A Citation Analysis of the Air Resources Laboratory (1999-2010)

Prepared by Chris Belter (Chris.Belter@noaa.gov), NOAA Central Library
April, 2011

Purpose and Scope

Citation analysis has rapidly become a popular tool for evaluating the research output of scientific institutions. US Government agencies (National Science Board 2010), academic ranking companies (ShanghaiRanking Consultancy 2010), and individual institutions have all begun to use bibliometric analyses to evaluate the quality of scientific research being done at scientific institutions around the world. While there are significant limitations on the validity and applicability of citation analyses for this purpose (e.g. Adler and Harzing 2009; Van Raan 2005), these analyses can be successfully used to supplement and/or validate the findings of peer review evaluations (e.g. Haeffner-Cavaillon and Graillot-Gak 2009; Moed 2005).

This citation analysis was prepared for the Air Resources Laboratory (ARL) in order to supplement its 2011 laboratory review. The purpose of this report is to evaluate the citations received by ARL papers published between 1999 and 2010 in order to determine if these publications received a below average, average, or above average number of citations compared to all of the articles published on the same subjects during the same years. While ARL was consulted to obtain an accurate and complete set of publications for analysis, ARL had no input or influence on the study design, methodology or results of this report. This report presents a condensed version of the bibliometric analysis initially provided to ARL by the author; the full analysis is available upon request to Chris.Belter@noaa.gov.

The analysis presented in this report was conducted using data from Web of Science (WoS). Analysis of ARL publications or citations received from publications not easily accessible through WoS is beyond the scope of this report. The analysis presented below relies on the use of subject categories. These subject categories were created by WoS and assigned to individual articles by WoS based on the journal in which those articles were published. These subject categories are not mutually exclusive, so a single article can be, and often is, assigned to multiple subject categories. Articles appearing in *Atmospheric Environment*, for instance, are assigned to both the “Meteorology and Atmospheric Sciences” subject category and the “Environmental Sciences” subject category. Such articles are counted multiple times in this report, rather than fractionally. The analysis presented in this report includes self-citations and has not been normalized for article type. All of the citation data used in this report were retrieved from WoS between January 17 and January 20, 2011 and are accurate as of January 20, 2011.

Methodology and Rationale

All papers by at least one author affiliated with ARL and published between January 1999 and December 2010 were identified in WoS by a series of searches for all variations of ARL's name and those of all of its field divisions in the "Address" field of WoS. In order to ensure that an accurate and complete set of these papers was compiled, the results of these searches in were compared to yearly bibliographies of ARL publications. Papers not published by ARL or its field divisions were excluded from the set, while articles not captured by the initial searches, but included in ARL's bibliographies, were added to the set through additional searches. The final result set includes publications produced by some divisions that were part of ARL at the time of publication, but are no longer affiliated with or managed by ARL. This final set of publications was analyzed to produce the basic bibliometric indicators summarized in Table 1. It was also used to identify the WoS subject categories in which ARL published the majority of its papers and from which ARL received the majority of its citations. The WoS subject categories identified were "Meteorological and Atmospheric Sciences" and "Environmental Sciences."

The actual distribution of citations received by all papers in these two subject categories were then used to evaluate the citation counts of ARL papers. For each subject category and year of publication from 2000 to 2009, a set of all papers published in that category and year was generated and ranked according to each paper's citation count. This ranking was then used to identify four citation percentiles: 99th, 90th, 50th, and below 50th. The 99th percentile includes papers in the top 1% of the field and year of publication based on their citation counts; the 90th percentile includes papers in the top 2% through the top 10% in that field for that year; and so on. Due to the indexing procedures of WoS, papers in these subject categories that were published in multidisciplinary journals like *Nature* and *Science* could not practically be counted in this process. Each ARL paper was then assigned to one of these four percentiles based on its citation count, subject category, and year of publication. ARL publications in *Nature* and *Science* were assigned to subject categories by the author and were given percentile rankings based on those assignments. The results of this process are summarized in Figures 1 (for Meteorological Sciences) and 2 (for Environmental Sciences).

This percentile evaluation method was used in order to partially correct for some of the limitations of current bibliometric indicators. Raw citation counts and institutional H-Indexes do not provide enough contextual information for their meanings to be easily interpreted; nor can they be compared across subject disciplines. Bibliometric indicators calculated by averages or ratios—indicators such as number of citations per paper, impact factor, the CWTS "crown indicator" (Van Raan 2006), etc—do not take the skewed distribution of citations among scientific papers (Seglen 1992) into consideration (Bornmann and Mutz 2011). The actual distribution of citations among all scientific papers seems to be best fit by a power law distribution (Peterson and others 2010; Redner 1998), such that a few papers receive a large number of citations and the majority of papers receive few to no citations. As a result of this skewed distribution, bibliometric indicators calculated by averages or ratios are artificially inflated by the presence of one or more highly cited papers.

