivity and impact, the p-index has been suggested (Prathap, 2010). By considering the total output of a scholar's papers and the number of citations, it takes into account all of the scholar's publications, thereby achieving a balance between quantity and quality of these publications and offering a more holistic evaluation of the scholar's comprehensive academic ability (Prathap, 2010; 2011). To assess the academic productivity, impact, and comprehensive abilities of scholars, we employ various metrics including publication count, h-index (Hirsch, 2005), and p-index (Prathap, 2010). By analyzing the temporal variation of these bibliometric indexes for scholars, dynamic indexes for evaluating scholars are derived, such as count (y), h (y), and p (y) as shown in Table 1. Specifically, a scholar's annual academic productivity indeed shows significant fluctuation and requires attention to year-by-year changes. Therefore, count (y) represents the total number of papers published by a scholar in year y. On the other hand, the academic impact and comprehensive ability of scholars are influenced by the cumulative effect of previously published academic works (Raan, 2006; Yang et al., 2022). Consequently, h (y) and p (y) are calculated based on all the publications a scholar has cumulatively published up to and including the career year y, which is similar to previous research (Egghe, 2013; Wu et al., 2011). The changes in count (y), h (y), and p (y) over time, as calculated, display the scholar's complete academic growth trajectory.

Additionally, considering that scholars may hold different author positions and collaborate with varying numbers of co-authors in publications, their contributions within each publication may vary (Fernandes et al., 2022). There are various approaches to calculate author contributions in publications, such as geometric counting (Egghe et al., 2000), arithmetic counting (also called proportional or positionwise counting) (Kalyane & Vidyasagar Rao, 1995; Van Hooydonk, 1997), and the harmonic algorithm (Hagen, 2008). However, the geometric counting method underestimates the contributions of authors ranked lower in the publication (Jian & Xiaoli, 2013), while the arithmetic counting method elevates the contributions of some authors at the cost of undervaluing primary authors (Hagen, 2008). In contrast, the harmonic algorithm produces author contributions to publications that robustly align with empirical data, making it more accurate (Hagen, 2010; Liu & Fang, 2012). Therefore, in this study, we utilized the following harmonic algorithm to determines the proportion of scholars' contributions (W_i) across various publications:

$$W_i = \frac{1/i}{1 + 1/2 + 1/3 + \ldots + 1/n}$$

In the formula, W_i represents the contribution of the i-th author to the publication. Here, i denotes the position of the author in the author list, and n represents the total number of authors for the paper. The weight assigned to authors decreases as their position in the author list increases. Furthermore, as the number of co-authors increases, the contribution value assigned to individual authors decreases. Based on the proportion of a scholar's contribution to each publication, when a scholar publishes an additional article, their total publication count should be increased by the contribution proportion multiplied by 1 for that article. Likewise, the scholar's total

Table 1
Bibliometric indexes for evaluating various abilities of scholars in static and dynamic assessments.

Academic ability	Evaluation index	Definition of index
Productivity	count	The total publications from scholar
	count (y)	The number of publications by scholars in the y-th year of their academic careers
Impact	h	$h = max\{min(c(k), k) \mid k = 1, 2, \ldots, n\}$ c(k) represents the number of citations for the k-th paper in the ranking sequence of paper citation counts in descending order, and k indicates its position in this sequence, and n represents the total publications from scholar
	h (y)	The h-index calculated based on the cumulative publications and citations of scholars up to the y-th year of their academic careers
Comprehensive ability	p	$p = (C^2/N)^{1/3}$ C and N respectively represent the total citation count and the number of publications from scholar.
	p (y)	The p-index calculated based on the cumulative publications and citations of scholars up to the y-th year of their academic careers

citation count should be increased by the contribution proportion multiplied by the number of citations for that article. This data processing calculates the proportion of each scholar's contribution to the publication and provides a more accurate representation of their individual impact. The application of the harmonic algorithm allows for the correction of the number of papers and citations, resulting in the recalculation of the h-index and p-index. Since the bibliography data obtained from DBLP does not include information about corresponding authors, the harmonic algorithm may underestimate the contributions of excellent scholars as corresponding authors or supervisors to publications (Buehring et al., 2007). However, the research data for this study all come from publications in the field of computer science. In this type of publication, the convention for authorship typically involves the first position in the author list being occupied by the researchers who have made the most intellectual contribution, while the last position is occupied by the researchers who have made the fewest contribution or are the most senior (Fernandes et al., 2022). Therefore, the harmonic algorithm used in this study to assess author contributions to publications based on their position order is, in most cases, reasonable. As a result, adjustments made to the number of publications, h-index, and p-index more accurately reflect the academic abilities of the scholars.

Finally, this study selected 287 scholars from a total sample of 1444 ACM Fellows, all of whom had at least 20 years of academic career experience. Their geographical distribution is consistent with the overall sample, with scholars from North America, Europe, and Asia in decreasing order, with North America being the predominant region (See Fig. 1A). As shown in Fig. 1B, the academic career

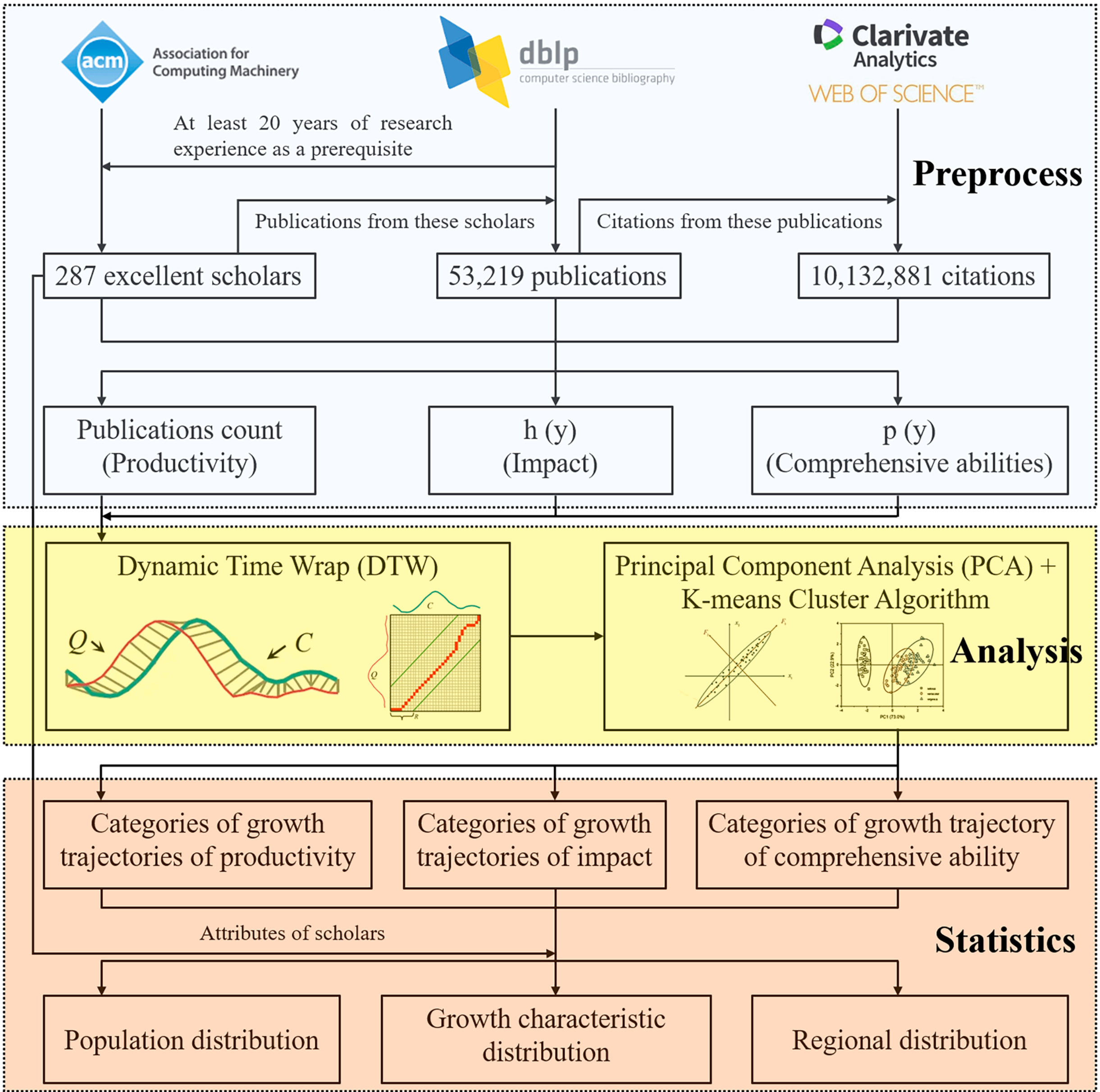

Fig. 2. Research framework of this study.

ages of the scholars range from 20 to 58, with the majority concentrated around 30. After adjustment with the harmonic algorithm, the number of publications for the scholars ranged from 8 to 22, with a concentration around 19. The h-index ranged from 2 to 129, with a concentration around 30, and the p-index ranged from 4 to 777, with a concentration around 110.

