Journal of Land Use Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tlus20

Gender and authorship patterns in urban land science

Tzu-Hsin Karen Chen & Karen C. Seto

To cite this article: Tzu-Hsin Karen Chen & Karen C. Seto (2022): Gender and authorship patterns in urban land science, Journal of Land Use Science, DOI: 10.1080/1747423X.2021.2018515

To link to this article: https://doi.org/10.1080/1747423X.2021.2018515

JOURNAL OF LAND USE SCIENCE
https://doi.org/10.1080/1747423X.2021.2018515

Taylor & Francis
Taylor & Francis Group

Gender and authorship patterns in urban land science

Tzu-Hsin Karen Chen [a,b] and Karen C. Seto [a]

[a]Yale School of the Environment, Yale University, New Haven, CT, USA; [b]Yale Institute for Biospheric Studies, Yale University, New Haven, Ct, USA

ABSTRACT

What are patterns of gender and authorship in urban land science? Our bibliometric analysis shows that the proportion of women shrinks among highly productive, impactful, and senior authors, akin to a pyramid shape. First, women are only one in ten researchers with an h-index above the 95th percentile. Second, women are first authors on 20% of all influential papers cited more than one hundred times. Third, women publish less frequently (1.6 papers/year) than men (2.2). Fourth, women have shorter career lengths (9.4 years) than men (11.8). Since the 2000s, citation rates for women and men have converged. For the generation starting careers since 2016, the proportion of women with an h-index above the 90th percentile increased to 25%. During the Covid-19 pandemic, there was a 51% increase in productivity for women. Despite these changes, gender disparities in urban land science are most pronounced among the most productive and impactful authors.

ARTICLE HISTORY
Received 10 July 2021
Accepted 9 December 2021

KEYWORDS
Gender inequality; productivity; citation impact; women in science; urban remote sensing

1. Introduction

Success in science and academia requires many things, including obtaining grants, publishing papers, and being productive and impactful. Numerous studies show that gender disparities are present in many scientific fields and aspects of science, including mentoring and hiring (Moss-Racusin et al., 2012), salaries (Shen, 2013), grants and funding (Hechtman et al., 2018; Tamblyn et al., 2018), and publications and authorship (Sarsons, 2017; West et al., 2013). Among these components, publications and citations are often used as research output and quality measures. Metrics on publications and citations are used for decisions about promotion, tenure, salary, and funding (Faria & Mixon, 2021).

Despite major advancements in women's participation in science and engineering over the past decades (Rivers, 2017), only 3.8% of environmental science articles for *Nature's News & Views* were authored by women despite making up 20% of the field (Conley & Stadmark, 2012). Although there are more female than male undergraduate students in many countries (Organisation for Economic Co-operation and Development, 2012), relatively few women possess a doctoral degree, and relatively few women are full professors in science, engineering, and medical fields (Rivers, 2017). Women tend to be better represented in social sciences than natural sciences (Henley, 2015). However, whether gender disparities exist in young, interdisciplinary fields that integrate approaches from natural and social sciences is unknown.

CONTACT Tzu-Hsin Karen Chen karen.t.chen@yale.edu Yale School of the Environment, Yale University, New Haven, Ct 06511, USA

Supplemental data for this article can be accessed here.

Urban land science is an emerging interdisciplinary field that brings together concepts from land system science, urbanization science, and global environmental change. This includes perspectives from a wide range of fields, including but not limited to geography, urban ecology, health sciences, anthropology, political science, and land system science. Central to urban land science is the use of remote sensing for characterizing and understanding urban land systems (Chen et al., 2020; Stokes & Seto, 2019). At its core, urban land science recognizes that urban systems both drive and are at the core of a range of global change issues, be it biodiversity loss (McDonald et al., 2020; Seto et al., 2012), urban heat Island (K. K. Huang et al., 2021; Stone et al., 2010), public health (Pinchoff et al., 2020), or climate change vulnerability (Shughrue et al., 2020).

The goal of this analysis is to examine gender and publication patterns in urban land science. Given that the field is relatively young and multidisciplinary, we hypothesized that gender disparities in publications and citations observed in other scientific fields will be less pronounced in urban land science. We developed a two-fold approach to identify authors' gender by first screening gender using an automatic API. There are tools online that identify gender based on names. One of the advantages of these tools (e.g. Genderize.io, Gender API) is that they allow for rapid and efficient processing of large datasets and thus provide crucial information about gender that can be coupled with bibliometric data. Genderize.io has an overall accuracy of 82%, with accuracies varying by country from 11% to 98% (Karimi et al., 2016). However, we used Gender API because it has been used by other studies on gender and publications (e.g. Bell & Fong, 2021). As a second step, we manually searched authors' biographies and photos in blogs, research media, and journal articles for genders not automatically recognizable using their names. We recognize that this method of assessing gender is not perfect and could miss other genders (e.g. non-binary gendered or trans-gender) and introduce biases. We conducted a bibliometric analysis for urban land science with remote sensing approaches worldwide, including 1,444 articles published between 1983 and 2020 from 1,157 first authors.

This study aims to answer the following questions concerning women's achievement and gender disparities in urban land science: What is the relative distribution of women for research output and citation impact, as measured by h-indices? Are there gender-specific differences in productivity and scientific career length? What percentage of highly influential papers are authored by women? How have gender disparities in publications changed over the past three decades? How has productivity during the recent Covid-19 pandemic varied by gender? Finally, do those gender disparities differ between continents and between sub-fields?

2. Previous studies on gender and authorship patterns in science

A main focus of research on gender disparities in science has been the 'productivity puzzle', namely that male scientists publish more papers than females over their entire careers (Aguinis et al., 2018; Bendels et al., 2018; Long, 1992). However, in relatively new fields, such as nanoscience, women and men publish an equal amount of research articles (Sotudeh & Khoshian, 2014). A recent study finds that women and men have similar rates of publications and citations across twelve scientific disciplines once differences in career lengths are controlled for (J. Huang et al., 2020).

Some studies find that while women publish fewer papers over their careers, they tend to publish in higher impact journals (Duch et al., 2012; Mauleón & Bordons, 2006). This has been found in fields incurring long lags between a doctoral degree and a faculty position, such as ecology (Duch et al., 2012). In material science, female research scientists are more likely than their male counterparts to publish their work in Science Citation Index journals (Mauleón & Bordons, 2006). However, it is important to note that there is no consistency across fields as to whether women publish higher quality articles than men, as measured by the number of citations received. Women's citation rate is

higher than men's in demographic economics (Grossbard et al., 2020) and biochemistry (Long, 1992), comparable to men's in nanoscience (Sotudeh & Khoshian, 2014), and lower than men's in life science and environmental science (Bendels et al., 2018).

Even though there has been an increase in the proportion of female authors in most scientific fields, the proportion of women as last or single authors remain lower than other positions (Bendels et al., 2018; Holman et al., 2018; West et al., 2013). Female first authors have collaborated with more co-authors and include co-authors of another sex than male first authors (Dworkin et al., 2020; Nunkoo et al., 2020). These designations in author order are important because they signal leadership and stature. For example, in cases where the first author is not the corresponding author, it is often the case that the first author has undertaken the study, whereas the corresponding author is the more senior scientist overseeing the study, although this varies significantly across disciplines. Women are represented the least in the last author position, indicating relatively few female laboratory or project leaders (West et al., 2013). The gender disparity in authorships is accentuated in areas where research is dependent on substantial funding (Duch et al., 2012). Women are also less often commissioned to write invited papers, consistent with observed gender bias by journal editors (Holman et al., 2018).

