Analysis of Cited References in Russian Publications on Web of Science

Dalibor Fiala[1],*, Daria Maltseva[2]

[1]Department of Computer Science and Engineering, Faculty of Applied Sciences, University of West Bohemia, Plzeň, CZECH REPUBLIC.
[2]International Laboratory for Applied Network Research, National Research University Higher School of Economics, Moscow, RUSSIA.

ABSTRACT

In this article we analyze the cited references in 1.38 million papers by Russian (co-)authors indexed in the Web of Science database until May 2022. Similarly, to the established processes in the so-called Reference Publication Year Spectroscopy (RPYS), we study the distribution of the references across the cited years and seek to identify the peak years with the publications that attracted the most attention of Russian scholars. In this way, the historical roots of Russian science may be traced and we take a closer look at these most influential works. In addition, we investigate the evolution of the mean age of references and of their average number per paper over time and inspect the most frequently cited sources. The results show that the average number of references in Russian papers has been steadily increasing, but the mean age of references has been declining in the most recent years. Also, the foundations of Russian science seem to be physics of particles and electrochemistry and have recently become based more internationally than in the past. This study is the first of its kind and may help better understand the character of Russian research.

Keywords: References, Web of Science, Russia, RPYS, Roots, Sources.

INTRODUCTION

After the collapse of the Soviet Union (USSR) in December 1991, Russia became the largest successor state of all the former Soviet republics. As a result, with the existence of an independent Russian Federation an independent Russian research system was also formed, which inherited many positive aspects of the huge and successful research infrastructure and capacities of the former USSR but also some structural weaknesses and a lack of funding, primarily in the transformation years of the 1990s. But in the following two first decades of the 21st century, Russian research has been trying to overcome the transformation obstacles and position itself firmly on the world stage again. Therefore, in this context, it appears to be important and interesting to study the various features of Russian scholarly literature, which may help us better understand the nature of modern Russian science.

One of such interesting aspects is the analysis of the cited references found in the scientific publications written by Russian authors or co-authors. Recently, this kind of analysis has often been linked to the term "Reference Publication Year Spectroscopy (RPYS)", which was coined in the study of cited references in research papers carried out by Marx et al.[1] and further elaborated in other analyses conducted by Barth et al.,[2] Bornmann and Marx,[3] Leydesdorff et al.,[4] Thor et al.,[5,6] Marx et al.,[7] or Thor et al.[8] One of the main goals of RPYS is to identify the so-called historical roots (origins), milestones, or landmark papers in a certain scientific discipline, journal, or region. To this end, Bornmann et al.[9] applied the RPYS method to the publications by a single researcher (Eugene Garfield), Haunschild[10] to a specific journal (*Information*), Ballandonne and Cersosimo[11] to a group of journals in management, economics, and finance, Comins and Hussey[12] to global positioning system (GPS) papers, Comins and Leydesdorff[13] to biomedical research, and Fiala and Bornmann[14] to Eastern European computer science research. Furthermore, by means of RPYS, Elango et al.[15] detected the historical roots of tribology research, Hou[16] of citation analysis, Li et al.[17] of various research fields in China, Khasseh and Mokhtarpour[18] of the field of knowledge management, Millán et al.[19] of social psychology in Brazil, Wray and Bornmann[20] of philosophy of science, and Yeung and Wong[21] of visual analogue scale in psychology.

However, none of the studies mentioned above dealt explicitly with an investigation into the characteristics of references cited in the scientific literature (co-)produced by Russian authors and no attempts were made to examine the historical origins of contemporary Russian science. Therefore, in this paper we aim to fill this gap and present an analysis of almost 33 million cited

references found in 1.38 million Russian publications indexed in the prestigious Web of Science (WoS) database as of May 2022 without any prejudice as to the nature of these publications or references. The main research questions are: a) How have the mean number of references per paper and the mean age of references in a paper evolved over time? b) What are the most frequently cited references and thus possibly the historical roots of Russian science? c) What are the most cited sources (journals) of Russian scholars and have they changed in the last three decades? We will try to find answers to these questions in the following sections of this article.

METHODOLOGY

For our analysis, we used a data acquisition approach similar to that in Fiala and Tutoky[22] and chose exactly the same data set as in Fiala and Maltseva,[23] where the process of acquiring this data collection is also described in greater detail. Here we reiterate only the most important facts: Between 27 April 2022 and 10 May 2022 the data on Russian publications were retrieved from the web interface of the Web of Science database in the form of plain-text files. The search query submitted to WoS was simply "CU=(Russia)" and we did not restrict the search to any time limits, document types, scientific categories, or any other characteristics. All citation indices being part of the standard "Core Collection" were included in the search and, finally, a total of 1.384 million bibliographic records were retrieved, with almost no papers published before 1992. This data set thus formed the basis for the investigation presented in this article and because Fiala and Maltseva,[23] whose relationship to this analysis is analogous to that of Fiala and Willett[24] to Fiala and Bornmann,[14] did not examine the cited references in Russian papers, their analysis was carried out only in this present study.

The exact total number of cited references we identified in the data was 32,747,382. (Let us note that, in the context of this article, references are represented by out-links in a citation graph of research papers, as opposed to citations which are represented by in-links.) The references usually include the surname and the first and middle name initials of the first author, publication year, abbreviated journal title, volume, and start page (see the tables in the next sections), which are the typical features sufficient to determine the cited papers. In addition, a large portion of the references (around 60%) also included a DOI (Digital Object Identifier), which allows for unambiguous identification of the cited work. In this way, 1.66 million references (or 8.5% of the references with a DOI) were clearly identified to be part of the core data set. In order to explore the almost 33 million cited references, however, we did not use the software developed specifically for RPYS by its inventors in Thor *et al*.[5] and called CRExplorer 1 https://andreas-thor.github.io/CRExplorer/ or any other specialized software for this purpose like in McLevey and

McIlroy-Young,[25] but, for efficiency reasons related to the much larger amount of analyzed data than in previous similar studies, except perhaps Thor *et al*.,[26] we took advantage of an approach of our own consisting in the combined use of a general relational database management system and a spreadsheet calculator, which enabled us to scale up the RPYS technique to tens of millions of data records quite comfortably.