3.2. Methods

By normalizing the number of publications, h-index and p-index of scholars, the study obtained the career development trajectory of scholars. After obtaining the developmental trajectories of each scholar, we utilize a computational method known as DTW to assess the similarity between these trajectories. DTW is a flexible method for measuring the similarity of time series, focusing on the morphological changes within trajectories. It has the capability to handle time series of varying lengths, address issues such as time drift and misalignment, and effectively handle noise and distortion within the sequences (Mueen et al., 2018). The fundamental concept behind DTW involves aligning two time series by mapping them onto a two-dimensional grid and identifying the shortest path on the grid, which minimizes the sum of distances between the corresponding points along the path (Rakthanmanon et al., 2012).

After obtaining the minimum distance matrix of each scholar's growth trajectory calculated using the DTW method (See more calculational detail in Fig. 1S), this study employs principal component analysis (PCA) to extract a two-dimensional representation of each scholar's growth trajectory. Subsequently, the K-means clustering algorithm is used to partition the scholars into categories, and the optimal number of clusters is determined by calculating within-cluster sum of squares (WCSS) (Zhao & Fränti, 2014) and silhouette coefficient (Rousseeuw, 1987). The silhouette coefficient represents the ratio of the distance within clusters to the distance between clusters, varying between -1 and 1. The silhouette coefficient closer to 1 indicates better clustering, while the silhouette coefficient closer to -1 suggests possible misclassification. The WCSS indicates the distance between the samples within the cluster and the cluster centroid. The higher WCSS suggests more similarity among the samples within the cluster, indicating better clustering performance. Therefore, the optimal number of clusters selected in this study needs to balance a sufficiently high silhouette coefficient and a sufficiently small sum of squared distances within clusters, for ensuring that the sum of distance values within the same category of scholars' growth trajectories is minimized, while maximizing the sum of distance values between different categories.

3.3. Framework

Overall, our research framework can be divided into three parts (see Fig. 2). The first part is the data preprocessing module, which involves collecting and preprocessing scholars' publications and citations to obtain time series on their publication count, h (y), and p (y). The second part is the data analysis module, where we calculate the distances between sequences using the DTW algorithm and classify scholars' growth trajectories using PCA and the K-means clustering algorithm. The third part is the category statistics section, where we discuss the geographical distribution, growth characteristic distribution, and population distribution of scholars with different growth trajectory categories, summarizing the growth patterns of outstanding scholars.

Based on this research framework, this study aims to address the following questions:

(1) Do excellent scholars exhibit consistent growth trajectories in their academic abilities, including academic productivity, impact, and overall capability? What are the specific similar growth trajectories of academic abilities?
(2) How are the growth trajectories of academic productivity, impact, and overall capability of excellent scholars combined with each other?
(3) Do excellent scholars with different growth patterns show regional differences? Is there a difference in the ease of peer recognition they receive?

4. Results

4.1. The distance of growth trajectories of different indexes

After analyzing the trajectory curves of publication count, h-index, and p-index for different scholars via the DTW method, this study obtains three types of distance matrices that measures the similarity between the bibliometric trajectories of these scholars (see Fig. 2S). In Fig. 2S, the color of each square represents the distance between the bibliometric curves of the corresponding scholars indicated by their respective horizontal and vertical coordinates. This distance is calculated based on the similarity of the bibliometric curves using the DTW method. The redder the color, the closer the distance between the trajectories of their bibliometric index changes, indicating a more similar growth pattern in that bibliometric index. In detail, the distances between the growth trajectories of publication count for excellent scholars mostly fall within the range of 0 to 5. The distances for h-index growth trajectories are even closer, typically ranging from 0 to 2. Besides, the distances for p-index growth trajectory curves are slightly larger, generally staying within the range of 0 to 10. In summary, the distances of these three types of growth trajectories are relatively small, indicating a certain degree of convergence in the bibliometric growth path among different excellent scholars.

4.2. The categories of bibliometric growth patterns

To facilitate the classification of scholars' bibliometric growth patterns based on the morphological changes of curves, this study employs the PCA method to extract two principal components from distance matrices of three different types of bibliometric growth paths. Subsequently, the K-means clustering algorithm is utilized to determine the categories of curve variations. Specifically, the range of cluster numbers is set between 2 and 50, and the optimal number of clusters for different types of trajectories is determined based on the WCSS and silhouette coefficient (see Fig. 3). To ensure the minimum WSSC, the maximum silhouette coefficient and the small number of clusters, the optimal cluster numbers for the growth paths of productivity, h-index, and p-index should be specified as 3, 4, and 6, respectively. Using these optimal cluster numbers, this study employs the K-means clustering method to analyze the two principal components extracted from different curves, resulting in scatter plots of labeled categories (see Fig. 4). In this plot, bibliometric trajectories of scholars within the same category exhibit close distances, while distinct boundaries are observed between different categories, demonstrating the effectiveness of the K-means clustering algorithm.

The study statistically depicts bibliometric growth patterns for various categories based on above clustering results (see Fig. 5). The peak of the scholars' bibliometric index of marks the most productive year of their academic career (Yair & Goldstein, 2020), after which their academic status typically advances to a relatively stable stage (Kwiek & Roszka, 2023), serving as a crucial turning point in their academic career. In Fig. 5, the gray lines within each box represent the trajectory of bibliometric changes for individual scholars in that category, while the red line represents the average trajectory of bibliometric changes for all scholars in that category. The positions of the peak values of various indicators divide the growth processes of a scholar's academic productivity, influence, and comprehensive ability into early development stages and mid-later maturity stages, denoted by green and yellow backgrounds, respectively. The time span and variation patterns of the trajectories in the two stages exhibit distinct characteristics across different categories. The results show that outstanding scholars' productivity growth patterns are A-II, A-I, and A-III in descending order of population statistics, h-index growth patterns are H-II, H-I, H-IV, and H-III in descending order of population statistics, and p-index growth patterns are P-IV, P-I, P-V, P-III, P-VI, and P-II in descending order of population statistics. Besides, the form of the growth curves of different categories differs significantly.

In the analysis of scholars' growth trajectory from a morphological perspective, specific variations have been observed in the curves of productivity, h-index, and p-index across different scholar categories. By examining the morphological characteristics of productivity curves within these categories, we discovered a decreasing rate of productivity increase during the developmental stage, with the order being A-III, A-II, and A-I. However, notable differences in productivity changes among the categories emerge during the maturity stage. Scholars in category A-I don't exhibit a significant decline in productivity, whereas those in categories A-II and A-III experience a noticeable decrease. By employing a similar analytical approach, we compare the trajectory of h-index changes among scholars across different categories. In the developmental stage, the rate of h-index increase follows the order: H-II, H-III, H-I, H-IV. The same analytical approach is applicable to comparing the trajectory of p-index changes among scholars in different categories. Scholars in categories P-II, P-III, and P-VI witness a rapid increase in p-index during the developmental stage, followed by a sharp decline. However, the baseline for the decline differs across categories, with the p-index decline baseline decreasing in the following order: P-III, P-VI, P-II. Scholars in the P-I, P-IV, and P-V categories have a slower increase in p-index during the developmental stage and only experience a minor decline in the subsequent maturity stage.