Women's authorship in Nature Index journals shows a slight increase from 29.7% in 2008 to 31.6% in 2016 (Bendels et al., 2018). Paradoxically, the increase of participation of women in science over the past 60 years was accompanied by an increase in gender differences in both the total number of publications and citation impact when comparing women and men's entire careers (J. Huang et al., 2020). While women and men have a comparable annual publication rate and equivalent career-wise impact for the same size body of work, women tend to drop out of the academic workforce earlier. Another paradox is present in geography: gender gaps in women's participation in STEMM (Science, Technology, Engineering, Mathematics, and Medicine) are larger in high income countries, such as Germany, Japan, and Switzerland, than lower income countries in South America (Holman et al., 2018). By continent, the proportion of female authorships is the lowest in Asia (Bendels et al., 2018).

The present state of knowledge on gender disparities in science is likely to be biased because of the challenges in identifying gender from names that do not originate from Indo-European languages (e.g. English; Bell & Fong, 2021; J. Huang et al., 2020; Jadidi et al., 2018). The mainstream tools for gender recognition, such as genderize.io and Gender API, predict gender based on country-specific name-gender associations, and their performance deteriorates with names from countries using non-alphabetic languages, such as China, Malaysia, Japan, Korea, and Singapore (Bell & Fong, 2021; Holman et al., 2018; J. Huang et al., 2020). Some countries use alphabetic languages, but the models were not well-trained for them (e.g. Brazilian names; Karimi et al., 2016). For instance, Jadidi et al. (2018) find only 11%, 28%, 44% of names gender-recognizable from South Korea, China, and Brazil using genderize.io, in contrast to 94%, 94%, and 83% recognizable for the UK, Germany, and the US among authors in computer science. Due to this difficulty, J. Huang et al. (2020) excluded all the authors from Brazil, Korea, Japan, China, Singapore, and Malaysia.

3. Methods

3.1 Assessing the literature and authors in urban land science

Our first task was to identify and catalog published papers in urban land science. Urban land science spans many disciplines, and as such, studies are published in journals across many fields. We used the PRISMA (i.e. preferred reporting items for systematic review and meta-analysis) methodology (Moher et al., 2015) to systematically review the English language peer-reviewed literature that examines urban land using a remote sensing approach published between January 1972 and December 2020. The citation count was collected on 28 May 2021. We searched the Web of Science database using topic keywords as follows:

TS= ("satellite data" OR "remote sensing" OR "images" OR "imagery" OR "LiDAR" OR "SAR" OR "aerial photos" OR "street view" OR "nighttime light" OR "night-time light" OR "nightlight") AND TS= ("urban extent" OR "urban land" OR "urban infrastructure" OR "urban form")AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=1972-2020

Our initial search yielded 1,582 articles and captured both methodological as well as application studies. We examined the bibliometric data of the database and removed articles without an abstract or affiliations, leading to 1,577 articles for gender prediction (Figure 1).

Considering the first authors of these articles, we measured their individual bibliometrics normalized by their career stage (detailed in section 3.3).

3.2 Determining gender of authors

An author's gender is neither explicit nor apparent. We therefore needed to develop a methodology to classify the gender of the first authors. We used a semi-automatic approach using a two-step process. First, we used an algorithm (by Gender API) that predicts gender based on the first name, along with an uncertainty estimate of the assigned gender accounting for country-of-origin of the individual (over 200 countries considered) (https://gender-api.com/). For instance, people named Kim are typically male in Denmark but female in the US (Holman et al., 2018), and therefore combined information of name and country leads to a more accurate prediction than using name alone. Gender API, a pay-for-service company, developed its algorithm using data from multiple

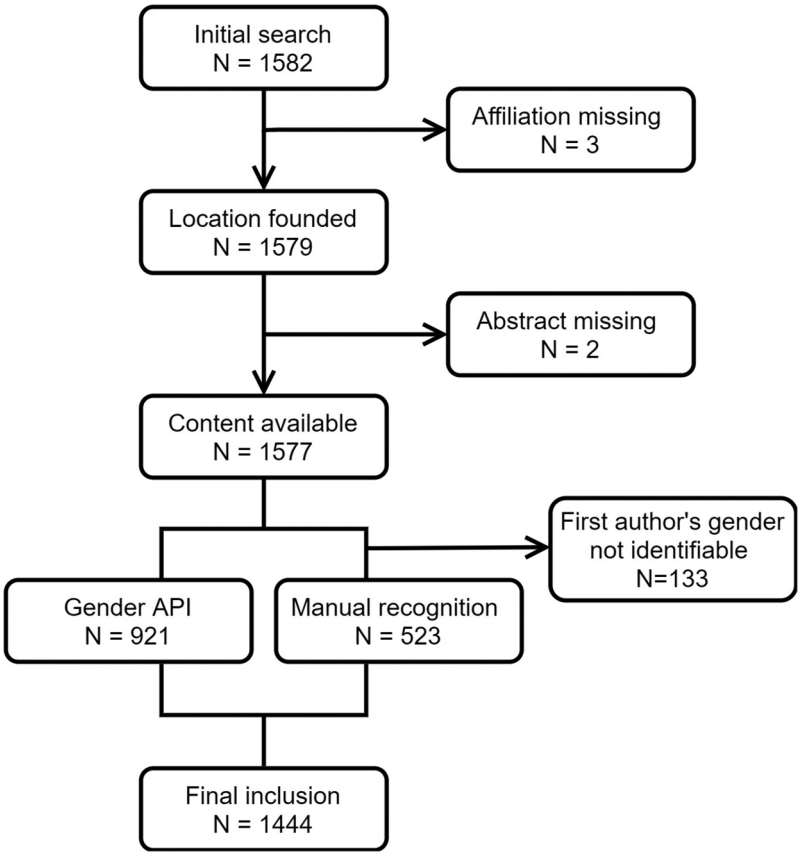

Figure 1. Flowchart of document inclusion.

publicly available and social network data sources and provides an accuracy estimate ranging from 50–100%. The accuracy is based on Gender API's database. A 90% accuracy means that of the country-name instances recorded in the Gender API database, 90% of the instances are of the same gender. We only used gender with an accuracy of 90% or more, which results in 921 articles' first authors identified with a binary gender.

Second, we manually classified the gender of authors that Gender API identified as having low accuracy (<90%). For these authors, we classified their gender based on the pronouns in their biography (e.g. attached in their affiliation web page or articles published by IEEE and SPIE) or photos on their public web page, such as their affiliation website, Google Scholar, or ResearchGate. As a result, we manually predicted gender for 523 articles' first author, in combination with 921 articles with API-predicted gender at accuracy higher than 90%, leading to 1,444 (91.6%) articles ready for further analysis.

3.3 Impact and productivity of individual scholars

Productivity and impact are two common measures that are used to evaluate scientists. They can affect hiring, promotion, and funding. There are many different metrics to determine productivity and impact. The h-index, one of the bibliometrics for comparing lifetime achievement of individual scientists, calculates the set of the scientist's most cited papers and the number of citations they have received. An h-index of 5 means that a scientist has published five papers that each has at least five citations. The h-index has been criticized as a metric of impact, which only increases and does not differentiate between past and current works that continue to shape the scientific field (Bornmann & Daniel, 2007). Critics also argue that the h-index does not take into account author order and the number of co-authors of each paper, so a scientist frequently publishing as a co-author may have an h-index as high as another scientist with many single-authored papers (Hirsch, 2010). Criticisms and limitations notwithstanding, the h-index correlates with academic promotion, merit increases, grant support, and awards (Bastian et al., 2017; Bornmann & Daniel, 2007).

