How to achieve gender parity in science? Providing global evidence on key educational and economic drivers

Elvira González-Salmón[1], Zaida Chinchilla-Rodríguez[2], Nicolás Robinson-García[3] and Gabriela F. Nane[4]

[1]elviragonzalez@go.ugr.es
ORCID 0000-0003-3826-766X
Información y Comunicación, Universidad de Granada, Spain
[2]zaida.chinchilla@csic.es;
0000-0002-1608-4478
Instituto de Políticas y Bienes Públicos (IPP), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
[3]elrobin@ugr.es;
0000-0002-0585-7359
Información y Comunicación, Universidad de Granada, Spain
[4]g.f.nane@tudelft.nl;
0000-0002-3614-1820
Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands

Abstract

Gender parity in science depends on a complex interplay of social, economic and educational variables. In this study, we compile a longitudinal dataset at the country level combining scientific bibliographic data from Dimensions, with the World Bank Open Data (WBOA), and the UNESCO Institute for Statistics (UIS). Our goal is to identify conditions and pathways that could lead to gender parity in different world regions, by applying time-series forecasting methods (ARIMA and Exponential Smoothing), along with correlation analysis and Bayesian networks. While results vary by region, one recurring recommendation emerging from our models is the need to increase the number of researchers and the percentage of women graduating in Engineering, Manufacturing, and Construction, as this appears to be a critical driver for reducing gender disparities in the scientific workforce.

Keywords: Gender parity, country-level indicators, Bayesian Networks, science national scientific systems

1. Introduction

Over the past decades, the representation of women[1] in academia has increased significantly. In countries such as Serbia (European Commission, 2024), nowadays there are more women than men working in science, and gender parity (often invoked as a prerequisite for gender equality) has been reached in disciplines such as Health Sciences in Spain (Aramayona et al., 2022). However, this positive trend is threatened by a backlash, a resistance to gender equality that is leading to a decline in support for feminist initiatives (Williamson, 2020). This has led to talk of "gender fatigue", which encapsulates the idea that we, as a society, have already dealt with some gender inequalities, thus we should move forward (Kelan, 2009).

Recent evidence does not support the statement that gender equality in science has been achieved (Larivière et al., 2013; Jaramillo et al., 2025; Zhao et al., 2023). While women are no longer systematically excluded from academia as in the past, discrimination has evolved into new forms of segregation (Walby, 1990). This implies that while women have gained access to academia, they do not experience the same opportunities or face the same constraints as their men counterparts. For instance, in the representation of women as last author (González-Álvarez & Sos-Peña, 2020) or corresponding author (Chinchilla-Rodríguez, et al., 2025), citations in STEM (Kuchanskyi et al., 2023), collaborations (Morillo et al., 2025), presence in higher academic ranks (e.g. Spoon et al., 2023), journal editorial boards presence (Burg et al., 2022) or conference participation (Jones et al., 2014), amongst others.

These disparities suggest that achieving gender parity is not a straightforward process, but rather one shaped by nuanced and systemic characteristics. Understanding gender parity therefore requires complementing this concept with contextual information that captures the mechanisms underlying it. In this paper, we aim to enrich the concept of gender parity by integrating it with other gender-related indicators that allow for a more nuanced analysis of its driving forces.

The present project has two main objectives. First, it aims to assess whether current country-level trends are moving towards gender parity in science and over what time horizon this may occur. Second, it seeks to identify the structural variables that are associated with higher or lower levels of gender parity across regions. Together, these objectives could inform science policy by highlighting structural patterns and conditions linked to different trajectories toward gender parity. Using data from Dimensions, World Bank Open Data (WBOA) and the UNESCO Institute for Statistics (UIS) from 1990 until 2020, alongside methodologies such as time-series forecasting techniques (ARIMA and Exponential Smoothing) and Bayesian

[1] Throughout this article, we use women/men to refer to gender rather than female/male. Quoted articles and names of variables are not modified; therefore, some of them may not follow this guideline.

Networks (BNs), we aim to examine how various factors intersect with the gender parity indicator and to estimate when gender parity may be reached in academia, and under which conditions.

2. Literature Review

Gender (in)equality in science

Notable gender differences exist in various aspects of the academic environment. These include differences in authorship positions (i.e. Demaine, 2021), in productivity (i.e. Campbell & Simberloff, 2022), citation impact (i.e. Wu, 2024), and the contributions that women authors tend to do in research (i.e. Macaluso et al., 2016), amongst others. Some authors have looked for the reasons as to why this happens and found that gender differences begin before entering academia (i.e. Ceci et al., 2014). One widely used indicator to assess the extent to which there are gender differences, is gender parity. In this paper, we focus specifically on this concept, which can be defined as "concerns relative equality in terms of numbers and proportions of women and men, girls and boys, and is often calculated as the ratio of female-to-male values for a given indicator." (EIGE, 2025). Building on this literature, our study focuses on gender parity as an aggregate indicator, examining how it varies geographically and over time.

Gender parity in science has been studied from different angles. Some research has focused on the consequences that lack of gender parity can have in science, including a homogeneous workforce that reinforces a vicious cycle that limits women's opportunities and negatively affects their professional trajectories (i.e. Aiston & Jung, 2015). It has also been studied how a lack of gender parity has implications for the nature and direction of the research produced, as evidence shows that women investigate different topics than men (i.e. Ceci et al., 2014; Ruiz-Baena et al., 2025) and are more inclined to integrate gender perspectives into their research (i.e. Key & Sumner, 2019).

Generally, however, gender parity has been studied in a descriptive way at the discipline level and at the country level. Since each discipline and country has their own cultural characteristics, it is usual that an analysis focuses exclusively on one discipline and/or country. When looking at research that focuses on gender parity in the different disciplines, although there are geographical and sub-disciplinal variations, we find that literature generally finds a lack of gender parity in STEM related fields, and a closer situation to parity in Social Sciences and Humanities and Biological Sciences (e.g. Casad et al., 2021; Nygaard et al., 2022; Sugimoto & Larivière, 2023).

For this research, we are particularly interested in the insight that the literature has on gender parity in countries, at a more macro level. When looking at parity in authorships, we find that women represent less than 30% of fractioned authorships globally (Larivière et al., 2013) and a third of the researchers (UNESCO, 2024). At the country level we find reports like the She Figures by the European Commission. The latest edition, from 2024, found that women in

Europe represent 34% of the total researchers, with data ranging from 50% (e.g. North Macedonia, Serbia, Montenegro), to around 27-29% (e.g. Germany, Hungary, Czechia), and even lower levels in countries such as Japan (17.8%) or South Korea (22.2%) (European Commission, 2024). This wide cross-country variation highlights the need for comparative approaches to better understand the structural conditions associated with parity trajectories.

Research also studies gender parity grouping countries according to different criteria. When studying gender and science, the aggrupation of countries is usually marked by geography (e.g. using the World Bank classification, such as Momeni et al., 2022, or continents, such as Frehill & Zippel, 2011), economic level (e.g. using OECD data, such as Stoet & Geary, 2018) or cultural similarities usually based in language (such as Thelwall, 2020, or Bauer et al., 2012). The She Figures itself studies the situation of gender in academia in the European Union. Interestingly, Narasimhan (2021) found that as the wealth of a country increases, so does the participation of women in science; and then it falls, marking a U-shaped curve, although there are differences between developed and developing countries (Narasimhan, 2021).

Gender parity predictions in science

Gender equality has not yet been globally achieved, and gender parity is often used as a reference point in assessing progress toward reducing gender imbalances in science. By examining how gender parity evolves across countries and over time, and by situating these trajectories within broader educational, economic, and scientific-system contexts, this study aims to explore the structural configurations associated with different pathways toward parity.

Thus, we first want to analyse country-level trends moving towards gender parity and the timeframe it may take to be reached. There are several methods to forecast expected growth. First, there are classical statistical time-series models, including time-series decomposition (used in Abdollahi, 2020 to forecast oil prices), exponential smoothing (used in Rao et al., 2023 to forecast energy demand in China), and ARIMA models (used in Njenga, 2024 to measure domestic credit growth in Kenya), which focus on trend and autocorrelation in univariate series. Then, there are regression-based approaches for longitudinal data, such as time-series regression (used in Hertzog et al., 2024 to study suicide deaths associated with climate change), vector autoregressive (VAR) models (used to measure the spread of COVID-19 in Rajab et al., 2022), and linear mixed models (used to study the ice retreat of the Arctic sea in Horvath et al., 2021). Third, there are machine-learning approaches, notably neural networks (used for rainfall forecasting in Darji et al., 2015), which aim at capturing non-linear patterns.

Using bibliometric methods, some authors have explored these methodologies as well, asking themselves when we will reach gender parity in science using predictive models. However, to the best of our knowledge, these efforts remain few (Table 1). The most notorious attempt is that of Holman et al. (2018). In their research, they use PubMed and arXiv databases to identify STEMM fields that will not reach gender parity without intervention. Using linear mixed models, they conclude that the gender gap is likely to persist for generations in some fields, such as Surgery, Computer Science, Physics and Maths. Jemielniak & Wilamowski (2025) forecasted gender parity in all fields using 2001-2021 Scopus data and saw that the pace of growth decreases over time. Additional studies have sought to predict gender parity in STEM. For instance, Wang et al. (2021) stated that the most optimistic predictions show that gender parity in Computer Science authorship will be reached after 2100, and Msosa et al. (2022) found that "the assumption that women are increasingly assuming positions once considered "male" roles, overcoming outdated stereotypes, and thriving and succeeding in the STEM profession on different continents is far from true" (p. 257). Lastly, López-Aguirre (2019) studied the post-war Colombian context and discovered that, all else being equal, gender parity will be reached in 10 years for the Humanities and it would take 150 years for Engineering. This study extends these works by combining forecasting with multivariate structural analysis which is not jointly addressed in prior studies.

Table 1. Non-exhaustive studies predicting gender parity in science and reasons leading to it.

Study	Geographical Focus	Fields	Methods	Causal Forces	Data
Holman et al. (2018)	Global	STEMM	Linear mixed models	No	PubMed, arXiv, 2002-2016
Wang et al. (2021)	Global	Computer Science	ARIMA	No	Semantic Scholar, 1970-2019
Msosa et al., (2022)	Global	STEM	Correlation, Regression Analyses	Yes	UNESCO, 2013-2017
López-Aguirre (2019)	Colombia	All	Correlation, linear regression models	No	UIS, RICYT, OCyT, COLCIENCIAS, World Bank Data, 2000-2015
Jemielniak & Wilamowski (2025)	Global	All	Trend extrapolations	No	Scopus and CrossRef, 2001-2021

However, these predictive efforts do not usually take into account factors external to academia and do not go beyond the number of men and women within it. It is crucial to acknowledge that science, and the dynamics surrounding gender parity within it, do not evolve in isolation. As a result, these models could run the risk of overlooking broader social, cultural, and economic contexts, and thus treating gender parity largely as an isolated academic outcome. Thus, as a second objective we want to identify key societal, educational and economic forces that lead to gender parity.

Bayesian Networks (Pearl, 1995) are popular graphical tools that allow for the exploration of structural interdependencies within multivariate data. They are typically used to explore these relationships, and for predictions. They are common methods in numerous fields, such as health (Witteveen et al., 2018; Sesen et al., 2013; Arora et al., 2019), energy transition (Borunda et al., 2016; Leicester et al., 2016; Shi et al., 2024) and engineering (Moradi et al., 2022; Misirli & Bener, 2014; del Águila & del Sagrado, 2015). Bayesian Networks are increasingly employed in bibliometric data. For instance, Ibáñez et al. (2011) applied BNs to explore relationships among bibliometric indicators and found them useful for addressing bibliometric questions. Furthermore, Robinson-Garcia et al. (2020) used this methodology to predict contributorships of researchers for given publications and Sun et al. (2023) employed it to examine the relationship between citation counts and potential influencing factors.

3. Methods

3.1 Data

In this study we analyse the percentage of women in science to forecast gender parity levels and understand the structural factors associated with such parity. To do so, we examine a total of 96,925,045 publications indexed in Dimensions between 1990 and 2021. Unlike Scopus or Web of Science, which emphasise the selection of certain scientific output, Dimensions prioritises comprehensiveness (Visser et al., 2021), offering, in principle, greater coverage of global scientific output. Based on this dataset, we identified 23,066,108 unique authors using the Dimensions Researcher ID. For analytical purposes, we focused exclusively on researchers with a minimum of five publications, ensuring the inclusion of authors with a consolidated publication record. The final sample comprises 8,247,978 authors.

To obtain data on various national statistics to enrich the data on the number of researchers by gender, we retrieved data from the UNESCO Institute for Statistics (UIS) and the World Bank Open Data (WBOD) platforms. The UIS database constitutes the official United Nations repository for country-level statistics in the fields of education, science, and culture. WBOD is an open-access database managed by the World Bank, which provides statistical data on various development indicators for countries worldwide. The datasets were collected in November 2023, covering the maximum available time span for each source: 1996–2021 for UIS and 1990–2021 for WBOD. Given the structure of the dataset, a maximum of 4,795 observations is possible for each indicator. We extracted a set of specific indicators from both databases, as detailed in Table 2. It is important to note that the data series are incomplete for certain years and countries (see Appendix 1 for information on handling missing data). The number of available observations and the corresponding available observation percentage is reported in the last two columns of Table 2. It is noteworthy the high proportion of complete data for length of maternity leave and labor force participation rate, whereas information regarding proportion of time spent on unpaid domestic and care work is severely lacking.

Table 2. List of variables collected from UIS and WBOD, with their number and percentage of observations out of all countries/years collected data.

Full list of variables	Abbreviation	Source	Observations (max. 4.795)	% of observations (100% = 4795)
Proportion of time spent on unpaid domestic and care work, female (% of 24 hour day)	Domestic	WBD	173	3.6%
Length of paid maternity leave (calendar days)	Mat. Leave	WBD	4,311	89.9%
Female share of graduates in Arts and Humanities programmes (%, tertiary)	Arts & Hum	WBD	1,227	25.6%

Female share of graduates in Business, Administration and Law programmes, tertiary (%)	Busi, Admin & Law	WBD	1,075	22.4%
Female share of graduates in Agriculture, Forestry, Fisheries and Veterinary programmes (%, tertiary)	Agri, Forest, Fish & Vete	WBD	1,158	24.2%
Female share of graduates from Science, Technology, Engineering and Mathematics (STEM) programmes, tertiary (%)	STEM	WBD	1,058	22.1%
Female share of graduates in education (%, tertiary)	Edu	WBD	1,228	25.6%
Female share of graduates in engineering, manufacturing and construction (%, tertiary)	Engi, Manu & Const	WBD	1,181	24.6%
Female share of graduates in health and welfare (%, tertiary)	Heal & Welf	WBD	1,229	25.6%
Female share of graduates in Information and Communication Technologies programmes, tertiary (%)	TIC	WBD	1,015	21.2%
Female share of graduates in Natural Sciences, Mathematics and Statistics programmes (%, tertiary)	Natu, Math & Stat	WBD	1,034	21.6%
Female share of graduates in other fields than Science, Technology, Engineering and Mathematics programmes, tertiary (%)	Others	WBD	1,055	22.0%
Female share of graduates in services (%, tertiary)	Servi	WBD	1,084	22.6%
Female share of graduates in Social Sciences, Journalism and Information programmes (%, tertiary)	Social, Jour & Info	WBD	1,029	21.5%
Adolescents out of school, female (% of female lower secondary school age)	Adolescents	WBD	1,650	34.4%
School enrollment, primary (gross), gender parity index (GPI)	School	WBD	3,596	75.0%
Trained teachers in upper secondary education, female (% of female teachers)	Teachers	WBD	530	11.1%
Labor force participation rate, female (% of female population ages 15+) (modeled ILO estimate)	Labor Force	WBD	4,321	90.1%
Researchers in R&D (per million people)	Researchers million	WBD/UIS	1,199	25.0%
Research and development expenditure (% of GDP)	R&D	WBD/UIS	2,005	41.8%

Based on the collected data, we constructed a dataset comprising the following variables: year, country, region, number of researchers, percentage of women researchers using Dimensions data, and the additional analytical variables from UIS and WOBD. Each row in the dataset corresponds to a unique year-country observation, with the relevant variables for that specific context. The number of years for which data are available varies per country, depending on the data availability in Dimensions. We included only those years for which researcher data were available, for a given country, in Dimensions. For instance, in cases where Dimensions data for a country were limited to three years, only those years were incorporated, regardless of data availability in the UIS or WBOD sources. For Dimensions data quality purposes, we had to limit our analysis until 2020.

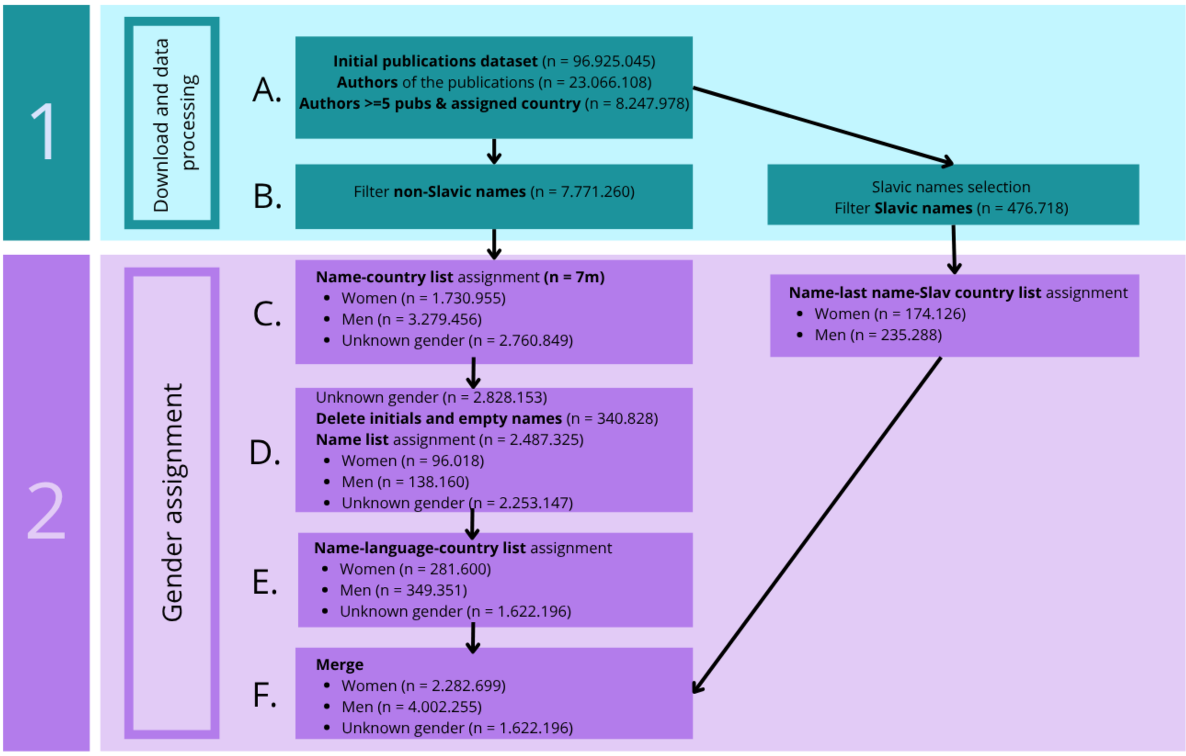
3.2 Data processing

To be able to perform analyses focused on gender, we first assigned a gender to researcher names extracted from the Dimensions database (Section 3.2.1). This step was necessary to transform author-level metadata into analytically usable, gender-disaggregated indicators. Another important step within the data processing was the interpolation and selection of statistical data to address missing values (Section 3.2.2). Both procedures are described in detail below.