4.3. The characteristic distributions from different categories of scholars

4.3.1. The duration of development and maturity stages from different categories of scholars

After classifying scholars based on their change trajectories of productivity, h-index and p-index, this study proceeds to calculate the duration of the development stage and maturity stage for scholars in different categories (see Fig. 6). Analyzing the changes in scholars' productivity, we observe that the development period for A-I, A-II, and A-III scholars consecutively decreases, ranging from 13 to 30 years, while the maturity period remains relatively constant at around 12 years, generally shorter than the development period. Moving on to the h-index changes, the maturity period for H-I, H-II and H-IV scholars remains stable at approximately 7 years, whereas scholars in the H-III category have a slightly longer maturity period, averaging about 14.5 years. Additionally, the growth period for H-IV, H-I, H-II and H-III scholars successively shortens, ranging from 14 to 30 years, all of which exceed the development period. Examining the changes in the p-index for scholars in different categories, we find that the maturity period is generally longer than the development period. Specifically, the development period for P-II, P-III, and P-VI scholars is very short, around 2 years, while the development period for P-V, P-I, and P-IV scholars exceeds that of the previous three categories, ranging from approximately 5 to 14 years. Besides, the maturity duration of scholars follows the order of P-IV, P-I, P-V, P-III, P-II, and P-VI, gradually increasing.

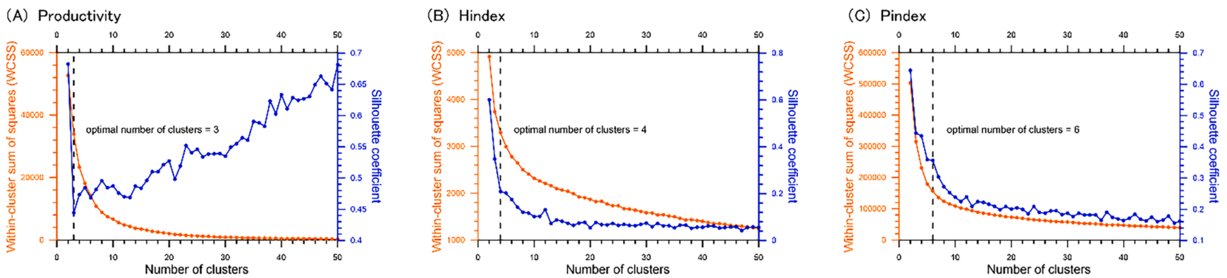

Fig. 3. Determination of optimal clustering number of bibliometric growth trajectories.

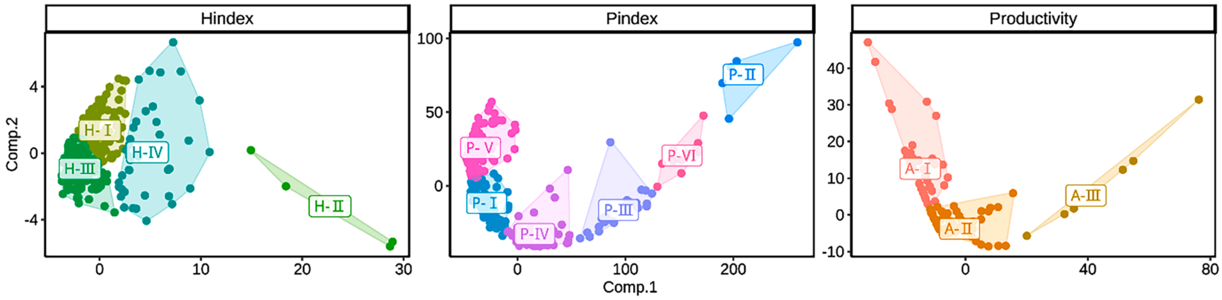

Fig. 4. K-means clustering of the growth trajectories based the results of PCA.

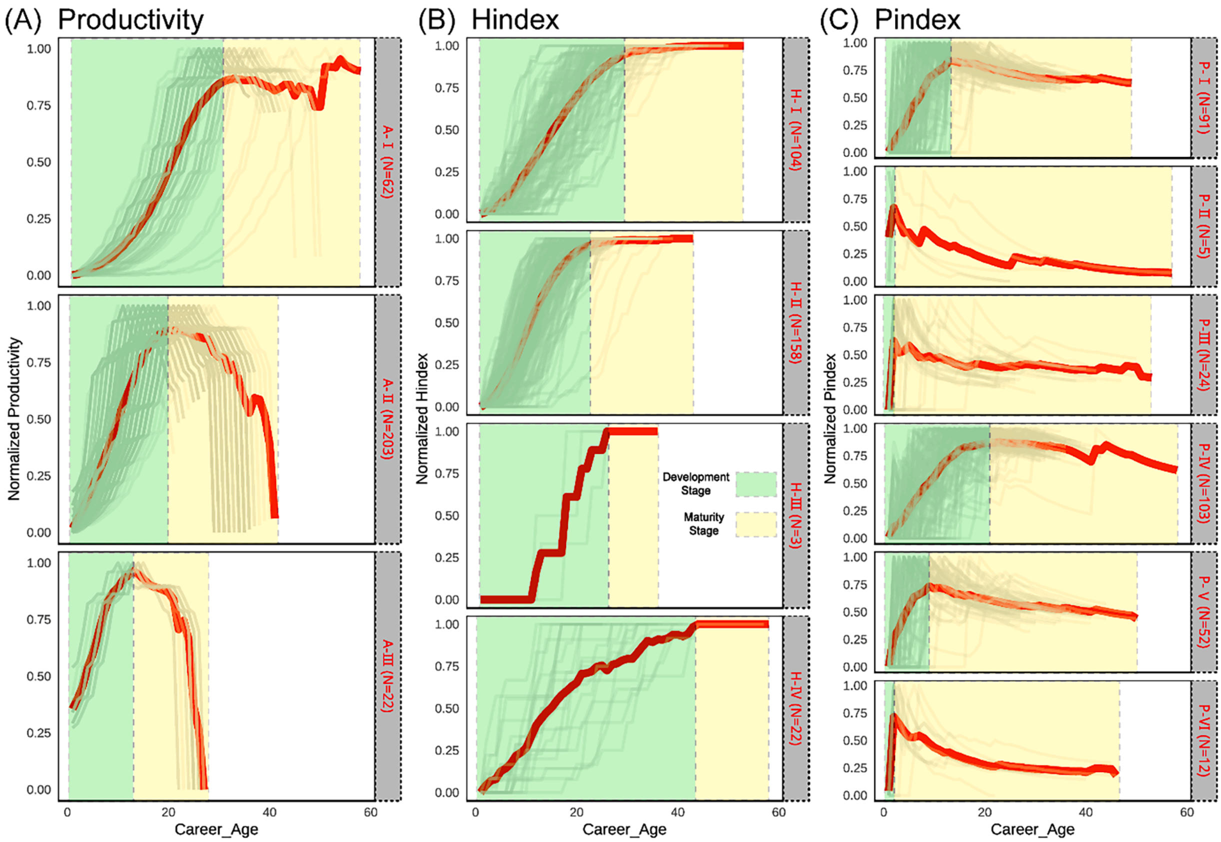

Fig. 5. The statistics and patterns of scholars' bibliometric growth trajectories for various categories.

