



ELSEVIER

Contents lists available at [ScienceDirect](#)

Journal of Informetrics

journal homepage: [www.elsevier.com/locate/joi](http://www.elsevier.com/locate/joi)

# Use of Benford's law on academic publishing networks

Aleksandar Tošić<sup>a,b,\*</sup>, Jernej Vičič<sup>a,c</sup><sup>a</sup> University of Primorska Faculty of Mathematics, Natural Sciences and Information Technologies, Koper, Slovenia<sup>b</sup> InnoRenew CoE Livade 6, 6310 Izola, Slovenia<sup>c</sup> Research Centre of the Slovenian Academy of Sciences and Arts, The Fran Ramovš Institute, Slovenia

## ARTICLE INFO

### Article history:

Received 4 August 2020

Received in revised form 26 March 2021

Accepted 6 April 2021

### Keywords:

Benford's law

Citation network

Bibliography

## ABSTRACT

Benford's law, also known as the first-digit law, has been widely used to test for anomalies in various data ranging from accounting fraud detection, stock prices, and house prices to electricity bills, population numbers, and death rates. Scientific collaboration graphs have been studied extensively as data availability increased. Most research was oriented towards analysing patterns and typologies of citation graphs and co-authorship graphs. Most countries group publications into categories in an attempt to objectively measure research output. However, the scientific community is complex and heterogeneous. Additionally, scientific fields may have different publishing cultures, which make creating a unified metric for evaluating research output problematic. In complex systems like these, it is important to regularly observe potential anomalies and examine them more carefully in an attempt to either improve the evaluation model or find potential loopholes and misuses. In this paper, we examine the potential application of Benford's law on the official research database of Slovenia. We provide evidence that metrics such as number of papers per researcher conform to Benford's distribution, while the number of authors per paper does not. Additionally, we observe some anomalies and provide potential reasoning behind them.

© 2021 Published by Elsevier Ltd.

## 1. Introduction

This paper describes a new application of Benford's law to the analysis of scientific research collaboration network, focusing on the example of Slovenian scientific publications and authors. It presents scientific foundations for the presented methodology with simplified usage directions. The method was tested on a case study of the Slovenian research collaboration network and shows that Benford's law holds. Further, several scientific and multidisciplinary fields were tested to see if the distribution of first digits, too, obey Benford's law.

The widely known phenomenon called Benford's law (Benford, 1938; Singleton, 2011), also referred to as the first-digit law, is an observation about the frequency distribution of leading digits in many real-life sets of numerical data. It describes the distribution of digits in natural and social processes. The numbers take the form of a logarithmic distribution. Though very old and extensively researched, Benford's law is still an interesting tool for finding anomalies in data. Further, the ever growing amount of data generated calls for simple and effective methods for anomaly detection. While Benford's law has defied many attempts at an easy derivation (Berger & Hill, 2011), many have focused on its application rather than its

\* Corresponding author at: University of Primorska Faculty of Mathematics, Natural Sciences and Information Technologies, Koper, Slovenia.

E-mail addresses: [aleksandar.tosic@innorenew.eu](mailto:aleksandar.tosic@innorenew.eu) (A. Tošić), [jernej.vicic@upr.si](mailto:jernej.vicic@upr.si) (J. Vičič).

URLs: <https://innorenew.eu> (A. Tošić), <https://osebje.famnit.upr.si/jetnej/> (J. Vičič).

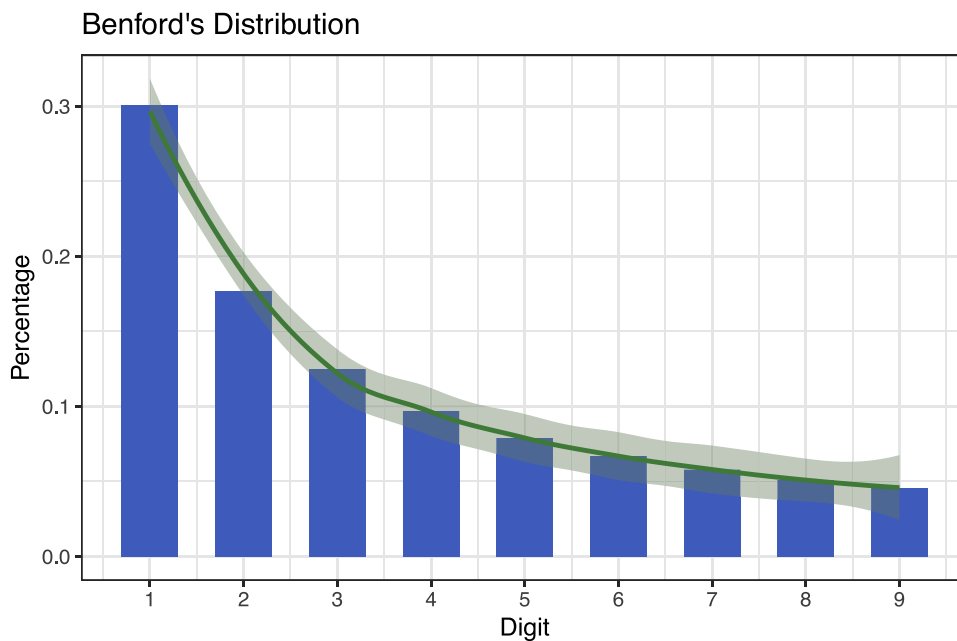


Fig. 1. The distribution of digits in accordance to Benford's law (Singleton, 2011).

theoretical background (Nigrini, 1996). The applications range from election fraud detection, detecting image manipulation, accounting fraud detection (Durtschi, Hillison, & Pacini, 2004), scientific fraud detection (Ranstam et al., 2000), etc. Benford's law has been effectively and frequently used in forensic accounting as presented in Bhattacharya and Kumar (2008) and Nigrini (2012). The same methodology has also been successfully used in other analyses such as the modelling of behavioural features for social network users (Golbeck, 2015) and meteorological events, for example, the travelled distances of tropical cyclones since 1842.

The objective of this paper is to evaluate the potential use of Benford's law on research networks. The method can be used to detect potential discrepancies from a macro level. Upon detection, the method can produce observations that skew the distribution referred to as suspects. These observations are a great entry point for further, more fine grained analysis in an attempt to explain and reason about them. The method is useful to self evaluate research communities as well as the maturity of specific research fields. It can serve as a feedback loop to the regulator in order to improve on the valuation system and respond to potentially unwanted shift in direction. Additionally, we show the method can be applied on temporal data. This is especially useful to examine when discrepancies occurred.

This paper is composed as follows: Section 2 describes Benford's law (Singleton, 2011) and its applications followed by a description of the state of the art in the scientific field with the assumption that the paper is multidisciplinary. Section 3 presents the Slovenian scientific/research network that was used as a case study for the presented methodology followed by methodology and results. The paper concludes with a discussion.

## 2. Benford's law

The first-digit law is an observation about the frequency distribution of leading digits. It is also known as the Newcomb-Benford law or Benford's law. It has been apparently first discovered by polymath Newcomb and published in Newcomb (1881) and later rediscovered by physicist F. Benford and presented in Benford (1938). The Benford's law (Singleton, 2011) defines a fixed probability distribution for leading digits of any kind of numeric data with the following properties:

- Data with values that are formed through a mathematical combination of numbers from several distributions.
- Data that has a wide variety in the number of figures (e.g., data with plenty of values in the hundreds, thousands, tens of thousands, etc.)
- Data set is fairly large, as a rule of a thumb at least 50–100 observations (Kenny, 2015).
- Data is right skewed (i.e., the mean is greater than the median), and the distribution has a long right-tail rather than being symmetric.
- Data has no predefined maximum or minimum value (with the exception of a zero minimum).