RESULTS AND DISCUSSION

Age of Cited References

Figure 1 shows the distribution of the age of cited references in Russian papers in years, split into 21 "buckets" from age 0 (i.e. published in the same year as the citing article) to 20+ (i.e. published 20 or more years earlier than the citing paper. We can clearly see in the chart that this last group of cited references is by far also the largest, containing almost eight million references, which is nearly 25% of all references in our data set. (This largest category is further broken down into subcategories in an inset pie chart in Figure 1, showing that about 44% of the 20-or-more-year-old references are younger than 30 years.) Russian science thus seems to still heavily build upon older research and as for the individual years, the most common age of cited literature is two years with around 7.5% of all references. It is also interesting to note from Figure 1 that close to 2% of references are of age 0 (i.e. the citing and cited paper are published in the same year) and that, after reaching the peak age of 2, the number of references becomes smaller with each additional year of age.

Top Cited References

As far as the absolute and relative numbers of references are concerned, Table 1 lists the 20 most cited publications in Russian science. The top three cited references with over 4,000 citations and more than a 0.3% share in citations each are Perdew's 1996 article "Generalized Gradient Approximation Made Simple" published in *PHYS REV LETT* (*Physical Review Letters*) on the calculation of electron spin densities energy, Sheldrick's 2008 paper "A short history of SHELX" in *ACTA CRYSTALLOGR A* (*Acta Crystallographica Section A*) on the development of classic computer software for the refinement of small-molecule structures, and Agostinelli's 2003 work "Geant4: a simulation toolkit" in NUCL INSTRUM METH A (*Nuclear Instruments & Methods in Physics Research Section A*) describing a software toolkit for the simulation of the passage of particles through matter. It is interesting to note that Agostinelli's paper is the only cited reference in the top 20 presented in Table 1 which is at the same time an article included in the core data set under study and thus partly a product of Russian research itself. Therefore, this reference (albeit being a multi-author paper written in a worldwide collaboration) may be considered as a self-citation of Russian science whereas all the other references shown in Table 1 refer to publications outside of the "Russian bubble".
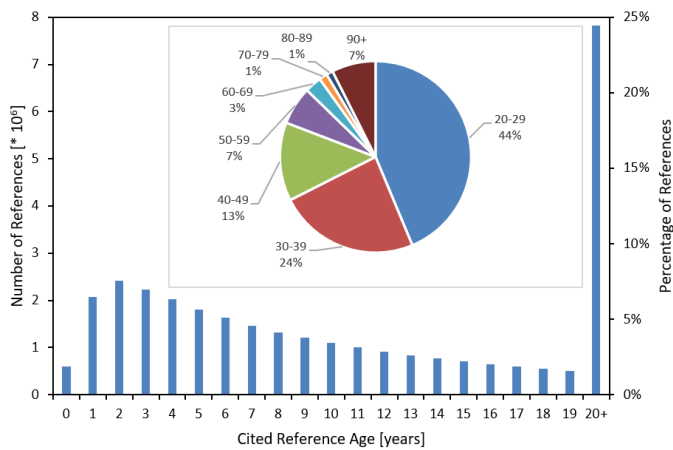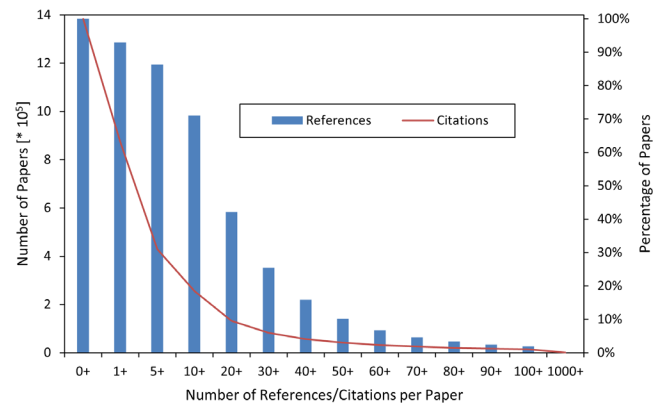
Figure 1: Distribution of the age (in years) of cited references (20+ category broken down in the inset pie chart).

Figure 2: Cumulative distribution of the number of references and citations per paper.

Distribution of References and Citations

Figure 2 shows the (cumulative) distribution of the number of references (bars) in Russian papers and, for comparison, also of citations (line) of those papers in terms of the number of papers exceeding a certain threshold value of references or citations. (The citation counts were obtained from the "Times Cited" data in the WoS core data set.) Thus, obviously, 100% of Russian papers (totalling almost 1.4 million) have zero or more references and the same holds for citations. But only about 93% of papers include at least one reference and 63% of papers have at least one citation (i.e. more than one third of papers remained uncited, which is actually less than the generally expected share of uncited papers being more than a half) and a further drop in the number of papers with five or more references or with five citations at least is more pronounced for citations: 86% versus 30%, respectively. Both series decline until negligible shares of papers having 1,000 or more references or 1,000 or more citations with these shares being 0.02% for references and 0.01% for citations using the threshold value of 100 or more. Interestingly, the sharpest decline in the number of references from 71% to 42% of papers occurs between the 10+ and 20+ "buckets", indicating that the relative majority of Russian papers include between 10 and 20 (excluding) cited references. (For citations, the biggest drop happens already between the 1+ and 5+ "buckets", meaning that Russian papers mostly receive from one to four citations).