3.2.1 Gender assignment

First, we assign gender to researchers' names from the Dimensions dataset. For each author we have the first name, last name and affiliation, thus enabling the application of a gender assignment algorithm. Gender was assigned to authors based on bibliographic data using the WikiGenDex algorithm (González-Salmón & Robinson-García, 2024a), which draws upon open data sources, primarily Wikidata and the World Gender Name Dictionary (WGND) (Raffo, 2021). This algorithm integrates four name lists: (i) name–surname–Slavic country–gender, (ii) name–country–gender, (iii) name–gender, and (iv) name–language–country–gender. By cross-referencing these lists with author names from Dimensions, the algorithm estimates the probability of a name being associated with a woman or a man. It takes into account the country of origin, language, and national particularities in gender assignment practices, including cases where surnames indicate gender, particularly in Slavic countries.

Figure 1. Flow of the gender assignment process.



WikiGenDex follows a sequential process (Figure 1). First, it separates author names based on whether they originate from Slavic or non-Slavic countries. For non-Slavic countries, it applies the name–country–gender list. For Slavic countries, it uses the name–surname–country–gender list, which incorporates both first names and surnames, as surnames can encode gender in Slavic countries. Names without assigned gender are subsequently processed by excluding missing data and applying a name–gender list that disregards country of origin, assigning gender when there is broad international consensus. Finally, a list based on WGND data is applied, using country languages and language-specific names to infer gender. All previously assigned names are merged in a final step.

Compared with other gender assignment algorithms, WikiGenDex favours precision over recall, with an estimated 80% recall (González-Salmón et al., 2024b). Nonetheless, any approach to gender assignment based on names is subject to limitations and potential biases. Algorithms that infer gender from names treat gender as a binary variable, excluding non-binary and other gender identities. Additionally, gender assignment algorithms exhibit uneven performance across countries, with a well-documented lower accuracy and coverage for Asian names.

3.2.2 Interpolation & data selection

We obtained the percentage of women researchers and the number of total researchers by country and year from the Dimensions data. For the rest of the variables, we turned to UIS and WBOD. When dealing with UIS and WBOD data we encountered issues related to missing data, as depicted by Table 1. It is important to note that the lack of data does not always reflect the same underlying cause. Several factors could contribute to data gaps, including limited resources, political conflict, or a lack of institutional interest. In the context of gender data, Criado Pérez, in her book Invisible Women, highlights how the failure to recognize the importance of gender-disaggregated data hinders the production of research that benefits society as a whole (Criado Pérez, 2019). Furthermore, some countries lack complete data because they did not exist during the entire period under study (World Bank Data, 2025). Finally, the absence of China's data is notable.

To deal with the missing data we implemented a natural cubic spline interpolation. With this kind of interpolation, a "series of unique cubic polynomials are fitted between each of the data points, with the stipulation that the curve obtained be continuous and appear smooth" (McKinley & Levine, 1998, p. 1). During this process, some interpolated values exceeded the bounds of the indicator's support; for example, some interpolated percentage values exceeded 100. These values were removed. Likewise, any interpolated values that produced negative numbers were also excluded from the dataset. Moreover, to assess the reliability of the interpolation and investigate whether the high percentage of interpolated values introduces too much noise, we calculated an error rate, which allowed us to investigate the performance of interpolation using test data, as explained in the Appendix 1 (A.1). We identified different patterns of missing data, and the performance analysis indicated different thresholds of interpolation that were appropriate for different patterns. In this way, for each country we interpolated a different proportion of data, and some data remained missing. Nonetheless, some countries had so little data, that no interpolation technique could compensate for, and were consequently excluded from the analysis.

Furthermore, to avoid excluding countries due to the absence of a single data point, we sought to retrieve the missing information from alternative sources, such as research articles or policy reports. This approach was applied only in a limited number of cases where the lack of one variable would otherwise have led to the complete removal of a country from the dataset, allowing us to preserve country coverage while maintaining data integrity. It is fully documented to ensure transparency in Appendix A.2, which shows the countries and variables for which this situation occurred, along with the method employed to address each case.

The process resulted in a final dataset comprising 17 variables and 81 countries, whose distribution across regions is listed in Table 3. For the analysis we merged countries in the regions, according to the World Bank classification. Notably, the regional distribution of countries was not uniform, with an overrepresentation of Europe & Central Asia, North America, and the Middle East & North Africa regions. In contrast, Sub-Saharan Africa is underrepresented, containing only 17.02% of the countries of that region in the dataset. This uneven regional representation reflects persistent inequalities in the availability of internationally comparable data (Pedró et al., 2025).

Table 3. Number of countries per region and percentage they represent from each region.

Region	Number of countries	Percentage from region
East Asia & Pacific	8	30.7%
Europe & Central Asia	39	71.7%
Latin America & Caribbean	10	31.3%
Middle East & North Africa	11	52.4%
North America	2	100.0%
South Asia	3	37.5%
Sub-Saharan Africa	8	17.0%

As we did not have sufficient data points in North America (which includes just two countries, resulting in 60 observations), we merged North America and Europe & Central Asia data. Similarly, we merged South Asia data with East Asia & Pacific data. In this way we ensure that the Bayesian Network Analysis is sufficiently robust.

3.3 Methodological design

To carry out the analysis, we first apply ARIMA and exponential smoothing models to Dimensions data to forecast progress towards gender parity in science at the regional level. We then incorporate additional statistical variables and, following a correlation analysis, estimate Bayesian networks. These networks are used to identify structural conditions and configurations associated with reaching a threshold of 50% women in science across regions. Below, we explain these steps in more detail.

First, we forecast the number of women by using historical data from Dimensions alone. To evaluate the models’ performance, the data was split into a training set (from 1990 until 2015) and a test set (from 2015 until 2020). We applied standard models: Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing methods (Hyndman &

Athanasopoulos, 2018). ARIMA models incorporate both past observations and errors from previous time steps to refine predictions. The selection of parameters and the number of lagged observations considered is typically determined through maximum likelihood estimation. Exponential smoothing methods also rely on past data and assign exponentially decreasing weights to past observations, prioritizing recent observations while diminishing the influence of older ones. We applied both modelling techniques for our data using the forecast package in R (Hyndman, 2024) and ended up choosing ARIMA for showing a better performance.

To account for the dependencies within the data, we computed a correlation matrix for the variables in the dataset (excluding the year column). In order to analyse the relationship between all variables and to further deepen the analysis, we used Bayesian Networks (BNs). BNs are graphical models that capture dependencies between multiple variables. These relationships are depicted by directed edges (arcs) connecting nodes, each corresponding to a random variable. The network structure can be inferred from data under specific modelling assumptions. BNs are widely used in various domains due to their flexibility and interpretability (Scutari & Denis, 2021).

In this article, BNs help incorporate additional information beyond the number of researchers by gender, such as women's enrolment in different levels and fields of education, investment in R&D, amongst many others. With their use, we examine whether additional information can enhance the understanding and prediction of gender parity. This approach captures not only the relationships between gender parity and its potential determinants, but also the interdependencies among those determinants themselves.

For implementation purposes, BN were estimated using Uninet and the visualization was done with iGraph. Uninet is a tool that facilitates the modelling and representation of BNs. It also enables the final predictive analysis central to our study. Uninet is an uncertainty analysis software package that focuses on dependence modelling for high dimensional distributions. The software is available at: https://lighttwist-software.com/uninet/. Uninet does not allow missing values. Therefore, we performed a final round of interpolation using the previously interpolated dataset, rather than the original data. Since this intermediate dataset already exhibited substantially fewer missing values, this additional interpolation step ensured compatibility with the software without introducing distortions associated with extensive imputation.

For the predictive modelling of the Bayesian Networks, we employed the software Uninet as well. Based on the fitted univariate distributions, this software also allowed us to condition on a 50% share of women, serving as an operational definition of gender parity. It also allowed us to analyse the configurations of variables associated with this threshold within a given region. As the distributions' support is determined in Uninet by the range of observations, the share of women in some regions could not have been conditioned at 50%. To overcome this challenge, we fitted Weibull distributions to the variable of interest (percentage of women researchers) for Latin America & the Caribbean, the Middle East & North Africa, and Sub-Saharan Africa, as these regions required distributional modelling due to data limitations. For the remaining regions, the empirical distribution functions were sufficient to perform the conditionalization directly, without the need for employing a parametric distribution and its fitting. We further confirmed that the differences between the original data and the conditionalized one were significant using the Welch's t-test.

4. Results

4.1 Overall description

Since 1990, the number of researchers has increased greatly (Figure 2). The share of women researchers has increased as well. In 1990, women represented 18.28% of researchers, in 2000 it was 23.38%, in 2010 it was 27.14% and in 2020 28.94%, while the share of researchers with unknown gender also increases over time, reflecting both data expansion and limitations of name-based gender assignment (Table 4). These numbers, however, vary considerably by country. If we take, for instance, 2020, proportion of women researchers can differ considerably, from countries with low proportions (such as Japan or Germany with 18.11% and 31.28% of women researchers), to countries where gender parity is approached (such as Romania and Tunisia with 49.37% and 49.71% of women researchers). Despite this substantial increase, women remain a minority among researchers globally.

Table 4. Number of women, men and unknown gender of researchers and percentage of women globally, by year.

	1990	**2000**	**2010**	**2020**
Women	100,487	267,454	675,001	1,158,159
Men	403,784	756,276	1,394,713	1,929,223
Unknown	45,491	120,145	417,543	913,951
% Women	18.3%	23.4%	27.1%	28.9%

Figure 2. Number of researchers over time by gender aggregated across all countries.

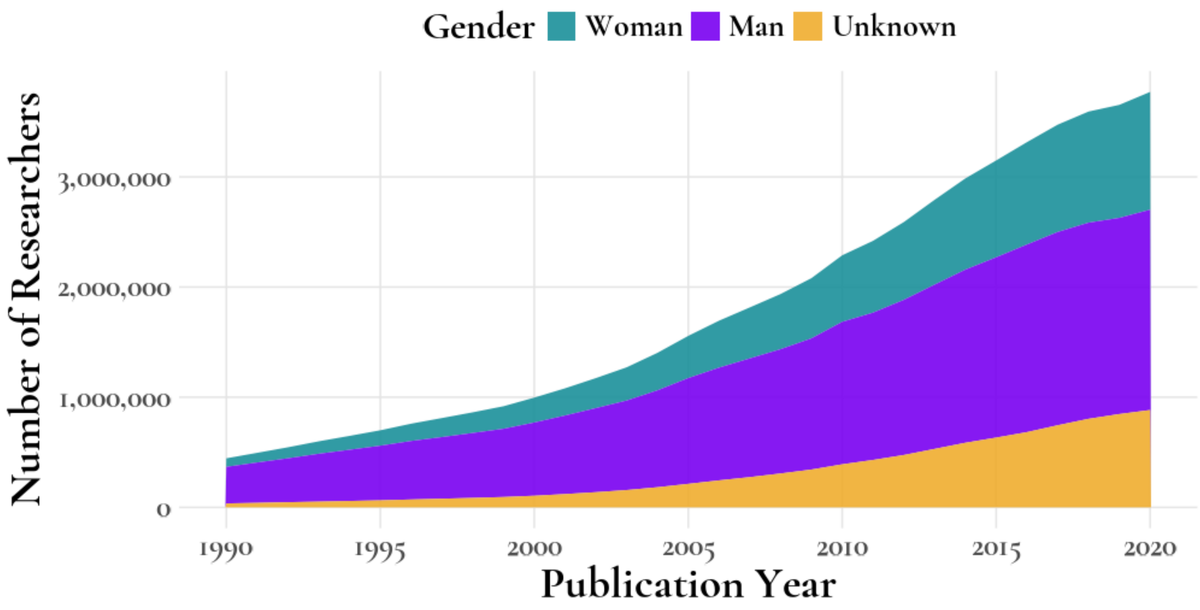
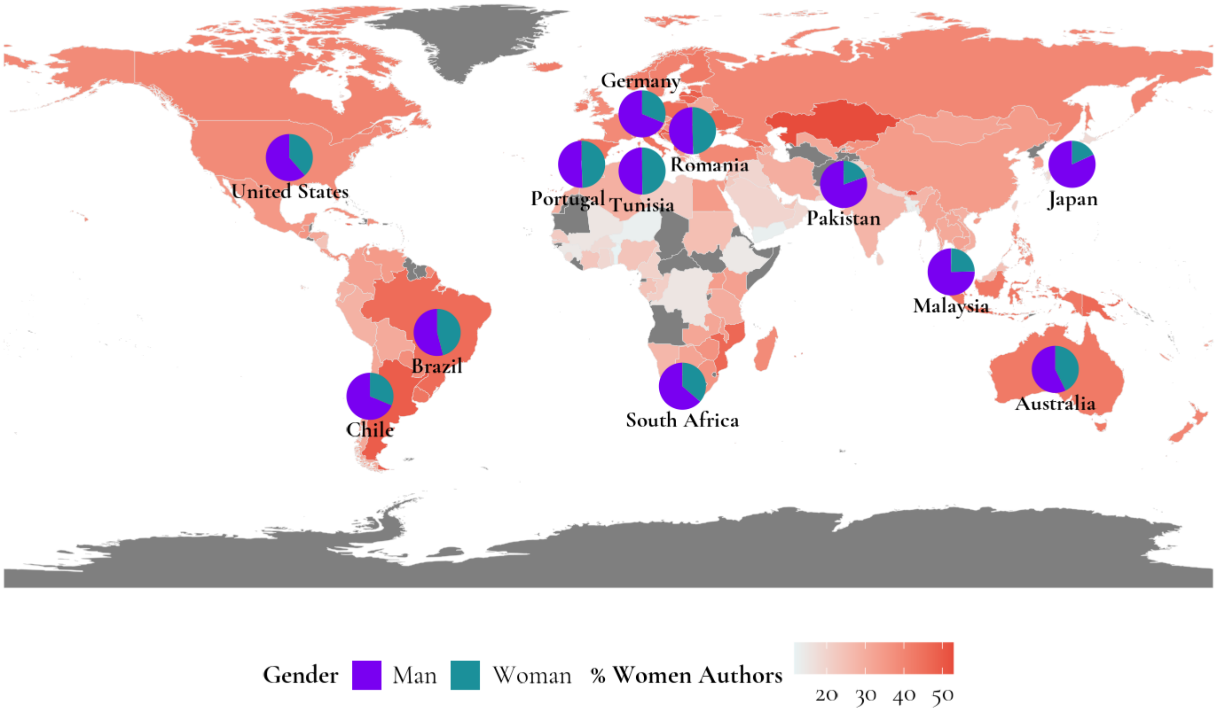


Figure 3 shows the global distribution of women researchers in 2020. Countries are shaded according to the proportion of women researchers, with darker tones indicating higher percentages. Overlaying pie charts in selected countries illustrate the distribution of men and women researchers in countries with proportions of women researchers above the global average of 28.94% (e.g., Portugal, Romania or Tunisia) or below it (e.g. Japan, Pakistan or Chile). This variation highlights the importance of adopting a global perspective, as focusing exclusively on Europe or the United States risks overlooking the distinct dynamics and challenges present in other countries and regions.

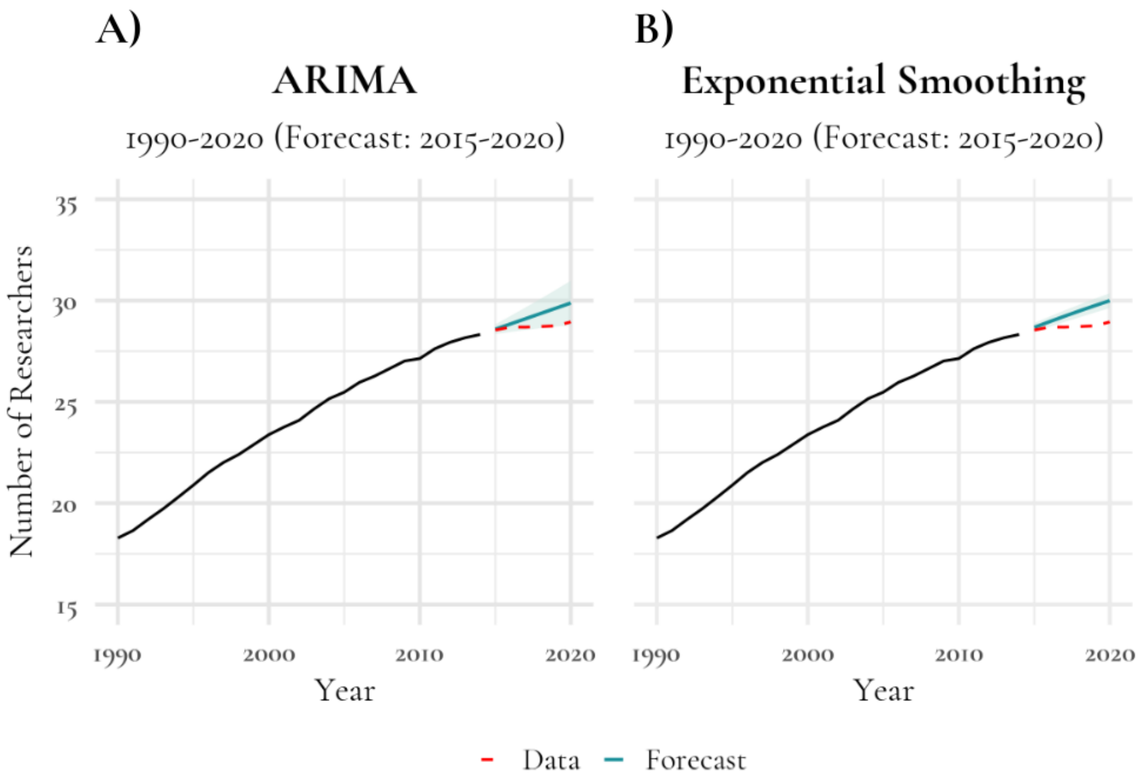
Figure 3. Percentage of women authors by country, 2020. Only those countries with >100 researchers in the year 2020 are presented.