The distribution of digits is presented in Fig. 1; the digit 1 occurs in roughly 30% of the cases, and the other digits follow in a logarithmic curve. It has been shown that this result applies to a wide variety of data sets (Singleton, 2011), including

electricity bills, street addresses, stock prices, house prices, population numbers, death rates, and lengths of rivers. The equation for the distribution of the first digits of observed data is presented in Eq. (1).

$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(1 + \frac{1}{d}\right) \quad (1)$$

### 3. Slovenian research network

In Slovenia, researchers are evaluated by a methodology issued by the Slovenian Research Agency (ARRS). The methodology is a transparent set of rules also implemented in the national research information system SICRIS (Korošec, 2014). Scientific contributions that get added to the database are classified into typologies and then credited with points upon verification. Special emphasis is given to scientific articles that are evaluated based on data from the Web of Science (WoS) with SCIE/SSCI/AHCI journal indexes, Journal of Citation Reports impact factor database (JCR), and WoS citations as well as Scopus Source Normalized Impact per Paper impact factor database (SNIP) for social sciences and humanities and Scopus citations (Falagas, Pitsouni, Malietzis, & Pappas, 2008). Such a robust system is needed in order to provide a holistic method for government funding of young researchers as well as national and foreign project proposals. Moreover, the system is used for progression of academic rank by enforcing minimal requirements from PhD onward. A detailed and systematic overview of Slovenian research network is given by (Curk, 2019).

Journal papers are covered by typologies 1.01, 1.02, and 1.03 and are ranked into four categories (quarters), corresponding to the journal's position in scientific field, that are used to compute the score, which is then divided equally amongst the authors. Additional points are added based on the number of citations to each author equally.

Another important aspect is the minimum requirements for PhD dissertations, which require candidates to publish at least one paper in an SCI indexed journal. (Heneberg, 2016) observed that bibliometric indicators increasingly affect careers, funding, and reputation, creating a new extreme: "where a scientist publishes has become much more important than what is published", which is in agreement with Holub, Tappeiner, and Eberharter (1991) and Shibayama and Baba (2015). While the need for a holistic and robust system for valuating research contributions is arguably necessary, it is equally important for the system to describe specifics for each research field. Some research fields have very different publishing habits and cultures, making it difficult to provide a unified framework of evaluation while maintaining comparability and simplicity. (Larivière, Archambault, Gingras, & Vignola-Gagné, 2006) concluded there are significant differences in publishing habits between social sciences and humanities (SSH) and natural sciences and engineering (NSE), creating a particular problem in the field of bibliometric valuations.

The discrepancy between valuation metrics and publishing habits can create unwanted incentives. Coupled with a systemic point-based funding system, this can create undesired incentives to abuse the valuation methodology in one's favour. Additionally, different research fields have different publishing cultures with respect to the topology of the contribution. It is important to observe, and monitor the publication network from a macro perspective to identify trends, and address potential discrepancies. A holistic view is of great importance both to academia to identify trends, as for the government to adjust the valuation metrics and guide the research towards predefined goals. Moreover, identifying unexpected changes in the publishing network on a macro level calls for further, more fine grained analysis on the micro level.

### 4. State of the art

Benford's law has been thoroughly researched and its theoretical grounds have been proved in many scientific papers. The phenomenon is discussed in greater detail in Berger and Hill, 2011 the article also provides strengthened versions of, and simplified proofs for, many key results in the literature. Many researchers have verified for themselves that the law is widely obeyed but have also noted that the popular explanations are not completely satisfying (Fewster, 2009). Bibliometric and infometric studies have addressed the issue of academic network and co-authorship network profusely going from global or national views Braun and Glaenzel (1996) or Leydesdorff and Wagner (2008) to the individual level Melin (2000) and Newman (2004). Ariel Xu and Chang (2020) show that the co-authorship network correlates well with the academic performance. There are numerous papers that found a positive relationship between the international collaboration and the research impact such as Narin and Whitlow (1990) or Katz and Hicks (1997). The background, the current status, and trends of academic social networks are researched and finds presented in Kong, Shi, Yu, Liu, and Xia (2019). A study on research collaboration Benavent-Pérez, Gorraiz, Gumpenberger, and de Moya-Anegón (2012) as well as some studies that date a few decades ago, such as Bordons, Gomez, Fernández, Zulueta, and Méndez (1996), focus on geographical impact on international and intra-national research collaboration. Ortega (2014) presents an analysis of the relationship between research impact and the structural properties of co-author networks. The methodology described in our paper is most suitable for the observation of maturity of the research area. There has been some research in this area such as Keathley-Herring et al. (2016); Pelacho, Ruiz, Sanz, Tarancón, and Clemente-Gallardo (2021) presents an analysis of the evolution of a targeted science field. One of our aims is to observe the changes through time in bibliographic network resulting in maturity of the network. (Batagelj & Maltseva, 2020) proposes a method to transform bibliographic networks, using the works' publication year, into corresponding temporal networks based on temporal quantities and then defines interesting temporal properties of nodes, links and their groups thus providing an insight into evolution of bibliographic networks.

The rest of the section presents state of the art on various fields and aspects that are connected with our research.

#### 4.1. Benford's law applications

The application of Benford's law is by far most prevalent in the accountant fraud detection and there has been a lot of research in the area, such as (Drake & Nigrini, 2000) who introduces students to Benford's Law and Digital Analysis (analysis of digit and number patterns of a data set), which can be used as an analytical procedure and fraud detection tool. Nigrini (2017) presents a current literature overview of the area. Durtschi et al. (2004) presents Benford's law as a simple and effective tool for the detection of fraud. The purpose of the paper is to assist auditors in the most effective use of digital analysis based on Benford's law by identifying data sets which can be expected to follow Benford's distribution, and presenting types of frauds that would be "detected/not detected" by such analysis.

Cleary and Thibodeau (2005) however, points out some inherent problems that potentially arise in the use of the Benford's law in the auditing process. The paper compares the merits of Benford's law and typical statistical test-by-test approach.

The simplicity of the Benford's law as a tool allows for a broad range of uses. Hickman and Rice (2010) examined crime statistics at the USA National, State, and local level in order to test for conformity to the Benford distribution. Burke and Kincanon (1991) observe the distribution of initial digits of physical constants, their results are inconclusive, though. One of the more recent researches involving Benford's law is Zhang (2020), where the authors propose a test of the reported number of cases of coronavirus disease 2019 in China with Benford's law and report that the reported numbers of affected people abide to Benford's law.

#### 4.2. Research integrity and unethical behaviour

Benford's law has been most successfully used in accounting fraud detection and quite a few research projects aimed at using the same tool to detect frauds in other fields such as Zhang (2020) in the detection of counterfeiting COVID-19 reports. It comes naturally to take the inspiration from financial area and just shift it to the new domain, but this is not our aim, the use of the presented tool as a fraud detection metric still needs to be explored.