Evolution of Mean Number of References over Time

In Figure 3 we can see how the average number of cited references in Russian papers evolved over time (vertical bars) and observe its steady increase from 14.59 in 1992 to 47.25 in 2022. Thus, a Russian article published in 2022 cited three-times more references on average than that appearing in 1992. In fact, the mean value throughout the whole period 1992-2022 is 25.48 references per paper (thin dotted line) and was first exceeded only in 2015, which means that the trend towards citing more references accelerated at the end of that time range. A little

different picture is produced when we have a look at the evolution of the mean age of cited references in Russian papers (thick solid line) and see it grow from 11.72 years in 1992 to its peak of 15.47 years in 2014 and then rapidly decline to 13.41 years in 2022. And although we did not have complete data for 2022 because of the May cut-off date, there seems to be a clear tendency to cite more recent papers (references) in publications appearing after 2014 with the mean age of cited references being below the overall average of 14.46 years (thin dashed line) in both 2021 and 2022. Yet still, the mean age of cited references increased by about 14% between 1992 and 2022.

Peak Years of Cited References

As can be observed in the following Figure 4, the number of references to papers published in individual years since 1990 grew rapidly (line with blank circles), especially after 1945, and reached its peak with almost a million references (exactly 972,776) to the literature appearing in 2013. (Even though there are references to papers even older than from 1900 in our data set, we chose this year as the starting point for the visualization in Figure 4 because those references are relatively few.) As far as the references to more recent papers are concerned, their numbers fall sharply after 2015, which is not unexpected, given the shortening time window for such papers to collect citations and the diminishing set of articles potentially referencing (citing) them. However, the absolute numbers of references to the individual cited years are perhaps less important than the difference between this value and the 5-year (rolling) median of the numbers of references to the cited year, the two preceding years, and the two succeeding years. (If the current cited year is t, the 5-year median is calculated based on the numbers of references to year's $t-2$, $t-1$, t, $t+1$, and $t+2$.) These differences are depicted in Figure 4 as a line with filled-in diamonds and for significantly positive deviations the corresponding cited years are shown. In these years, important scientific achievements were published which made the papers more frequently cited than those from the surrounding years. These publications had obviously a great impact on Russian

Table 1: Top 20 cited references. Only Agostinelli's 2003 paper is in the core data set with 14,888 citations. DOI of Lowry's 1951 article was missing in the original data.

Cited reference	Count	%
Perdew J.P., 1996, PHYS REV LETT, V77, P3865. Doi:10.1103/PHYSREVLETT.77.3865	4,989	0.36%
Sheldrick G.M., 2008, ACTA CRYSTALLOGR A, V64, P112. Doi:10.1107/S0108767307043930	4,898	0.35%
Agostinelli S., 2003, NUCL INSTRUM METH A, V506, P250. Doi:10.1016/S0168-9002(03)01368-8	4,290	0.31%
Laemmli U.K., 1970, NATURE, V227, P680. doi:10.1038/227680A0	3,998	0.29%
Shannon R.D., 1976, ACTA CRYSTALLOGR A, V32, P751. Doi:10.1107/S0567739476001551	3,325	0.24%
Becke A.D., 1993, J CHEM PHYS, V98, P5648. doi:10.1063/1.464913	3,252	0.23%
Lee C.T., 1988, PHYS REV B, V37, P785. Doi:10.1103/PHYSREVB.37.785	2,793	0.20%
Sheldrick G.M., 2015, ACTA CRYSTALLOGR C, V71, P3. Doi:10.1107/S2053229614024218	2,741	0.20%
Schmidt M.W., 1993, J COMPUT CHEM, V14, P1347. Doi:10.1002/JCC.540141112	2,721	0.20%
Bradford M.M., 1976, ANAL BIOCHEM, V72, P248. doi:10.1016/0003-2697(76)90527-3	2,433	0.18%
Lowry O.H., 1951, J BIOL CHEM, V193, P265. doi:10.1016/S0021-9258(19)52451-6	2,354	0.17%
Sjostrand T., 2006, J HIGH ENERGY PHYS. doi:10.1088/1126-6708/2006/05/026	2,267	0.16%
Kresse G., 1996, PHYS REV B, V54, P11169. Doi:10.1103/PHYSREVB.54.11169	2,254	0.16%
Dolomanov O.V., 2009, J APPL CRYSTALLOGR, V42, P339. Doi:10.1107/S0021889808042726	1,784	0.13%
Lange D.J., 2001, NUCL INSTRUM METH A, V462, P152. Doi:10.1016/S0168-9002(01)00089-4	1,618	0.12%
Kresse G., 1999, PHYS REV B, V59, P1758. Doi:10.1103/PHYSREVB.59.1758	1,610	0.12%
Monkhorst H.J., 1976, PHYS REV B, V13, P5188. Doi:10.1103/PHYSREVB.13.5188	1,542	0.11%
Allen F.H., 1987, J CHEM SOC PERK T 2, PS1. Doi:10.1039/P298700000S1	1,477	0.11%
Blochl P.E., 1994, PHYS REV B, V50, P17953. Doi:10.1103/PHYSREVB.50.17953	1,421	0.10%
Cacciari M., 2008, J HIGH ENERGY PHYS. doi:10.1088/1126-6708/2008/04/063	1,364	0.10%