We use the h-index as a metric of research output and citation impact. In addition, we measure total citations to reflect one's cumulative scientific impact, career length to understand how long the author has been in the academic workforce and to normalize productivity. Career length is defined as the number of years between the author's first and last publications. We acknowledge this measure emphasizes a publishing career and may exclude those who stop publishing but retain teaching or administrative duties. Productivity is defined as the number of publications per year during one's career. We used Scopus to collect bibliometric data because Scopus uses unique author IDs to differentiate authors of the same name, while Web of Science does not.

We measured the 25th, 50th, 75th, and 95th percentiles of each metric and calculated the proportion of women at each of the five levels (i.e. ≤ 25th, 25–50th, 50–75th, 75–95th, > 95th). This way, we examined how women's representation is distributed. We also measured the mean value for women and men and implemented a t-test to determine whether there is a difference between genders.

3.4 Authorship and impact of individual papers

In academia, the number of publications and the author order influence hiring and promotion decisions. Usually, the first and corresponding authorships indicate major contributions to a paper and are given more credit when evaluating scientific achievement. Therefore, we measured the proportion of female first authors and whether the first author and corresponding author are the same person.

Highly cited papers in any discipline become canons in the literature. They are often idea-shaping and highly influential papers that become common knowledge in a discipline. As such, we examined the gender composition of these influential papers. We categorized *influential papers* as those with citations above the 90th percentile in the field, and *highly influential papers* as those above the 95th percentile.

At their most basic level, paper citations confer intellectual contributions. We cite papers to give due credit to a finding, an idea, a method previously published. Citations also reflect the quality of the work. Thus, we examined whether there are differences in citation rates between female- and male-led papers. We measured the number of citations and annual citation rate (citations per year) for each paper. We also measured the annual citation rate, which normalizes the citation rate by how long a paper has been published, accounting for year and month.

While most studies on gender disparities in science do not examine regional differences, some have found large variations by country (Holman et al., 2018; J. Huang et al., 2020). We expect those gender disparities to be less prevalent in the Global North compared to the Global South due to differences in gender norms and women in the workforce. We tested this hypothesis by attributing each paper to a region using the first author's affiliation. We then grouped them into the following regions: Africa, Asia, Europe, North America, Oceania, and South America.

3.5 Pre-pandemic vs. COVID-19 pandemic productivity

The Covid-19 pandemic has altered the nature of work, and scientists are adapting to working from home and the lack of access to laboratories. Recently, there are arguments claiming enlarged gender disparities in scientific productivity due to gender imbalance in housework and childcare (Bell & Fong, 2021). To determine whether there are gender differences in productivity change due to the pandemic, we compared the publication rates of each gender and region between the pre-pandemic and Covid-19 pandemic periods. We first split the period January 2019-December 2020 into two time periods by the start (breakpoint) of the pandemic. To determine the breakpoint was not straightforward, as the start date of the pandemic varies across countries. The start date is defined as the date when 50 or more new Covid-19 cases were confirmed in the first author's affiliation country (Bell & Fong, 2021). For instance, the start date of the US is 5 March 2020, while the start date of Taiwan is 15 May 2021. Since the date is after the end of our study period (i.e. 31 December 2020), articles from Taiwan are all classified as pre-pandemic publications. We applied the same rule to all publications to classify them into pre-pandemic or during Covid-19 pandemic publications. As the bibliometric data only record the month and year of publications, we used the 15th of this month as the publication date. For example, a publication marked as published in March 2020 would be considered published on 15 March 2020. We calculated the average number of publications per day for women and men based on the number of days in the pre-pandemic and Covid-19 pandemic periods.

3.6 Gender disparity in authorship by sub-field

Women tend to be represented more in the social sciences than natural sciences. Given the interdisciplinary nature of urban land science, women's representation may vary by sub-field. We examined the proportion of female authorships in each area to determine women's authorships by sub-field. Instead of classifying the sub-disciplines subjectively, we performed an automatic content analysis on the abstract to digest the major research topics. The analysis is based on non-negative matrix factorization, an algorithm that considers the co-occurrence, frequency, and uniqueness of the terms used among the dataset (Lee & Seung, 1999).

The words which frequently co-occur in abstracts but are not ubiquitous tend to derive a topic. For instance, 'urban' is present in all included documents, thus not contributing to a topic. In contrast, the co-occurrence of 'optical' and 'SAR' is unique and thus derives a topic of image fusion. Each topic consists of a series of keywords (stemmed by co-occurrence), a topic name (assigned by

our inspection of keywords and related documents), and a percentage of authorships by gender (describing the percentage of the topic contributed by female and male first authors). Although a publication may relate to multiple topics, we assigned only one topic to each article based on the highest factor value among all the topics. Topics may be similar in some cases owing to different nomenclature within the same subject area (e.g. impervious surface versus built-up area). The content analysis algorithm was implemented using scikit-learn in Python software, including sub-packages sklearn.decomposition and sklearn.feature_extraction.text. The code and data for content analysis are available at https://urbanization.yale.edu/data.

4. Results

4.1 What is the distribution of women's productivity, impact, and career length?

We analyzed a total of 1,582 papers published from 1983 to 2020 from 73 countries and territories. Our primary analysis considered 91% of the papers with the gender of the first author identified (1,444 manuscripts from 1,157 first authors) (Table S1, Appendix A). We measured bibliometric and affiliation regions for those 1,157 scientists who have published as first authors (Table S2, Appendix A). Overall, 27.6% of the first authors were female, among which 43% were from Asia, 32% from North America, 20% from Europe, 3% from South America, 2% from Africa, and 1% from Oceania.

Women's research output is disproportionate to the proportion of female scientists. Figure 2 shows that for any metric, the proportion of female scientists decreases for the top tier group. For instance, we found that women represent 37.6% of scientists with an h-index below the 25th percentile (h-index = 3) but represent only 10.5% of those with an h-index above the 95th percentile (h-index = 31). Considering total citations over their careers, women represent 37% in the group with total citations fewer than the 25th percentile (N = 55) but only represent 16% in the group with total citations more than the 95th percentile (N = 4,380).

Gender disparities were also present in career lengths and productivity (Figure 2). Our results show fewer female scientists at senior levels and more female scientists at junior levels. On average, men had significantly longer careers (11.8 years) than women (9.4 years) ($P < 0.001$), measured as the time between the first and last publication. Women represent around 35% of the scientists publishing less than two years, including those juniors who just started publishing and those who dropped out of academia within two years of their first publication. In contrast, women represent only 17% of the senior researchers who have been publishing for at least 20 years. To disentangle the two reasons that may lead to this result (high female drop-out rate vs. increasing women's participation in recent years), we repeated the analysis on a subset of those ceasing to publish before 2018 to focus on the drop-out (n = 162 from 1157 scientists). Still, the results show that women have a significantly shorter academic career age of 5.1 years than men (8.1 years) before ceasing to publish ($P < 0.05$), indicating that women drop out of their academic careers earlier than men.

Over their careers, on average, women published 1.6 papers per year, which is significantly fewer than men, who published 2.2 papers per year ($P < 0.05$). In urban land science, the upper quantile of productivity is to publish at least 2.4 articles per year, with only 18% of scientists reaching this productivity level female. We calculated how many years of maternity leave or other forms of leave women might have taken that could explain the productivity gap: the result is one year and five months. That is, the productivity difference between male and female authors would be comparable if subtracting one year and five months in the career length of female authors.