The problem of research integrity and unethical behaviour has spread in new forms and dimensions as observed by many scholars such as Martin (2013) and Bohannon (2013). There is now a growing body of research on scientific integrity and misconduct as well as presentation of guidelines and best practices such as Fanelli (2013) or Peterson (2007). There are whole courses devoted to this issue at the university level using textbooks such as Macrina (2014) and Sponholz (2000). Many research papers and opinions have been published in recent years expressing concerns over research integrity such as Godecharle, Nemery, and Dierickx (2014) and Bernstein (1984), which examines ethical issues raised in one example case. Edwards and Roy (2017) argues that scientists have become increasingly perverse in terms of competition for research funding and development of quantitative metrics to measure performance. The peer-review system's effectiveness as a means of preventing misconduct in science is challenged in papers such as Van der Heyden, van de Derks Ven, and Opthof (2009). Surveys indicate increasing numbers and extremes of misconduct (John, Loewenstein, & Prelec, 2012). Our paper does not try to newly define the delicate subject of academic integrity, nor does it present a tool that pinpoints misconduct of a single author or a single poor research contribution. Instead, the article proposes a methodology for following changes and assessing the maturity of research system.

#### 4.3. The research of the Slovenian research network

Each new bibliographic source, should be tested for its suitability for bibliometric analyses such as the case of Microsoft Academic Search (Ortega, 2014). The Slovenian academic/research network has been examined extensively through different perspectives including a quantitative and qualitative methodological approach to scientific cooperation by Mali, Pustovrh, Cugmas, and Ferligoj (2018), a study of community structures through scientific co-authorship (Cugmas, Ferligoj, & Kronegger, 2019). Pisanski, Pisanski, and Pisanski (2020) present two methods to ease visualization of large networks such as bibliographic networks. They showcase the methods on Slovenian research network. Ferligoj, Kronegger, Mali, Snijders, and Doreian (2015) examine the collaboration structures and dynamics of the co-authorship network of all Slovenian researchers. Its goal is to identify the key factors driving collaboration and the main differences in collaboration behavior. A new measure for interdisciplinarity that takes into account graph content and structure is proposed in Karlovčec and Mladenič (2015). The proposed new measure is applied in exploratory analysis of research community in Slovenia; a commentary to this paper (Rodela, 2016) addresses two shortcomings while still supporting the weight of the paper. Lužar, Levnajič, Povh, and Perc (2014) presents a study of the dynamics of interdisciplinary sciences in the case of Slovenian scientific network.

## 5. Methodology

As mentioned in Section 1, this paper proposes a methodology for following changes and assessing the maturity of research system. As such, the purpose is to present scientific grounds that allow feasibility and usefulness of the method as well as to propose a set of usage guidelines and a use case where our hypotheses were confirmed. The observation sets need to conform to all the basic prerequisites for Benford's law as described in Section 2. The Slovenian research network

described in Section 3 was used as a use case for a network of scientific publications. The methods and premises are easily applicable to other European and non-European countries. The data identified in our study includes, but it is not limited to, the following examples:

- the number of publications per author,
- the number of co-authors per author,
- authors are usually categorized in one or more scientific fields,
- authors are associated to one or more research institutions that can be further geo-located,
- publications are all tagged with the year they were published.

Another comparison comes naturally when observing the presented data: “the number of authors per publication”, but although, in theory, this number is unlimited, in practice they are not, eliminating just a few of the most populated publications. We are left mostly with numbers 1 or low 2-digit numbers. We hypothesise, and confirm the validity in Section 7, that the following comparisons are all subject to Benford’s law:

- the number of co-authors per author,
- the number of publications per author.

The highest number of publications per author in our dataset was 8483 and there were more than 0.5% that were higher than 1000, which satisfies (although borderline) the second prerequisite described in Section 2. Each of the assumptions can be further distributed on selected sets:

1. the whole (Slovenian) scientific network,
2. broken into scientific areas (e.g., natural sciences, social sciences, etc.),
3. grouping publications in discrete time periods,
4. grouping authors into geographical areas.<sup>1</sup>

The rationale behind selection of the presented sets follows: #1 all publications, #2 is there a scientific area that has is in starting phases of development or are the numbers of publications and connections skewed by some other property, #3 was there a time period when the observed set did not behave according to Benford’s law, #4 are there geographical areas (usually corresponding to national borders – ethnic or institutional perceptions of research publications) where, again, the numbers do not conform to Benford’s law.

Testing that data conforms to Benford’s distribution has been done with many goodness of fit tests ranging from Pearson’s Chi squared, Kolmogorov–Smirnov  $D$  statistics, Freedman’s modification of Watson  $U^2$  statistics, euclidean distance  $d$  statistics, and many others. However, no real data will ever follow the exact distribution; hence, most analysis supplements statistical testing with graphical representations that help in pointing out suspicious patterns in the data for further investigation. Additionally, different tests have different reactions on sample sizes. The Chi square test suffers from an excess power problem in that when the number of observations becomes large (above 5000 records estimated by Nigrini (2012)) it becomes more sensitive to insignificant spikes, leading to the conclusion that the data does not conform. Nigrini (2001) suggested some statistical tests can render misleading results when applied to large number of observations. On the other hand, Druica, Oancea, and Vâlsan (2018) conclude that MAD test is reliably with as low as 200 observations. Alexander (2009) proposed the Mantissa Arc test, which is a very interesting geometrical test. Unfortunately, it tolerates little deviation from Benford’s distributions. Nigrini (2012) concluded that the best test is Mean Absolute Deviation (MAD), also setting critical objective scores for conformity (0.000), acceptable conformity (0.006), Marginally acceptable conformity (0.012), and nonconformity (0.015). The adapted MAD is used to measure the average deviation between the heights of the bars and the Benford line. The higher the MAD, the larger the average difference between actual and expected proportions. In our use case, we perform all conformity tests using all three of the aforementioned tests as our sample sizes are well within the acceptable ranges. We supplement the statistical tests with graphical representations; the results are presented in Section 7.

**Basic recipe:** Select a big enough set of aggregated data that conforms to Benford’s law prerequisites. Gather data and count aggregated values. Count leading values and perform Mean Absolute Deviation (MAD) on the gathered data. Plot simple bar charts with the numbers for each leading digit and observe the distribution. If the data does not conform to Benford’s law, investigate further.

<sup>1</sup> this set was not addressed in the paper but the authors do not expect any discrepancy from the presented results as long as we keep the number high enough to conform to Benford’s law prerequisites.

### 5.1. Description of the data set and data acquisition process (Slovenian academic network)

The Slovenian academic network is represented by two public services: Slovenian Current Research Information System (SICRIS) (Curk, 2019), which stores data about scientific research including scientist/researcher education, degree, scientific field, affiliation, type of employment, and national funded research projects, and Cooperative Bibliographic System and Services (COBISS) (Seljak & Bošnjak, 2006), which stores data about all publications (including research publications that are our primary concern in this paper). A research classification scheme is used for unique identification of researchers. Researchers' bibliographies are created in the shared cataloguing process; however, the use of a uniform methodology of documents/works is mandatory for classification of bibliographic items. The two services are seamlessly combined and publicly available through a querying interface. Although the data is publicly available and can be acquired for research purposes, we opted for crawling the available data. In March 2020, a local database was constructed by slowly crawling the service in order to avoid load problems. Each researcher is represented by a unique Researcher ID (mstid) issued by the Slovenian Research Agency. The foreign co-authors were not disambiguated; a new entity was created for each co-author of an observed bibliographic entry. The data gathered in this process is valid as long as we observe Slovenian authors and use foreign co-authorship only in aggregated form. The local database's structure is presented in Fig. 2.

## 6. Misuse of the proposed metric

This section presents an example where the proposed metric will give misleading results. Benford's law is not applicable for the number of authors per research contribution case as it violates the second constraint outlined in Section 2 that requires a wide variety in the number of figures, ranging into hundreds and thousands. The highest number of coauthors on a single contribution in Slovenian academic network is 40 and there are only 92 contributions with 30 or more co-authors and only 323 with 20 or more. Fig. 3 shows a skewed distribution of first digits. The first digits were analyzed with a sample size of 788,410 observations. The MAD Conformity (Nigrini, 2012) was classified as Nonconformity with all statistical tests having  $P$ -values near 0. Since the number of authors per paper does not conform to Benford's law.