4.2 Gender parity forecast

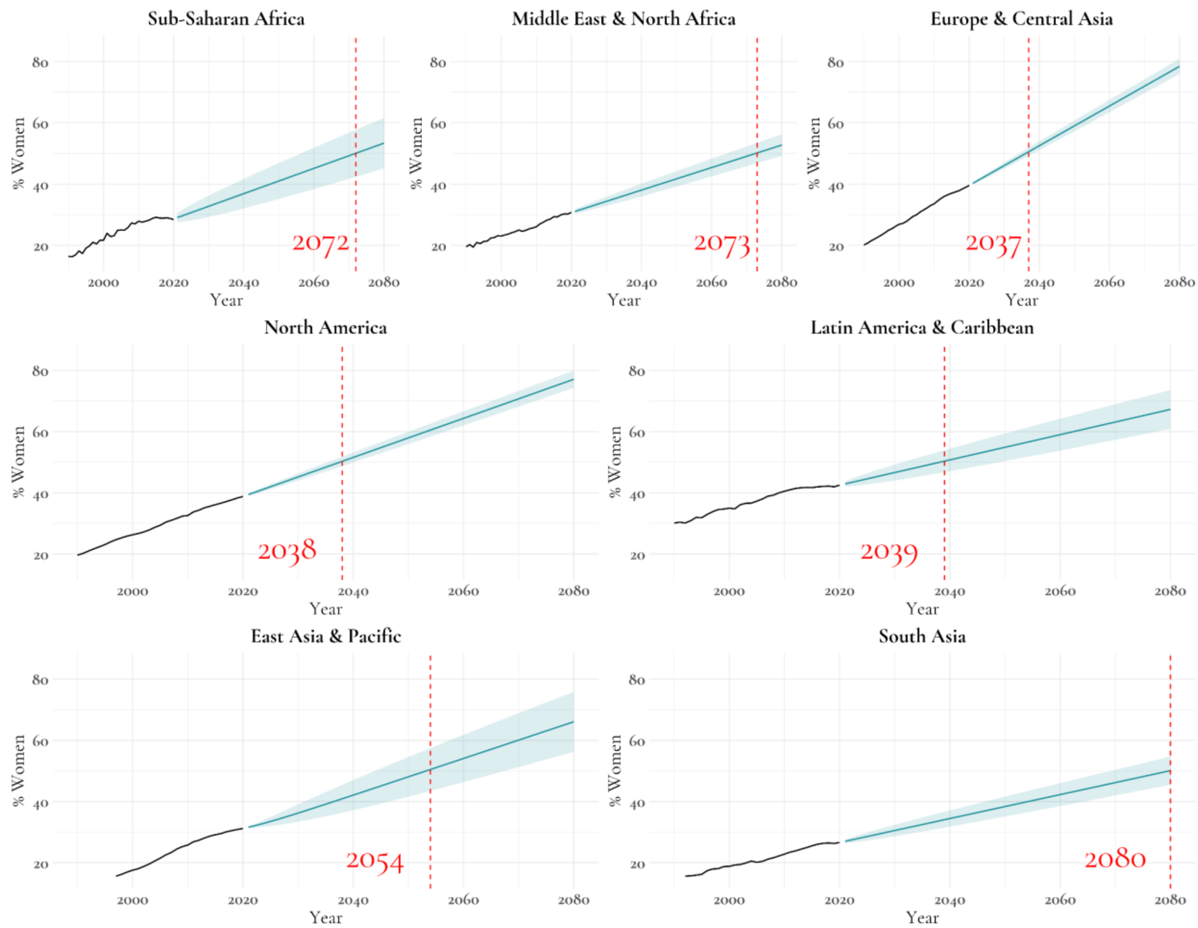
Given our objective of estimating when gender parity in science may be reached, we used ARIMA and Exponential Smoothing on our Dimensions data, first globally to test the models and then by region. To evaluate the performance of the models, the data have been split into a training and a test set, for which both the forecast (in blue) and the actual data (in red) is presented. The models also present the confidence intervals that capture the uncertainty of the predictions. Based on the historical data until 2015, both models tend to slightly overestimate the number of female researchers, as we can see in Figure 4 (to see figures corresponding to all regions, see Appendix A.3). The true values are, however, captured by the uncertainty intervals. The interval in ARIMA is relatively wide and accommodates both modest increases as well as stronger increases. Thus, we rely on ARIMA to produce regional forecasts.

Figure 4. ARIMA (A) and Exponential Smoothing (B) for all women researchers globally, 1990-2020 (Forecast from 2015 until 2020).



Now we apply ARIMA models to regional percentages (Figure 5). This approach allows us to estimate when gender parity may be reached in each region, assuming the continuation of observed trends in the share of women researchers. The results indicate that gender parity (50% women researchers) is projected to be reached in slightly more than a decade in Europe & Central Asia (2037), North America (2038), and Latin America & Caribbean (2039). The model further forecasts that gender parity in East Asia & Pacific is projected around 2054. For the remaining regions, nearly half a century remains before parity is achieved, with projected dates around 2072 for Sub-Saharan Africa, around 2073 for the Middle East & North Africa, and around 2080 for South Asia.

Figure 5. ARIMA forecasts, along with uncertainty bands, for all regions.



We note that, although informative, these forecasts do not provide information on underlying factors that could lead to an increase/decrease of women in science or in gender parity. Thus, we aim at enriching these predictions with other variables, such as the percentage of graduates in STEM that are women or expenditure in R&D. We investigate whether this information can lead to additional insights into the potential increase of women in science.

4.3. Structural Determinants of Gender Parity

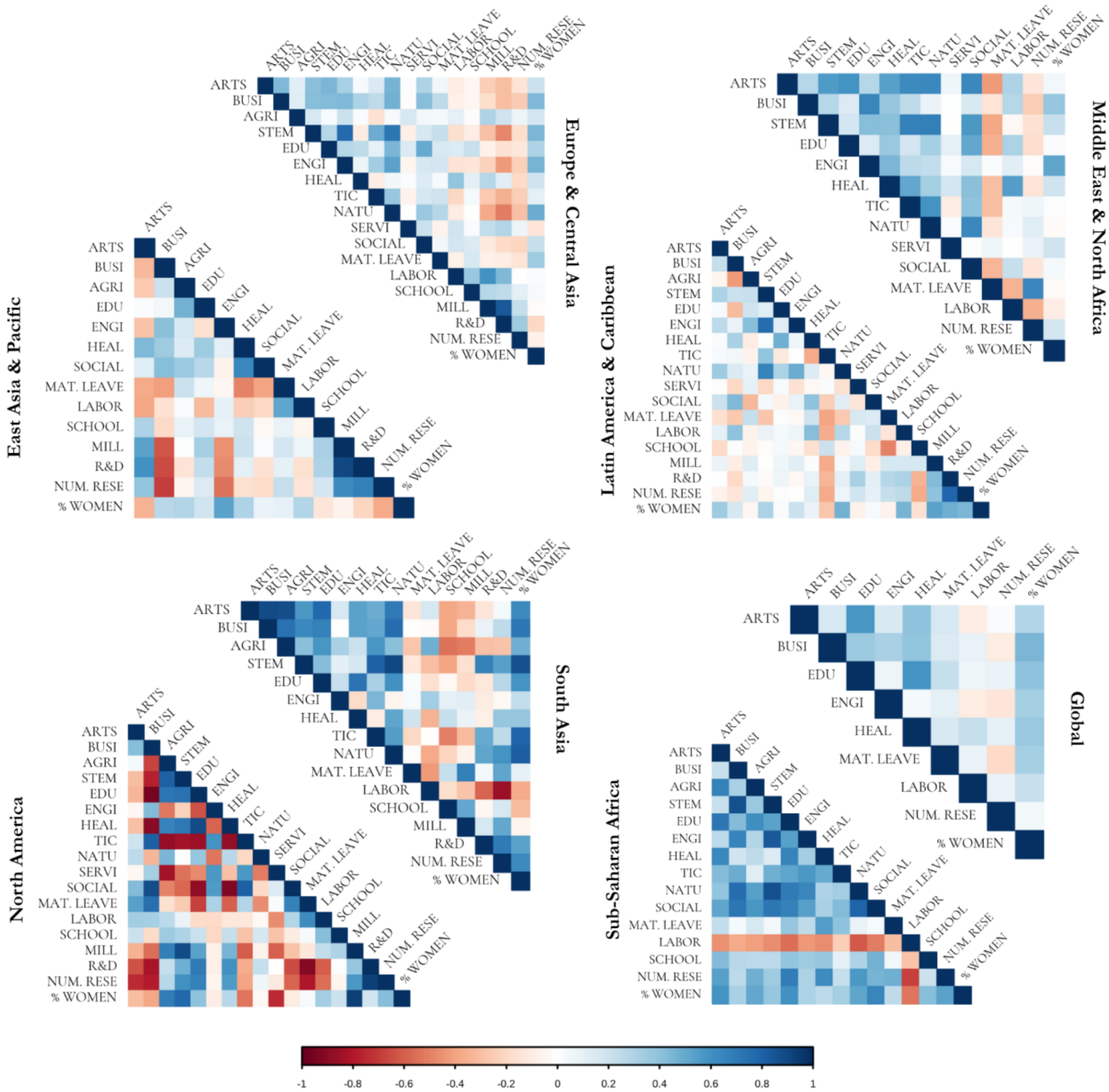
4.3.1. Correlation structures and multivariate dependencies

We first report correlation matrices using all variables for each region (Figure 6). The results reveal correlations among the variables across all regions, with stronger associations in North America and weaker ones in Latin America and the Caribbean. This gradient suggests that the structural relationships among the indicators vary by context, leading to more tightly connected patterns in the northern part of the American continent. The percentage of women also shows several positive and negative associations with other variables, indicating that gender composition interacts with broader demographic, economic, or institutional factors rather than operating independently. However, although these correlations are statistically observable, they are not sufficiently strong to justify the exclusion of any variable from the analysis.

Regional patterns reveal distinct correlation structures. A recurrent feature is the presence of strong negative correlations across multiple regions. For instance, in East Asia, the number of researchers is negatively correlated with the percentages of graduates in Business, while in South Asia, it shows a negative correlation with women's labor force participation. In the Middle East and North Africa, the number of days with paid maternity leave exhibits the weakest association with the remaining variables, followed by the total number of researchers. In Europe and Central Asia, the least connected variable is the share of GDP devoted to R&D. In East Asia and the Pacific, the percentage of women graduating in Business shows a notably strong relationship with several variables, while in Latin America and the Caribbean the weakest links appear for the percentage of women graduates in Natural Sciences. North America presents marked negative correlations for Business graduates, Natural Sciences graduates, R&D expenditure, and researcher counts. In Sub-Saharan Africa, the most striking pattern is the negative association between women's labor force participation and the rest of the variables. At the global level, the structure aligns with these regional tendencies: the most negatively correlated indicators are women's labor force participation and the number of researchers, highlighting their distinct behaviour across contexts.

Overall, the analysis shows that the variables are interdependent, not only in relation to the percentage of women but also among themselves. This multidimensional dependency structure provided a clear justification for adopting Bayesian Networks, as they are specifically designed to model complex systems where multiple variables are interrelated.

Figure 6. Correlation matrix by region for variables in Table 2 (abbreviations are in this table as well).



4.3.2. Bayesian Network results and conditional configurations

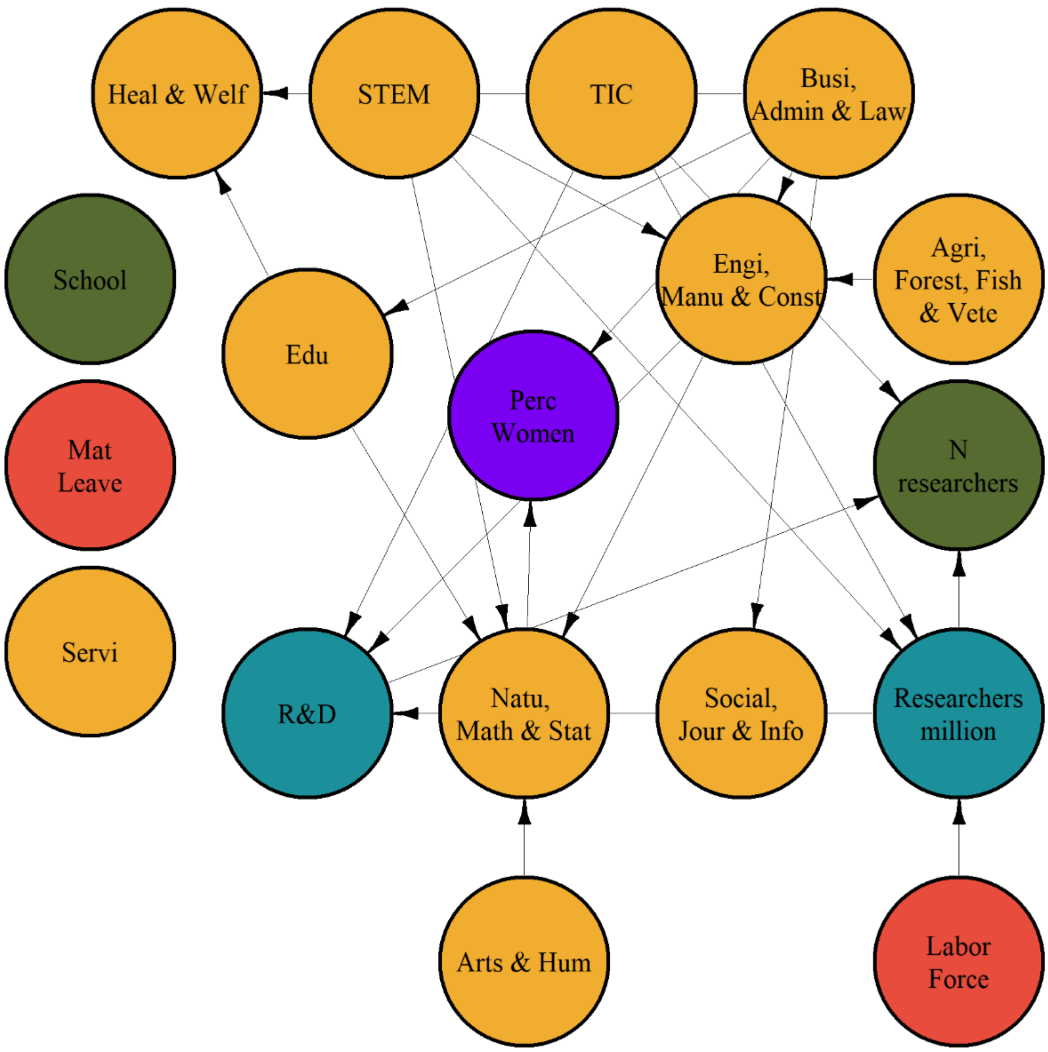
To further investigate the relationships among all variables, we modelled the data with Bayesian Networks (BNs, Table 5). We created BNs for each region, using the available data. As we can see in Figure 7, for the Bayesian Network for Europe & Central Asia (other networks can be found in Appendix A.4), each node corresponds to a variable, such as the percentage of women researchers (Perc Women), percentage of graduates from a field that are women (e.g., Arts & Humanities, STEM, Health & Welfare), and socioeconomic factors (e.g., R&D, Mat. Leave). The meaning of the abbreviations are provided in table 2. The variable of interest, the percentage of women researchers (Perc Women), is in purple, educational variables are in yellow, gender-sensitive labour indicators in red, innovation systems metrics are in blue and scientific capacity variables are in green. Arcs indicate directed conditional dependencies (showing which conditional information contributes best to the joint distribution) and are learned from the data. Table 5 displays the mean and standard deviation for each variable. The network's dense structure reflects the strong interconnectivity between policy variables and gender metrics, underscoring the structural complexity associated with gender parity.

Table 5. Mean and Standard deviation of variables on the Bayesian Network produced by Uninet.

	East Asia & Pacific		Europe & Central Asia		Latin America & Caribbean		Middle East & North Africa		North America		South Asia		Sub-Saharan Africa	
Variables	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Perc_Women_Researchers	27.1	9.23	35.5	9.18	32.1	8.61	23.2	9.24	31.4	9.05	24.7	9.14	23	9.13
N of researchers	21,000	35,700	16,500	28,900	7,200	17,400	2,310	5,940	26,600	68,800	18,700	33,500	1,050	1,990
Arts & Hum	51.6	55.4	61.6	44.9	48.4	90.9	42.3	60.5	61.6	43.8	50.1	48.4	30.2	116
Busi, Admin & Law	50.9	21.4	52.8	56.7	62.7	53.3	46.1	28.5	52.8	55.3	46.8	27.7	43.5	52.5
Agri, Forest, Fish & Vete	35.5	37.6	52	85.7	39.6	51.6			51.9	83.5	35.3	32.9	29.4	48.3
STEM			31.7	13.2	44.8	55.2	35.9	43.2	31.6	12.9			31.6	54.8
Edu	71.2	29.7	75.7	42.4	85.6	43.3	47.4	62.6	75.7	41.3	60.3	35.1	47.4	32.9
Engi, Manu & Const	44.5	78.7	25.3	40.6	44.9	81.9	31.5	93.6	25.1	39.6	43.8	72.4	7.88	83.5
Heal & Welf	58.7	32.4	62.4	87.9	71.3	43.7	72.4	60.1	63.3	85.7	39.6	57.5	69.2	83.9
Natu, Math & Stat			52.7	23.6	55.1	42.4	57.6	61.3	52.7	23			45	33.2
Servi			45.9	82.8	71.8	174	12.1	90.9	46.3	80.7				
Social, Jour & Info	61.5	34.7	72.3	61	59.4	62.2	63.3	59.7	72	70.9			51.4	82.6
TIC			31.8	41.8	53.2	73.8	30.3	53.7	31.3	41			33.4	112
Mat. Leave	80.6	34.2	143	68.8	103	18	67.6	35.7	139	70.9	83.9	31.7	84.7	27.9
Labor Partic	53.5	8.03	51.5	8.25	46.3	6.53	25.3	12.1	51.9	8.21	47.4	12.3	64.5	7.25
School	1.0	0.0	1.0	0.0	1.0	0.0			1.0	0.0	1.0	0.0	1.0	0.0
Researchers million	2,310	2,200	2,420	2,230	378	336			2,490	2,200	1,720	2,100		
R&D	1.3	1.2	1.3	1.0	0.4	0.3			1.3	1.0	1.0	1.1		

In the case of Europe & Central Asia, we see that the percentage of women researchers receives incoming arrows from the percentage of graduates in Business, Administration & Law and Natural, Mathematics & Statistics that are women. Within the BN structure, this indicates associations between these educational variables and the share of women researchers. Since variables of other types do not show direct conditional links to the percentage of women researchers, this pattern suggests that gender parity in this region is more closely related to horizontal field distribution than to overall participation in academia. That is, the horizontal distribution rather than the vertical (Tomassini, 2021). We observe that variables that influence others are mainly the percentage of women graduating in STEM, Business, Administration & Law, and Information & Communication sciences. On the other hand, those that exhibit more dependencies with other variables are the percentage of women graduating in Natural, Mathematics & Statistics and variables related to the scientific capacity of a country (percentage of GDP spent in R&D and researchers per million inhabitants). It is noteworthy that the variables referring to the percentage of women graduating in Services, the weeks of paid maternity leave and school enrolment of girls are not connected to any other nodes, suggesting a lack of conditional association with the percentage of women researchers or other variables.