4.3.2. *The population of different categories of scholars*

To find out the combination form of different categories of scholars based on three bibliometric indexes, we analyze the population distribution of scholars in the three combinations of productivity-h-index, productivity-p-index, and h-index-p-index (see Fig. 7). In Fig. 7, each pie chart represents the proportion and number of scholars categorized based on different indices, with the x-axis and y-axis representing specific combinations. For instance, Fig. 7A illustrates a pie chart at coordinates A-I and H-I, indicating the number of scholars belonging to category A-I in terms of productivity trajectory and category H-I in terms of h-index trajectory, along with their percentage in the total population. According to the combination of productivity and h-index, scholars are categorized into different groups. The A-III category is predominantly associated with the H-II combination, while the A-II category is mostly associated with the H-II and H-I combinations. The A-I category is predominantly associated with the H-I combination. Similarly, when considering the combination of productivity and p-index, the A-III category is predominantly associated with the P-IV combination, while the A-II category is mostly associated with the P-I, P-IV, and P-V combinations. The A-I category is predominantly associated with the P-IV, P-I, P-V, and P-III combinations. Analyzing from a similar perspective, with respect to the combination of h-index and p-index, the H-I category is predominantly associated with the P-IV, P-I, and P-V combinations, while the H-II category is mostly associated with the P-

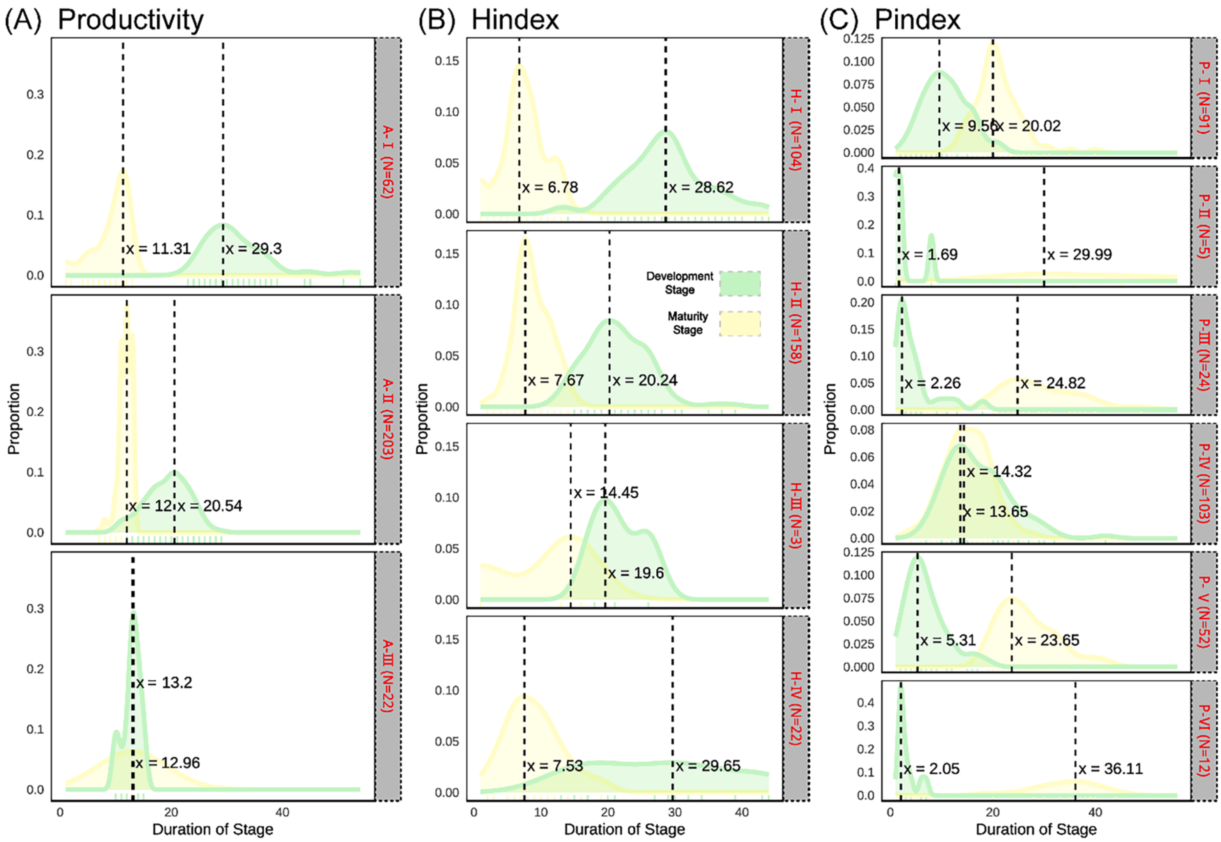

Fig. 6. The duration distribution of the early development and the mid-later maturity stages from different categories of scholars.

IV and P-I combinations. Due to a small number of scholars in the H-III category, it is not suitable for analysis. The H-IV category is predominantly associated with the P-I combination.

4.3.3. The election time and regional distribution of different categories of scholars

According to the information on the timeline of scholars being elected as Fellow provided by the ACM website, it is evident that scholars are typically recognized as ACM Fellow during the maturity stage of their careers. By further examining the proportion of a scholar's career dedicated to their election as ACM Fellow, we can analyze the distribution of this proportion across different scholar categories (see Fig. 8A). In Fig. 8A, the distribution curve of a certain category of scholars is more concentrated towards the reddish region, indicating that more individuals in this category were elected as ACM Fellows during the later stages of their careers. Conversely, the distribution curve of another category of scholars is more concentrated towards the bluish region, suggesting that more individuals in this category were elected as ACM Fellows during the earlier stages of their careers. From the perspective of productivity-based categorization, we observe a decreasing trend in the proportion of career time spent before being elected as ACM Fellow for the A-I, A-II, and A-III categories, ranging from approximately 80 % to 100 %. Analyzing the four categories based on the h-index, except for the H-III category where scholars are primarily recognized as ACM Fellow around the 90 % mark of their careers, the proportions of career time before election as ACM Fellow for H-I, H-II, and H-IV scholars show a relatively dispersed distribution, ranging from approximately 80 % to 100 %. Moreover, regarding the six categories based on the p-index, the proportions of career time before scholars' election as ACM Fellow exhibit a similar dispersed distribution, ranging from approximately 80 % to 100 %, without significant differences among the various categories.

According to the scholar's regional information provided on the ACM website, we analyze the distribution of growth trajectory types among scholars in four regions: Asia, Europe, North America, and the Pacific Rim (see Fig. 8B). However, we exclude the Pacific Rim region from our analysis due to the limited number of scholars. In Asia, the growth trajectory types among scholars show an equal distribution between A-II and A-III categories. Specifically, the predominant trajectory types for h-index and p-index changes are H-II and P-IV, respectively. In contrast, the distribution of growth trajectory types among scholars in Europe and North America differs from that in Asia. However, the distribution of growth trajectory types in Europe and North America is highly similar, with only slight differences. For instance, the productivity trajectory types decrease in the following order of proportion: A-II, A-I, A-III in Europe and North America. The h-index trajectory types decrease in the following order: H-II, H-I, H-IV, H-III. The p-index trajectory types decrease in the following order: P-IV, P-I, P-V. Additionally, the proportion of scholars in the P-III category is higher in North America compared to Europe, while the proportion of scholars in the P-VI category is lower in North America compared to Europe.