## 7. Results and discussion

This section presents the results of the methods applied on identified sets (all described in Section 5. Fig. 4 shows that distribution of first digits for number of contributions per author for the whole dataset (e.g., the whole Slovenian scientific network) conforms to Benford's law. The MAD Conformity according to (Nigrini, 2012) is Acceptable.

Table 2 summarizes the results of testing individual research fields on conference papers and journal papers, separately. The main statistics used to determine if the data conforms to Benford's law was mean absolute deviation (MAD Conformity) (Nigrini, 2012). The distortion factor model indicates whether the digit patterns are over- or understated and extent of the distortion. Additionally, Pearson's Chi-squared and Mantissa Arc tests are given for clarity and confirmation in cases with marginally acceptable conformity. Based on the aforementioned tests, data sets get classified as either nonconformity, marginal, acceptable, or close conformity. We observe that most fields conform very strongly, with exceptions being number of journal papers published in social sciences and conference papers published in humanities and biotechnology. Further analysis can be drawn from the digit distribution charts for individual data set.

Figures 5, and 6 provide more insight into individual research fields, and contribution typologies. Suspects are marked (red) if the squared difference between the observed digit distribution and Benford's distribution is greater than 4. Finding the correct threshold is an empirical process in which most suspect observations can be obtained by starting with the highest threshold (observation with the highest squared difference) and decreasing it in a step by step process to obtain a larger set of observations that skew the fit the most. These observations can provide more insight into the objective MAD conformity results. The threshold was chosen empirically solely for a meaningful graphical representation. Social sciences have the biggest offset in the first three digits, which could be explained by the different publishing culture or an increase in the number of researchers that only publish the single journal paper required for academic title or PhD thesis.

We establish that, in general, the number of published papers per author conforms to Benford's distribution with some variances between different research fields, which could be attributed to different publishing cultures. However, using Benford's law to possibly identify anomalies in publishing is not very useful if performed without taking into account the temporal aspect of the data. We show that Benford's law can be used on shorter time frames to narrow the search considerably. Analyzing temporal data using Benford's distribution has been shown before (Sambridge, Tkaličič, & Jackson, 2010), where samples are divided into time intervals and tested individually. In Fig. 7, we show the digit distribution of the same data split by decade. We observe that from 1960 until 1990 the digit distribution does not conform. However, after the 1990s we see a strong conformation. The data from 1960–1970 only has 38 observations, which is well below the threshold. The contributions published between 1970–1980 only amount to 493 observations, suggesting the minimum sample size could be increased considerably for this particular use case.

Table 1 shows the numerical results for conformity tests. The key observation is that Benford's law can be used on lower time frames as long as the sample size is large enough. Additionally, we observe a sizable increase in the number of papers published between 1990–2000; this is possibly a result of increased Internet access availability, which bolstered international collaboration that resulted in more published papers.

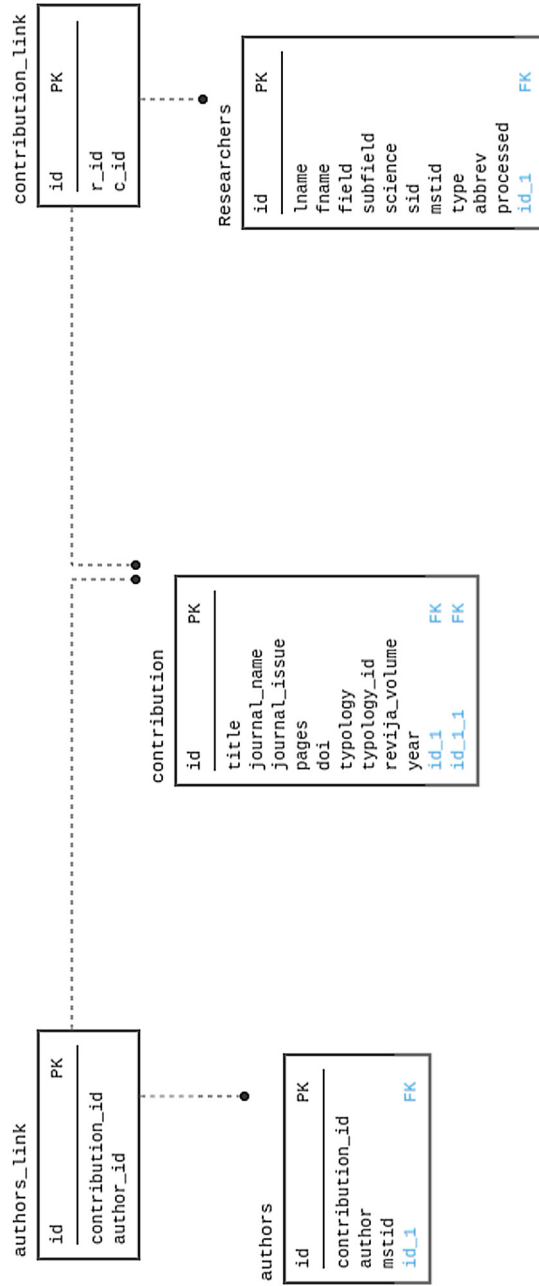
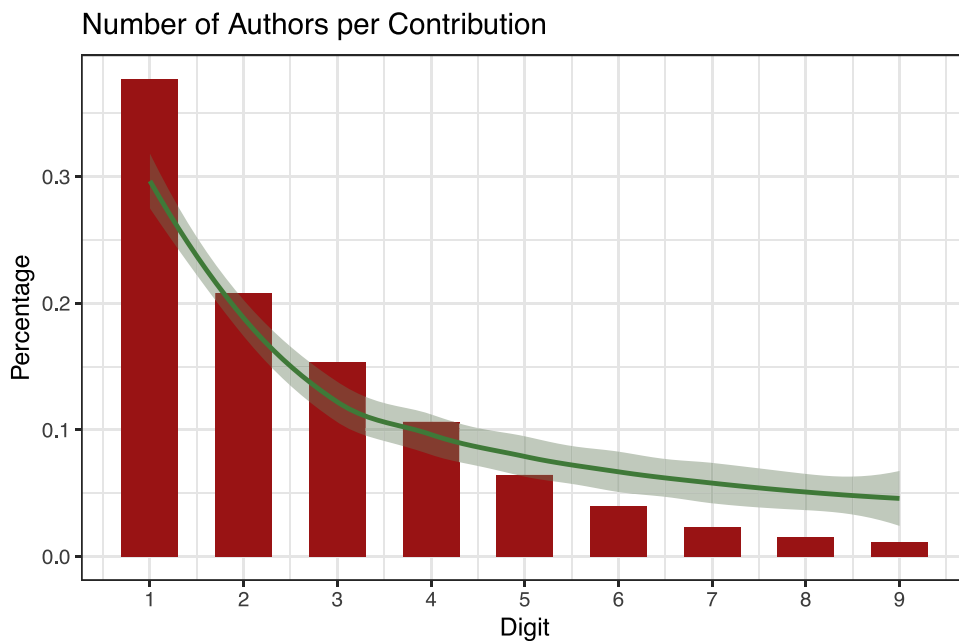
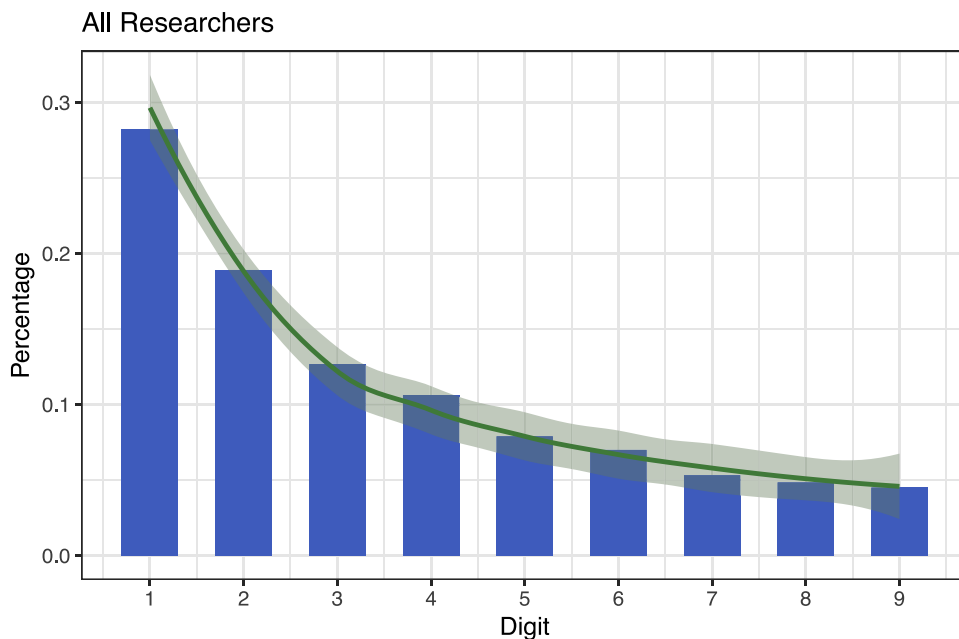