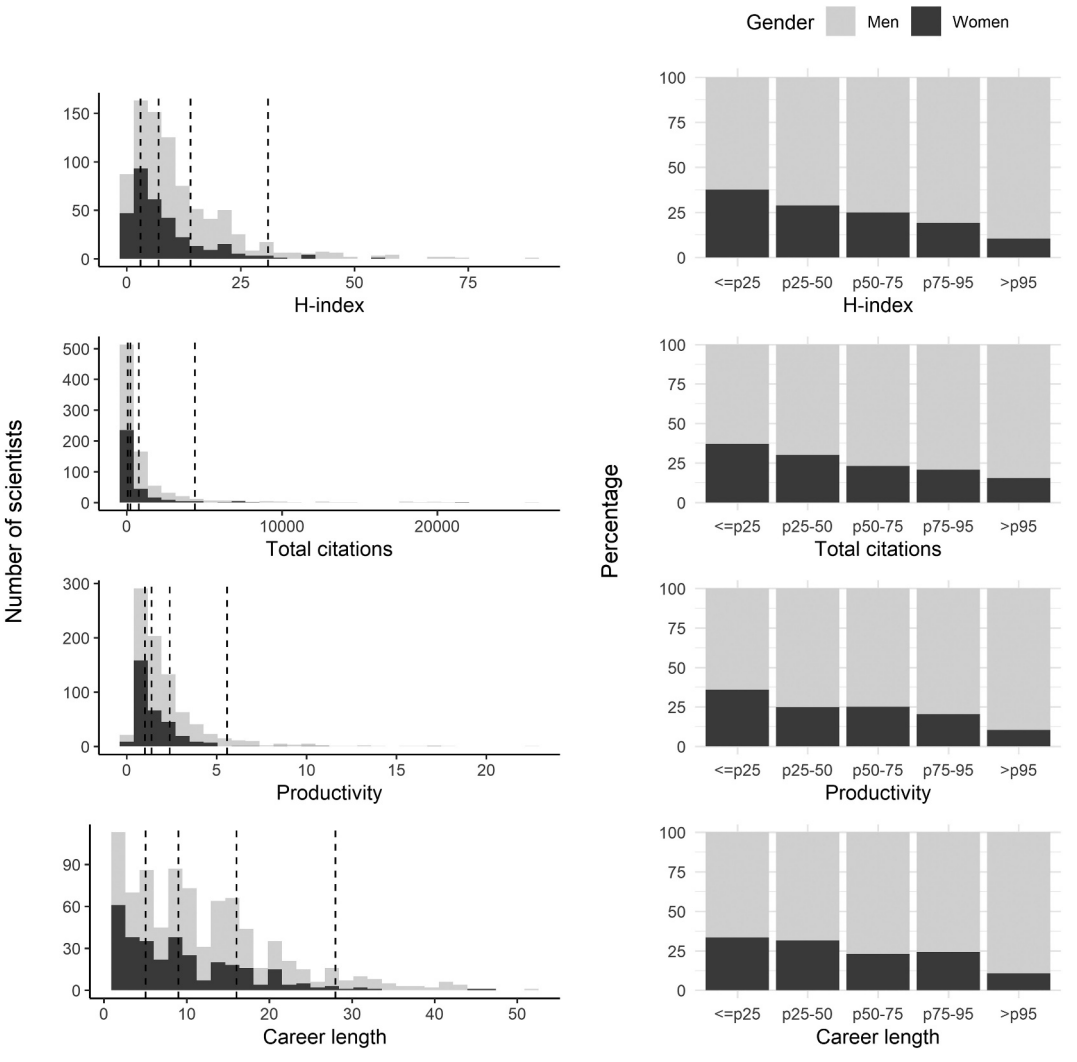

Figure 2. Women's distribution (left) and percentages (right) across the levels of h-index, total citations, productivity, and career length. The dashed lines indicate the 25th, 50th, 75th, and 95th percentiles. Productivity is measured as the number of papers published per year during the researcher's career (spanning from the first to the last publication).

4.2 What percentage of highly influential papers are authored by women?

We consider papers with extraordinarily high citations in the field to be idea-shaping. In urban land science, influential papers (above the 90th percentile) are cited 100 times, and highly influential papers (above the 95th percentile) are cited 165 times (Figure 3, left). In contrast, papers at the 25th and 50th percentiles are just cited 5 and 15 times.

Although women are first authors on 26% of all urban land science papers, they are first authors on 20% and 18% of the influential and highly influential papers, respectively (Figure 3, right). On average, women's papers received 32 citations (SD = 77), which is fewer than 41 citations received by men's papers (SD = 75) ($P < 0.05$).

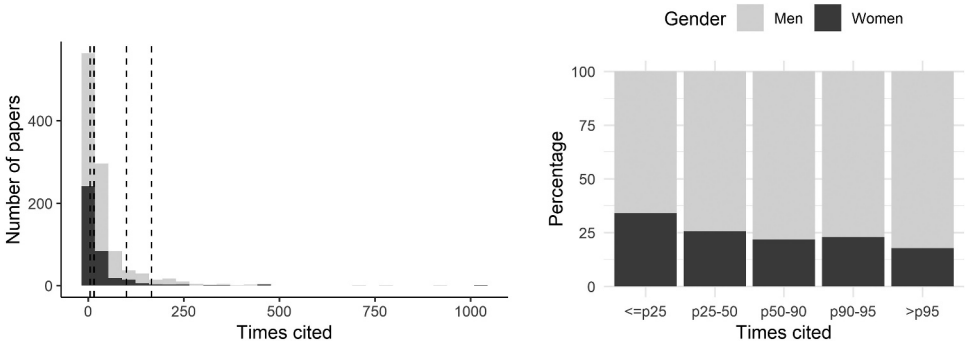

Figure 3. Proportion of female first authorships of influential (above the 90th percentile) and highly influential (above the 95th percentile) papers.

4.3 Have gender disparities changed over time?

Long-term trends

The gender gap between female and male scientists' citation rates and the h-index has decreased in the past two decades, yet female first authorships have increased by only 2% during the same period. We grouped the publication years into five-year periods and measured citation rates and authorships by gender per period. There was only one female-led paper before 1995. Women's papers published during 1996–2000 have been significantly less cited than men's work from the same period (mean values 1.3 and 5.3 for women and men's annual citation rate, respectively, $P < 0.05$). Since the 2000s, women's citation rates have increased substantially (Figure 4). In the following periods until 2020, papers produced from women's works have been cited at frequencies that are not significantly different from men's. The h-index gap has been reduced too. For scientists who started their career before 2006, only 7% of those with an h-index above the 95th percentile are female. In contrast, 25% of those who started publishing after 2015 and have an h-index above the 95th percentile are women. Despite the narrowing citation gap, women have been first authors on fewer than 30% of the papers in the recent two decades (26.3% of first authors were women in 1996–2000 and 28.1% in 2016–2020). The percentage of females among the high h-index and junior research group (25%) is still lower than the percentage of females among the total junior researcher group (33.8%).