scholars and particularly those appearing prior to the 1960s might be considered as the historical roots of contemporary Russian science.
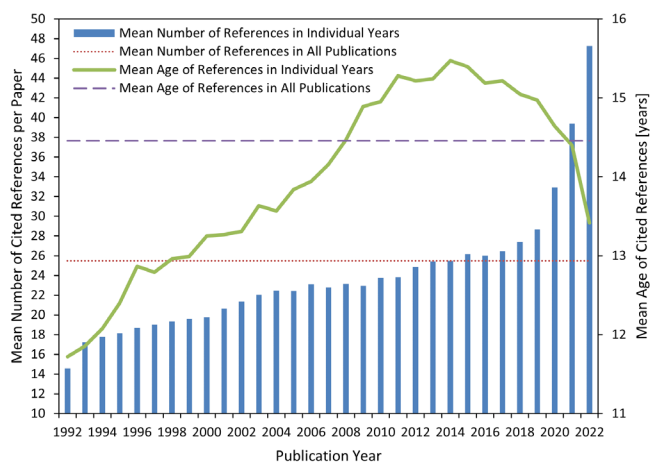
Top Cited References in Peak Years

Regarding the peak years highlighted in Figure 4, they are mainly driven by several highly-cited references which we present in Table 2. Again, these are typically given by the surname and given names' initials of the first author, followed by publication year, journal or book title in abbreviated form, and possibly also by volume and first page number. Most references have a DOI too, except some books, for which a DOI seems not to exist. As we can see both in Figure 4 and Table 2, the peak year 1928 is not particularly pronounced and the three most cited references are Fowler's article in *P R SOC LOND A-CONTA* (*Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*), Gamow's paper in *Z PHYS* (*Zeitschrift für Physik* or *Journal for Physics*), and Dirac's publication also in *P R SOC LOND A-CONTA*, which have over 1.0% share in total cited references to 1928 literature each. As for the topics of these articles, "Electron emission in intense electric fields" by Fowler, "Zur Quantentheorie des Atomkernes (On the quantum theory of atomic nucleus)" by Gamow, and "The quantum theory of the electron" by Dirac all deal with subatomic particles such as electrons and a mathematical description of their behaviour.

Table 2: Top cited references in peak years indicated in Figure 4.

Cited reference	Count	%
1928		
Fowler R.H., 1928, P R SOC LOND A-CONTA, V119, P173. doi:10.1098/RSPA.1928.0091	185	1.69%
Gamow G., 1928, Z PHYS, V51, P204. doi:10.1007/BF01343196	128	1.17%
Dirac P.A.M., 1928, P R SOC LOND A-CONTA, V117, P610. doi:10.1098/RSPA.1928.0023	113	1.03%
Nyquist H., 1928, PHYS REV, V32, P110. doi:10.1103/PHYSREV.32.110	103	0.94%
1932		
Wigner E., 1932, PHYS REV, V40, P0749. doi:10.1103/PHYSREV.40.749	471	4.13%
Zener C., 1932, P R SOC LOND A-CONTA, V137, P696. doi:10.1098/RSPA.1932.0165	192	1.68%
Lamb H., 1932, HYDRODYNAMICS.	158	1.38%
von Neumann J., 1932, MATH GRUNDLAGEN QUAN.	137	1.20%
1937		
Fisher R.A., 1937, ANN EUGENIC, V7, P355. doi:10.1111/J.1469-1809.1937.TB02153.X	158	1.00%
Jahn H.A., 1937, PROC R SOC LON SER-A, V161, P220. doi:10.1098/RSPA.1937.0142	125	0.79%
Kolmogorov A., 1937, AKAD NAUK SSSR IZV M, V1, P355.	104	0.66%
1951		
Lowry O.H., 1951, J BIOL CHEM, V193, P265.	2,354	7.55%
Schwinger J., 1951, PHYS REV, V82, P664. doi:10.1103/PHYSREV.82.664	584	1.87%
Zener C., 1951, PHYS REV, V82, P403. doi:10.1103/PHYSREV.82.403	344	1.10%
1965		
Kohn W, 1965, PHYS REV, V140, P1133. doi:10.1103/PHYSREV.140.A1133	1,310	1.12%
Abramowitz M., 1965, HDB MATH FUNCTIONS.	491	0.42%
2000		
Valiev R.Z., 2000, PROG MATER SCI, V45, P103. doi:10.1016/S0079-6425(99)00007-9	971	0.12%
Berman H.M., 2000, NUCLEIC ACIDS RES, V28, P235. doi:10.1093/NAR/28.1.235	580	0.07%
2013		
Frisch M.J., 2013, GAUSSIAN 09 REVISION.	733	0.08%
Reimer P.J., 2013, RADIOCARBON, V55, P1869. doi:10.2458/AZU_JS_RC.55.16947	517	0.05%
2015		
Sheldrick G.M., 2015, ACTA CRYSTALLOGR C, V71, P3. doi:10.1107/S2053229614024218	2,741	0.28%
Sjostrand T., 2015, COMPUT PHYS COMMUN, V191, P159. doi:10.1016/J.CPC.2015.01.024	629	0.07%

Figure 3: Mean number of cited references in individual publication years (vertical bars) compared to the mean in all publication years (thin dotted line) and mean age of cited references in individual publication years (thick solid line) compared to the mean in all publication years (thin dashed line).
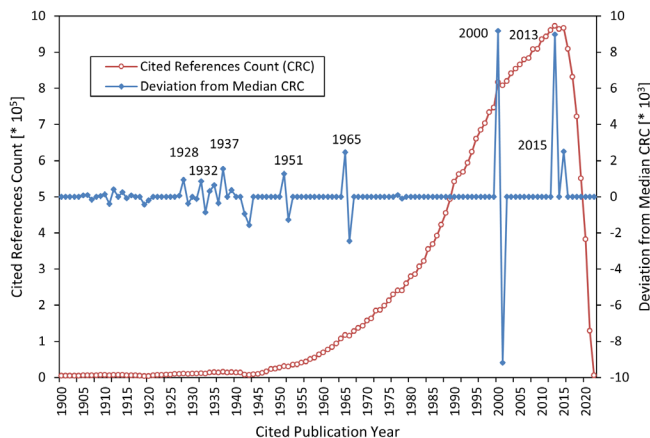
Figure 4: Number of cited references from individual cited publication years (line with blank circles and left-hand Y-axis) and its deviation from a 5-year (rolling) median thereof (line with filled-in diamonds and right-hand Y-axis).