Fig. 2. Structure of the gathered data set.



**Fig. 3.** Example of a misused metric. The graph shows that the observed data do not conform to Benford's law. The data does not conform to Benford's law because the number of authors per publication rarely exceeds 10. The Benford's law is not applicable to the presented data.



**Fig. 4.** The distribution of first digits for number of contributions per author for the whole Slovenian scientific network conforms to Benford's law. The MAD Conformity (Nigrini, 2012) is Acceptable.

Fig. 8 shows the results of tests performed on a yearly basis for each individual field. We observe that for the first 20 years the sample sizes were below the threshold  $T = 50$ , and we have marked those with “Insufficient data” accordingly. We observe that after the 1990s the Slovenian scientific community matured as confirmed by the general conformity of all fields with the exception of humanities, which lagged behind a few years. From these results, we conclude Benford's law can be effectively used on a yearly basis to identify potential discrepancies, which could be attributed to a number of factors ranging from unethical behaviour, different publishing cultures, inefficiencies in the valuation metric, underfunded research fields, etc. A more detailed analysis should be performed on those observations that skew the distribution the most, referred to as suspects.



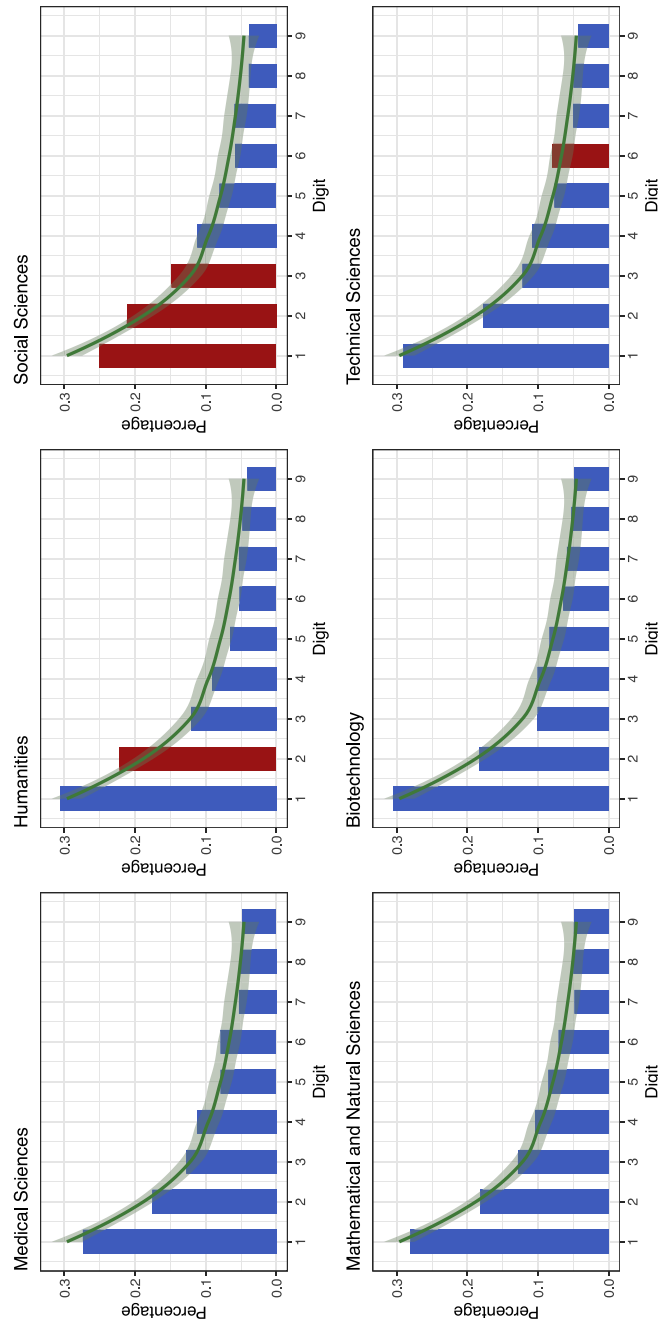


Fig. 5. Leading number distribution of journal articles (1,01).

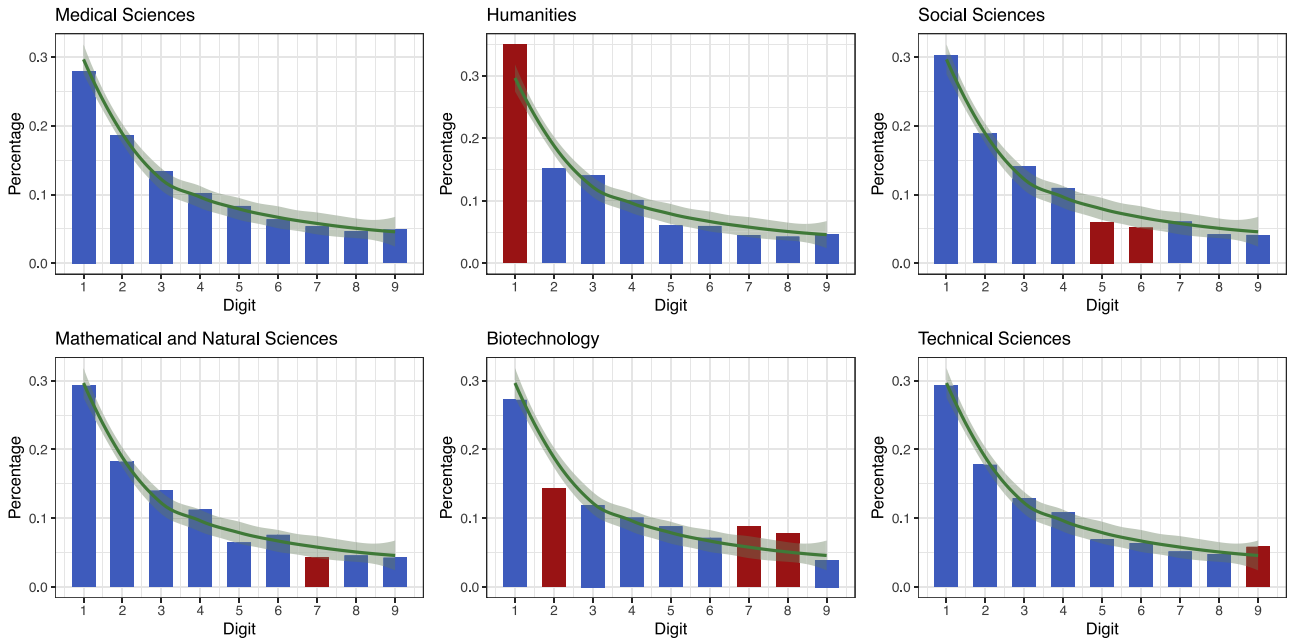


Fig. 6. Leading number distribution of conference articles (1.08).