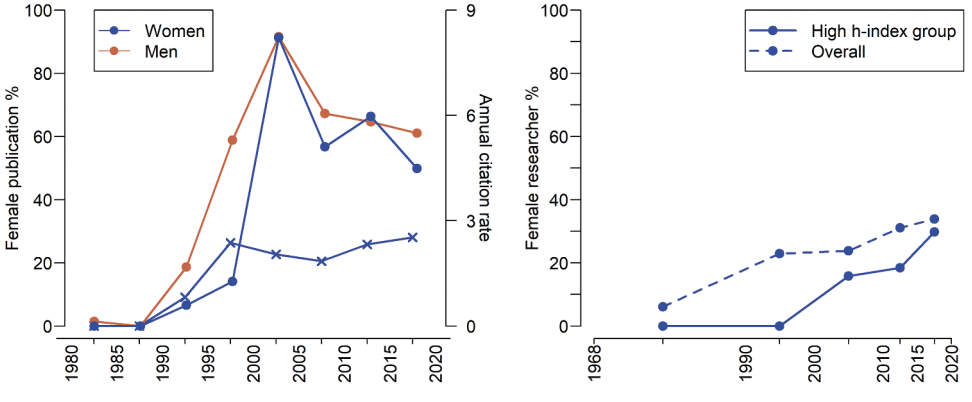

Figure 4. Percentage of female first-authorships (crossed marks), citation rates (round dots) in the literature of urban land science and remote sensing, by five-year period (left). Percentage of female authors in high h-index group (≥ the 90th percentile), by career starting periods (before 1990, 1991–2000, 2001–2010, 2011–2015, and 2016–2020). The dotted line indicates the percentage of females in each generation (right).

Pre-pandemic vs. Covid-19 pandemic periods

We compared publication rates in the pre-pandemic and Covid-19 pandemic periods by gender on a subset of 380 papers published between January 2019 and 2020 December. Overall, more papers per day were published in the pandemic period than previously, with an increase of 6%. However, the pattern of publications differed by gender (Table 1). For women, the publication rate (papers/day) for first authors went up 51%, whereas for men they decreased 11%, indicating an opposite change in productivity for women and men. This pattern was found to be persistent in the subsets of Asia and Europe. For North America, while men's publication rate increased 14%, women's publication rate went up 55%, indicating a more than three-times higher increase in productivity for women compared to men. For authors from Africa and Oceania, the productivity patterns were opposite for women and men, but their sample size was small (i.e. < 3 pre-pandemic papers from women in each region, Table 1).

4.4 How do gender disparities vary by continent?

Women's first authorships range from 13% to 31% across continents. Women in North America are first authors on 31% of the papers (highest), whereas women in Africa and Oceania are first authors on 14% and 13% of papers (lowest) from the region, respectively (Table 2). Overall, women are only 22% of corresponding authors, despite representing 26% of first authors. The largest discrepancy between the first and corresponding authors is in Asia, where women are 24% of the first authors but 15% of the corresponding authors.

The first author and corresponding authors may be the same person, but not always. In total, 67% of the male first authors also led the corresponding role, while 3% fewer (64%) of the female first authors did the same. Sorting the percentages from the lowest, we find that only 9% and 12% of the papers affiliated with an institute in Asia and Africa were led by a woman simultaneously being the first and corresponding author. In North America, more than 55% of the papers were led by a male first and corresponding author, but fewer than one-third of papers were with a female author leading both positions.

Table 1. Percentage change in publication productivity (publications/day) comparing pre-pandemic period to COVID-19 pandemic period, by gender.

	Women change in rates	Men change in rates	Women's pre-pandemic papers	Men's pre-pandemic papers
Overall	51%	−11%	56	154
Africa	−22%	149%	2	6
Asia	59%	−21%	26	91
Europe	55%	−46%	11	31
North America	55%	14%	14	19
Oceania	49%	315%	1	2
South America	−25%	−69%	2	5

Table 2. Gender composition by region: percentages of first authorships, corresponding authors, and same first and corresponding authors relative to the whole dataset.

		% First authors		% Corresponding authors		% Same first and corresponding authors	
Region	Publications	Female	Male	Female	Male	Female	Male
Overall	1444	26	74	22	78	17	49
Asia	668	24	76	15	85	9	36
North America	380	31	69	31	69	28	56
Europe	276	27	73	25	75	22	63
Africa	51	14	86	14	86	12	71
South America	38	26	74	26	74	16	63
Oceania	31	13	87	15	85	10	74

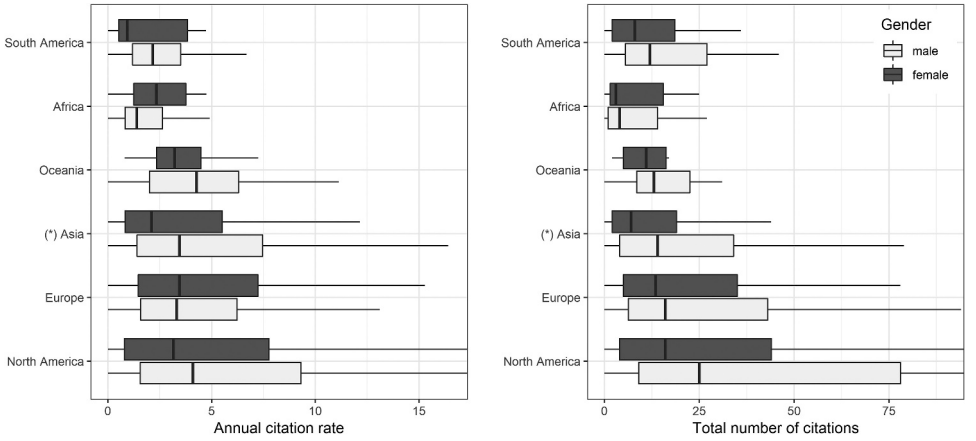

Figure 5. The distribution of annual citation rates and the total number of citations by region and gender. (*) indicates that the mean values for women and men are different at significance level $P < 0.05$.

For citations, we consider the annual citation rate as an estimate of how fast a study has impacted on subsequent studies and the total number of citations as an estimate of the total impacts. The distribution of annual citation rates and total citations by gender is shown in Figure 5. Women's papers from North America have the highest mean of annual citation rates (6.5), followed by Europe (5.2), Asia (4.2), Oceania (3.6), Africa (2.4), and South America (2.3). The annual rates are not significantly different between women and men across all regions except Asia (female = 4.2 vs. male = 6.0, $P < 0.001$). The same pattern was found when we repeated the t-test on the total number of citations; the significant difference was only present in Asia (mean values 19 and 32 for women and men in Asia, respectively, $P < 0.001$).

4.5 How do gender disparities vary by sub-field?

Even though women are first authors on around one-third of the papers, they are disproportionately distributed in subfields. We identified 26 topics in urban land science and remote sensing literature using automated content analysis. There was a cluster of topics about environmental issues related to land cover change (e.g. agricultural land competition, water quality, urban heat Island, and energy emission), a cluster of methodological topics (e.g. object-based classification, SAR and optical data fusion, and machine learning), and a cluster of topics focusing on urbanization patterns (e.g. urban sprawl and planning, urban form, and urban growth simulation; Figure 6). Proportions of female first authors range from the highest in green space studies (43%) to the lowest in energy and emission studies (11%). Five topics where women represent greater than 30% of first authorships are green space, built-up area, global mapping, urban form, and temperature vs. vegetation. In contrast, five topics where women have lower than 20% of first authorships are energy and emission, surface temperature, impervious surface, urban heat Island, and US land change studies. Some STEM-oriented subjects are disproportionally authored by males, such as GHG emissions and urban growth simulation. However, in machine learning, women are 31% of all first authors, higher than the percentage of women first authors in the whole field (27.6%).

5. Discussion

The results show that the proportion of women shrinks in groups with higher research output, citation impact, productivity, and career lengths, akin to a pyramid. Gender disparities are most pronounced among the most productive and impactful authors: women only represent 10% of the scientists with an h-index above the 95th percentile, despite making up 28% of urban land science.