The percentile method used in this analysis seems to correct for many of these limitations. The use of percentiles allows for direct and accurate comparison of the papers under evaluation to the actual distribution of all papers in the same subject category—making interpretation of the results relatively straightforward. The results of this method are not inflated by highly cited papers and, moreover, provide a more detailed depiction of the citations received by all articles under evaluation than any other current bibliometric indicator. Finally, because these percentiles take the differences in citation activity between subject areas into account, it may be that the results of this method could be compared across subject areas. For examples of similar methodologies, see Bornmann and others (2010), Bornmann and Mutz (2011), the National Science Board (2010), and Pudovkin and Garfield (2009).

Results

In all, 598 papers by ARL-affiliated authors and published from 1999 to 2010 were identified in WoS and analyzed. These papers received a total of 16,443 citations in that time period, yielding an average citation rate of 27.5 citations per paper. The H-Index (Hirsch 2005) for this group of publications is 62, meaning that the group includes 62 articles that have each been cited at least 62 times. Since the majority of scientific papers receive few to no citations (Seglen 1992), these results indicate that papers published by ARL are cited significantly more often than average.

This indication is supported by percentile analysis. 6% of the publications produced by ARL authors were ranked in the 99th percentile in their subject categories. Approximately 25% of their publications were ranked in the 90th percentile in those same categories. Conversely, only 24% of ARL publications were ranked below the 50th percentile in Meteorology and Atmospheric Sciences, while only 20% of ARL papers had that rank in Environmental Sciences. These figures indicate that ARL not only publishes a significantly higher number of highly-cited articles than average, but ARL also publishes significantly fewer lowly-cited articles than average. Taken together, these results indicate that the overall quality of ARL papers, as measured by citation counts, is significantly higher than that of papers produced by the majority of their peers.

Bibliometric Indicator	Value
Number of Papers (p)	**598**
Number of Citations Received (c)	**16,443**
Average Number of Citations per Paper (c/p)	**27.5**
H-Index	**62**

Table 1*: Basic bibliometric indicators for ARL publications (1999-2010) in WoS.*

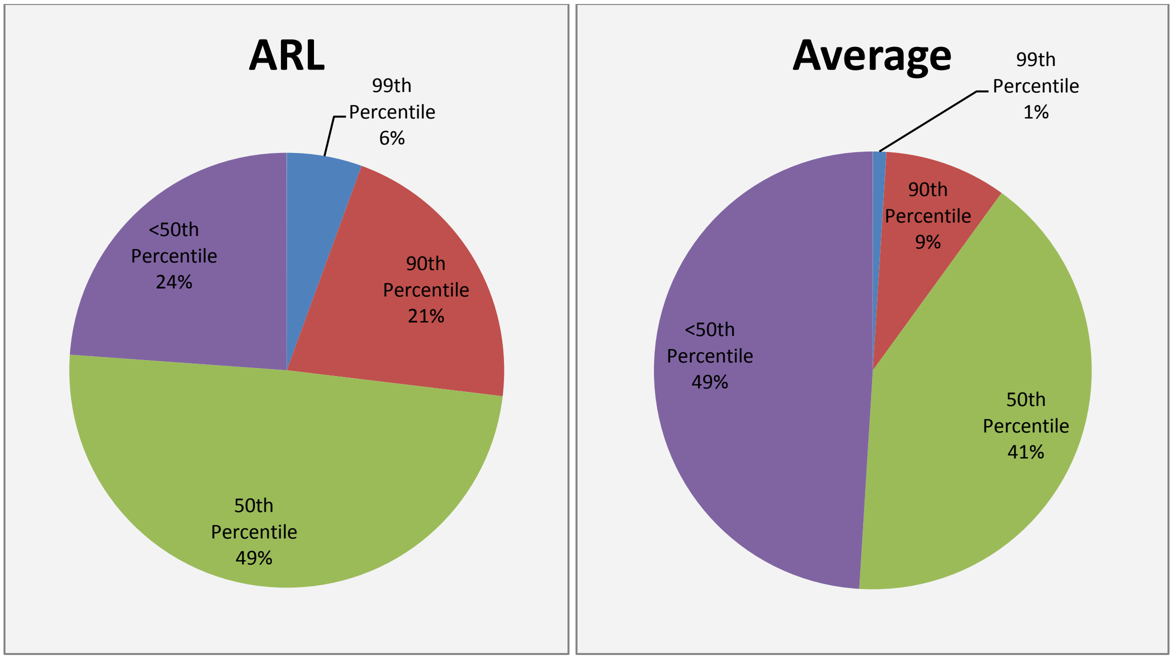

Figure 1: *The percentile distribution, based on citation counts, of ARL publications in Meteorology and Atmospheric Sciences (2000-2009) compared to an average distribution.*