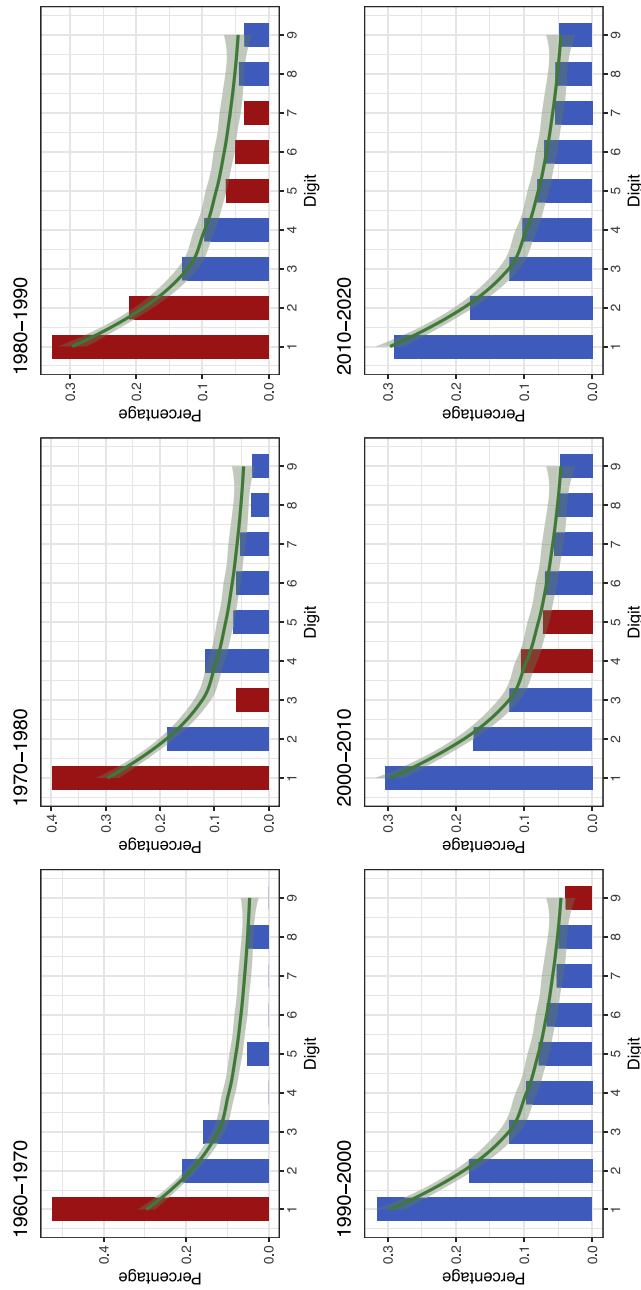
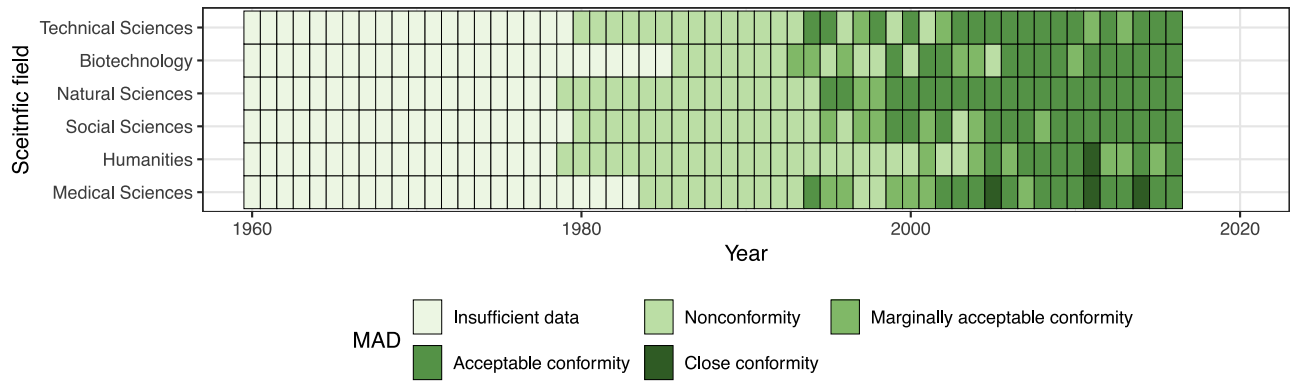


Fig. 7. Number of papers per author grouped by decade.



**Fig. 8.** Benford's conformity for each research field on a yearly basis.

**Table 1**  
Conformity tests for number of contributions per author per decade.

Data set	Observations	Pearson's Chi-squared test		Mantissa Arc Test		MAD	MAD Conformity	Distortion Factor
		X-squared	P-value	L2	P-value			
1960–1970	38	17.502	0.02529	0.12376	0.00907	0.0653	Nonconformity	–51.41354
1970–1980	403	35.344	2.31E–05	0.046227	8.12E–09	0.0284	Nonconformity	–44.27548
1980–1990	1898	49.522	5.05E–08	0.0047154	0.0001298	0.0149	Marginal conformity	–31.19197
1990–2000	5176	13.556	0.09411	0.0011942	0.002068	0.0042	Close conformity	–12.50242
2000–2010	8819	13.872	0.08515	0.00040233	0.02878	0.0032	Close conformity	–7.73507
2010–2020	9660	13.909	0.08418	0.00024952	0.08978	0.0039	Close conformity	–6.949466

**Table 2**

Benford's conformity by field and typology. Abbreviations: H=Humanities, SS=Social Sciences, TS=Technical Sciences, B=Biotechnical Sciences, MNS=Mathematics and Natural Sciences, and MS=Medical Sciences.