In contrast, references to publications from 1932 (which is itself a smaller peak than 1928 in Figure 4) have a clear winner in Wigner's article in *PHYS REV* (*Physical Review*) with a 4.13% share and 471 citations, while, after a gap, the second most cited reference is Zener's paper in *P R SOC LOND A-CONTA* with a share in citations of just above 1.5%. The first article ("On the quantum correction for thermodynamic equilibrium") is concerned with the thermodynamics of quantum mechanical systems and the second one ("Non-adiabatic crossing of energy levels") deals with the calculation of the transition probability of electrons between energy levels. The other two most frequently cited works published in 1932 are Lamb's book on hydrodynamics, which is actually already the sixth edition of the original 1879 text and in further editions still in use today as a classic handbook for fluid dynamicists, and von Neumann's "Mathematische Grundlagen der Quantenmechanik (Mathematical Foundations of Quantum

Mechanics)", in which the principles of the quantum theory were explained.

As far as the references to 1937 publications are concerned, there is no outstanding work at all with each of the most frequently cited pieces of literature having a 1.0% share in citations or less. The most cited articles are Fisher's "The wave of advance of advantageous genes" on the calculations of the diffusion of mutated genes, which appeared in *ANN EUGENIC* (*Annals of Eugenics*), Jahn's "Stability of polyatomic molecules in degenerate electronic states – I: Orbital degeneracy" on the stability of nuclear configurations in *PROC R SOC LON SER-A* (*Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*), and Kolmogorov's "Zur Statistik der Kristallisationsvorgänge in Metallen (On the statistical theory of crystallization of metals)" with a probabilistic solution to the problem of metal crystallization in space and time published in *AKAD NAUK SSSR IZV M* (*Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya* or *Bulletin of the Academy of Sciences of the USSR: Mathematics Series*).

A completely different situation arises with respect to the references citing 1951 literature with the number one paper attracting over 7.5% of citations to publications from that year alone. This top paper is Lowry's article "Protein Measurement with the Folin Phenol Reagent" published in J BIOL CHEM (*Journal of Biological Chemistry*). The other two most frequently cited papers following after a big gap are Schwinger's "On Gauge Invariance and Vacuum Polarization" and Zener's "Interaction between the d-Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure", which both appeared in *PHYS REV* and received between 1.0% and 2.0% of citations. Lowry's article, with its more than 2,300 citations from Russian publications as of May 2022, gives simple and sensitive directions for the measurement of proteins in solution or after precipitation with acid or other agents, while the other two papers are concerned with the behaviour of subatomic particles in electromagnetic fields and the relation between electrical conductivity and ferromagnetism, respectively.

Kohn's article "Self-Consistent Equations Including Exchange and Correlation Effects" in *PHYS REV* on the interaction of electrons in electronic systems at finite temperatures and in magnetic fields and Abramowitz's book "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables", which is a classic source of reference in applied mathematics even nowadays, are the most cited works from 1965 and they are also the most recent cited references that we might still consider as the historical roots of modern Russian science. And because the following peak years (2000, 2013, and 2015) are too recent for the references to be called historical roots, we will only enumerate the most referenced publications having relatively low citation shares each: Valiev's paper "Bulk nanostructured materials from severe plastic deformation" in *PROG MATER SCI* (*Progress in Materials*

Table 3: Top 20 cited references belonging to the 10% most frequently cited papers in each year throughout the period 1992-2022, i.e. from 31 citing years (2022 being incomplete). DOIs of Lowry's 1951 article, Folch's 1957 article, and Bligh's 1959 article were missing in the original data.

Cited Reference	Count
Laemmli U.K., 1970, NATURE, V227, P680. doi:10.1038/227680A0	3,983
Shannon R.D., 1976, ACTA CRYSTALLOGR A, V32, P751. doi:10.1107/S0567739476001551	3,314
Bradford M.M., 1976, ANAL BIOCHEM, V72, P248. doi:10.1016/0003-2697(76)90527-3	2,427
Lowry O.H., 1951, J BIOL CHEM, V193, P265. doi:10.1016/S0021-9258(19)52451-6	2,334
Allen F.H., 1987, J CHEM SOC PERK T 2, PS1. doi:10.1039/P298700000S1	1,447
Becke A.D., 1988, PHYS REV A, V38, P3098. doi:10.1103/PHYSREVA.38.3098	1,322
Kohn W., 1965, PHYS REV, V140, P1133. doi:10.1103/PHYSREV.140.A1133	1,307
Bondi A., 1964, J PHYS CHEM-US, V68, P441. doi:10.1021/J100785A001	957
Altarelli G., 1977, NUCL PHYS B, V126, P298. doi:10.1016/0550-3213(77)90384-4	693
Folch J., 1957, J BIOL CHEM, V226, P497. doi:10.1016/S0021-9258(18)64849-5	682
Dewar M.J.S., 1985, J AM CHEM SOC, V107, P3902. doi:10.1021/JA00299A024	642
Murashige T., 1962, PHYSIOL PLANTARUM, V15, P473. doi:10.1111/J.1399-3054.1962.TB08052.X	639
Vosko S.H., 1980, CAN J PHYS, V58, P1200. doi:10.1139/P80-159	615
Fano U., 1961, PHYS REV, V124, P1866. doi:10.1103/PHYSREV.124.1866	573
Dicke R.H., 1954, PHYS REV, V93, P99. doi:10.1103/PHYSREV.93.99	536
Bligh E.G., 1959, CAN J BIOCHEM PHYS, V37, P911. doi:10.1139/o59-099	529
Dubois M., 1956, ANAL CHEM, V28, P350. doi:10.1021/AC60111A017	489
Kosterlitz J.M., 1973, J PHYS C SOLID STATE, V6, P1181. doi:10.1088/0022-3719/6/7/010	471
Brown I.D., 1985, ACTA CRYSTALLOGR B, V41, P244. doi:10.1107/S0108768185002063	443
Reynolds E.S., 1963, J CELL BIOL, V17, P208. doi:10.1083/JCB.17.1.208	430