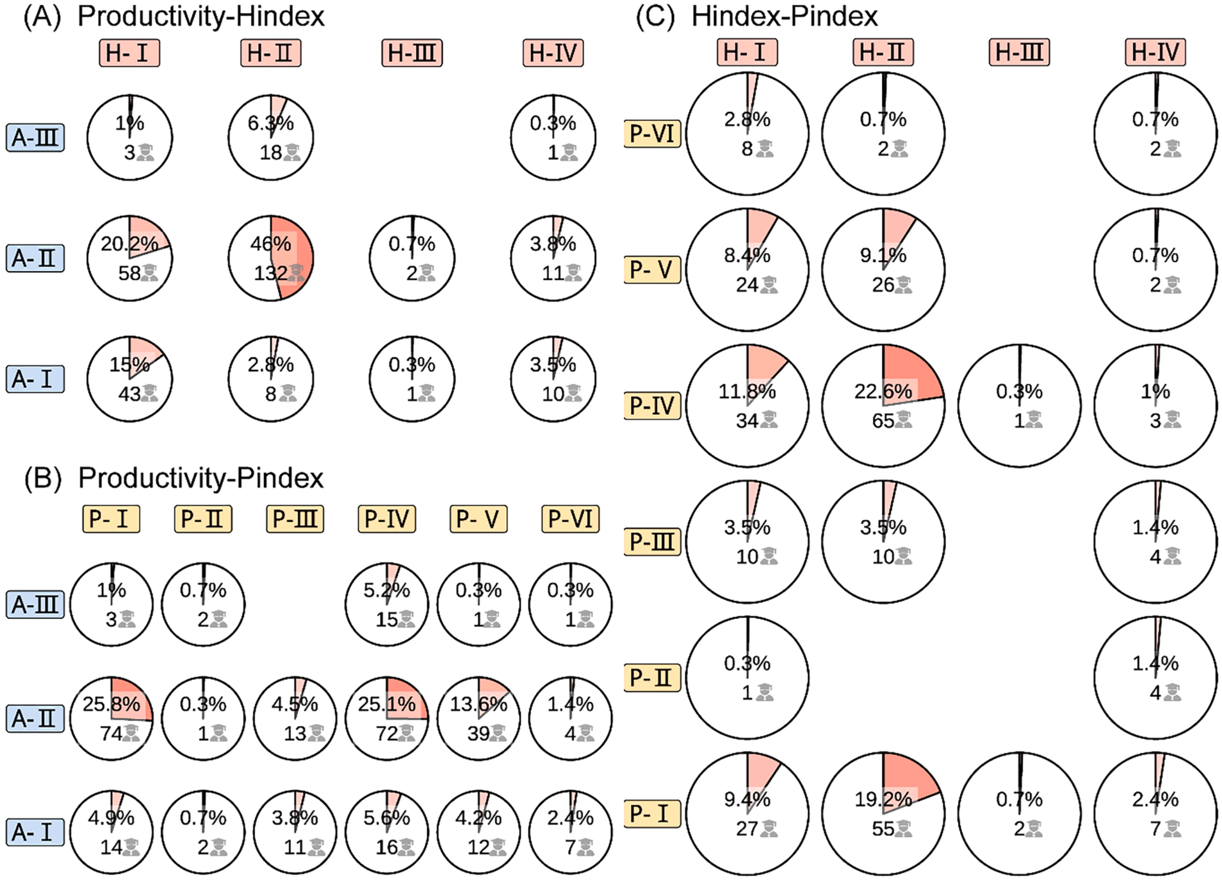

Fig. 7. the population distribution of scholars in the three combinations from different categories.

5. Discussion

5.1. The growth patterns of different categories of excellent scholars

Most outstanding scholars, classified into three types of growth paths based on the productivity curve, predominantly belong to the A-II category. In this category, productivity experiences a rapid rise followed by a decline, with the decline rate accelerating. The growth pattern of productivity among these scholars of the A-II category is similar to that of the A-III category, which comprises a smaller number of scholars. However, A-III scholars experience rapid fluctuations in productivity, and the duration of their high productivity is short-lived. Additionally, the growth pattern of productivity for A- I scholars significantly differs from that of the other two categories. They exhibit slow growth initially and maintain high productivity over an extended period in the mid-later maturity stage. Based on the analysis of three growth modes of outstanding scholars classified by the productivity curve, their productivity continues to increase in the early development stage, typically following a pattern of initial growth followed by decline. The state of high productivity tends to last for a considerable duration, although a minority of scholars experience a shorter duration, characterized by rapid productivity increase in the early development stage. From these characteristics of the productivity growth path, outstanding scholars have two remarkable characteristics: (1) high productivity with sustained growth in the early development stage, aligning with previous research findings (Barabási & Wang, 2021; Schilling & Green, 2011); (2) Scholars who maintain a high level of productivity for an extended duration in the middle and late maturity stages, but those who have extremely rapid increasing productivity in their early development stage cannot possess this feature, although such cases are relatively uncommon.

The shape of the h-index curve illustrates the trajectory of scholars' impact. Based on the similarity of h-index curves, scholars' impact can be categorized into four growth modes. Most scholars fall into the H-I and H-II categories, indicating a continuous rise in their impact during the developmental stage, followed by a prolonged period of stagnation and maturity after reaching the peak. In contrast, H-III and H-IV scholars experience intermittent growth and stagnation during the developmental stage, with a very short period of stagnant maturity after reaching the peak. Apart from a small number of scholars in the H-III category, the developmental stage for the H-IV category is considerably long, spanning almost their entire career. By summarizing the development trajectory of scholars' impact, it is observed that their impact generally increases consistently in the initial two decades of their career, although for a small number of exceptional scholars, their impact may follow a more convoluted and protracted path.

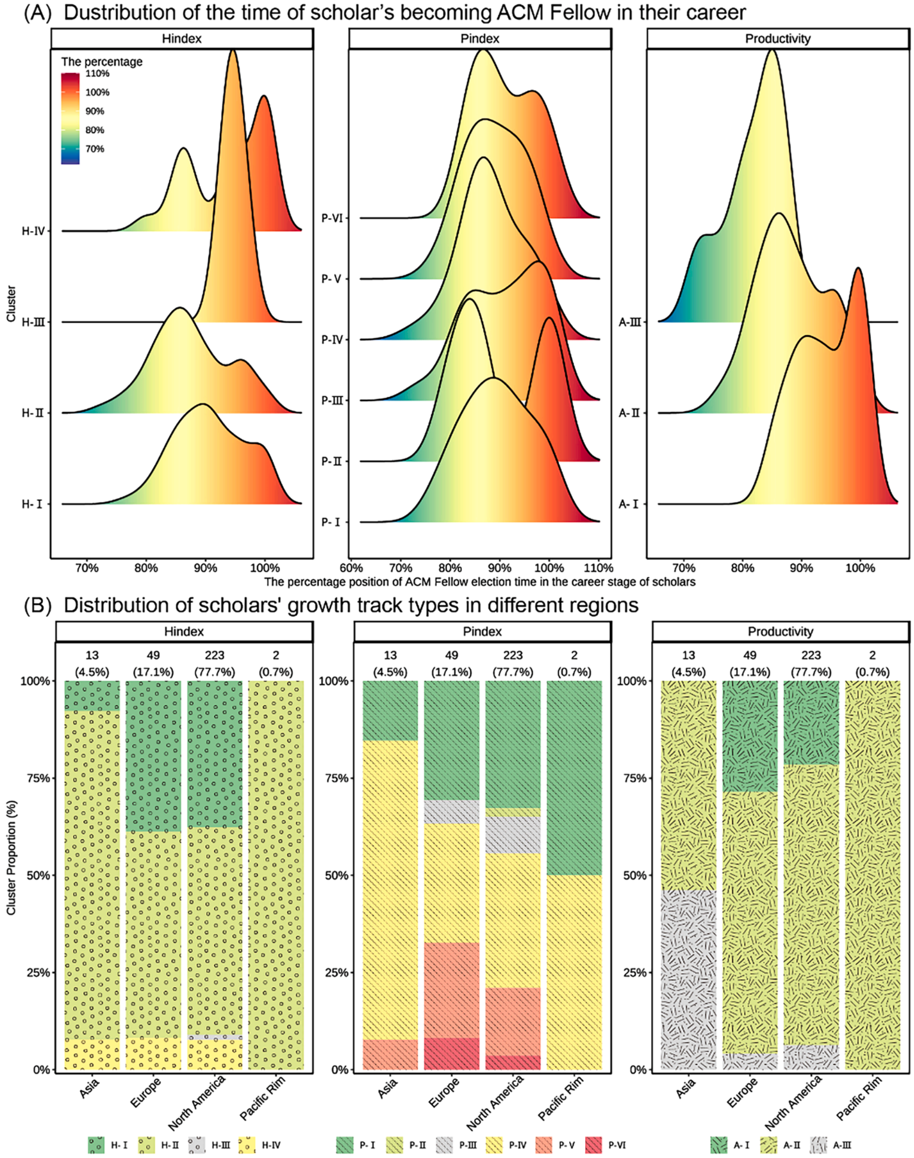

Fig. 8. The election time and regional distribution of different categories of scholars.