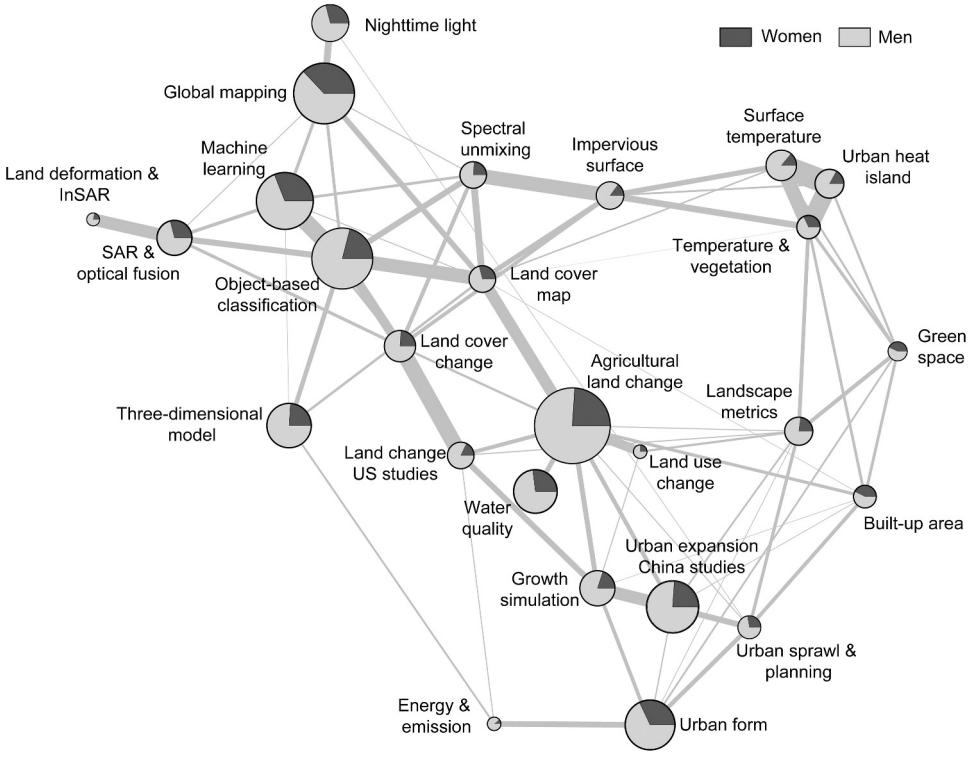

Figure 6. Prevalent cluster and correlation structure of urban land science and remote sensing topics. Each node is a topic with a pie chart indicating the proportion of women-led research, scaled by the number of papers; each line represents a correlation based on the co-occurrence of two themes within article abstracts. Lines are scaled from the highest correlation of 0.23 (between surface temperature and urban heat Island) to the lowest of 0.0003 (between land change US studies and built-up areas). The visualization is generated using the force-directed algorithm forceatlas2 in Gephi.

We also find gender differences in career length. On average, men had significantly longer careers (11.0 years) than women (9.3 years), which is similar to findings in other long-standing scientific fields (J. Huang et al., 2020). Despite large variations, women on average publish less frequently (1.6 papers per year) over their careers than men (2.2 papers per year).

Our findings point to gender disparities in three key domains. First, there are relatively few female authors of highly influential papers. While citation rates between women and men are similar, women are first authors on 20% of all influential papers cited more than one hundred times. Our results support previous findings of a considerable gender difference in highly-cited papers in life sciences, multidisciplinary, environmental science, chemistry, and physics (Bendels et al., 2018). Many reasons have been proposed for why there is a gender gap in citation impacts. For instance, male scientists are more likely to self-cite (Andersen et al., 2019; King et al., 2017). Female scientists are less likely to present their research positively through terms such as 'novel' and 'excellent' in titles and abstracts (Lerchenmueller et al., 2019). Women are also less likely to be involved in collaborative research projects, which is one of the characteristics hallmarking competitive publications in science (Fox, 2001).

Second, women have lower publication rates over their careers than men. Our results show that men on average published 2.2 articles per year when they are actively publishing, which is 37% more than women's average annual publication. Productivity needs focus time, and housework is a trade-off. A study on 1,222 tenure-track scientists who are partnered with someone of another sex revealed that female scientists spend more time on housework than their partner, especially for food preparation and housekeeping, resulting in their spending four times more hours per week on housework than male scientists (Schiebinger & Gilmartin, 2010). In particular, mothers in academia

are more likely to experience a 'motherhood penalty', the decline in professional activities such as publishing and patenting after having a child, compared to fathers and scientists with no children (Lutter & Schröder, 2020). Other studies show that women have more administrative and teaching responsibilities (Aiston & Jung, 2015), and hence have less time for writing and publishing. There are also other factors that determine productivity, including predoctoral affirmation, mentoring provided during early career stages, as well as faculty office space, where subtle biases favoring males have been found (Hopkins, 1999; Moss-Racusin et al., 2012). A study investigating 43,235 invited commentaries revealed that women are 21% less likely to author a commentary, regardless of seniority, expertise, and citation impact (Thomas et al., 2019).

Third, women have shorter careers than men. Counting periods when researchers are actively publishing, we find women only represent 17% of the careers lasting for more than twenty years, despite making up 38% of the junior level publishing less than two years. Another result from our analysis echoes this finding: fewer women are in the corresponding author position than the first author position, indicating that there are fewer women ascending the academic ladder to hold funding and lead a research group. Eliminating currently active careers, we find the length of women's career shorter than men's. The 18% discrepancy between women in the junior and senior level may reflect a 'leaky pipeline' – women disproportionately leave academia when ascending the academic ladder (e.g. from graduate students, postdoctoral researchers, to tenure tracks). For instance, women with child and doctoral degrees are more likely to leave academia or be adjunct faculty than men (Wolfinger et al., 2009). Around the world, 96% of countries offered some pay to mothers during leave, but only 44% offered the same for fathers (Heymann, 2013), which prevents fathers from leading the roles in childcare. Although the higher female proportion in the junior than the senior level may indicate an increasing trend of women's participation in science, it takes time for younger female scholars to become leaders and role models in science.

Importantly, our results show women had increasingly led the first author position since the early 2000s (i.e. 18% over the period 1980–1999 vs. 26% over the period 2000–2020). Women represent 25% high-impact (h-index higher than the 90th percentile) in the generation who just started their career after 2015, in contrast to no women with this metric before 2000. Alongside this trend, women's papers in urban land science have received significantly fewer citations compared to men's papers before 2000, while the citation rate of women's and men's papers have become similar after 2000. However, the increase in female first authorships seems to have plateaued (26% female in 2011–2015 and 28% in 2016–2020).

Recent studies in biomedical science (Muric et al., 2021), public health (Bell & Fong, 2021), and broader STEM fields (Krukowski et al., 2021; Lerchenmüller et al., 2021) find that the Covid-19 pandemic exacerbated gender disparities in publications. Surprisingly, we find the opposite in the field of urban land science: the publication rate for women increased more than for men during the pandemic. Several reasons could explain the differences in results. First, Bell and Fong (2021) were based on a dataset comprising 79% of publications originating from North America, and Muric et al. (2021)'s dataset comprised 67% of publications from Europe and North America. During the gender identification process, Bell and Fong (2021) excluded a number of studies from China and Korea (around 50% of the papers from these countries) because the surnames from these countries are not phonetically related to gender. In urban land science, 50% of the studies were affiliated with an institute in Asia, 25% North America, 18% Europe, and 2–3% for each of Africa, South America, and Oceania. Excluding studies from Asia due to unknown gender is likely to change the conclusion for the urban land science field. Our use of authors' public biography allows us to preserve 85% of the studies from Asia in the analysis (Table S3, Appendix A). Our pandemic analysis shows the same patterns in the subsets of Asia, North America, and Europe that the increase in women's publication rate was larger than men's. We are uncertain about the reason, but one hypothesis is that men are less accustomed to the home working environment, including distractions and interruptions from children and other household members. Thus far, there is no evidence to support the claim that men were affected more than women in terms of productivity when working from home. Second, our

sample selection was based on an interdisciplinary subfield of science and might result in the observed difference compared to broader STEM fields. Moreover, we note that the Covid-19 pandemic patterns presented here might largely reflect the work started before the pandemic because a publication can take more than a year to get published, and the real pandemic effects might still be unknown. Future studies in 2022–2023 are critical to validate the effects of the pandemic. Studies examining a wide range of fields would be helpful to disentangle differences between disciplines in terms of gender and productivity change during the Covid-19. Further studies considering different policy factors (e.g. lockdown, whether schools were open) and individual factors (e.g. marriage, whether having children) will help to contextualize the gender disparity results.