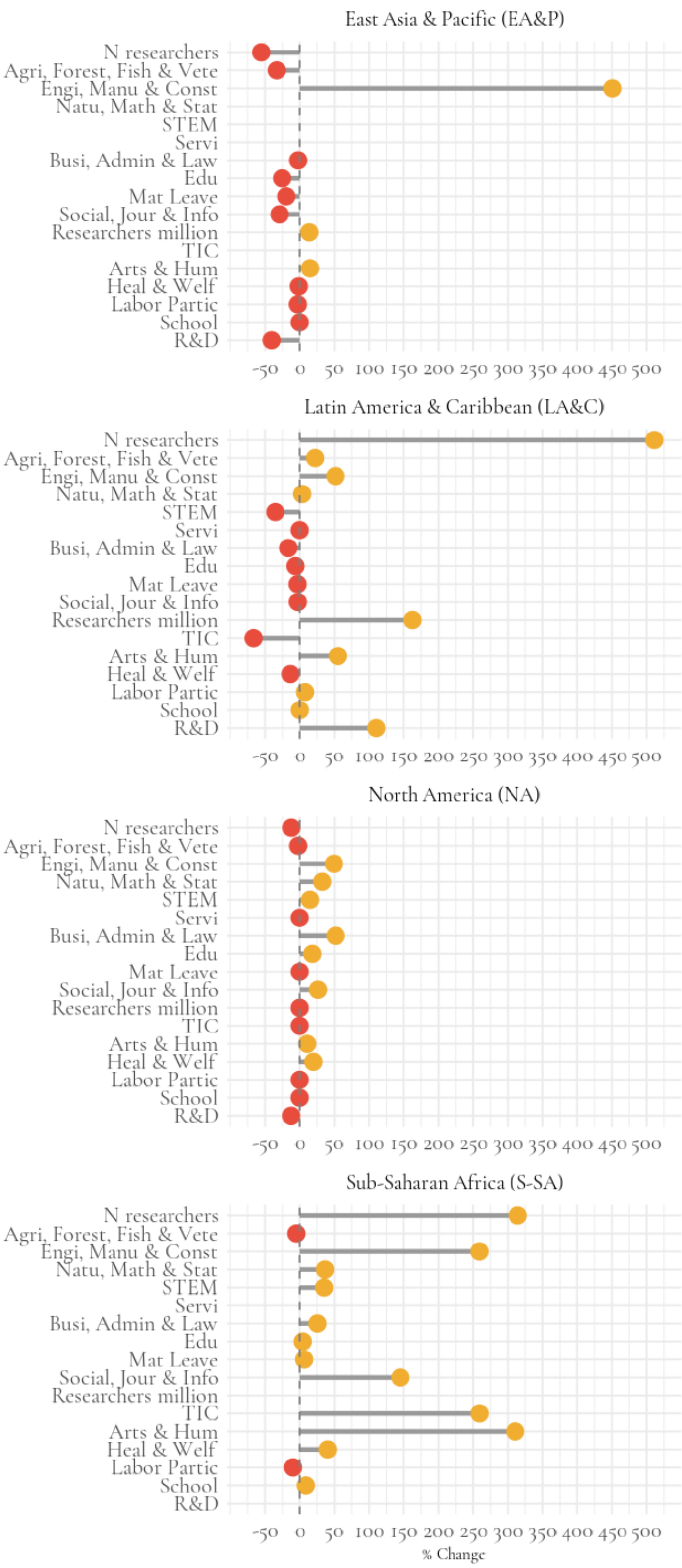
Figure 7. Bayesian Network for Europe & Central Asia. Abbreviations can be found in table 2.



In Appendix A.4 we can see the BNs for the other regions. Overall, we find that educational variables show the strongest conditional associations with the percentage of women in science. This may partly reflect the larger number of educational indicators included in our dataset. However, it is noteworthy that this pattern does not hold in Sub-Saharan Africa or Latin America & Caribbean. In these regions, the results point to a stronger role of vertical segregation, whereas in the other regions the main challenge lies in the horizontal segregation of women across fields. It is also interesting to see how Business, Administration & Law and Engineering, Manufacturing & Constructions had a big presence in terms of relationships within the network in most regions, suggesting that the percentage of women that graduate from those areas of knowledge have a high impact in the remainder of the socioeconomic sphere. Finally, the percentage of women researchers appears primarily as an outcome variable within the networks, rather than as a variable that exerts conditional influence on the rest of the system.

After fitting the BNs, we were able to conditionalize the variable of interest, that is, the percentage of women researchers in each region, at a specified level. This implies a diagnostic analysis, where we investigate the setting under which the variable of interest exhibits the parity threshold level. Thus, we were able to see how other variables are estimated to shift in order to reach 50% of women researchers (Table 6). Figure 8 illustrates how the values of various variables could change, according to our model under the parity threshold scenario in the region. Yellow dots indicate variables that would need to increase, while red dots indicate variables that would need to decrease to reach 50% women researchers. And the distance from the current values (dashed line) represents the amplitude of the estimated change for each variable. The bottom-right area of the graph displays the variables that, globally, on average, exhibit the largest deviations from current values under the parity scenario.

Figure 8. Percentual change after conditioning, with respect to the original data. Red dots are negative changes; yellow dots are positive.



Variables with biggest average % change

	EA&P	E&CA	LA&C	ME&NA	NA	SA	S-SA
N researchers	−55.29	−13.33	511.11	493.07	−12.03	185.03	314.29
Engi, Manu & Const	450.56	49.80	51.67	0.00	49.40	384.02	259.14
Arts & Hum	14.92	11.04	55.17	48.23	10.88	0.00	310.60
Researchers million	13.85	−9.92	162.70	NA	0.00	57.56	NA
TIC	NA	0.00	−66.35	0.00	0.00	NA	259.28

Table 6. Mean and Standard deviation of variables on the Bayesian Network produced by Uninet after conditionalizing to 50% of women researchers.

	East Asia & Pacific		Europe & Central Asia		Latin America & Caribbean		Middle East & North Africa		North America		South Asia		Sub-Saharan Africa	
Variables	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
N of researchers	9,390	2,290	14,300	26,600	44,000	38,700	13,700	15,000	23,400	62,900	53,300	50,500	4,350	3,810
Arts & Hum	59.3	39.2	68.4	32.7	75.1	93.5	62.7	41.8	68.3	31.7	50.1	48.3	124	128
Busi, Admin & Law	49.8	21.5	80.8	49.8	52.4	44.2	46.1	28.5	80.2	49.2	70.7	31.3	54.6	51.2
Agri, Forest, Fish & Vete	23.8	47.6	50.8	84.5	48.4	51.9			50.8	82.5	47.8	21.5	28	47.3
STEM			36.4	13	29.1	26	35.9	43.1	36.3	12.7			42.7	57.4
Edu	53.2	33.6	89.9	51.9	80.4	38.3	88.5	28.3	89.5	51.1	93.9	42.9	49.5	33.4
Engi, Manu & Const	245	171	37.9	38	68.1	105	31.5	93.5	37.5	37.4	212	160	28.3	80.3
Heal & Welf	58	33	75.9	48.9	61.8	45.7	72.4	60	75.8	49.1	80.3	5.9	97.1	111
Natu, Math & Stat			70.1	27.7	57.1	42.9	81.2	67.6	69.8	27.3			61.4	35.6
Servi			45.9	82.7	71.8	174	11.6	91.1	46.3	80.6				
Social, Jour & Info	43.7	28.4	91.6	93.6	57.8	61.8	83.7	73.4	91	70.9			126	111
TIC			31.8	41.7	17.9	19.6	30.3	53.6	31.3	41			120	115
Mat. Leave	65	35.1	143	68.7	99.9	17.6	67.6	35.6	139	70.9	75.8	32	90.2	24.1
Labor Partic	52.1	7.43	51.5	8.24	49.9	5.66	30.2	13.5	51.9	8.2	63.9	9.66	58.3	7.13
School	1.0	0.0	1.0	0.0	1.0	0.0			1.0	0.0	1.0	0.0	1.0	0.0
Researchers million	2,630	2,250	2,180	2,290	993	481			2,490	2,200	2,710	2,380		
R&D	0.8	1.0	1.1	0.9	0.7	0.3			1.1	1.0	2.0	1.4		

The employed BN regional models suggest that regions that exhibit the largest estimated deviations from current values under the conditional parity scenario are Latin America & Caribbean, South Asia, and Sub-Saharan Africa. In contrast, conditional estimates indicate that East Asia & Pacific, Europe & Central Asia, the Middle East & North Africa, and North America require, comparatively, modest changes. For the Europe & Central Asia dataset, it is noteworthy that, since some countries have already surpassed the threshold of having more than 50% of women researchers (such as Republic of North Macedonia in 2019 and 2020, Kazakhstan in 2020), achieving parity may imply decreasing current levels of other variables. Although the unconditional mean is not higher than 50% in any region, it is important to take that into account when interpreting the results.

The model presents a similar arrangement in Europe & Central Asia and North America. In these regions, variables requiring adjustment appear to increase slightly. For Europe & Central Asia and North America these include mainly an increase in the shares of women graduates in Business, Administration & Law and Engineering, Manufacturing & Construction.

Then, Latin America & Caribbean and the Middle East & North Africa stand out for the requirement to considerably increase the number of researchers, more than in any other region. In the Middle East & North Africa, bigger increases also appear necessary in the share of graduates in Education and Arts & Humanities that are women. Latin America & Caribbean exhibits a more complex scenario. Conditional estimates suggest that some indicators would need to increase, while others would need to decline to reach gender parity. The shares of women graduates in TIC and STEM appear to require decreases, whereas the number of researchers, researchers per million inhabitants, R&D expenditure, and the shares of women graduates in Arts and Humanities and Engineering, Manufacturing and Construction would need to increase.

In South Asia and Sub-Saharan Africa, achieving gender parity appears to be associated with a substantial rise in the number of researchers, although the increase would not need to be as large as in the aforementioned regions. South Asia would require marked increases in the number of researchers and percentage of graduates in Engineering, Manufacturing & Construction. Moreover, increases would be needed in the percentage of graduates in Health &Welfare that are women, percentage of GDP spent in R&D, number of researchers per million inhabitants, percentage of graduates in Education and in Business, Administration & Law that are women. Sub-Saharan Africa shows a similar pattern, with notable projected increases in the number of researchers and in the proportion of women graduates in Arts and Humanities, TIC, Engineering, Manufacturing & Construction, and Social Sciences, Journalism & Information.

East Asia & Pacific shows a different pattern. In this region, most indicators appear to reflect minimal changes necessary to achieve gender parity, with the exception of the number of graduates in Engineering, Manufacturing & Construction that are women, which would need to increase substantially. Conditional projections indicate that higher percentages of female researchers are associated with reductions in the number of researchers, R&D expenditure, and the share of women graduates in Agriculture, Forestry, Fisheries & Veterinary fields. These inferences result from the negative correlation between these variables and the percentage of female researchers. The negative correlations reflect, for example, the challenges in sustaining high(er) percentages of female researchers when the number of researchers increases, as captured by the historical data.

When analysing by variables instead of by regions, the largest projected adjustments concern the number of researchers, the shares of women graduates in Engineering, Manufacturing & Construction, Arts & Humanities, and TIC, as well as researchers per million inhabitants. It is noteworthy that some variables have not appeared during this conditional analysis: the share of women graduates in Services, the duration of paid maternity leave, women's labour force participation, and girls' school enrolment. These indicators appear less central in the model's projections, and it is striking that most of them are not directly linked to educational pathways. That is, those indicators that would need to change less in general are gender-sensitive labour and scientific capacity ones.

5. Discussion

In this study, we have employed predictive models to estimate when gender parity in research may be achieved, based on available data. However, we identified the need to enrich the data on the number of women researchers with additional indicators in order to obtain a more nuanced and comprehensive understanding. To address this, we developed Bayesian Networks incorporating gender parity alongside other variables, enabling us to identify which factors may be associated with progress toward gender parity in each region, and the extent of their potential influence.

ARIMA offers useful forecasts, showing how gender parity in science could be achieved in around 15 years in Europe & Central Asia, Latin America & Caribbean and North America; in approximately 30 years in East Asia & Pacific and in around 50 years in Middle East & North Africa, South Asia and Sub-Saharan Africa. These forecasts, however, answer the "when" question, but not the "how" question. These limitations have also been noted in earlier applications, such as Wang et al. (2021) and Nane et al. (2022), where temporal models alone could not capture the structural complexity of the scientific system. By incorporating correlation matrices and Bayesian Networks, we obtain a more detailed representation in which educational variables show particularly strong associations in shaping women's participation in research in most regions.

We find that some variables have a direct conditional association within the BN structure on the percentage of women researchers in a region. In particular, it is remarkable how the share of women graduating in Business, Administration & Law and Engineering, Manufacturing & Constructions play a key role in shaping the percentage of women researchers and a central role in BNs in general in most regions. It is also interesting to see how the percentage of women researchers is a variable that is influenced more than it influences, and this configuration results solely from data-driven structure learning techniques rather than from a design choice. This suggests that changes in other variables could have, via the employed BNs, influence on gender parity.

When examining the changes required to achieve gender parity, the global model indicates that educational and innovation-system indicators would require the most substantial adjustments at the global level. In contrast, gender-sensitive labour variables and scientific capacity indicators appear less influential in our global model's projections. This pattern echoes Walby's (1990) argument that exclusion has largely been replaced by discrimination within the labour system: vertical segregation giving way to horizontal segregation. Thus, the issue nowadays lies less in admission of women to the scientific system than in the dynamics that shape how women are distributed across its fields.

Regional results further suggest that trajectories toward parity differ: in Sub-Saharan Africa and South Asia, educational indicators appear to have weaker predictive power, and conditional estimates point to a stronger role for innovation and scientific-capacity variables. This should not be interpreted as a sequential progression in which regions must first address one set of inequalities and then advance to another. Rather, the patterns suggest that regions may follow distinct pathways toward gender parity. We emphasize that this interpretation does not rely on a developmentalist logic that assumes all regions move through the same stages or that replicating the experiences of already industrialized regions would yield similar outcomes elsewhere (Temin, 2023). Instead, the results indicate that each region faces a different configuration of challenges, not a different level of development.

There are several limitations to this study. First, the dataset contains substantial gaps, and while we made every effort to interpolate missing values where possible, we were constrained by the availability of reliable data. A significant portion of the missing data pertains to countries in the Global South, which limits the generalizability of our findings. Moreover, the choice to work with national and regional aggregates necessarily obscures within-country disparities that future research could explore in greater detail. Second, there are many variables that could be included here. We have done a selection of variables that could affect the percentage of women researchers and were not redundant, but many more could be included. Third, gender assignment algorithms present inherent limitations, as their accuracy varies across countries and regions, and they rely on binary classifications that exclude non-binary identities, thereby reinforcing a limited conceptualization of gender. Moreover, the fact that the share of researchers with unknown gender increases over time deserves further investigation and should be considered when analysing the results. Fourth, since the analysis relies on historical data, Dynamic Bayesian Networks (DBNs) could also be applied. DBNs extend traditional Bayesian Networks by explicitly incorporating temporal dependencies between variables across time steps. However, this methodology remains relatively recent, and the resulting graphical representations often present interpretability and stability challenges. Future work may consider the implementation of DBNs as the underlying methods and tools are further developed.

Despite these limitations, we believe our results provide valuable insights into the measures needed to increase the representation of women researchers across regions. These findings could inform policy decisions (for instance, by encouraging greater investment in increasing the number of women entering Business, Engineering and related fields) and, in turn, contribute to the development of a more inclusive and representative scientific workforce. Such a transformation is essential if science is to reflect the diversity of humanity and fully leverage its collective potential.

5.1 Policy implications and concluding remarks

Closing the gender gap in science is a complex endeavour rooted in subtle and structural forms of bias. The study of gender in academia therefore presents significant conceptual and methodological challenges. For this reason, this research shows that context needs to be explicitly considered in at least two crucial ways. First, indicators such as gender parity require contextualised interpretation. Analyses based solely on the absolute numbers or percentages of women in science are insufficient when examined in isolation. Instead, these indicators should be assessed alongside complementary measures that capture structural dynamics. Only through such an integrated approach is it possible to obtain a more accurate understanding of the mechanisms shaping gender inequalities and to identify where targeted interventions are most likely to advance gender equity in academia.

Secondly, there is no universal measure for reaching gender equality in science, as such measures must be sensitive to contextual and geographical differences. Bibliometric research with a gender lens has traditionally focused on Europe and the United States, yet the resulting findings are often implicitly treated as globally applicable, despite substantial regional differences. From a policy perspective, closing the gender gap in science requires targeted and context-specific policy interventions. For instance, our results highlight the relevance of educational pathways in Engineering, Manufacturing & Construction studies in Europe & Central Asia and North America; the role of R&D investment in South Asia and Latin America & the Caribbean; and increasing the overall number of researchers in Latin America & the Caribbean, the Middle East & North Africa, South Asia, and Sub-Saharan Africa.