The p-index encompasses the number of published papers and citations of scholars, thereby considering their productivity and impact, and providing a measure of their overall scientific research capability. By analyzing and comparing the six growth path types of exceptional scholars based on the change curve of the p-index, these types can be categorized into two groups. The first group represents the mainstream model of comprehensive ability development among scholars, including P-V, P-I, and P-IV. In general, it indicates a gradual increase in overall competency over the first 10 to 20 years, with a decreasing rate of increase and proportion in the order of P-V, P-I, and P-IV. In the subsequent maturity stage, there is a slight decline in overall competency. The second group consists of P-II, P-VI, and P-IV, which includes a smaller number of individuals. These scholars exhibit a rapid initial increase in comprehensive ability during the first 2 years of their career, followed by a gradual decline. Specifically, their comprehensive ability declines in the order of P-II, P-VI, and P-IV, with a corresponding increase in the proportion of individuals. In summary, most excellent scholars take 10 to 20 years to reach the peak of their comprehensive ability, while a small number demonstrate rapid growth. However, the comprehensive ability of outstanding scholars in the middle and late stages generally remains at a medium or superior level, with the exception of the relatively small categories of P-II and P-VI.

5.2. The combination of scholar's different growth patterns

According to changes in productivity, influence, and comprehensive ability, outstanding scholars are classified into three, four, and six categories respectively. The matching of scholar categories from different perspectives can be determined by examining their distribution. The impact of A-III scholars, whose productivity increases rapidly and then declines, follows the pattern of H-II. Their impact experiences rapid growth during the development stage, while their comprehensive ability follows the pattern of P-IV scholars, with slow and sustained growth at a high level in the mid-later maturity stages. A-II scholars, with slow and gradual changes in productivity, primarily exhibit impact changes in the form of H-II and H-I, with H-II scholars being the predominant group. This indicates that the impact of outstanding scholars with slow productivity growth continues to rise rapidly. Furthermore, A-II scholars predominantly show changes in comprehensive ability categorized as mainstream P-IV and P-I. A-I scholars, on the other hand, experience a slow increase in productivity but maintain a long period of high productivity. Their impact changes mainly fall into the category of H-I, characterized by a slow rise, while their comprehensive ability changes are more scattered, distributed across various categories. Furthermore, the variations in scholars' impact are predominantly classified into two categories: H-II and H-I. The variations in comprehensive ability are mainly classified into three categories: P-IV, P-I, and P-V, with a decreasing number of individuals in that order. Scholars' combination categories, based on their impact and comprehensive ability, are primarily composed of combinations of their mainstream categories.

According to the analysis of productivity, impact, and comprehensive ability, the growth pattern of excellent scholars can be classified into several categories, often characterized by combinations of two or more mainstream categories. Most outstanding scholars exhibit a consistent and gradual increase in productivity, with a prolonged period of high productivity. Their impact tends to grow steadily and rapidly in the development stage, while their comprehensive ability shows a slower initial growth, minimal decline in the mid-later maturity stages, and a sustained high level throughout their careers.

5.3. The characteristics analysis for different categories of excellent scholars

The moment when excellent scholars are elected as ACM Fellow represents widespread recognition from the industry. Analyzing the timing of this moment in their careers can reveal the difficulty or ease with which scholars of different growth types receive recognition in the research community. Based on the distribution of the timing of ACM Fellow elections for different types of scholars, it is observed that nearly all scholars are elected during the 80 % to 100 % period of their careers (maturity stage). However, there exists a relative variation in the timing of elections among different types of scholars. Firstly, examining the growth patterns of scholars categorized based on their productivity, scholars in categories A-III, A-II, and A-I are elected as ACM Fellow in progressively later stages of their careers. Scholars who experience a rapid increase in productivity during the early stages of their careers tend to be elected as ACM Fellow earlier, indicating a significant association between the rate of early productivity growth and the timing of widespread recognition later on. Secondly, considering the growth patterns of scholars categorized based on their impact, scholars in categories H-I and H-II demonstrate smoother changes in impact compared to scholars in categories H-III and H-IV. Scholars in categories H-I and H-II experience sustained growth during the developmental stage and are elected as ACM Fellow earlier, suggesting that scholars who face more obstacles in the rise of their impact often achieve recognition later in their careers. Regarding the six categories of scholars classified based on their overall abilities, there is no significant difference in the timing of their election as ACM Fellow. The elections are predominantly concentrated around 90 % of their careers, indicating that the level of difficulty in achieving widespread recognition among scholars with different growth patterns of overall abilities is not significantly distinct.

Furthermore, due to the limited number of scholars from the Pacific Rim region in the sample, this study focuses on the growth patterns of scholars in Asia, Europe, and North America in terms of productivity, impact, and comprehensive abilities. The statistical results indicate that excellent scholars from North America and Europe show a relatively slow initial rise in productivity, followed by an extended period of high productivity. In contrast, excellent scholars from Asia exhibit rapid growth in productivity during the early development stages, but their high productivity period is relatively short-lived, demonstrating significant regional differences in the growth patterns of scholar productivity. The regional characteristics observed in the growth patterns of scholar impact are similar to those of productivity. Excellent scholars from Asia also experience a rapid increase in impact during the early development stages, while scholars from Europe and North America show a slightly slower rise in impact. The regional characteristics are also pronounced in the growth patterns of scholar comprehensive abilities. Excellent scholars from Asia demonstrate a relatively slower growth process in comprehensive abilities, while scholars from Europe and North America exhibit a slightly faster growth rate. Additionally, there are cases of individual scholars from Europe and North America experiencing a rapid initial growth in comprehensive abilities followed by a gradual decline.

This result suggests that excellent scholars in Asia, while publishing a significant number of articles in the early development stages of their academic careers, also produce high-impact publications. However, they also have a considerable number of low-impact publications, leading to a rapid increase in their publication output and h-index during the early development stages of their academic careers. For the p-index, which comprehensively considers the quality and quantity of each publication, the growth in academic comprehensive ability is relatively slow. In comparison, excellent scholars in Europe and North America typically publish fewer papers in the early development stages of their academic careers, but they are often of high quality. This results in a slower growth in their publication output and h-index during the early development stage, but a rapid increase in their p-index. The reasons for this phenomenon may stem from different scientific policy systems (Korytkowski & Kulczycki, 2019), educational training models (Yuret, 2017, 2018), and research resource allocations (Shu et al., 2021; Yue et al., 2020) in different regions, and require further analysis in conjunction with the specific characteristics of the research environment in which scholars work.

6. Conclusion

In summary, excellent scholars in the field of computer science generally exhibit common patterns in the trajectory of their bibliometric indexes. Most of them experience continuous growth in productivity during the early development stages of their careers and maintain a prolonged period of high productivity in the mid-later maturity stages. Their impact tends to rise smoothly, demonstrating sustained growth over time. Furthermore, the development of comprehensive abilities for most excellent scholars requires 10 to 20 years to reach its peak, and they can maintain a consistently high level in the mid-later maturity stages of their careers. Additionally, there are regular patterns in the combination of growth patterns in productivity, influence, and comprehensive abilities. Excellent scholars who exhibit a steady and slow increase in productivity with a prolonged high productivity period often experience a smoother and more rapid growth in impact. Their comprehensive abilities show a gradual increase in the early development stages and a slight decline in the mid-later maturity stages, while maintaining a high level for an extended period in their careers.

Excellent scholars with different growth patterns are elected as ACM Fellows at different times, indicating variations in the level of recognition they receive within the scientific research community. Specifically, the most significant differences in growth patterns can be observed between scholars categorized based on productivity and impact. Excellent scholars who exhibit rapid growth in productivity during the early development stages and experience smooth and rapid increases in impact are often elected as ACM Fellows at an earlier stage, as they are more likely to receive widespread recognition from their peers in the scientific research community. Furthermore, there are notable differences in the growth patterns of excellent scholars between the Western regions (Europe and North America) and Asia. Scholars from the Western regions tend to have slower growth in productivity and influence during the early development stages compared to their counterparts in Asia, while their comprehensive abilities develop more rapidly, and they maintain a longer period of high productivity.

This study analyzes the consistency of bibliometric indexes changes among excellent scholars in the field of computer science and further categorizes their growth patterns. It investigates the characteristics, distribution, election timing, and regional differences of excellent scholars in different categories. The results indicate that the growth trajectories of excellent scholars in the field of computer science exhibit similarities and can be classified into three categories based on productivity, four categories based on impact, and six categories based on comprehensive abilities. The level of recognition within the scientific research community varies among different categories of excellent scholars, and there are certain regional distribution characteristics.