Table 4: Top 20 cited sources in the whole period (1992-2022) and in several sub-periods with unique sources highlighted.

1992-2022	Before 2000	2000-2004
PHYS REV LETT	PHYS REV LETT	PHYS REV LETT
PHYS REV B	PHYS REV B	PHYS REV B
PHYS REV D	PHYS LETT B	PHYS LETT B
NATURE	J CHEM PHYS	J CHEM PHYS
ASTROPHYS J	J AM CHEM SOC	PHYS REV D
J AM CHEM SOC	NATURE	NATURE
J CHEM PHYS	DOKL AKAD NAUK SSSR+	J AM CHEM SOC
SCIENCE	PHYS REV D	ASTROPHYS J
PHYS LETT B	NUCL PHYS B	SCIENCE
P NATL ACAD SCI USA	ASTROPHYS J	NUCL PHYS B
APPL PHYS LETT	J BIOL CHEM	PHYS REV A
J APPL PHYS	PHYS REV A	APPL PHYS LETT
PHYS REV A	J APPL PHYS	J APPL PHYS
ASTRON ASTROPHYS	J PHYS CHEM-US	J BIOL CHEM
J BIOL CHEM	P NATL ACAD SCI USA	P NATL ACAD SCI USA
NUCL PHYS B	APPL PHYS LETT	PHYS REV
PHYS REV C	SCIENCE	ASTRON ASTROPHYS
MON NOT R ASTRON SOC	PHYS REV	PHYS REV C
PHYS REV	ZH OBSHCH KHIM+	J PHYS CHEM-US
PHYS REV E	BIOCHEMISTRY-US	NUCL INSTRUM METH A
2005-2009	**2010-2014**	**Since 2015**
PHYS REV LETT	PHYS REV LETT	PHYS REV LETT
PHYS REV B	PHYS REV B	PHYS REV B
PHYS REV D	PHYS REV D	PHYS REV D
NATURE	NATURE	NATURE
PHYS LETT B	ASTROPHYS J	ASTROPHYS J
ASTROPHYS J	J AM CHEM SOC	SCIENCE
J CHEM PHYS	SCIENCE	J AM CHEM SOC
J AM CHEM SOC	PHYS LETT B	P NATL ACAD SCI USA
SCIENCE	J CHEM PHYS	J CHEM PHYS
APPL PHYS LETT	APPL PHYS LETT	APPL PHYS LETT
PHYS REV A	ASTRON ASTROPHYS	J APPL PHYS
J APPL PHYS	P NATL ACAD SCI USA	PLOS ONE
P NATL ACAD SCI USA	J APPL PHYS	ASTRON ASTROPHYS
J BIOL CHEM	PHYS REV A	MON NOT R ASTRON SOC
NUCL PHYS B	J BIOL CHEM	PHYS LETT B
ASTRON ASTROPHYS	PHYS REV C	SCI REP-UK
PHYS REV C	NUCL PHYS B	PHYS REV A
NUCL INSTRUM METH A	MON NOT R ASTRON SOC	J HIGH ENERGY PHYS
PHYS REV E	NUCL INSTRUM METH A	J BIOL CHEM

Science), Frisch's software "Gaussian 09 Revision", and Sheldrick's article "Crystal structure refinement with SHELXL" in *ACTA CRYSTALLOGR C* (*Acta Crystallographica Section C: Structural Chemistry*).

Top Consistently Cited References

Table 3 is yet another presentation of the most frequently cited references in Russian publications, but unlike Table 1 with the top 20 references that could be cited very unevenly throughout the whole period 1992-2022, the references in Table 3 were cited quite consistently in that time range and appeared among the 10% of the most cited references in each year of the investigated period, i.e. in 31 citing years from 1992 to 2022 (with the last year being incomplete, of course). Thus, all the papers mentioned in Table 3 were published prior to 1992 and have been consistently highly cited in Russian literature since then. Therefore, none of them can be considered as a one-year excess which may be true for some of the cited references in Table 1. On the other hand, Laemmli's 1970 article in *Nature* "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4" on an improved method of gel electrophoresis, appearing in the first place in Table 3 is also very well ranked in Table 1, which means that it has always been very highly cited by Russian scientists.