Data set	Observations	Pearson's Chi-squared test		Mantissa Arc Test		MAD	MAD Conformity	Distortion Factor
		X-squared	P-value	L2	P-value			
H (1.01)	977	18.843	0.01572	0.002711	0.07074	0.0113	Acceptable conformity	–30.43495
H (1.08)	817	18.406	0.01838	0.0079669	0.00149	0.0156	Nonconformity	–44.13039
SS (1.01)	1401	38.374	6.42E–06	0.0067329	8.01E–05	0.0171	Nonconformity	–15.35622
SS (1.08)	1346	19.807	0.01109	0.0006882	0.396	0.0104	Acceptable conformity	–421.12554
TS (1.08)	2567	19.203	0.01381	0.00001248	0.9685	0.006	Close conformity	–14.40058
BS (1.01)	775	4.5918	0.8002	0.0011387	0.4137	0.005	Close conformity	–10.07059
BS (1.08)	713	29.099	0.0003046	0.012097	0.0001795	0.016	Nonconformity	–2.469761
MNS(1.08)	1477	19.085	0.01441	0.00020869	0.7347	0.01	Acceptable conformity	–19.83444
MNS(1.01)	1919	9.2685	0.3202	0.00064159	0.2919	0.0068	Acceptable conformity	–13.24065
MS (1.01)	1583	12.084	0.1475	0.0015583	0.08485	0.007	Acceptable conformity	–10.55757
MS (1.08)	1388	6.033	0.6435	0.00015432	0.8072	0.007	Acceptable conformity	–19.89672

## 8. Conclusion and further work

This paper proposes a methodology for following changes and assessing the maturity of research system using Benford's law. The paper identifies the type of data sets that can be used for this purpose. The presented method was evaluated on a real-world test case: the Slovenian research network. Research findings suggest that the method can be used to identify possible non-isolated cases of deviations. The method identifies groups and does not concentrate on individuals. To identify individual cases, a more granular inspection is needed by analysing suspect observations. We identify two approaches for narrowing the search space, namely analyzing individual research fields and temporal analysis. We conclude the method can be reliably used for individual research fields and shows some discrepancies in the social sciences and humanities, which can be attributed to field-specific publishing culture that requires a more fine-grained method to conform. Additionally, temporal analysis confirms that the method can be used reliably on 12-month intervals, provided the sample meets number of observation threshold. Temporal analysis hints at the possibility of using annual research reports to continuously monitor the behaviour of individual research fields.

Additionally, we observe an unexpected increase in the number of papers written between 1990–2000. We identified two possible hypothesis:

- In 1991 Slovenia became the first republic that split from Yugoslavia. The change in political structure had large implications in the funding, promotion, and general structure of the scientific network.
- In the 90s, the Internet became a widely used media of information exchange. This allowed researchers to bridge the physical gaps that existed before, and allowed for more international collaboration.

Additionally, temporal data suggests that in this decade the conformity improved considerably.

The paper also presents a case where the method was intentionally misused, thus producing misleading results. The importance of which is emphasised to avoid false positives. However, we are fully aware of potential shortcomings of using smaller data samples, and suggest the method be verified on countries with larger data sets.

All aggregated (and anonymized) data is available on Zenodo:[dataset] <https://doi.org/10.5281/zenodo.3935770> to provide researchers the ability to replicate our experiment and reuse our data for additional research.

In the future, we would like to test the proposed method on a new test case; a selection of possible candidates are the French national repository HAL,<sup>2</sup> Current Research Information System in Norway (Cristin),<sup>3</sup> Hungarian Scientific Bibliography,<sup>4</sup> and Information system for research, experimental development and innovation research for Czech Republic.<sup>5</sup>

The authors also hypothesize that the method would perform as predicted on a group of authors that extends across national borders such as authors in a global scientific fields using methodology similar to (Zdravevski et al., 2019).

## Author contributions

**Conceived and designed the analysis:** Aleksandar Tošič, Jernej Vičič

**Collected the data:** Aleksandar Tošič

**Contributed data or analysis tools:** Aleksandar Tošič, Jernej Vičič

**Performed the analysis:** Aleksandar Tošič, Jernej Vičič

**Wrote the paper:** Aleksandar Tošič, Jernej Vičič

## Acknowledgment

The authors gratefully acknowledge the European Commission for funding the InnoRenew project (Grant Agreement #739574) under the Horizon2020 Widespread-Teaming program and the Republic of Slovenia for investment funding from the Republic of Slovenia and the European Union's European Regional Development Fund.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.joi.2021.101163>.

## References

- Alexander, J. C. (2009). *Remarks on the use of Benford's law*. Available at SSRN 1505147
- Ariel Xu, Q., & Chang, V. (2020). Co-authorship network and the correlation with academic performance. *Internet of Things*, 12, 100307. <https://doi.org/10.1016/j.iot.2020.100307> <https://www.sciencedirect.com/science/article/pii/S2542660520301396>
- Batagelj, V., & Maltseva, D. (2020). Temporal bibliographic networks. *Journal of Informetrics*, 14, 101006. <https://doi.org/10.1016/j.joi.2020.101006> <http://www.sciencedirect.com/science/article/pii/S1751157719301439>
- Benavent-Pérez, M., Gorraiz, J., Gumpenberger, C., & de Moya-Anegón, F. (2012). The different flavors of research collaboration: A case study of their influence on university excellence in four world regions. *Scientometrics*, 93, 41–58.
- Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 551–572.
- Berger, A., & Hill, T. P. (2011). A basic theory of Benford's Law. *Probability Surveys*, 8, 1–126. <https://doi.org/10.1214/11-PS175> <http://projecteuclid.org/euclid.ps/1311860830>
- Bernstein, G. S. (1984). Scientific rigor, scientific integrity: A comment on sommer and sommer. *American Psychologist*, 39, 1316. <https://doi.org/10.1037/0003-066X.39.11.1316.a>
- Bhattacharya, S., & Kumar, K. (2008). Forensic Accounting and Benford's Law [In the Spotlight]. *IEEE Signal Processing Magazine*, 25. <https://doi.org/10.1109/MSP.2007.914724>, 152–150; <http://ieeexplore.ieee.org/document/4472258/>
- Bohannon, J. (2013). Who's afraid of peer review? *Science*, 342, 60–65. <https://doi.org/10.1126/science.342.6154.60>
- Bordons, M., Gomez, I., Fernández, M., Zulueta, M., & Méndez, A. (1996). Local, domestic and international scientific collaboration in biomedical research. *Scientometrics*, 37, 279–295.
- Braun, T., & Glaenzel, W. (1996). International collaboration: Will it be keeping alive east European research? *Scientometrics*, 36, 247–254.
- Burke, J., & Kincanon, E. (1991). Benford's law and physical constants: The distribution of initial digits. *American Journal of Physics*, 59, 952. <https://doi.org/10.1119/1.16838>
- Cleary, R., & Thibodeau, J. C. (2005). Applying digital analysis using Benford's law to detect fraud: The dangers of type i errors. *AUDITING: A Journal of Practice & Theory*, 24, 77–81. <https://doi.org/10.2308/aud.2005.24.1.77>
- Cugmas, M., Ferligoj, A., & Kronegger, L. (2019). Scientific co-authorship networks. *Advances in Network Clustering and Blockmodeling*, 363–387.
- Curk, L. (2019). Implementation of the evaluation of researchers' bibliographies in Slovenia. *Procedia Computer Science*, 146, 72–83. <https://doi.org/10.1016/j.procs.2019.01.082>, 14th International Conference on Current Research Information Systems, CRIS2018, FAIRness of Research Information. <http://www.sciencedirect.com/science/article/pii/S1877050919300870>
- Drake, P. D., & Nigrini, M. J. (2000). Computer assisted analytical procedures using Benford's law. *Journal of Accounting Education*, 18, 127–146.
- Druica, E., Oancea, B., & Válsán, C. (2018). Benford's law and the limits of digit analysis. *International Journal of Accounting Information Systems*, 31, 75–82. <https://doi.org/10.1016/j.accinf.2018.09.004> <http://www.sciencedirect.com/science/article/pii/S146708951730101X>
- Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting*, 5, 17–34.
- Edwards, M. A., & Roy, S. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, 34, 51–61.
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of pubmed, scopus, web of science, and google scholar: Strengths and weaknesses. *The FASEB Journal*, 22, 338–342.
- Fanelli, D. (2013). Redefine misconduct as distorted reporting. *Nature*, 494, 149–149.