When we began this project, our working hypothesis was that the gender disparities in scientific publishing would be less pronounced in urban land science because the field is both nascent and comprised of relatively young scientists. However, the results were surprising. Together, the findings raise two crucial interrelated questions: Why do these gender differences exist and how can we close the gender gap? Although it is beyond the scope of this analysis to explore the underlying root causes of gender gaps in publishing, we hypothesize that they exist in urban land science because of structural inequities in academia that create and perpetuate systematic bias. A number of studies have pointed to structural sexism affecting women scientists at every stage in research and publishing, and that many social and research practices explicitly and implicitly disadvantage women (Lundine et al., 2018; Nafade et al., 2019). If we are to close the gender gap in scientific publishing in the field of urban land science, we must first start with dismantling structural and systemic bias in research. Institutional and policy changes will be critical to accelerate the transition to gender parity.

6. Concluding remarks

To date, the dominant topic for studies of gender disparities in science is productivity. However, citation rates and highly influential papers provide a metric to shed light on the gender disparities in research impact. We find that while citation rates between women and men have been similar since 2000, women are first authors on only 20% of all influential papers cited more than one hundred times. The results indicate that idea-shaping papers are disproportionately led by men. One clear conclusion from this analysis is that a gender gap exists – the proportion of women shrinks among the highly productive, impactful, and senior authors – those who train the junior scientists and serve as role models for them.

These gender disparity patterns are not only present in long-standing scientific fields but also in a new and multidisciplinary field, urban land science. The results are contrary to our expectation that gender disparities would be less prevalent in this field. Even though women's publications have increased since the 2000s and citation rates between women and men have been similar in the recent decade, the gender gap at the top-tier academic group remains notable – only one in ten is female above the 95th percentile of the h-indices. Our results underscore that the enlarging gender gap among the highly productive and impactful authors is a notable characteristic in urban land science, aligned with other long-standing scientific fields.

Acknowledgments

We thank Thu Ho and Raghav Srivastava for helping to prepare the data set, and Meredith Reba for comments on earlier drafts. T.H.K. Chen's effort was partly supported by NASA LCLUC grant NNX17AH98G.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the NASA grant [NNX17AH98G].

ORCID

Tzu-Hsin Karen Chen http://orcid.org/0000-0002-0343-6147
Karen C. Seto http://orcid.org/0000-0002-4928-2446

References

Aguinis, H., Ji, Y.H., & Joo, H. (2018). Gender productivity gap among star performers in STEM and other scientific fields. *Journal of Applied Psychology*, *103*(12), 1283. https://doi.org/10.1037/apl0000331

Aiston, S.J., & Jung, J. (2015). Women academics and research productivity: An international comparison. *Gender and Education*, *27*(3), 205–220. https://doi.org/10.1080/09540253.2015.1024617

Andersen, J.P., Schneider, J.W., Jagsi, R., & Nielsen, M.W. (2019). Meta-Research: Gender variations in citation distributions in medicine are very small and due to self-citation and journal prestige. *Elife*, *8*(p), e45374. https://doi.org/10.7554/eLife.45374

Bastian, S., Ippolito, J.A., Lopez, S.A., Eloy, J.A., & Beebe, K.S. (2017). The use of the h-index in academic orthopaedic surgery. *JBJS*, *99*(4), e14. https://doi.org/10.2106/JBJS.15.01354

Bell, M.L., & Fong, K.C. (2021). Gender differences in first and corresponding authorship in public health research submissions during the COVID-19 pandemic. *American Journal of Public Health*, *111*(1), 159–163. https://doi.org/10.2105/AJPH.2020.305975

Bendels, M.H., Müller, R., Brueggmann, D., Groneberg, D.A., & Lozano, S. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PloS One*, *13*(1), e0189136. https://doi.org/10.1371/journal.pone.0189136

Bornmann, L., & Daniel, H.D. (2007). What do we know about the h index? *Journal of the American Society for Information Science and Technology*, *58*(9), 1381–1385. https://doi.org/10.1002/asi.20609

Chen, T.H.K., Qiu, C., Schmitt, M., Zhu, X.X., Sabel, C.E., & Prishchepov, A.V. (2020). Mapping horizontal and vertical urban densification in Denmark with Landsat time-series from 1985 to 2018: A semantic segmentation solution. *Remote Sensing of Environment*, *251*, 112096. https://doi.org/10.1016/j.rse.2020.112096

Conley, D., & Stadmark, J. (2012). A call to commission more women writers. *Nature*, *488*(7413), 590-590. https://doi.org/10.1038/488590a

Duch, J., Zeng, X.H.T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T.K., Amaral, L.A.N., & Perc, M. (2012). The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PloS One*, *7*(12), e51332. https://doi.org/10.1371/journal.pone.0051332

Dworkin, J.D., Linn, K.A., Teich, E.G. Zurn, P., Shinohara, R.T., & Bassett, D.S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*(8), 918–926. https://doi.org/10.1038/s41593-020-0658-y

Faria, J.R., & Mixon, F.G. (2021). The marginal impact of a publication on citations, and its effect on academic pay. *Scientometrics 1269* , 8217–8226. https://doi.org/10.1007/s11192-021-04073-z

Fox, M.F. (2001). Women, science, and academia: Graduate education and careers. *Gender & Society*, *15*(5), 654–666. https://doi.org/10.1177/089124301015005002

Grossbard, S., Yilmazer, T., & Zhang, L. (2020). The gender gap in citations of articles published in two demographic economics journals. *Review of Economics of the Household*, 19, 677–697. https://doi.org/10.1007/s11150-020-09508-9

Hechtman, L.A., Moore, N.P., Schulkey, C.E., Miklos, A.C., Calcagno, A.M., Aragon, R., & Greenberg, J.H. (2018). NIH funding longevity by gender. *Proceedings of the National Academy of Sciences*, *115*(31), 7943–7948. https://doi.org/10.1073/pnas.1800615115

Henley, M.M. (2015). Women's success in academic science: Challenges to breaking through the ivory ceiling. *Sociology Compass*, *9*(8), 668–680. https://doi.org/10.1111/soc4.12291

Heymann, J. (2013). *Children's Chances*. Harvard University Press.

Hirsch, J.E. (2010). An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics*, *85*(3), 741–754. https://doi.org/10.1007/s11192-010-0193-9

Holman, L., Stuart-Fox, D., & Hauser, C.E. (2018). The gender gap in science: How long until women are equally represented? *PLoS Biology*, *16*(4), e2004956. https://doi.org/10.1371/journal.pbio.2004956

Hopkins, N. (1999). MIT and gender bias: Following up on victory. *The Chronicle of Higher Education*, *45*(40), B4–B5 http://www-sop.inria.fr/members/Juliette.Leblond/Parite/doc-mixite/Women%20Faculty%20at%20MIT.pdf .