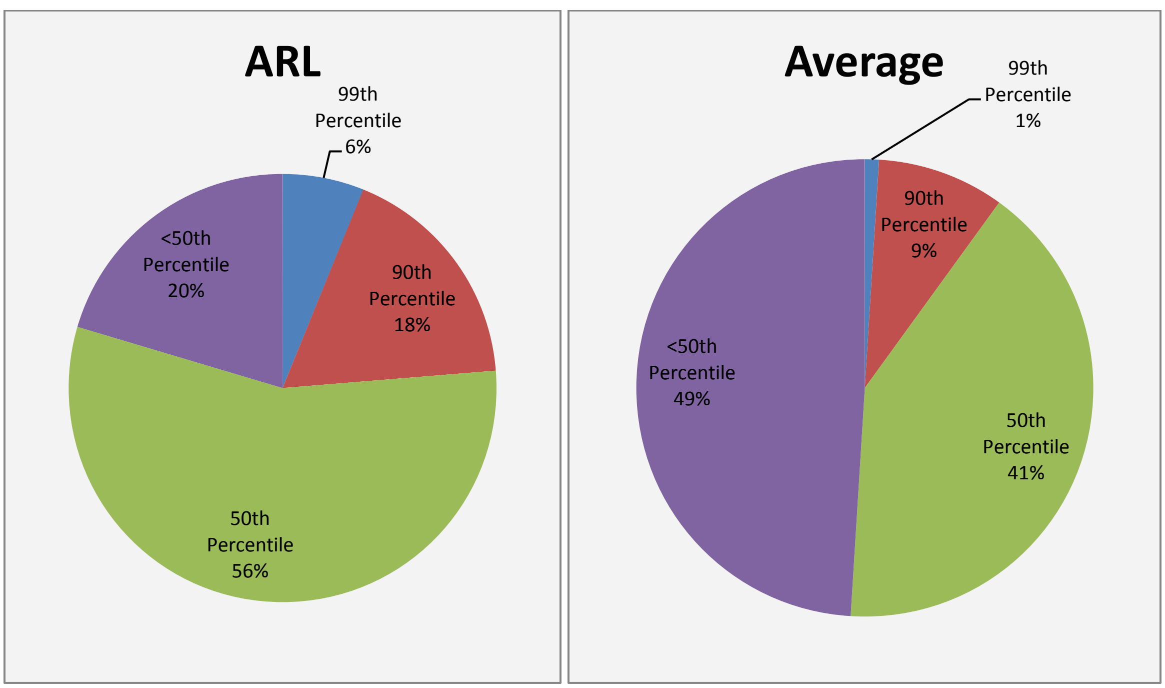

Figure 2: *The percentile distribution, based on citation counts, of ARL publications in Environmental Sciences (2000-2009) compared to an average distribution.*

References

Adler NJ, Harzing A-W. 2009. When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. Academy of Management Learning and Education 8(1):72-95.

Bornmann L, de Moya Anegón F, Leydesdorff L. 2010. Do Scientific Advancements Lean on the Shoulders of Giants? A Bibliometric Investigation of the Ortega Hypothesis. PLoS ONE 5(10):e13327. doi:10.1371/journal.pone.0013327

Bornmann L, Mutz R. 2011. Further steps towards an ideal method of measuring citation performance: The avoidance of citation (ratio) averages in field-normalization. Journal of Informetrics 5(1):228-230. doi:10.1016/j.joi.2010.10.009

Haeffner-Cavaillon N, Graillot-Gak C. 2009. The use of bibliometric indicators to help peer-review assessment. Archivum Immunologiae Et Therapiae Experimentalis 57(1):33-38. doi:10.1007/s00005-009-0004-2

Hirsch JE. 2005. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America 102(46):16569-16572. doi:10.1073/pnas.0507655102

Moed H. 2005. Citation Analysis in Research Evaluation. Dordrecht: Springer.

National Science Board. 2010. Science and Education Indicators 2010. Arlington, VA: National Science Foundation.

Peterson GJ, Pressé S, Dill KA. 2010. Nonuniversal power law scaling in the probability distribution of scientific citations. Proceedings of the National Academy of Sciences 107(37):16023-16027. doi:10.1073/pnas.1010757107

Pudovkin AI, Garfield E. 2009. Percentile Rank and Author Superiority Indexes for Evaluating Individual Journal Articles and the Author's Overall Citation Performance. Collnet Journal of Scientometrics and Information Management 3(2):3-10.

Redner S. 1998. How popular is your paper? An empirical study of the citation distribution. European Physical Journal B 4(2):131-134. doi:10.1007/s100510050359

Seglen PO. 1992. The skewness of science. Journal of the American Society for Information Science 43(9):628-638. doi:10.1002/(sici)1097-4571(199210)43:9<628::aid-asi5>3.0.co;2-0

ShanghaiRanking Consultancy. Academic Ranking of World Universities - 2010 [Internet]. [cited]. Available from: http://www.arwu.org/ARWU2010.jsp

Van Raan AFJ. 2005. Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. Scientometrics 62(1):133-143. doi:10.1007/s11192-005-0008-6

Van Raan AFJ. 2006. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. Scientometrics 67(3):491-502. doi:10.1556/Scient.67.2006.3.10