7. Limitations and future research

This study has several limitations, specifically in terms of external validity and construct validity. Firstly, the external validity of this study is limited, and therefore the conclusions drawn from it are only applicable to outstanding scholars in the field of computer science. It has not been confirmed in other academic disciplines, nor has it explored whether the academic growth trajectories of outstanding scholars differ from those of ordinary scholars. Additionally, there is room for improvement in the construct validity of this study, as the measurement of scholars' academic capabilities is somewhat biased, focusing excessively on changes in bibliometric indexes and not considering other aspects of their academic abilities, such as academic innovative capabilities. Furthermore, this study has not conducted in-depth analysis on the reasons for the impact of bibliometric changes and how to use this growth pattern of scholars to evaluate the potential of young scholars.

As a result, future research will test the applicability of the above conclusions in other academic disciplines by incorporating samples of excellent and ordinary scholars from other fields to determine whether this academic growth pattern is unique to outstanding scholars. This will enrich the evaluation system used to assess scholars' academic abilities, comprehensively display scholars' academic growth trajectories in multiple dimensions, and analyze the potential factors influencing the growth of different types of scholars by combining information on the gender, research direction, nationality, collaborators, educational background, and work experience of outstanding scholars. This will further reveal the growth patterns of outstanding scholars.

Data availability

The key data and code used in this study are available in the github platform (https://github.com/Pengxz-tm/Bibliometric-growth-patterns-of-excellent-scholars).

CRediT authorship contribution statement

Xianzhe Peng: Data curation, Formal analysis, Writing – original draft. **Huixin Xu:** Data curation, Writing – original draft. **Jin Shi:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

This research is supported by the National Social Science Foundation, People's Republic of China (No. 21BTQ012) and China

Mobile Communications Corporation (No. 2024200249).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.joi.2024.101543.

References

Barabási, A. L., Wang, D., Barabási, A. L., & Wang, D. (2021). The Science of science. *The science of science* (pp. i–ii). Cambridge University Press. https://doi.org/10.1017/9781108610834

Bornmann, L., & Williams, R. (2017). Can the journal impact factor be used as a criterion for the selection of junior researchers? A large-scale empirical study based on ResearcherID data. *Journal of Informetrics, 11*(3), 788–799. https://doi.org/10.1016/j.joi.2017.06.001

Buehring, G. C., Buehring, J. E., & Gerard, P. D. (2007). Lost in citation: Vanishing visibility of senior authors. *Scientometrics, 72*(3), 459–468. https://doi.org/10.1007/s11192-007-1762-4

Egghe, L. (2013). On the correction of the h-index for career length. *Scientometrics, 96*(2), 563–571. https://doi.org/10.1007/s11192-012-0926-z

Egghe, L., Rousseau, R., & Van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for information Science, 51*(2), 145–157. https://doi.org/10.1002/(SICI)1097-4571(2000)51:2<145 AID-ASI6>3.0.CO;2-9

Feichtinger, G., Grass, D., & Kort, P. M. (2019). Optimal scientific production over the life cycle. *Journal of Economic Dynamics and Control, 108*, Article 103752. https://doi.org/10.1016/j.jedc.2019.103752

Fernandes, J. M., Costa, A., & Cortez, P. (2022). Author placement in computer science: A study based on the careers of ACM Fellows. *Scientometrics, 127*(1), 351–368. https://doi.org/10.1007/s11192-021-04035-5

Fernandes, J. M., & Monteiro, M. P. (2017). Evolution in the number of authors of computer science publications. *Scientometrics, 110*(2), 529–539. https://doi.org/10.1007/s11192-016-2214-9

Fiala, D. (2014). Current index: A Proposal for a dynamic rating system for researchers. *Journal of the Association for Information Science and Technology, 65*(4), 850–855. https://doi.org/10.1002/asi.23049

Fortunato, S. (2014). Growing time lag threatens nobels. *Nature, 508*(7495), 186. https://doi.org/10.1038/508186a. -186.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., et al. (2018). Science of science. *Science, 359*(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science, 178*(4060), 471–479. https://doi.org/10.1126/science.178.4060.471

Győrffy, B., Csuka, G., Herman, P., & Török, Á. (2020). Is there a golden age in publication activity?—an analysis of age-related scholarly performance across all scientific disciplines. *Scientometrics, 124*(2), 1081–1097. https://doi.org/10.1007/s11192-020-03501-w

Hagen, N. T. (2008). Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis. *PloS One, 3*(12), e4021. https://doi.org/10.1371/journal.pone.0004021

Hagen, N. T. (2010). Harmonic publication and citation counting: Sharing authorship credit equitably – not equally, geometrically or arithmetically. *Scientometrics, 84*(3), 785–793. https://doi.org/10.1007/s11192-009-0129-4

Haunschild, R., & Bornmann, L. (2023). Identification of potential young talented individuals in the natural and life sciences: A bibliometric approach. *Journal of Informetrics, 17*(3), Article 101394. https://doi.org/10.1016/j.joi.2023.101394

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

Jian, D., & Xiaoli, T. (2013). Perceptions of author order versus contribution among researchers with different professional ranks and the potential of harmonic counts for encouraging ethical co-authorship practices. *Scientometrics, 96*(1), 277–295. https://doi.org/10.1007/s11192-012-0905-4

Jin, Y., Yuan, S., Shao, Z., Hall, W., & Tang, J. (2021). Turing Award elites revisited: Patterns of productivity, collaboration, authorship and impact. *Scientometrics, 126*(3), 2329–2348. https://doi.org/10.1007/s11192-020-03860-4

Jones, B. F., & Weinberg, B. A. (2011). Age dynamics in scientific creativity. *Proceedings of the National Academy of Sciences, 108*(47), 18910–18914. https://doi.org/10.1073/pnas.1102895108

Kalyane, VL., & Vidyasagar Rao, K. (1995). Quantification of credit for authorship. *ILA Bulletin, 30*(3–4), 94–96.

Kim, J. (2018). Evaluating author name disambiguation for digital libraries: A case of DBLP. *Scientometrics, 116*(3), 1867–1886. https://doi.org/10.1007/s11192-018-2824-5

Korytkowski, P., & Kulczycki, E. (2019). Examining how country-level science policy shapes publication patterns: The case of Poland. *Scientometrics, 119*(3), 1519–1543. https://doi.org/10.1007/s11192-019-03092-1

Kwiek, M., & Roszka, W. (2023). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*. https://doi.org/10.1007/s10734-023-01022-y

Laurance, W. F., Useche, D. C., Laurance, S. G., & Bradshaw, C. J. A. (2013). Predicting Publication Success for Biologists. *BioScience, 63*(10), 817–823. https://doi.org/10.1525/bio.2013.63.10.9

Lee, D. H. (2019). Predicting the research performance of early career scientists. *Scientometrics, 121*(3), 1481–1504. https://doi.org/10.1007/s11192-019-03232-7

Ley, M. (2009). DBLP: Some lessons learned. *Proceedings of the VLDB Endowment, 2*(2), 1493–1500. https://doi.org/10.14778/1687553.1687577

Li, J., Yin, Y., Fortunato, S., & Wang, D. (2020). Scientific elite revisited: Patterns of productivity, collaboration, authorship and impact. *Journal of the Royal Society Interface, 17*(165), Article 20200135. https://doi.org/10.1098/rsif.2020.0135

Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications, 10*(1), 5170. https://doi.org/10.1038/s41467-019-13130-4

Lindahl, J. (2018). Predicting research excellence at the individual level: The importance of publication rate, top journal publications, and top 10% publications in the case of early career mathematicians. *Journal of Informetrics, 12*(2), 518–533. https://doi.org/10.1016/j.joi.2018.04.002

Liu, L., Dehmamy, N., Chown, J., Giles, C. L., & Wang, D. (2021). Understanding the onset of hot streaks across artistic, cultural, and scientific careers. *Nature Communications, 12*(1), 5392. https://doi.org/10.1038/s41467-021-25477-8