To conclude, we emphasise the continued need for explicitly gender-sensitive policies in science. The results demonstrate that gender differences persist and that the passage of time alone and an increase in resources are insufficient to address existing inequalities. We can see that in the East Asia & Pacific case, where the fact that having a high GDP spent in R&D or a substantial number of researchers does not necessarily mean a higher share of women in science, as politics have to be intentional and not gender-neutral to have positive gendered consequences, in line with Steinþórsdóttir et al. (2020). Consequently, gender must be incorporated as a central analytical and decision-making variable in science policy, as failing to do so risks obscuring structural dynamics that shape participation and outcomes.

6. References

Abdollahi, H. (2020). A novel hybrid model for forecasting crude oil price based on time series decomposition. *Applied Energy, 267*, 115035. https://doi.org/10.1016/j.apenergy.2020.115035

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

Aramayona, J., Cruz Castro, L., Sanz Menéndez, L., & Timón García-Longoria, Á. A. (2022). *La desafección por la carrera investigadora en matemáticas: Diferencias entre hombres y mujeres*. https://doi.org/10.13039/501100004837

Arora, P., Boyne, D., Slater, J. J., Gupta, A., Brenner, D. R., & Druzdzel, M. J. (2019). Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. *Value in Health, 22*(4), 439–445. https://doi.org/10.1016/j.jval.2019.01.006

Bauer, H. P. W., Schui, G., von Eye, A., & Krampen, G. (2013). How does scientific success relate to individual and organizational characteristics? A scientometric study of psychology researchers in the German-speaking countries. *Scientometrics, 94*(2), 523–539. Scopus. https://doi.org/10.1007/s11192-012-0760-3

Borunda, M., Jaramillo, O. A., Reyes, A., & Ibargüengoytia, P. H. (2016). Bayesian networks in renewable energy systems: A bibliographical survey. *Renewable and Sustainable Energy Reviews, 62*, 32–45. https://doi.org/10.1016/j.rser.2016.04.030

Burg, M. L., Sholklapper, T., Kohli, P., Kaneko, M., Maria Autran, A., Teoh, J., Murphy, D. G., Samplasky, M., Psutka, S. P., Loeb, S., Ribal, M. J., & Cacciamani, G. E. (2022). Gender Disparities Among Editorial Boards of International Urology Journals. *European Urology Focus*, *8*(6), 1840–1846. https://doi.org/10.1016/j.euf.2022.04.007

Casad, B. J., Franks, J. E., Garasky, C. E., Kittleman, M. M., Roesler, A. C., Hall, D. Y., & Petzel, Z. W. (2021). Gender inequality in academia: Problems and solutions for women faculty in STEM. *Journal of Neuroscience Research, 99*(1), 13–23. https://doi.org/10.1002/jnr.24631

Campbell, S. E., & Simberloff, D. (2022). The Productivity Puzzle in Invasion Science: Declining but Persisting Gender Imbalances in Research Performance. *BioScience, 72*(12), 1220–1229. https://doi.org/10.1093/biosci/biac082

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest, 15*(3), 75–141. https://doi.org/10.1177/1529100614541236

Chinchilla-Rodríguez, Z., Escabias-Machuca, M., & Morillo, F. (2025). Meeting evaluation criteria based on authorship positions by gender, academic age, and research field. The Spanish case. *Research Evaluation,* https://doi.org/10.1093/reseval/rvaf008

Criado Pérez, C. (2019). *Invisible Women: Exposing Data Bias in a World Designed for Men*. Ballantine Books.

Darji, M. P., Dabhi, V. K., & Prajapati, H. B. (2015). Rainfall forecasting using neural network: A survey. *2015 International Conference on Advances in Computer Engineering and Applications*, 706–713. https://doi.org/10.1109/ICACEA.2015.7164782

del Águila, I. M., & del Sagrado, J. (2016). Bayesian networks for enhancement of requirements engineering: A literature review. *Requirements Engineering, 21*(4), 461–480. https://doi.org/10.1007/s00766-015-0225-3

Demaine, J. (2021). Trends in authorship by women at Canadian universities 2006 to 2019. *The Canadian Journal of Information and Library Science, 44*(2/3), 1–11. https://doi.org/10.5206/cjilsrcsib.v44i2.13687

Gender parity | European Institute for Gender Equality. (n.d.). Retrieved 17 November 2025, from https://eige.europa.eu/publications-resources/thesaurus/terms/1258?language_content_entity=en

European Commission. (2024). *She figures 2024: Gender in research and innovation : statistics and indicators*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/6847557.

Ferdaous, J., & Rahman, M. M. (2017). The effects of research and development expenditure on firm performance: an examination of pharmaceuticals industry in Bangladesh. *Business & Entrepreneurship Journal*, *6*(2), 1-20.

Frehill, L. M., & Zippel, K. (2011). Gender and International Collaborations of Academic Scientists and Engineers: Findings from the Survey of Doctorate Recipients, 2006. *Journal of the Washington Academy of Sciences, 97*(1), 49–69.

González-Alvarez, J., & Sos-Peña, R. (2020). Women Publishing in American Psychological Association Journals: A Gender Analysis of Six Decades. *Psychological Reports*, *123*(6), 2441–2458. https://doi.org/10.1177/0033294119860257

González-Salmón, E., & Robinson-García, N. (2024a). * WikiGenDex: Un nuevo algoritmo de identificación de género basado en fuentes abiertas. *Infonomy*, *2*(1).

González-Salmón, E., Chinchilla-Rodríguez, Z., Nane, G. F., & Robinson García, N. (2024b). What contributes to gender parity in science? A Bayesian Network analysis.International Conference on Science, Technology and Innovation Indicators, 2024.

Global Expansion (2023). Global Hiring Guide: Cuba. https://www.globalexpansion.com/hubfs/Countrypedia%20PDFs/In%20use/Cuba%20-%20%20Global%20Employer%20Guide.pdf

Hertzog, L., Charlson, F., Tschakert, P., Morgan, G. G., Norman, R., Pereira, G., & Hanigan, I. C. (2024). Suicide deaths associated with climate change-induced heat anomalies in Australia: A time series regression analysis. *BMJ Mental Health, 27*(1), e301131. https://doi.org/10.1136/bmjment-2024-301131

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

Horvath, S., Stroeve, J., & Rajagopalan, B. (2021). A linear mixed effects model for seasonal forecasts of Arctic sea ice retreat. *Polar Geography, 44*(4), 297–314. https://doi.org/10.1080/1088937X.2021.1987999

Hyndman, R.J. (2024). forecast: Forecasting Functions for Time Series and Linear Models. R package version 8.22.0. https://cran.r-project.org/web/packages/forecast/index.html

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* OTexts.

Ibáñez, A., Larrañaga, P., & Bielza, C. (2011). Using Bayesian networks to discover relationships between bibliometric indices. A case study of computer science and artificial intelligence journals. *Scientometrics*, *89*(2), 523-551.

Inter-American Commission on Human Rights (1983). The situation of human rights in Cuba. Seventh Report. Chapter XI. The right to social security. https://cidh.oas.org/countryrep/Cuba83eng/chap.11.htm

International Labour Organization (1998). More than 120 Nations Provide Paid Maternity Leave. https://www.ilo.org/resource/news/more-120-nations-provide-paid-maternity-leave

Jaramillo, A. M., Macedo, M., Oliveira, M., Karimi, F., & Menezes, R. (2025). *Systematic comparison of gender inequality in scientific rankings across disciplines* (No. arXiv:2501.13061). arXiv. https://doi.org/10.48550/arXiv.2501.13061

Jemielniak, D., & Wilamowski, M. (2025). Gender Gap in All Academic Fields Over Time. *Social Science Computer Review, 43*(4), 888–896. https://doi.org/10.1177/08944393241270633

Jones, T. M., Fanson, K. V., Lanfear, R., Symonds, M. R. E., & Higgie, M. (2014). Gender differences in conference presentations: A consequence of self-selection? *PeerJ, 2*, e627. https://doi.org/10.7717/peerj.627

Kelan, E. K. (2009). Gender fatigue: The ideological dilemma of gender neutrality and discrimination in organizations. *Canadian Journal of Administrative Sciences / Revue*

Canadienne Des Sciences de l'Administration, 26(3), 197–210. https://doi.org/10.1002/cjas.106

Key, E. M., & Sumner, J. L. (2019). You Research Like a Girl: Gendered Research Agendas and Their Implications. *PS: Political Science & Politics, 52*(4), 663–668. https://doi.org/10.1017/S1049096519000945

Kuchanskyi, O., Andrashko, Y., Biloshchytskyi, A., Omirbayev, S., Mukhatayev, A., Biloshchytska, S., & Faizullin, A. (2023). Gender-Related Differences in the Citation Impact of Scientific Publications and Improving the Authors' Productivity. *Publications*, *11*(3), Article 3. https://doi.org/10.3390/publications11030037

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature, 504*(7479), Article 7479. https://doi.org/10.1038/504211a

Leicester, P. A., Goodier, C. I., & Rowley, P. N. (2016). Probabilistic analysis of solar photovoltaic self-consumption using Bayesian network models. *IET Renewable Power Generation, 10*(4), 448–455. https://doi.org/10.1049/iet-rpg.2015.0360

López-Aguirre, C. (2019). Women in Latin American science: gender parity in the twenty-first century and prospects for a post-war Colombia. *Tapuya: Latin American Science, Technology and Society*, *2*(1), 356-377.

Macaluso, B., Larivière, V., Sugimoto, T., & Sugimoto, C. R. (2016). Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship. *Academic Medicine, 91*(8), 1136. https://doi.org/10.1097/ACM.0000000000001261

McKinley, S., & Levine, M. (1998). Cubic spline interpolation. *College of the Redwoods, 45*(1), 1049-1060.

Misirli, A. T., & Bener, A. B. (2014). Bayesian Networks For Evidence-Based Decision-Making in Software Engineering. *IEEE Transactions on Software Engineering, 40*(6), 533–554. https://doi.org/10.1109/TSE.2014.2321179

Momeni, F., Karimi, F., Mayr, P., Peters, I., & Dietze, S. (2022). The many facets of academic mobility and its impact on scholars' career. *Journal of Informetrics, 16*(2), 101280. https://doi.org/10.1016/j.joi.2022.101280

Moradi, R., Cofre-Martel, S., Lopez Droguett, E., Modarres, M., & Groth, K. M. (2022). Integration of deep learning and Bayesian networks for condition and operation risk monitoring of complex engineering systems. *Reliability Engineering & System Safety, 222*, 108433. https://doi.org/10.1016/j.ress.2022.108433

Morillo, F., Escabias-Machuca, M., & Chinchilla-Rodríguez, Z. (2025). How do women and men differ in research collaborations based on authorship positions? The Spanish case. *Frontiers in Research Metrics and Analytics,* https://doi.org/10.3389/frma.2025.1631931

Msosa, S. K., Ntshangase, B. A., & Mlambo, C. (2022). Gender parity among researchers in science, technology, engineering and mathematics. *Entrepreneurship and Sustainability Issues*, *9*(4), 243.

Nane, G. F., Robinson-Garcia, N., van Schalkwyk, F., & Torres-Salinas, D. (2022). COVID-19 and the scientific publishing system: Growth, open access and scientific fields. Scientometrics, 128(1), 345–362. https://doi.org/10.1007/s11192-022-04536-x

Narasimhan, S. (2021). Participation of women in science in the developed and developing worlds: Inverted U of feminization of the scientific workforce, gender equity and retention. *Pure and Applied Chemistry, 93*(8), 913–925. https://doi.org/10.1515/pac-2021-0101

Njenga, J. K. (2024). Domestic Credit Growth Analysis using ARIMA Technique: A Case Study of Kenya. *Journal of Economics, Management and Trade, 30*(5), 12–21. https://doi.org/10.9734/jemt/2024/v30i51206

Nygaard, L. P., Aksnes, D. W., & Piro, F. N. (2022). Identifying gender disparities in research performance: The importance of comparing apples with apples. *Higher Education, 84*(5), 1127–1142. https://doi.org/10.1007/s10734-022-00820-0

Pearl, J. (1995). From Bayesian Networks to Causal Networks. In G. Coletti, D. Dubois, & R. Scozzafava (Eds.), *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence* (pp. 157–182). Springer US. https://doi.org/10.1007/978-1-4899-1424-8_9

Pedró, F., Bouckaert, M. & Gutović, V. (2025, June 23). *Bridging the data gap in higher education policymaking: Challenges and opportunities for evidence-based governance.* World Education Blog. https://world-education-blog.org/2025/06/23/bridging-the-data-gap-in-higher-education-policymaking-challenges-and-opportunities-for-evidence-based-governance/

Powell, C. (2017). Postcard from Havana: A Lack of Childcare Leaves Cuban Women in Quandary. *Council on Foreign Relations.* https://www.cfr.org/blog/postcard-havana-lack-childcare-leaves-cuban-women-quandary

Raffo, J. (2021). WGND 2.0. *Harvard Dataverse*. DOI: https://doi.org/10.7910/DVN/MSEGSJ.

Rao, C., Zhang, Y., Wen, J., Xiao, X., & Goh, M. (2023). Energy demand forecasting in China: A support vector regression-compositional data second exponential smoothing model. *Energy, 263*, 125955. https://doi.org/10.1016/j.energy.2022.125955

Rajab, K., Kamalov, F., & Cherukuri, A. K. (2022). Forecasting COVID-19: Vector Autoregression-Based Model. *Arabian Journal for Science and Engineering, 47*(6), 6851–6860. https://doi.org/10.1007/s13369-021-06526-2

Robinson-Garcia, N., Costas, R., Sugimoto, C. R., Larivière, V., & Nane, G. F. (2020). Task specialization across research careers. *eLife*, *9*, e60586. https://doi.org/10.7554/eLife.6058

Ruiz-Baena, E., Vargas-Quesada, B., & Chinchilla-Rodríguez, Z. (2025). Gendered Knowledge Production in García Lorca Studies: A Bibliometric Mapping. *Zenodo*. https://doi.org/10.5281/zenodo.18258961

Scutari, M., & Denis, J.-B. (2021). *Bayesian Networks | With Examples in R | Marco Scutari, Jean-Baptiste*. Chapman and Hall/CRC. https://www.taylorfrancis.com/books/mono/10.1201/9780429347436/bayesian-networks-marco-scutari-jean-baptiste-denis

Sesen, M. B., Nicholson, A. E., Banares-Alcantara, R., Kadir, T., & Brady, M. (2013). Bayesian Networks for Clinical Decision Support in Lung Cancer Care. *PLOS ONE, 8*(12), e82349. https://doi.org/10.1371/journal.pone.0082349

Shi, H., Li, X., & Wang, S. (2024). How Bayesian networks are applied in the subfields of climate change: Hotspots and evolution trends. *Environmental Modelling & Software, 172*, 105921. https://doi.org/10.1016/j.envsoft.2023.105921

Spoon, K., LaBerge, N., Wapman, K. H., Zhang, S., Morgan, A., Galesic, M., Fosdick, B., Larremore, D., & Clauset, A. (2023). *Gender and retention patterns among U.S. faculty*. SocArXiv. https://doi.org/10.31235/osf.io/u26ze

Steinþórsdóttir, F. S., Einarsdóttir, Þ., Pétursdóttir, G. M., & Himmelweit, S. (2020). Gendered inequalities in competitive grant funding: An overlooked dimension of gendered power relations in academia. *Higher Education Research & Development, 39*(2), 362–375. https://doi.org/10.1080/07294360.2019.1666257

Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science, 29*(4), 581–593. https://doi.org/10.1177/0956797617741719

Sugimoto, C. R., & Larivière, V. (2023). Equity for women in science. Harvard University Press

Sun, M., Ma, T., Zhou, L., & Yue, M. (2023). Analysis of the relationships among paper citation and its influencing factors: A Bayesian network-based approach. *Scientometrics*, *128*(5), 3017–3033. https://doi.org/10.1007/s11192-023-04697-3

Temin, D. M. (2023). Development in Decolonization: Walter Rodney, Third World Developmentalism, and "Decolonizing Political Theory". *American Political Science Review, 117*(1), 235–248. https://doi.org/10.1017/S0003055422000570

The Business Standard (2024). Bangladesh's R&D expenditure among lowest globally, reveals BBS survey. https://www.tbsnews.net/economy/bangladeshs-rd-expenditure-among-lowest-globally-reveals-bbs-survey-861971

Thelwall, M. (2020). Gender differences in citation impact for 27 fields and six English-speaking countries 1996–2014. *Quantitative Science Studies, 1*(2), 599–617. https://doi.org/10.1162/qss_a_00038

Tomassini, C. (2021). Gender gaps in science: Systematic review of the main explanations and research agenda. *Education in the Knowledge Society (EKS), 22*. https://doi.org/10.14201/eks.25437

UNESCO (2024). *UNESCO Call to Action to Close the Gender Gap in Science*. Retrieved 16 December 2025, from https://www.unesco.org/en/science-technology-and-innovation/cta

UNPD (2024). *Sesgos codificados: La subrepresentación de las mujeres en STEM en América Latina y el Caribe*. Retrieved 17 November 2025, from https://www.undp.org/es/latin-america/blog/sesgos-codificados-la-subrepresentacion-de-las-mujeres-en-stem-en-america-latina-y-el-caribe

Visser, M., Van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies, 2*(1), 20–41. https://doi.org/10.1162/qss_a_00112

Walby, S. (1990). *Theorizing Patriarchy.* Oxford: Blackwell.

Wang, L.L., Stanovsky, G., Weihs, L., & Etzioni, O. (2021). Gender trends in computer science authorship. *Commun. ACM, 64*(3), 78-84. https://doi.org/10.1145/3430803

Williamson, S. (2020). Backlash, gender fatigue and organisational change: AIRAANZ 2019 presidential address. *Labour and Industry, 30*(1), 5–15. https://doi.org/10.1080/10301763.2019.1677202

Witteveen, A., Nane, G. F., Vliegen, I. M., Siesling, S., & IJzerman, M. J. (2018). Comparison of logistic regression and Bayesian networks for risk prediction of breast cancer recurrence. *Medical decision making*, *38*(7), 822-833.