<sup>2</sup> HAL: <https://hal.archives-ouvertes.fr/>.

<sup>3</sup> Cristin: <https://www.cristin.no/english/>.

<sup>4</sup> MTMT database: <https://m2.mtmt.hu/gui2/>.

<sup>5</sup> Information system for research, experimental development and innovation research: <https://www.rvvi.cz/riv>.

- Ferligoj, A., Kronegger, L., Mali, F., Snijders, T. A., & Doreian, P. (2015). Scientific collaboration dynamics in a national scientific system. *Scientometrics*, *104*, 985–1012.
- Fewster, R. M. (2009). A simple explanation of Benford's law. *The American Statistician*, *63*, 26–32. <https://doi.org/10.1198/tast.2009.0005>
- Godecharle, S., Nemery, B., & Dierickx, K. (2014). Heterogeneity in european research integrity guidance: Relying on values or norms? *Journal of Empirical Research on Human Research Ethics*, *9*, 79–90.
- Golbeck, J. (2015). *Benford's law applies to online social networks..* CoRR abs/1504.0. arXiv:1504.04387
- Heneberg, P. (2016). From excessive journal self-cites to citation stacking: Analysis of journal self-citation kinetics in search for journals, which boost their scientometric indicators. *PLOS ONE*, *11*.
- Van der Heyden, M. A., van de Derks Ven, T., & Ophof, T. (2009). Fraud and misconduct in science: The stem cell seduction. *Netherlands Heart Journal*, *17*, 25–29.
- Hickman, M. J., & Rice, S. K. (2010). Digital analysis of crime statistics: Does crime conform to Benford's law? *Journal of Quantitative Criminology*, *26*, 333–349. <https://doi.org/10.1007/s10940-010-9094-6>
- Holub, H. W., Tappeiner, G., & Eberharter, V. (1991). The iron law of important articles. *Southern Economic Journal*, *317*–328.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*, 524–532.
- Karlovčec, M., & Mladenič, D. (2015). Interdisciplinarity of scientific fields and its evolution based on graph of project collaboration and co-authoring. *Scientometrics*, *102*, 433–454.
- Katz, J., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, *40*, 541–554.
- Keathley-Herring, H., Van Aken, E., Gonzalez-Aleu, F., Deschamps, F., Letens, G., & Orlandini, P. C. (2016). Assessing the maturity of a research area: Bibliometric review and proposed framework. *Scientometrics*, *109*, 927–951.
- Kenny, D. A. (2015). *Measuring model fit.* <http://davidakenny.net/cm/fit.htm>
- Kong, X., Shi, Y., Yu, S., Liu, J., & Xia, F. (2019). Academic social networks: Modeling, analysis, mining and applications. *Journal of Network and Computer Applications*, *132*, 86–103. <https://doi.org/10.1016/j.jnca.2019.01.029> <https://www.sciencedirect.com/science/article/pii/S1084804519300438>
- Korošec, A. (2014). Sicris, v3. *Organizacija Znanja*, *19*, 22–26. <http://www.dlib.si/?URN=URN:NBN:SI:DOC-KY3LXZHR>
- Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, *57*, 997–1004.
- Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, *2*, 317–325.
- Lužar, B., Levnajić, Z., Povh, J., & Perc, M. (2014). Community structure and the evolution of interdisciplinarity in Slovenia's scientific collaboration network. *PLOS ONE*, *9*, e94429.
- Macrina, F. L. (2014). *Scientific integrity: Text and cases in responsible conduct of research.* John Wiley & Sons.
- Mali, F., Pustovrh, T., Cugmas, M., & Ferligoj, A. (2018). The personal factors in scientific collaboration: Views held by slovenian researchers. *Corvinus Journal of Sociology and Social Policy*, *9*, 3–24.
- Martin, B. R. (2013). Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment. *Research Policy*, *42*, 1005–1014. <https://doi.org/10.1016/j.respol.2013.03.011> <http://www.sciencedirect.com/science/article/pii/S004873331300067X>
- Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, *29*, 31–40.
- Narin, F., & Whitlow, E. S. (1990). *Measurement of scientific cooperation and coauthorship in CEC-related areas of science.* Commission of the European Communities Directorate-General.
- Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, *4*, 39–40. <https://doi.org/10.2307/2369148>
- Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, *101*, 5200–5205.
- Nigrini, M. (2001). Digital analysis using benford's law: Tests and statistics for auditors. *EDPACS*, *28*, 1–2. <https://doi.org/10.1201/1079/43266.28.9.20010301/30389.4>
- Nigrini, M. J. (1996). A taxpayer compliance application of Benford's law. *The Journal of the American Taxation Association*, *18*, 72.
- Nigrini, M. J. (2012). *Benford's law: Applications for forensic accounting, auditing, and fraud detection.* John Wiley & Sons. <https://www.wiley.com/en-us/Benford%27s+Law%3A+Applications+for+Forensic+Accounting%2C+Auditing%2C+and+Fraud+Detection-p-9781118152850>
- Nigrini, M. J. (2017). Audit sampling using benford's law: A review of the literature with some new perspectives. *Journal of Emerging Technologies in Accounting*, *14*, 29–46.
- Ortega, J. L. (2014). Influence of co-authorship networks in the research impact: Ego network analyses from microsoft academic search. *Journal of Informetrics*, *8*, 728–737. <https://doi.org/10.1016/j.joi.2014.07.001> <https://www.sciencedirect.com/science/article/pii/S1751157714000613>
- Pelacho, M., Ruiz, G., Sanz, F., Taracón, A., & Clemente-Gallardo, J. (2021). Analysis of the evolution and collaboration networks of citizen science scientific publications. *Scientometrics*, *126*, 225–257.
- Peterson, M. (2007). Best practices for ensuring scientific integrity and preventing misconduct. *Essays on Ethical Standards*, *21*.
- Pisanski, T., Pisanski, M., & Pisanski, J. (2020). A novel method for determining research groups from co-authorship network and scientific fields of authors. *Informatica*, *44*.
- Ranstam, J., Buysse, M., George, S. L., Evans, S., Geller, N. L., Scherrer, B., et al. (2000). Fraud in medical research: An international survey of biostatisticians. *Contemporary Clinical Trials*, *21*, 415–427.
- Rodela, R. (2016). On the use of databases about research performance: Comments on karlovčec and mladenič (2015) and others using the sicris database. *Scientometrics*, *109*, 2151–2157.
- Sambridge, M., Tkalčić, H., & Jackson, A. (2010). Benford's law in the natural sciences. *Geophysical Research Letters*, *37*, 1–5.
- Seljak, T., & Bošnjak, A. (2006). Researchers' bibliographies in cobiss. si. *Information Services & Use*, *26*, 303–308.
- Shibayama, S., & Baba, Y. (2015). Impact-oriented science policies and scientific publication practices: The case of life sciences in Japan. *Research Policy*, *44*, 936–950.
- Singleton, T. W. (2011). IT audit basics: Understanding and applying Benford's law. *Isaca Journal*, *3*, 6.
- Sponholz, G. (2000). Teaching scientific integrity and research ethics. *Forensic Science International*, *113*, 511–514.
- Zdravevski, E., Lameski, P., Trajkovik, V., Chorbev, I., Goleva, R., Pombo, N., et al. (2019). Automation in systematic, scoping and rapid reviews by an nlp toolkit: A case study in enhanced living environments. *Enhanced living environments*, 1–18.
- Zhang, J. (2020). *Testing case number of coronavirus disease 2019 in china with Newcomb-Benford law.* arXiv:2002.05695