Liu, L., Wang, Y., Sinatra, R., Giles, C. L., Song, C., & Wang, D. (2018). Hot streaks in artistic, cultural, and scientific careers. *Nature, 559*(7714), 396–399. https://doi.org/10.1038/s41586-018-0315-8

Liu, X. Z., & Fang, H. (2012). Fairly sharing the credit of multi-authored papers and its application in the modification of h-index and g-index. *Scientometrics, 91*(1), 37–49. https://doi.org/10.1007/s11192-011-0571-y

Milojević, S. (2014). Principles of scientific research team formation and evolution. *Proceedings of the National Academy of Sciences, 111*(11), 3984–3989. https://doi.org/10.1073/pnas.1309723111

Mueen, A., Chavoshi, N., Abu-El-Rub, N., Hamooni, H., Minnich, A., & MacCarthy, J. (2018). Speeding up dynamic time warping distance for sparse time series data. *Knowledge and Information Systems, 54*(1), 237–263. https://doi.org/10.1007/s10115-017-1119-0

Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences, 118*(7), Article e2012208118. https://doi.org/10.1073/pnas.2012208118

Pan, R. K., & Fortunato, S. (2014). Author Impact Factor: Tracking the dynamics of individual scientific impact. *Scientific reports, 4880*. https://doi.org/10.1038/srep04880

Petersen, A. M. (2015). Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences, 112*(34), E4671–E4680. https://doi.org/10.1073/pnas.1501444112

Petersen, A. M. (2018). Multiscale impact of researcher mobility. *Journal of The Royal Society Interface, 15*(146), Article 20180580. https://doi.org/10.1098/rsif.2018.0580

Prathap, G. (2010). The 100 most prolific economists using the p-index. *Scientometrics, 84*(1), 167–172. https://doi.org/10.1007/s11192-009-0068-0

Prathap, G. (2011). Qualifying scholarly impact using an iCX (impact-Citations-Exergy) analysis. *DESIDOC Journal of Library & Information Technology, 31*(5), 382–386. https://doi.org/10.14429/djlit.31.5.1196

Raan, A. F. J. V. (2006). Performance-related differences of bibliometric statistical properties of research groups: Cumulative advantages and hierarchically layered networks. *Journal of the American Society for Information Science and Technology, 57*(14), 1919–1935. https://doi.org/10.1002/asi.20389

Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., et al. (2012). Searching and mining trillions of time series subsequences under dynamic time warping. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. https://doi.org/10.1145/2339530.2339576

Ram, S. K., Nandan, S., Boulebnane, S., & Sornette, D. (2022). Synchronized bursts of productivity and success in individual careers. *Scientific reports, 12*(1), 7637. https://doi.org/10.1038/s41598-022-10837-1

Rosenfeld, A. (2023). Is DBLP a good computer science journals database? *Computer, 56*(3), 101–108. https://doi.org/10.1109/MC.2022.3181977

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Saad, G. (2010). Applying the h-index in exploring bibliometric properties of elite marketing scholars. *Scientometrics, 83*(2), 423–433. https://doi.org/10.1007/s11192-009-0069-z

Schilling, M. A., & Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy, 40*(10), 1321–1331. https://doi.org/10.1016/j.respol.2011.06.009

Shu, F., Sugimoto, C. R., & Larivière, V. (2021). The institutionalized stratification of the Chinese higher education system. *Quantitative Science Studies, 2*(1), 327–334. https://doi.org/10.1162/qss_a_00104

Sinatra, R., Wang, D., Deville, P., Song, C., & Barabási, A.-L. (2016). Quantifying the evolution of individual scientific impact. *Science, 354*(6312), aaf5239. https://doi.org/10.1126/science.aaf5239

Sun, Y., Caccioli, F., & Livan, G. (2023). Ranking mobility and impact inequality in early academic careers. *Proceedings of the National Academy of Sciences, 120*(34), Article e2305196120. https://doi.org/10.1073/pnas.2305196120

Tol, R. S. J. (2013). Identifying excellent researchers: A new approach. *Journal of Informetrics, 7*(4), 803–810. https://doi.org/10.1016/j.joi.2013.06.003

van Dijk, D., Manor, O., & Carey, L. B. (2014). Publication metrics and success on the academic job market. *Current Biology, 24*(11), R516–R517. https://doi.org/10.1016/j.cub.2014.04.039

Van Hooydonk, G. (1997). Fractional counting of multiauthored publications: Consequences for the impact of authors. *Journal of the American Society for information Science, 48*(10), 944–945. https://doi.org/10.1002/(SICI)1097-4571(199710)48:10<944::AID-ASI8>3.0.CO;2-1

Van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. J. (2003). The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics, 57*(2), 257–280. https://doi.org/10.1023/A:1024141819302

Vinkenburg, C. J., Connolly, S., Fuchs, S., Herschberg, C., & Schels, B. (2020). Mapping career patterns in research: A sequence analysis of career histories of ERC applicants. *PloS One, 15*(7), Article e0236252. https://doi.org/10.1371/journal.pone.0236252

Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. *Nature Communications, 10*(1), 4331. https://doi.org/10.1038/s41467-019-12189-3

Way, S. F., Morgan, A. C., Clauset, A., & Larremore, D. B. (2017). The misleading narrative of the canonical faculty productivity trajectory. *Proceedings of the National Academy of Sciences, 114*(44), E9216–E9223. https://doi.org/10.1073/pnas.1702121114

Way, S. F., Morgan, A. C., Larremore, D. B., & Clauset, A. (2019). Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences, 116*(22), 10729–10733. https://doi.org/10.1073/pnas.1817431116

Wu, J., Lozano, S., & Helbing, D. (2011). Empirical study of the growth dynamics in real career h-index sequences. *Journal of Informetrics, 5*(4), 489–497. https://doi.org/10.1016/j.joi.2011.02.003

Yair, G., & Goldstein, K. (2020). The Annus Mirabilis paper: Years of peak productivity in scientific careers. *Scientometrics, 124*(2), 887–902. https://doi.org/10.1007/s11192-020-03544-z

Yang, J., Bu, Y., Lu, W., Huang, Y., Hu, J., Huang, S., et al. (2022). Identifying keyword sleeping beauties: A perspective on the knowledge diffusion process. *Journal of Informetrics, 16*(1), Article 101239. https://doi.org/10.1016/j.joi.2021.101239

Yue, M.-l., Li, R.-n., Ou, G.-y., Wu, X., & Ma, T.-c. (2020). An exploration on the flow of leading research talents in China: From the perspective of distinguished young scholars. *Scientometrics, 125*(2), 1559–1574. https://doi.org/10.1007/s11192-020-03562-x

Yuret, T. (2017). An analysis of the foreign-educated elite academics in the United States. *Journal of Informetrics, 11*(2), 358–370. https://doi.org/10.1016/j.joi.2017.02.008

Yuret, T. (2018). Path to success: An analysis of US educated elite academics in the United States. *Scientometrics, 117*(1), 105–121. https://doi.org/10.1007/s11192-018-2850-3

Zeng, A., Fan, Y., Di, Z., Wang, Y., & Havlin, S. (2022). Impactful scientists have higher tendency to involve collaborators in new topics. *Proceedings of the National Academy of Sciences, 119*(33), Article e2207436119. https://doi.org/10.1073/pnas.2207436119

Zhang S., LaBerge N., Way S.F., Larremore D.B., & Clauset A. (2023). Scientific productivity as a random walk. *arXiv e-prints*, arXiv:2309.04414. doi:10.48550/arXiv.2309.04414.

Zhang, S., Wapman, K. H., Larremore, D. B., & Clauset, A. (2022). Labor advantages drive the greater productivity of faculty at elite universities. *Science Advances, 8*(46), eabq7056. https://doi.org/10.1126/sciadv.abq7056

Zhao, Q., & Fränti, P. (2014). WB-index: A sum-of-squares based index for cluster validity. *Data & Knowledge Engineering, 92*, 77–89. https://doi.org/10.1016/j.datak.2014.07.008