World Bank Data (2025). Why are some data not available?. Retrieved 18 November 2025, from https://datahelpdesk.worldbank.org/knowledgebase/articles/191133-why-are-some-data-not-available

Wu, C. (2024). The gender citation gap: Approaches, explanations, and implications. *Sociology Compass, 18*(2), e13189. https://doi.org/10.1111/soc4.13189

Zhao, X., Akbaritabar, A., Kashyap, R., & Zagheni, E. (2023). A gender perspective on the global migration of scholars. *Proceedings of the National Academy of Sciences, 120*(10), e2214664120. https://doi.org/10.1073/pnas.2214664120

7. Acknowledgments

We thank the participants of the 28th International Conference on Science, Technology and Innovation Indicators (2024) and of the RESPECT workshop at Georgia Institute of Technology in July 2024 for their valuable comments and constructive feedback on earlier versions of this work.

8. Author contribution statements

Elvira González-Salmón: Methodology, Software, Validation, Investigation, writing - Original Draft, Visualization

Zaida Chinchilla-Rodríguez: Conceptualization, Writing - Review & Editing, Supervision

Gabriela F. Nane: Methodology, software, Writing - Review & Editing, Supervision

Nicolas Robinson-Garcia: Conceptualization, Methodology, Writing - Review & Editing, Supervision

9. Competing interest statements

Authors of this research have no competing interests.

10. Funding information

This work is part of the STITCH project (Ref: PID2024-155412OB-I00) and the RESPONSIBLE project (Ref: PID2021-128429NB-I00), both funded by the Spanish Ministry of Science (Ref: MCIN/AEI /10.13039/501100011033 FSE invierte en tu futuro). Elvira González-Salmón is currently supported by an FPU grant from the Spanish Ministry of Science (Ref: FPU2021/02320). Nicolas Robinson-Garcia is currently supported by a Ramón y Cajal grant from the Spanish Ministry of Science (Ref: RYC2019-027886-I).

11. Data availability statements

Data underlying this study are available at https://doi.org/10.5281/zenodo.18403324

12. Appendixes

A.1 Calculation of error rate for interpolation

To assess how much data could be interpolated, we first calculated an error rate. This calculation was based on two variables with complete data across all years: the percentage of female researchers in Spain and the percentage of female labor force participation in Spain. We manually removed one value from each variable, interpolated it, and then calculated the error rate as the absolute value of the difference between the interpolated and actual values. We repeated this process by removing two values, then three, and continued incrementally up to 31 observations.

However, we did not remove values randomly; instead, we followed patterns previously identified in the data. We observed that missing data could occur in seven distinct patterns:

A. Latest: Data is missing only for the most recent year(s). If multiple years are missing, they must be consecutive and immediately precede the last year.
    a. Examples:
        i. Only 2020 is missing
        ii. 2020 and 2019 are missing
        iii. 2020, 2019, 2018 and 2017 are missing
    b. Non-examples:
        i. 2020, 2019 and 2017 are missing
B. Earliest: Data is missing only for the earliest year(s). If more years are missing, they must be consecutive and immediately follow the first year.
    a. Examples:
        i. Only 1990 is missing
        ii. 1990 and 1991 are missing
        iii. 1990, 1991, 1992 and 1993 are missing
    b. Non-examples:
        i. 1990, 1991 and 1993 are missing.
C. Middle: Data is missing for years in the middle of the series, but both the earliest and latest years are present.
    a. Examples:
        i. 2015 is missing
        ii. 2000, 2001, 2002, 2003 are missing
    b. Non-examples:
        i. 2015 and 1990 are missing
        ii. 2015 and 2020 are missing
D. Latest & Middle: Data is missing for years in the middle of the series, and the most recent year (or consecutive recent years) is also missing.
    a. Examples:

i. 2018 and 2020 are missing

ii. 1993, 1999, 2000, 2007, 2018, 2019 and 2020 are missing

E. Earliest & Middle: Data is missing for the earliest year (or consecutive earliest years) and for some middle years, but the latest year is present.

a. Examples:

i. 1990 and 2017 are missing

ii. 1990, 1991, 1993, 1999, 2000, 2007 and 2018 are missing

F. Latest, Earliest & Middle: Data is missing for the earliest, latest, and some middle years.

a. Examples:

i. 1990, 2001 and 2020 are missing

ii. 1990, 1991, 2002, 2004, 2019 and 2020 are missing

G. Latest & Earliest: Both the earliest and latest year(s) are missing, but no middle years are missing randomly. Missing years must be consecutive to either the first or last year.

a. Examples:

i. Both 1990 and 2020 are missing.

ii. 1990, 1991, 2019, and 2020 are missing.

Below we provide a visual representation of each type of missing data (table A). We calculated the error rate for each of the seven patterns of missing data. As the analysis was conducted using two variables, we obtained fourteen different error rates. This variation occurred because the error rates differed slightly between the two variables examined: the percentage of women researchers and the percentage of female labor force participation. In cases where the error rates diverged, we selected the lower error rate as the standard conservative criterion to guide the interpolation of other variables. Additionally, we established a threshold: when the error rate reached 10, it represented the maximum acceptable limit for interpolation.

Table A. Patterns of missing data (A - latest years missing, B - earliest years missing, C - middle years missing, D - latest and middle years missing, E - earliest and middle years missing, F - earliest, latest and middle years missing, G - latest and earliest years missing).

	A	B	C	D	E	F	G
	Latest	Earliest	Middle	Latest & Middle	Earliest & Middle	Latest, Earliest & Middle	Latest & Earliest
2020	Missing	Data	Data	Missing	Data	Missing	Missing
2019	Missing	Data	Data	Data	Data	Data	Missing
2018	Missing	Data	Data	Data	Data	Data	Data
2017	Data	Data	Missing	Missing	Missing	Missing	Data
2016	Data	Data	Missing	Missing	Missing	Missing	Data
2015	Data	Data	Data	Data	Data	Data	Data

2014	Data	Data	Missing	Missing	Missing	Missing	Data
2013	Data	Data	Data	Data	Data	Data	Data
2012	Data	Missing	Data	Data	Data	Data	Missing
2011	Data	Missing	Missing	Missing	Missing	Missing	Missing
2010	Data	Missing	Data	Data	Missing	Missing	Missing

Below, we present an example of the error rate calculated when values were missing randomly in the middle of the time series (Middle). As shown in Figure A.1.1, the error rate remained relatively stable until more than 20 values were interpolated. For the percentage of women researchers, it was possible to interpolate up to 26 missing values without the error rate exceeding 2. However, for the percentage of female labor force participation, the error rate surpassed 15 after interpolating 24 missing values. Therefore, we established 24 as the maximum number of values that could be interpolated when data were missing randomly in the middle (Middle pattern).

Figure A.1.1 Error rate (y axis) for pattern C (middle values missing), depending on the number of excluded observations (x axis) for percentage female researchers in Spain (left) and percentage of Female labor force participation in Spain (right).

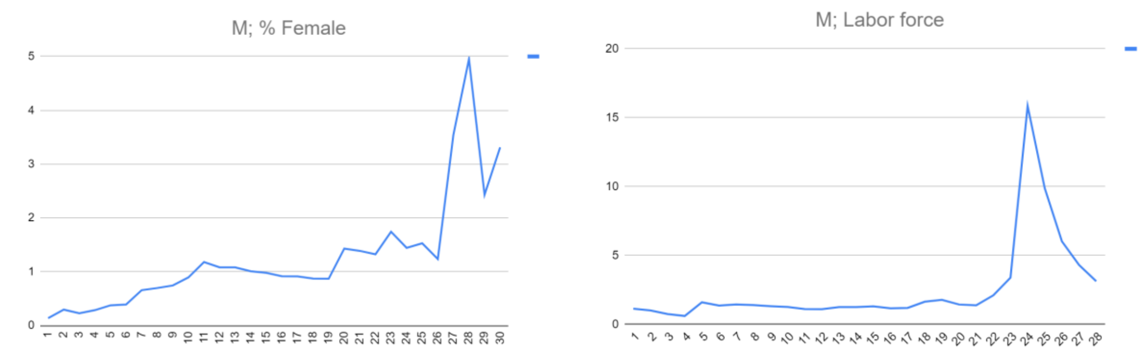


Table B reflects the maximum number of values that could be interpolated for the percentage female researchers and percentage female labor force participation, assuming a dataset spanning 32 years (1990-2021). For each type of missing data, we chose the lowest value as the maximum number of data points that we could interpolate. For instance, for the pattern “Latest”, we could interpolate 11 values for percentage of female researchers and 12 for the percentage of female labor force participation. To keep it robust, we chose 11 as the maximum number of years we could interpolate. We did the same for all missing data patterns. This ensured that interpolation limits were not driven by the most favorable case.

Table B. Number of values interpolated in percentage of female researchers and percentage of female force participation variables before +-10 error rate was reached. The lowest number is shown in bold, and the final column reports the corresponding percentage of interpolable data points.

Type of missing data	Nº of values interpolated before +-10 error rate in % Female researchers	Nº of values interpolated before +-10 error rate in % Female labor force	% of values that represents

		participation	
Latest	**11**	12	34.37%
Earliest	**6**	8	18.75%
Middle	30	**23**	71.87%
Latest & Middle	**30**	**30**	93.75%
Earliest & Middle	30	**15**	46,87%
Latest, earliest & middle	**30**	**30**	93.75%
Latest & earliest	**20**	**20**	62.5%

However, as previously mentioned, not all countries had complete data for the full period of analysis, due to gaps in the Dimensions database. In some cases, data were available for only 10 years. Therefore, we calculated the percentage of observations that could be interpolated for each country, based on the number of years with available data. Table C presents these results. For example, if a country had all years of data available and it was possible to interpolate 11 values, this represented 34.37% of the total observations (Table B). Then, for countries with 16 years of data, a maximum of 5 interpolated values was established, and so on (Table C)

Table C. Number of data points we could interpolate for each variable given the number of total years and the missing pattern (L - latest years missing, E - earliest years missing, M - middle years missing, L, E - latest and middle years missing, L, E - latest and earliest years missing, E, M - earliest and middle years missing, L, E, M - earliest, latest and middle years missing).

Nº of years	L (34,37%)	E (18,75%)	M (71,87%)	L, E (62,5%)	L, M (93,75%)	E, M (46,87%)	L, E, M (93,75%)
1	0,34	0,19	0,72	0,63	0,94	0,47	0,94
2	0,69	0,38	1,44	1,25	1,88	0,94	1,88
3	1,03	0,56	2,16	1,88	2,81	1,41	2,81
4	1,38	0,75	2,88	2,50	3,75	1,88	3,75
5	1,72	0,94	3,59	3,13	4,69	2,34	4,69
6	2,06	1,13	4,31	3,75	5,63	2,81	5,63
7	2,41	1,31	5,03	4,38	6,56	3,28	6,56
8	2,75	1,50	5,75	5,00	7,50	3,75	7,50
9	3,09	1,69	6,47	5,63	8,44	4,22	8,44
10	3,44	1,88	7,19	6,25	9,38	4,69	9,38
11	3,78	2,06	7,91	6,88	10,31	5,16	10,31
12	4,13	2,25	8,63	7,50	11,25	5,63	11,25
13	4,47	2,44	9,34	8,13	12,19	6,09	12,19
14	4,81	2,63	10,06	8,75	13,13	6,56	13,13
15	5,16	2,81	10,78	9,38	14,06	7,03	14,06
16	5,50	3,00	11,50	10,00	15,00	7,50	15,00

17	5,84	3,19	12,22	10,63	15,94	7,97	15,94
18	6,19	3,38	12,94	11,25	16,88	8,44	16,88
19	6,53	3,56	13,66	11,88	17,81	8,91	17,81
20	6,88	3,75	14,38	12,50	18,75	9,38	18,75
21	7,22	3,94	15,09	13,13	19,69	9,84	19,69
22	7,56	4,13	15,81	13,75	20,63	10,31	20,63
23	7,91	4,31	16,53	14,38	21,56	10,78	21,56
24	8,25	4,50	17,25	15,00	22,50	11,25	22,50
25	8,59	4,69	17,97	15,63	23,44	11,72	23,44
26	8,94	4,88	18,69	16,25	24,38	12,19	24,38
27	9,28	5,06	19,41	16,88	25,31	12,66	25,31
28	9,63	5,25	20,13	17,50	26,25	13,13	26,25
29	9,97	5,44	20,84	18,13	27,19	13,59	27,19
30	10,31	5,63	21,56	18,75	28,13	14,06	28,13
31	10,66	5,81	22,28	19,38	29,06	14,53	29,06

Once we decided how much data we could interpolate in each case, we did a selection of countries. We excluded territories that were not recognized as countries by the United Nations. In addition, we removed countries for which there were fewer than 3000 researchers recorded in the Dimensions database, as it was not feasible to draw reliable conclusions from such limited data, and country-year combinations which resulted in less or equal 20 researchers. After applying these criteria, the final sample included 106 countries, which were grouped by region as seen in table D. We use the World Bank division of countries.

Table D. Number of countries by region after excluding those with less than 3000 researchers.

	Total countries	Final count (>3000 researchers)	Final percentage
East Asia & Pacific	28	11	39,29%
Europe & Central Asia	52	42	80,77%
Latin America & Caribbean	33	14	42,42%
Middle East & North Africa	20	16	80%
North America	2	2	100%
South Asia	8	5	62,50%
Sub-Saharan Africa	46	16	34,78%
Total UN countries	189	106	56,08%

We decided to exclude variables for which data were available for fewer than 60% of countries. These variables included: Proportion of time spent on unpaid domestic and care work, female (% of 24 hour day), Adolescents out of school, female (% of female lower secondary school age) and Trained teachers in upper secondary education, female (% of female teachers).

Additionally, we removed countries that had data for fewer than 70% of the 18 selected variables (i.e., fewer than 12 variables) from the 106 countries mentioned above. The countries excluded were as follows:

- Sub-Saharan Africa: Cote d'Ivoire (CI), Cameroon (CM), Malawi (MW), Nigeria (NG), Senegal (SN), Seychelles (SY), Tanzania (TZ), Zambia (ZM)
- East Asia & Pacific: China (CN), Singapore (SG), Thailand (TH)
- Middle East & North Africa: Egypt (EG), Israel (IL), Iraq (IQ), Kuwait (KW)
- South Asia: Nepal (NP), Pakistan (PK)
- Europe & Central Asia: Russian Federation (RU)
- Latin America & Caribbean: Peru (PE), Venezuela (VE), Jamaica (JM), Trinidad and Tobago (TT)

In the data-cleaning process, each variable could only be retained within a region if all countries had valid observations for it. This requirement inevitably involved trade-offs between coverage and variable completeness: in several instances, preserving the maximum number of variables meant removing a country with missing data; in others, retaining the maximum number of countries required excluding a variable that was incomplete for too many countries and years. To prevent excessive loss on either side, we opted for a balanced approach. Some variables were excluded when they imposed an unreasonable reduction in country coverage in a region, and some countries were excluded when their missingness prevented the retention of otherwise informative variables within that region. The excluded countries and variables are:

- East Asia & Pacific: STEM, TIC, Natural, services
- Europe & Central Asia: Luxembourg (LU) and Bosnia (BA)
- Middle East & North Africa: Agriculture, School enrollment, million researchers per habitants and expenditure
- South Asia: services, social sciences
- Sub-Saharan Africa: Services, million inhabitants and expenditure; Sudan (SD)

We ended up with 81 countries and 17 variables.

A.2 Other sources for exceptional cases

In a small number of exceptional cases, additional data-handling strategies were applied to preserve country coverage while maintaining consistency with the general interpolation rules described in Appendix A.1. Those tailored approaches to interpolate values are presented in Table E below.

Table E. Sources of data used to obtain information for the selected countries and variables.

Country	Variable	Strategy	Additional sources
Indonesia (ID)	% GDP	Pattern transformation	Missing 20 out of 29 values, following the EM pattern. By transforming the pattern to LEM, interpolation was

			possible.
Moldova (MD)	School Enrollment	Pattern transformation	Missing 24 out of 31 values, following the E pattern. By transforming the pattern to LEM, interpolation was possible.
Serbia (RS)	School Enrollment	Conservative threshold extension	Missing 9 out of 31 values, following the E pattern. In theory, up to 6 values could be interpolated; however, greater flexibility was applied in this case.
Armenia (AM)	Researchers per million	Use of regional averages	Regional average data provided by WBOD were used to fill missing values.
Azerbaijan (AZ)	Researchers per million	Use of regional averages	Regional average data provided by WBOD were used to fill missing values.
Belarus (BY)	Researchers per million	Use of regional averages	Regional average data provided by WBOD were used to fill missing values.
Estonia (EE)	% GDP	Conservative threshold extension	Missing 8 out of 31 values, following the E pattern. In theory, up to 6 values could be interpolated; however, greater flexibility was applied.
Croatia (HR)	% GDP	Conservative threshold extension	Missing 9 out of 31 values, following the E pattern. In theory, up to 6 values could be interpolated; however, greater flexibility was applied.
Ireland (IE)	% GDP	Pattern transformation	Missing 16 out of 31 values, following the EM pattern. By transforming the pattern to LEM, interpolation was possible.
Cuba (CU)	STEM	Use of documented external sources	Data for 2016 indicated 39% of women in STEM. Additional evidence confirmed a value of 41% in 2024 (UNPD, 2024). Thus, we could expect 2017-2020 to be 40% and further interpolate.
Cuba (CU)	Maternity leave	Use of documented external sources	Evidence confirmed that, since at least 1974, maternity leave has been 18 weeks (126 days) (Powell 2017; International Labour Organization, 1998; Global Expansion, 2023; Inter-American Commission on Human Rights, 1983)
Cuba (CU)	Researchers per million	Use of regional averages	Regional average data provided by WBOD were used to fill missing values.
Brazil (BR)	% GDP	Conservative threshold extension	Missing 10 out of 31 values, following the E pattern. In theory, up to 6 values could be interpolated; however, greater flexibility was applied.
Bangladesh (BD)	Researchers per million	Use of regional averages	Regional average data provided by WBOD were used to fill missing values.
Bangladesh (BD)	% GDP	Use of documented external sources	Data were found for 2000–2016 and 2020 (Ferdaous & Rahman, 2017; The Business Standard, 2024).

1

A.3 ARIMA and Exponential Smoothing for all regions

Figure A.3.1. ARIMA and Exponential Smoothing for East Asia.