Huang, J., Gates, A.J., Sinatra, R., & Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, *117*(9), 4609–4616. https://doi.org/10.1073/pnas.1914221117

Huang, K., Lee, X., Stone, B., Jr, Knievel, J., Bell, M.L., & Seto, K.C. (2021). Persistent increases in nighttime heat stress from urban expansion despite heat Island mitigation. *Journal of Geophysical Research: Atmospheres 126* (4), e2020JD033831. https://doi.org/10.1029/2020JD033831

Jadidi, M., Karimi, F., Lietz, H., & Wagner, C. (2018). Gender disparities in science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, *21*(3n04), p 1750011. https://doi.org/10.1142/S0219525917500114

Karimi, F., Wagner, C., Lemmerich, F., Jadidi, M., & Strohmaier, M. (2016). Inferring gender from names on the web: A comparative evaluation of gender detection methods. *Proceedings of the 25th International conference companion on World Wide Web* Montréal Québec Canada (International World Wide Web Conferences Steering Committee).

King, M.M., Bergstrom, C.T., Correll, S.J., Jacquet, J., & West, J.D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, *3*, 1–22. https://doi.org/10.1177/2378023117738903

Krukowski, R.A., Jagsi, R., & Cardel, M.I. (2021). Academic productivity differences by gender and child age in science, technology, engineering, mathematics, and medicine faculty during the COVID-19 pandemic. *Journal of Women's Health*, *30*(3), 341–347. https://doi.org/10.1089/jwh.2020.8710

Lee, D.D., & Seung, H.S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791. https://doi.org/10.1038/44565

Lerchenmueller, M.J., Sorenson, O., & Jena, A.B. (2019). Gender differences in how scientists present the importance of their research: Observational study. *bmj*, *367*, l6573. https://doi.org/10.1136/bmj.l6573

Lerchenmüller, C., Schmallenbach, L., Jena, A.B., & Lerchenmueller, M.J. (2021). Longitudinal analyses of gender differences in first authorship publications related to COVID-19. *BMJ Open*, *11*(4), e045176. https://doi.org/10.1136/bmjopen-2020-045176

Long, J.S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, *71*(1), 159–178. https://doi.org/10.2307/2579971

Lundine, J., Bourgeault, I.L., Clark, J., Heidari, S., & Balabanova, D. (2018). The gendered system of academic publishing. *The Lancet*, *391*(10132), 1754–1756. https://doi.org/10.1016/S0140-6736(18)30950-4

Lutter, M., & Schröder, M. (2020). Is there a motherhood penalty in academia? The gendered effect of children on academic publications in German sociology. *European Sociological Review*, *36*(3), 442–459. https://doi.org/10.1093/esr/jcz063

Mauleón, E., & Bordons, M. (2006). Productivity, impact and publication habits by gender in the area of materials science. *Scientometrics*, *66*(1), 199–218. https://doi.org/10.1007/s11192-006-0014-3

McDonald, R.I., Mansur, A.V., Ascensão, F., Colbert, M., Crossman, K., Elmqvist, T., Güneralp, B., Haase, D., Hamann, M., Hillel, O., Huang, K., Kahnt, B., Maddox, D., Pacheco, A., Pereira, H.M., Seto, K.C., Simkin, R., Walsh, B., Werner, A.S., … Gonzalez, A. (2020). Research gaps in knowledge of the impact of urban growth on biodiversity. *Nature Sustainability*, *3*(1), 16–24. https://doi.org/10.1038/s41893-019-0436-6

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L.A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4* (1), 1–9. https://doi.org/10.1186/2046-4053-4-1

Moss-Racusin, C.A., Dovidio, J.F., Brescoll, V.L., Graham, M.J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479. https://doi.org/10.1073/pnas.1211286109

Muric, G., Lerman, K., & Ferrara, E. (2021). Gender disparity in the authorship of biomedical research publications during the covid-19 pandemic: retrospective observational study. *Journal of Medical Internet Research*, *23*(4), e25379. https://doi.org/10.2196/25379

Nafade, V., Sen, P., & Pai, M. (2019). Global health journals need to address equity, diversity and inclusion. *BMJ Global Health*, *4*(5), e002018. https://doi.org/10.1136/bmjgh-2019-002018

Nunkoo, R., Thelwall, M., Ladsawut, J., & Goolaup, S. (2020). Three decades of tourism scholarship: Gender, collaboration and research methods. *Tourism Management*, *78*(p), 104056. https://doi.org/10.1016/j.tourman.2019.104056

Organisation for Economic Co-operation and Development. (2012). *Education at a Glance 2012* (Paris: OECD).

Pinchoff, J., Mills, C.W., Balk, D., & Liu, S.Y. (2020). Urbanization and health: The effects of the built environment on chronic disease risk factors among women in Tanzania. *PloS One*, *15*(11), e0241810. https://doi.org/10.1371/journal.pone.0241810

Rivers, E. (2017). *National Science Foundation* (National Science Foundation) https://ncses.nsf.gov/pubs/nsf21321/downloads

Sarsons, H. (2017). Recognition for group work: Gender differences in academia. *American Economic Review*, *107*(5), 141–145. https://doi.org/10.1257/aer.p20171126

Schiebinger, L., & Gilmartin, S.K. (2010). Housework is an academic issue. *Academe*, *96*(1), 39–44. https://www.jstor.org/stable/20694619 .

Seto, K.C., Güneralp, B., & Hutyra, L.R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, *109*(40), 16083–16088. https://doi.org/10.1073/pnas.1211658109

Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature News*, *495*(7439), 22. https://doi.org/10.1038/495022a

Shughrue, C., Werner, B., & Seto, K.C. (2020). Global spread of local cyclone damages through urban trade networks. *Nature Sustainability*, *3*(8), 606–613. https://doi.org/10.1038/s41893-020-0523-8

Sotudeh, H., & Khoshian, N. (2014). Gender differences in science: The case of scientific productivity in nano science & technology during 2005–2007. *Scientometrics*, *98*(1), 457–472. https://doi.org/10.1007/s11192-013-1031-7

Stokes, E.C., & Seto, K.C. (2019). Characterizing urban infrastructural transitions for the sustainable development goals using multi-temporal land, population, and nighttime light data. *Remote Sensing of Environment*, *234*, 111430. https://doi.org/10.1016/j.rse.2019.111430

Stone, B., Hess, J.J., & Frumkin, H. (2010). Urban form and extreme heat events: Are sprawling cities more vulnerable to climate change than compact cities. *Environmental Health Perspectives*, *118*(10), 1425–1428. https://doi.org/10.1289/ehp.0901879

Tamblyn, R., Girard, N., Qian, C.J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. *CMAJ*, *190*(16), E489–E499. https://doi.org/10.1503/cmaj.170901

Thomas, E.G. Jayabalasingham, B., Collins, T., Geertzen, J., Bui, C., & Dominici, F. (2019). Gender disparities in invited commentary authorship in 2459 medical journals. *JAMA Network Open*, *2*(10), e1913682–e1913682. https://doi.org/10.1001/jamanetworkopen.2019.13682

West, J.D., Jacquet, J., King, M.M., Correll, S.J., & Bergstrom, C.T. (2013). The role of gender in scholarly authorship. *PloS One*, *8*(7), e66212. https://doi.org/10.1371/journal.pone.0066212

Wolfinger, N.H., Mason, M.A., & Goulden, M. (2009). Stay in the game: Gender, family formation and alternative trajectories in the academic life course. *Social Forces*, *87*(3), 1591–1621. https://doi.org/10.1353/sof.0.0182