The same holds for Shannon's 1976 article "Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides" published in ACTA CRYSTALLOGR A (*Acta Crystallographica Section A*) on the length of bonds with localized and delocalized electrons in atoms, Bradford's paper from the same year entitled "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding" with a description of a new protein determination method, which appeared in *ANAL BIOCHEM* (*Analytical Biochemistry*), Lowry's 1951 publication previously discussed and shown additionally also in Table 2, Allen's 1987 paper "Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds" on the average length of bonds involving select chemical elements in *J CHEM SOC PERK T 2* (*Journal of the Chemical Society, Perkin Transactions 2*), and Becke's 1988 article "Density-functional exchange-energy approximation with correct asymptotic behavior" from *PHYS REV A* (*Physical Review A*) presenting a gradient-corrected one-parameter functional for the computation of exchange-energy densities in various atomic and molecular systems. Finally, the last work with over 1,300 citations in Table 3 is Kohn's 1965 article which was already discussed in the context of Table 2, but the others have citation counts lying below this threshold and will thus not be further described.

Top Cited Sources

Table 4 presents the most frequently cited sources (journals) in publications appearing in the whole period under study (i.e. roughly 1992-2022) and in several of its sub-periods (before 2000, 2000-2004, 2005-2009, 2010-2014, and since 2015) with unique sources in the top 20 of each sub-period being highlighted in boldface. Thus, the consistently cited journals across the whole time range seem to be well-known specialty physics or chemistry journals like *PHYS REV LETT* (*Physical Review Letters*), *PHYS REV B* (*Physical Review B: Condensed Matter and Materials Physics*), *ASTROPHYS J* (*Astrophysical Journal*), and *J AM CHEM SOC* (*Journal of the American Chemical Society*), but also prestigious multidisciplinary sources such as *NATURE* (*Nature*), *SCIENCE* (*Science*), or *P NATL ACAD SCI USA* (*Proceedings of the National Academy of Sciences of the United States of America*). However, in the oldest sub-period (prior to 2000) there are three outlets that appear uniquely in that time range and are no longer among the top cited sources in the later sub-periods: *DOKL AKAD NAUK SSSR+* (*Doklady Akademii Nauk SSSR* or *Proceedings of the USSR Academy of Sciences*), *ZH OBSHCH KHIM+* (*Zhurnal obshcheĭ khimii* or *Journal of General Chemistry*), and *BIOCHEMISTRY-US* (*Biochemistry*). Similarly, since 2015 there have emerged four sources that had not been among the top cited journals until then: *PLOS ONE* (*PLoS One*), *SCI REP-UK* (*Scientific Reports*), *J HIGH ENERGY PHYS* (*Journal of High Energy Physics*), and *ANGEW CHEM INT EDIT* (*Angewandte Chemie-International Edition* or *Applied Chemistry-International Edition*). We may thus draw the conclusion that the focus of Russian scientists shifted from citing domestic journals to some extent in the 1990s to referencing predominantly foreign research outlets in the late 2010s.

CONCLUSION

In this article, we presented a study of the cited references in 1.38 million research papers published by scholars (authors or co-authors) affiliated with Russian institutions and indexed in the Web of Science database until May 2022. We considered all document types of these papers exactly like in Fiala and Maltseva,[23] but the vast majority of the papers were journal articles and conference proceedings papers. No explicit publication year range was set for the papers, but almost all of them appeared between 1992 and 2022. Based on this large data set, we made the following contributions to the bibliometric analysis of Russian science by an approach analogous to the so-called Reference Publication Year Spectroscopy (RPYS):

- We extracted more than 32.7 million cited references from the "Russian" publications and imported them into a relational database for further analysis.

- We inspected the distribution of the mean number of references per paper and of the average age of cited references in the individual years from 1992 to 2022.

- We identified the peak years by citation intensity and determined the key publications for Russian science, thus tracing the historical origins of Russian research, and examined the most frequently cited sources.

The main findings of our analysis are the following:

- The mean number of cited references in Russian papers has been constantly growing over time (14.59 in 1992 and 47.25 in 2022), but the average age (in years) of the references has been declining in the most recent years (from 15.47 in 2014 to 13.41 in 2022), meaning that Russian scientists now cite more often and more recent research than in the past.

- The most cited references in Russian publications are papers from the fields of physics of particles and electrochemistry, which thus appear to be the most visible pillars for modern Russian science.

- Finally, the top cited sources (journals) in Russian works have become more international after 2000 and the previously impactful domestic outlets have lowered their impact on Russian scholars in recent years.

To the best of our knowledge, this study is the first of its kind in that it analyzes the cited references in Russian publications in Web of Science "as is", i.e. without any restrictions as to the time range, document types, or any others. An inherent limitation of this analysis is the cut-off date of May 2022 when the data for this investigation were collected, which means that papers published in 2022 are covered only to a small extent. Given the well-known indexation delay in the Web of Science database (especially with conference proceedings papers), it will likely be needed to repeat this study in a few years again in order to verify that the additional newly indexed papers published in 2022 or in any future years do not change the main results of the present analysis. Also, in our future work it might be interesting to try to map USSR's papers from before 1992 onto Russian papers (by analyzing the affiliations of the individual authors) and thus substantially further enlarge the investigated data set. Finally, it should be noted too that our results show the dominance of some disciplines in natural sciences, but we are well aware of various citation practices in different scientific fields and, therefore, we would like to have a closer look at the differences between the disciplines in the future.

ACKNOWLEDGEMENT

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Marx W, Bornmann L, Barth A, Leydesdorff L. Detecting the historical roots of research fields by reference publication year spectroscopy (RPYS). J Assoc Inf Sci Technol. 2014; 65(4): 751-64. doi: 10.1002/asi.23089.

2. Barth A, Marx W, Bornmann L, Mutz R. On the origins and the historical roots of the Higgs boson research from a bibliometric perspective. Eur Phys J Plus. 2014; 129(6): 111. doi: 10.1140/epjp/i2014-14111-6.