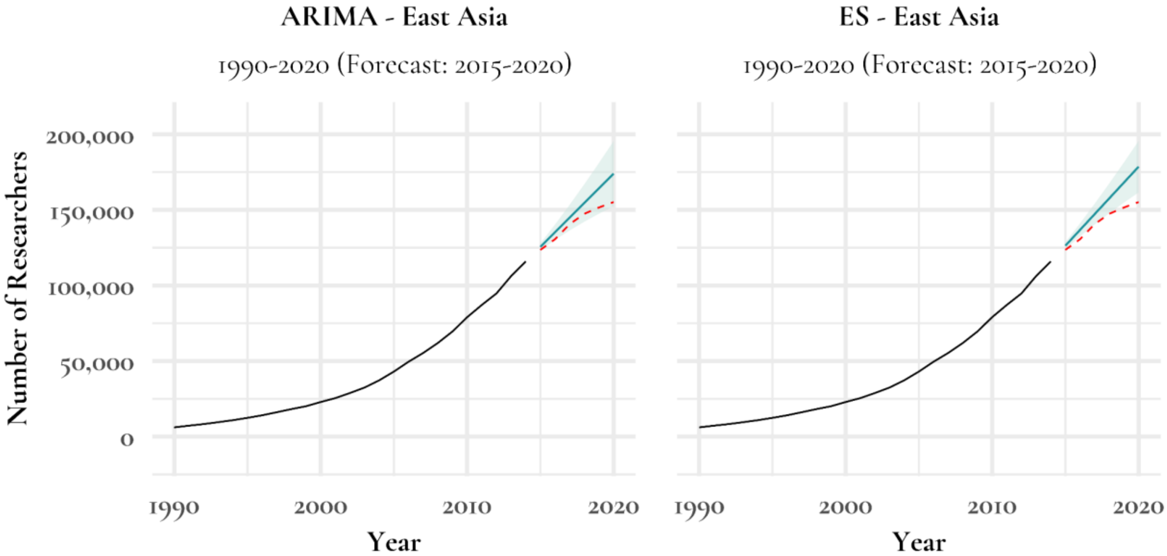


Figure A.3.2. ARIMA and Exponential Smoothing for Europe & Central Asia.

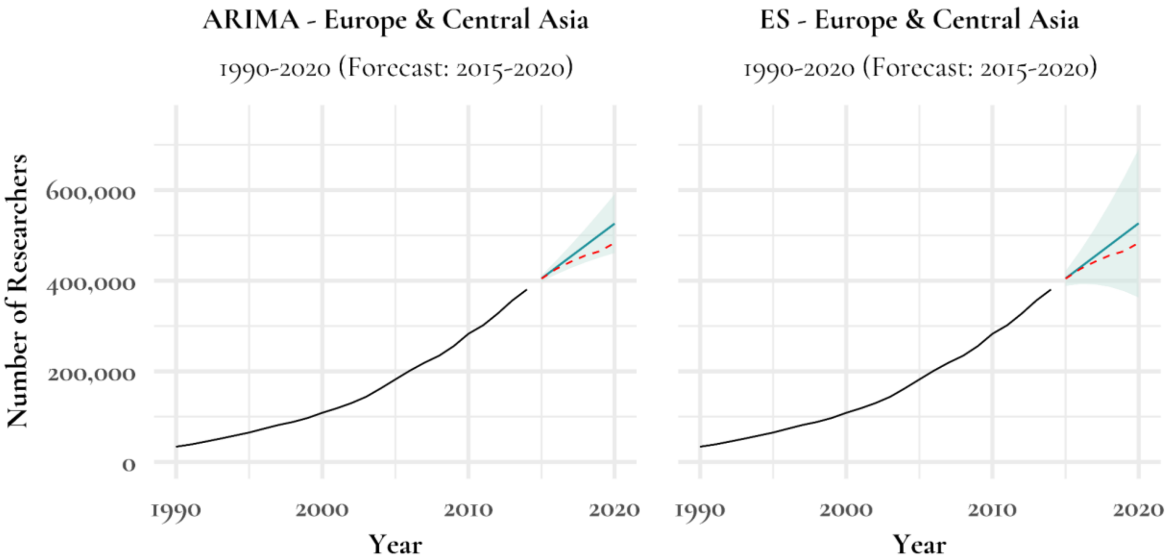


Figure A.3.3. ARIMA and Exponential Smoothing for Latin America & Caribbean.

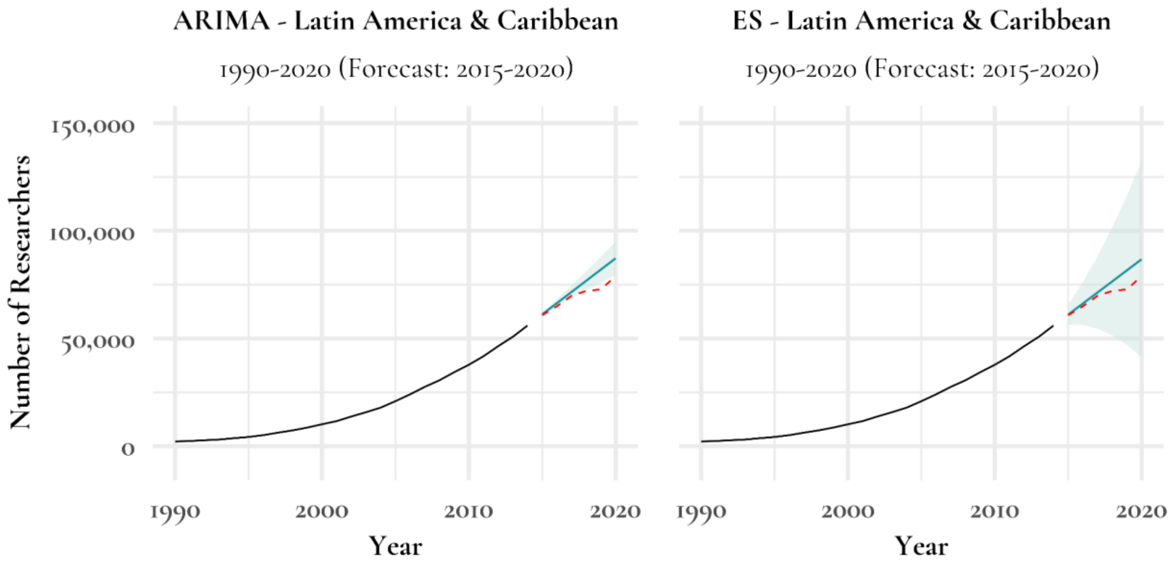


Figure A.3.4. ARIMA and Exponential Smoothing for Middle East & North Africa.

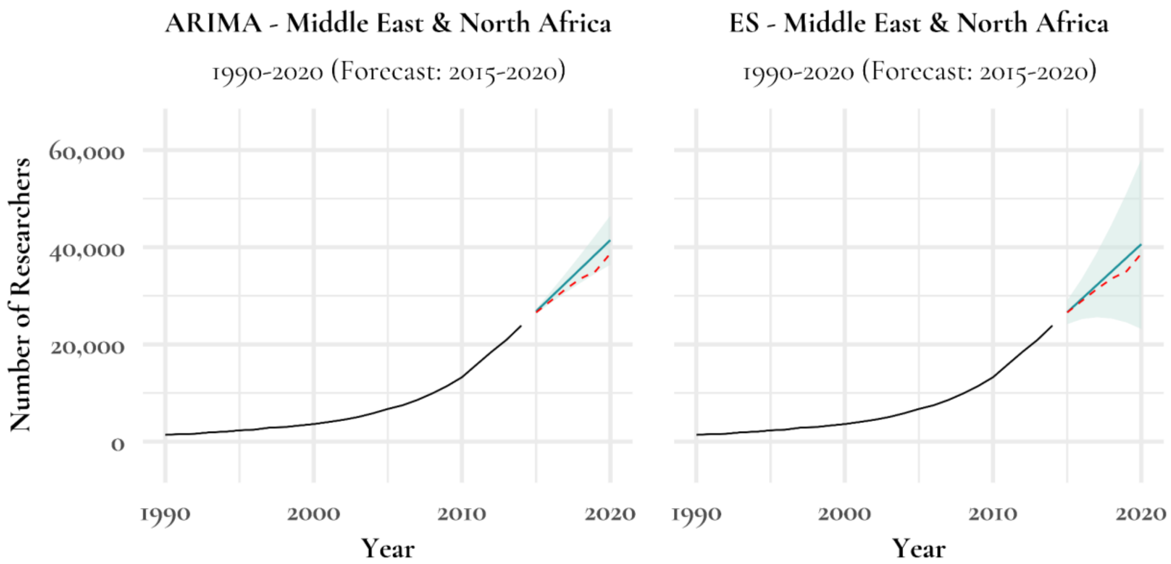


Figure A.3.5. ARIMA and Exponential Smoothing for North America.

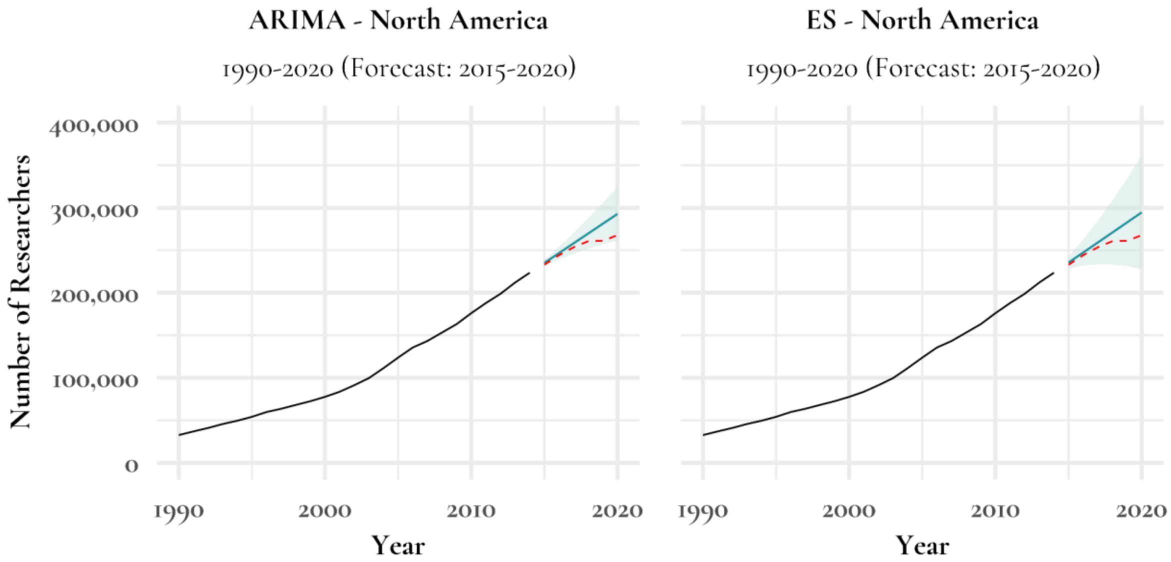


Figure A.3.6. ARIMA and Exponential Smoothing for South Asia.

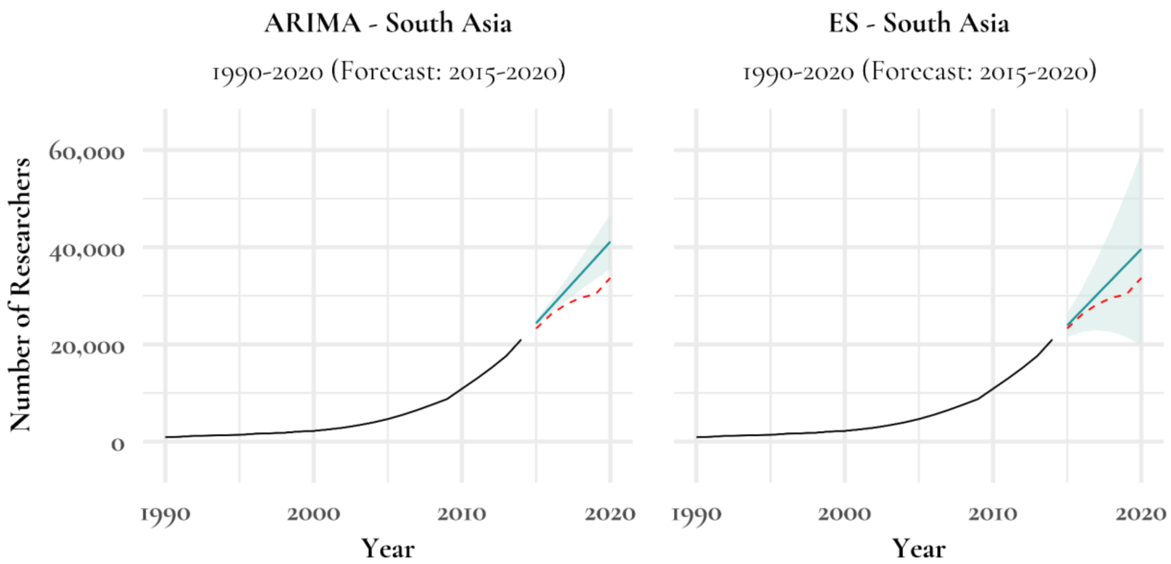
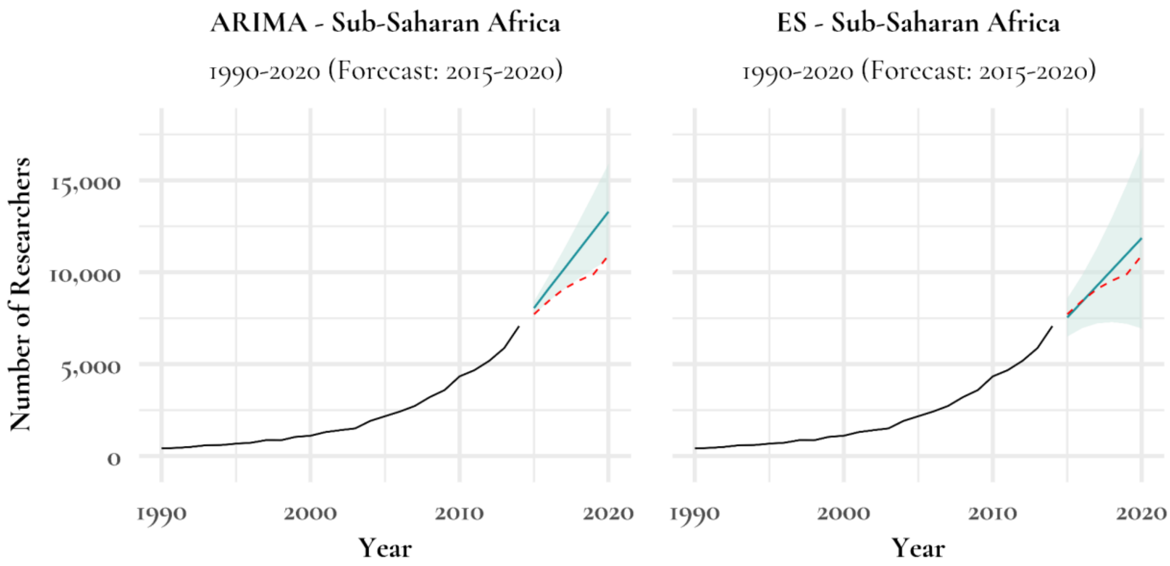


Figure A.3.7. ARIMA and Exponential Smoothing for Sub-Saharan Africa.

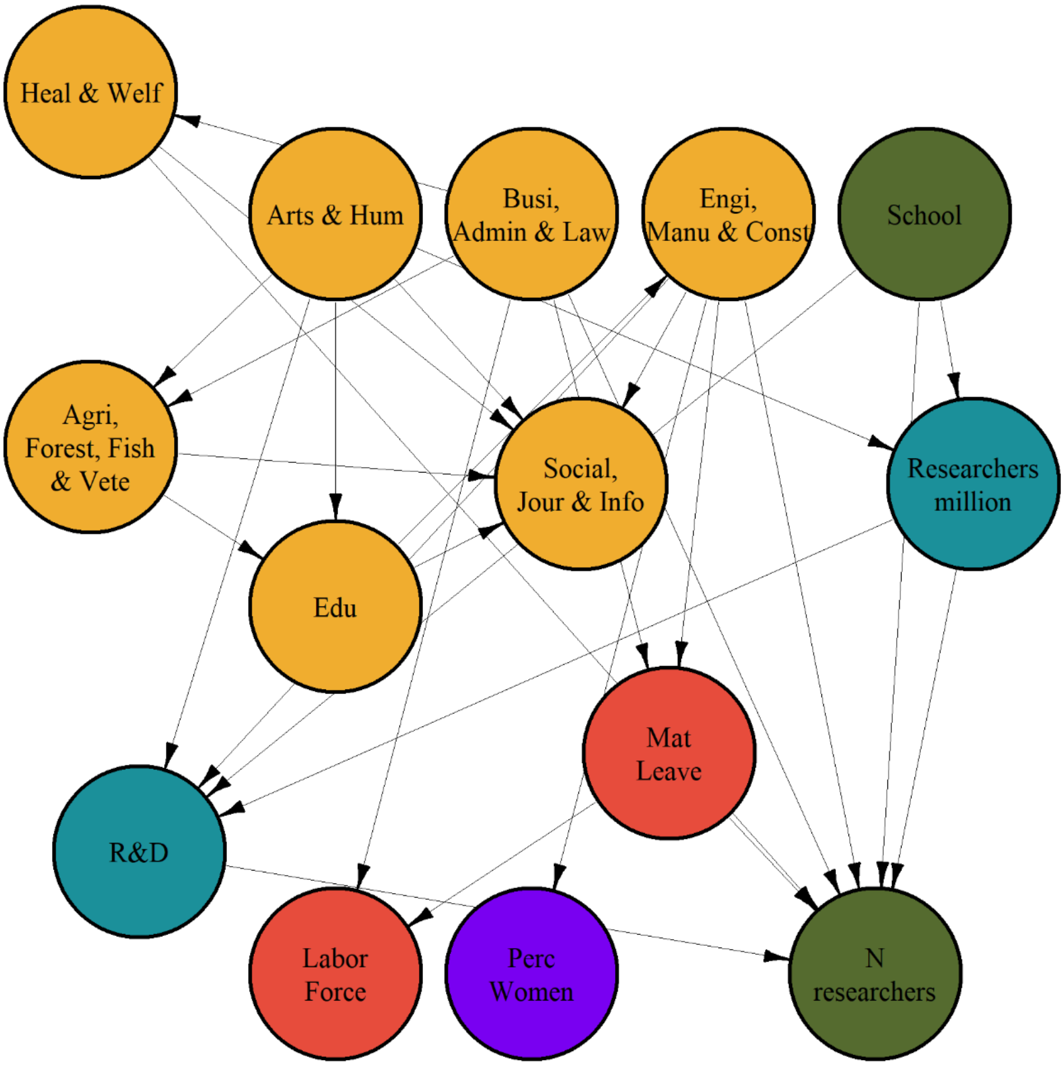


A.4 Bayesian Networks by region

East Asia & Pacific

In East Asia & Pacific we find that the percentage of women researchers is strongly connected to the percentage of women graduating in Engineering, Manufacturing & Construction. We also see that those variables that are most strongly associated with the percentage of women researchers are, again, percentage of women graduating in Business, Administration & Law, in Arts & Humanities and in Engineering, Manufacturing & Constructions. On the other hand, those that show the strongest associations with other variables others are the percentage of GDP spent in R&D, the number of researchers and percentage of graduates in Social Sciences, Journalism & Information that are women.

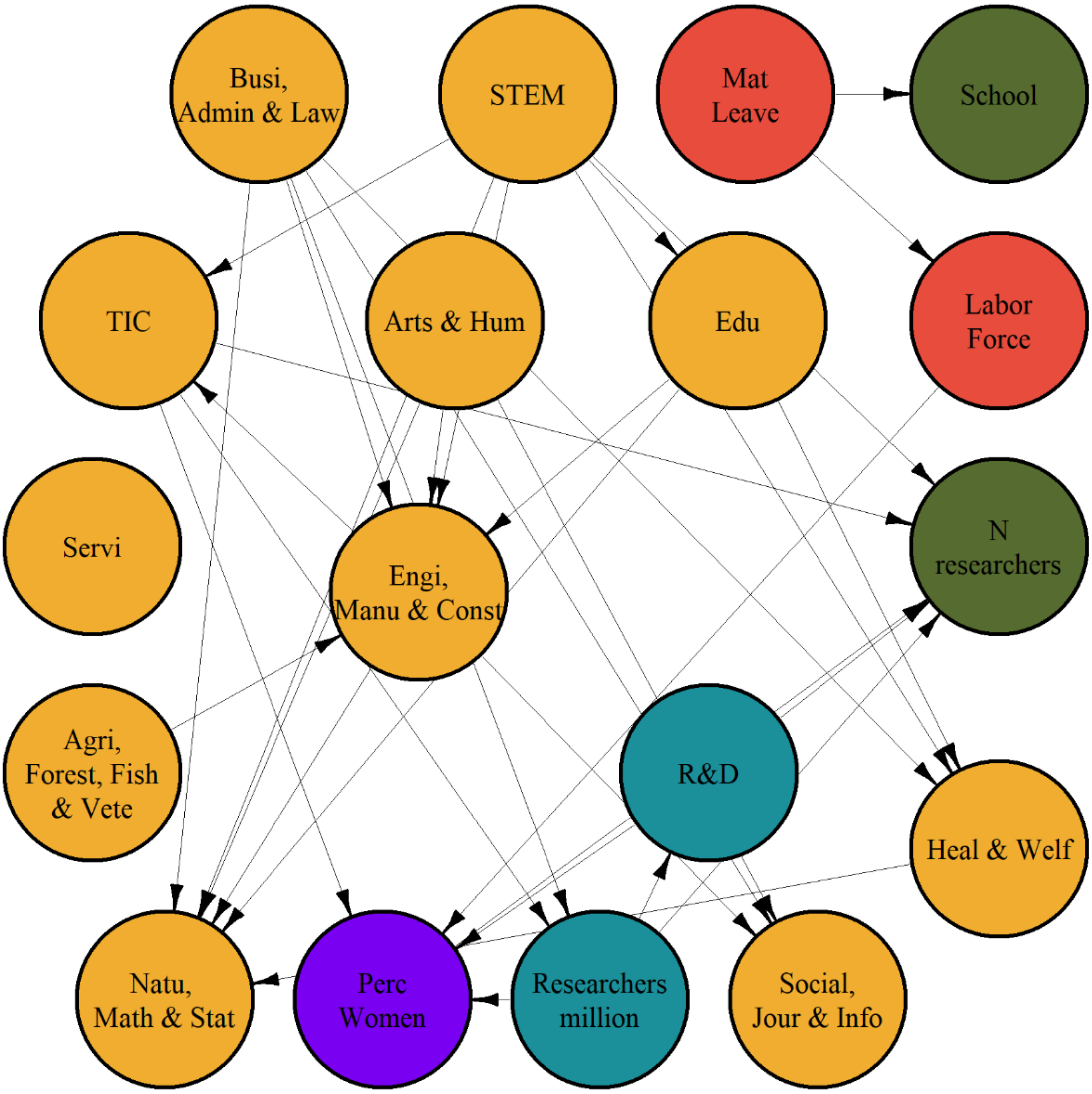
Figure A.4.1. Bayesian Network for East Asia & Pacific.



Latin America & Caribbean

The percentage of women researchers in Latin America is closely associated with more factors: percentage of women graduating in TIC, labour force participation of women, percentage of GDP spent in R&D and the number of researchers per million inhabitants. The fact that those variables that are associated with the percentage of women researchers are mostly variables that are not educational is surprising, compared with other regions. Those variables that are most strongly connected to others are Business, Administration & Law, STEM and Arts & Humanities. Conversely, variables that are more influenced by other variables include the number of researchers, Engineering, Manufacturing & Constructions and Natural, Mathematics & Statistics.

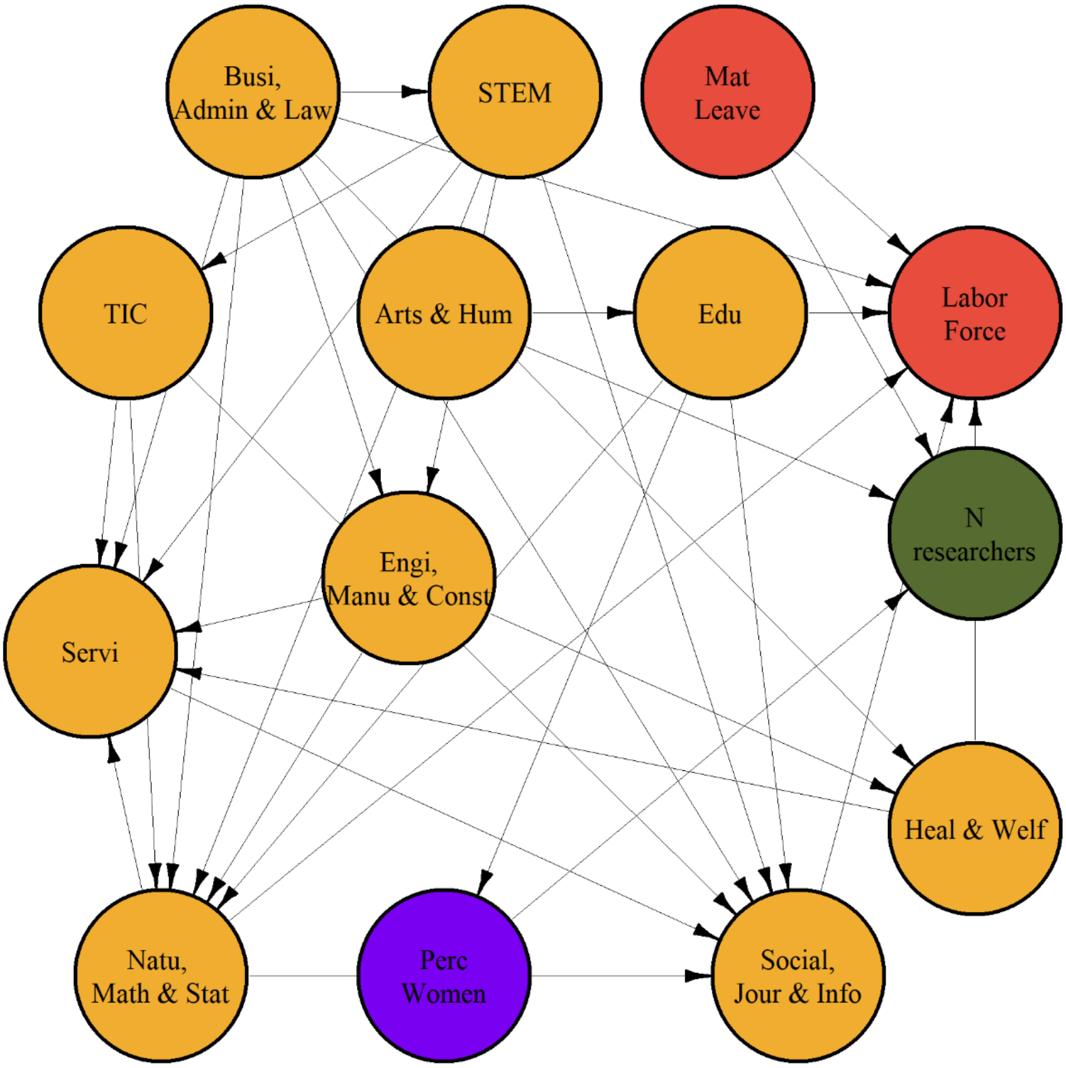
Figure A.4.2. Bayesian Network for Latin America & Caribbean.



Middle East & North Africa

In the case of the Middle East, the percentage of women researchers is directly influenced by the percentage of graduates in Education that are women. This region is interesting, as the percentage of women researchers is modelled as to be indicative of another variable in return: that of the number of researchers. Most shaping variables are Business, Administration & Law and STEM. Those that, in turn, are heavily shaped by others include Services and Natural, Mathematics & Statistics, Social Sciences, Journalism & Information and labor force participation of women.

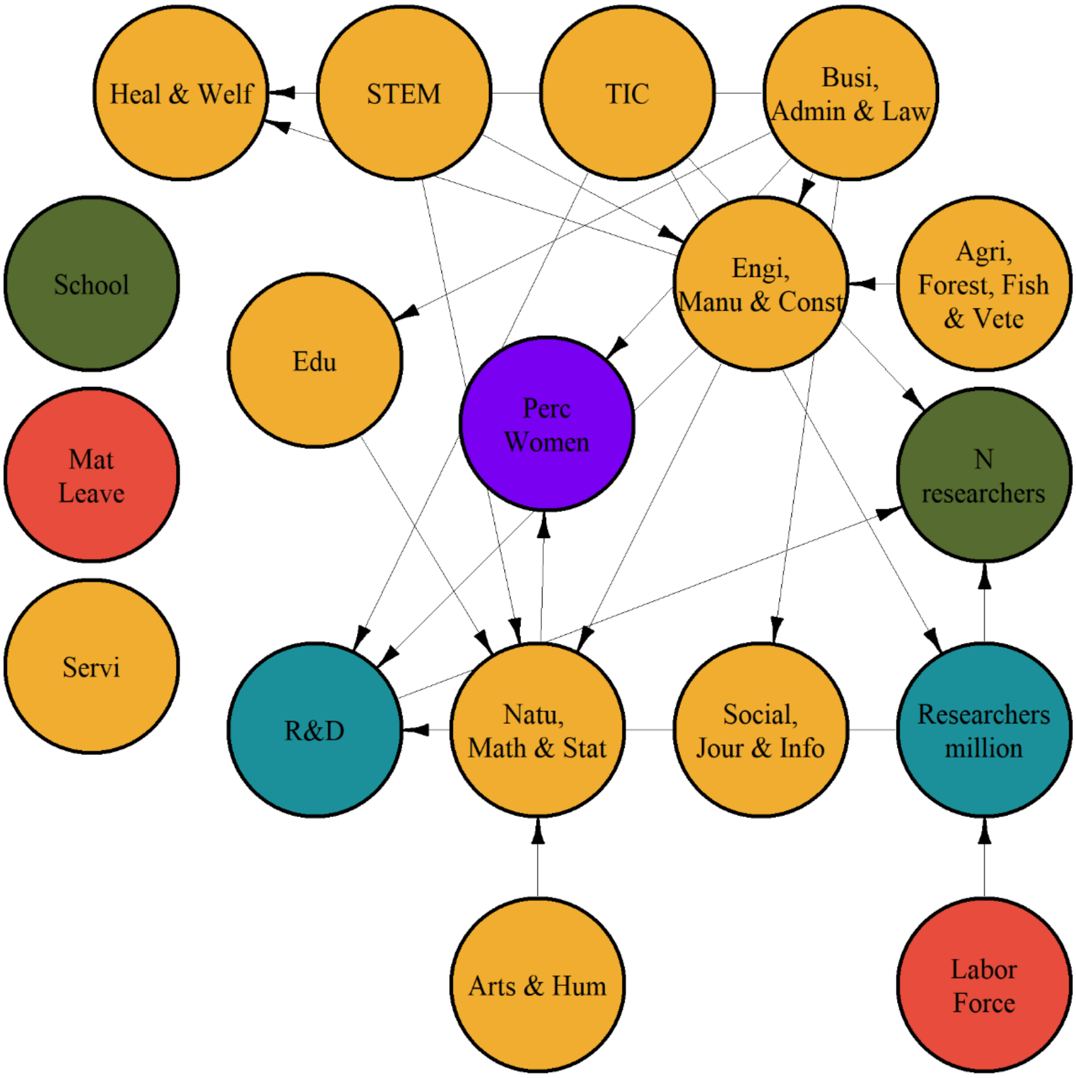
Figure A.4.3. Bayesian Network for Middle East & North Africa.



North America

North America data was merged with Europe & Central Asia to be able to perform the analysis. However, the BN differs slightly from the Europe & Central Asia one, suggesting that there are strong patterns in the North American data. We can see that the percentage of women researchers is associated with Business, Administration & Law and Natural, Mathematics & Statistics, a shift from the European and Central Asia BN. The main difference with the Europe & Central Asia network is the importance that the percentage of women graduates in Engineering, Manufacturing & Constructions has on shaping the number of researchers per million inhabitants, and the association that percentage of graduates in Education that are women has with Health & Welfare. School, Mat Leave and Servi appear to be unconnected with any other considered variable.

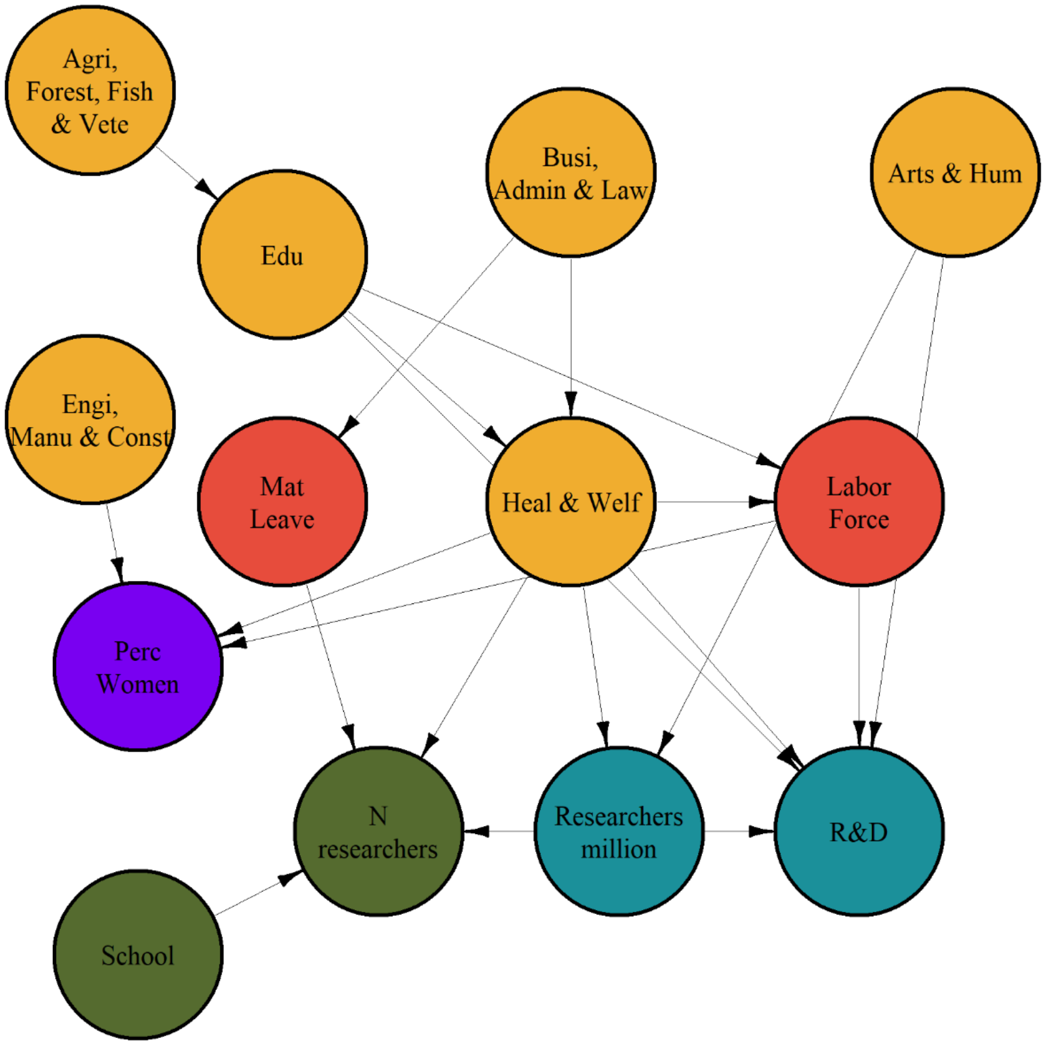
Figure A.4.4. Bayesian Network for North America.



South Asia

Data for South Asia also had to be merged, in this case with data from East Asia & Pacific. However, it shows its own patterns. In this case, the percentage of women researchers was not only associated with Engineering, Manufacturing & Constructions, but also by Health & Welfare and labour force participation of women. Moreover, this BN shows fewer relationships in general, with those variables that shape the network the most being percentage of graduates in Education that are women and Health & Welfare, and those more shaped by others being number of researchers and percentage of GDP spent on R&D. The fewer arcs however, can also be an artifact of the smaller sample size.

Figure A.4.5. Bayesian Network for South Asia.



Sub-Saharan Africa

The percentage of women researchers in Sub-Saharan Africa appears to be associated with many variables: Arts & Humanities, labor force participation of women and school enrollment of girls. In return, it shapes the number of researchers. Those variables that are more strongly connected to others are Arts & Humanities, Business, Administration & Law and Engineering, Manufacturing & Constructions. Those that are shaped by many variables are labor force participation of women, Health & Welfare and number of researchers.

Figure A.4.6. Bayesian Network for Sub-Saharan Africa.