3. Bornmann L, Marx W. The wisdom of citing scientists. J Assoc Inf Sci Technol. 2014; 65(6): 1288-92. doi: 10.1002/asi.23100.

4. Leydesdorff L, Bornmann L, Marx W, Milojević S. Referenced publication years spectroscopy applied to iMetrics: scientometrics, Journal of Informetrics, and a relevant subset of JASIST. J Inf. 2014; 8(1): 162-74. doi: 10.1016/j.joi.2013.11.006.

5. Thor A, Marx W, Leydesdorff L, Bornmann L. Introducing CitedReferencesExplorer (CRExplorer): A program for Reference Publication Year Spectroscopy with Cited References Standardization. J Informetr. 2016; 10(2): 503-15. doi: 10.1016/j.joi.2016 .02.005.

6. Thor A, Marx W, Leydesdorff L, Bornmann L. New features of CitedReferencesExplorer (CRExplorer). Scientometrics. 2016; 109(3): 2049-51. doi: 10.1007/s11192-016-2082-3.

7. Marx W, Haunschild R, Thor A, Bornmann L. Which early works are cited most frequently in climate change research literature? A bibliometric approach based on Reference Publication Year Spectroscopy. Scientometrics. 2017; 110(1): 335-53. doi: 10.1007/s11192-016-2177-x, PMID 28190904.

8. Thor A, Bornmann L, Marx W, Mutz R. Identifying single influential publications in a research field: new analysis opportunities of the CRExplorer. Scientometrics. 2018; 116(1): 591-608. doi: 10.1007/s11192-018-2733-7.

9. Bornmann L, Haunschild R, Leydesdorff L. Reference publication year spectroscopy (RPYS) of Eugene Garfield's publications. Scientometrics. 2018; 114(2): 439-48. doi: 10.1007/s11192-017-2608-3, PMID 29449749.

10. Haunschild R. Which are the most influential cited references in Information? Information. 2019; 10(12): 395. doi: 10.3390/info10120395.

11. Ballandonne M, Cersosimo I. A note on reference publication year spectroscopy with incomplete information. Scientometrics. 2021; 126(6): 4927-39. doi: 10.1007/s11192-021-03976-1.

12. Comins JA, Hussey TW. Detecting seminal research contributions to the development and use of the global positioning system by reference publication year spectroscopy. Scientometrics. 2015; 104(2): 575-80. doi: 10.1007/s11192-015-1598-2.

13. Comins JA, Leydesdorff L. Citation algorithms for identifying research milestones driving biomedical innovation. Scientometrics. 2017; 110(3): 1495-504. doi: 10.1007/s11192-016-2238-1.

14. Fiala D, Bornmann L. Reference publication year spectroscopy (RPYS) of computer science papers from Eastern Europe. Aslib J Inf Manag. 2020; 72(3): 305-19. doi: 10.1108/AJIM-06-2019-0142.

15. Elango B, Bornmann L, Kannan G. Detecting the historical roots of tribology research: a bibliometric analysis. Scientometrics. 2016; 107(1): 305-13. doi: 10.1007/s11192-016-1877-6.

16. Hou J. Exploration into the evolution and historical roots of citation analysis by referenced publication year spectroscopy. Scientometrics. 2017; 110(3): 1437-52. doi: 10.1007/s11192-016-2206-9.

17. Li X, Yao Q, Tang X, Li Q, Wu M. How to investigate the historical roots and evolution of research fields in China? A case study on iMetrics using RootCite. Scientometrics. 2020; 125(2): 1253-74. doi: 10.1007/s11192-020-03659-3.

18. Khasseh AA, Mokhtarpour R. Tracing the historical origins of knowledge management issues through referenced publication years spectroscopy (RPYS). J Knowl Manag. 2016; 20(6): 1393-404. doi: 10.1108/JKM-01-2016-0019.

19. Millán JD, Cudina JN, Ossa JC, Vega-Arce M, Scholten H, Salas G. Academic networks of critical social psychology in Brazil. an analysis of the impact and the intellectual roots. Curr Psychol. 2022; 41(5): 3062-77. doi: 10.1007/s12144-020-00827-9.

20. Wray KB, Bornmann L. Philosophy of science viewed through the lense of "Referenced Publication Years Spectroscopy" (RPYS). Scientometrics. 2015; 102(3): 1987-96. doi: 10.1007/s11192-014-1465-6.

21. Yeung AW, Wong NS. The historical roots of visual analog scale in psychology as revealed by reference publication year spectroscopy. Front Hum Neurosci. 2019; 13: 86. doi: 10.3389/fnhum.2019.00086, PMID 30914939.

22. Fiala D, Tutoky G. Computer science papers in Web of Science: A bibliometric analysis. Publications. 2017; 5(4): 23. doi: 10.3390/publications5040023.

23. Fiala D, Maltseva D. Russian publications in Web of Science: A bibliometric study. COLLNET J Scientometr Inf Manag. 2023; 17(2): 217-45. doi: 10.47974/CJSIM-2022-0 006.

24. Fiala D, Willett P. Computer science in Eastern Europe 1989-2014: a bibliometric study. Aslib J Inf Manag. 2015; 67(5): 526-41. doi: 10.1108/AJIM-02-2015-0027.

25. McLevey J, McIlroy-Young R. Introducing metaknowledge: software for computational research in information science, network analysis, and science of science. J Informetr. 2017; 11(1): 176-97.

26. Thor A, Bornmann L, Haunschild R, Leydesdorff L. Which are the influential publications in the Web of Science subject categories over a long period of time? CRExplorer software used for big-data analyses in bibliometrics. J Inf Sci. 2021; 47(3): 419-28. doi: 10.1177/0165551520913817.