



12-2010

Citation Accuracy in Environmental Science Journals

Robert Lopresti

Western Washington University, robert.lopresti@wwu.edu

Follow this and additional works at: http://cedar.wwu.edu/library_facpubs



Part of the [Library and Information Science Commons](#)

Recommended Citation

"Citation accuracy in environmental science journals." *Scientometrics*. (a peer-reviewed journal) 85 (3) 2010. 647-655, DOI: 10.1007/s11192-010-0293-6

This Article is brought to you for free and open access by the Western Libraries and the Learning Commons at Western CEDAR. It has been accepted for inclusion in Western Libraries Faculty & Staff Publications by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.

CITATION ACCURACY IN ENVIRONMENTAL SCIENCE JOURNALS

Robert Lopresti
Wilson Library
Western Washington University
Bellingham WA
98225-9103
Fax: 360-650-3044
Rob.lopresti@wwu.edu

ABSTRACT

Citations in five leading environmental science journals were examined for accuracy. 24.41% of the 2,650 citations checked were found to contain errors. The largest category of errors was in the author field. Of the five journals Conservation Biology had the lowest percentage of citations with errors and Climatic Change had the highest. Of the citations with errors that could be checked in Web of Science, 18.18% of the errors caused a search for the cited article to fail. Citations containing electronic links had fewer errors than those without.

INTRODUCTION

Citations are a basic part of the system of scholarly communication and are the standard way of acknowledging credit in science. (CRONIN 1984) The field of ecology is no exception, as TODD, et.al. (2007) noted: "An essential component of most ecology papers is a clear, well-crafted argument that builds upon the existing research base within the subject area in question and substantiates important assumptions, technical information and opinions by accurately identifying (i.e. citing) the source material." (Parentheses in the original)

Citations serve several purposes. They connect the current work to the framework of research that has gone before. As PLACE (1916) observed, providing good references relates to the scientific method: "Substantiate your statement by proof, either of your own or by the work that others have done before you." Citations also permit the reader to confirm (or refute) the author's claims.

While those two purposes have existed for as long as citations have been used, there is a third, more recent development. Citation indexes, such as ISI's Web of Science, permit researchers to evaluate how often an author, article, or journal has been cited by other scholars, and in which fields. This is a useful criterion for determining how significant a role the cited object is playing in the ongoing scientific discussion, and can also become important in questions of tenure and promotion.

All of these purposes are only served if the citations are accurate. Many articles have revealed that this is frequently not the case. BOOTH (2004) examined 36 published studies on citation accuracy and reported that they found error rates ranging from 8% to 66%. He found "a clear trend for between 25% and 40% of citations to be inaccurate."

This study examines a random sample of articles in the 2006 issues of five leading environmental journals in order to compare their error rates to each other and to the trend BOOTH (2004) found. 24.41% of the citations had errors, placing these journals at the low end of the trend, but it means, disturbingly, that each citation has a one in four chance of being incorrect.

METHODOLOGY

The five environmental science journals ranked highest by the CENTER FOR JOURNAL RANKING (2007) were examined. They are, in alphabetical order: Climatic Change, Conservation Biology, Global Biogeochemical Cycles, Remote Sensing of Environment, and Water Resources Research. Five percent of the articles that appeared in each journal during 2006 were selected using a random number generator. Each citation in those articles was examined for accuracy.

Citations were generally compared to online journals or (in the case of citations from other than journal articles) to library catalog records. If a discrepancy appeared the publication was checked in paper or in a full-text digital version, such as PDF. Out of 2,724 citations 2,650 were checked. The rest (2.717%) were either inherently uncheckable (in press, personal communication, etc.) or proved to be beyond the reach of Interlibrary Loan (mostly government or corporate reports, foreign theses, or computer programs).

Certain web pages could not be found and, when there was an absence of obvious typographical errors, this was assumed to be due to "link decay." (GOH and NG 2006) and categorized as uncheckable, rather than as an error. All percentages in the article from this point on refer to the searchable citations (100%=2,650).

Table 1

Following O'CONNOR and KRISTOF (2001) a discrepancy was considered an error if it hindered the locating of the item or if a missing bibliographic element was required by the journal's style guide. For example, changing the order of co-author's names or the recommended order of elements was not treated as an error.

Another category of problem arose that does not appear to have been mentioned in other articles on this subject: In 51 citations to books the publisher and/or place cited disagreed with the one found in the book. This happened often and in many cases consistently. For example, six different authors citing different books published between 1974 and 1993 said that Prentice-Hall is located in Upper Saddle River, New Jersey, while the copies of those same books examined for this article said it is located in Englewood Cliffs, NJ. It appears that different editions of the same book may list different publishers and/or places. Therefore, these 51 cases were not treated as errors.

Errors were categorized as follows:

- Name errors (including authors and editors)
- Title errors (including titles of articles, chapters, journals, books, and series)
- Page errors
- Date errors (including year, month, and day)
- Volume errors (including issue numbers and digital article numbers)
- Publisher or place errors (not counting the 51 errors mentioned above)
- Other

Figure 1.

For each citation, only one mistake was counted per category. For example, if a citation left out one author's name, misspelled a second and missed the middle initial of a third, this was treated as a single author error, and listed under what was judged to be the most serious error.

RESULTS

647 of the citations contained errors (24.41%). There were a total of 792 errors. 124 (4.679% of all citations) contained errors in more than one category. Citations that contained any errors averaged 1.224% errors. No articles contained errors in more than three categories.

Table 2.

Name Errors

By far the largest category of errors was authors' names (352, or 44.4% of all errors). Of these, 192 (24.242% of all errors) were missing middle initials. The similar category of errors in editors' names accounts for another 53 problems (6.692%). 11 of these (1.389%) were missing middle initials. 157 citations (24.268% of those containing errors) contain no mistakes except for leaving out the middle initials of authors and/or editors.

Treating author and editor name errors as one category, 48 (7.419% of all citations with errors) dropped at least one name entirely. 54 (8.346%) contained a mistake in a last name. 52 (8.037%) reversed the initials, or used incorrect ones. 35 (5.409%) added a middle initial that didn't appear in the journal. 10 (1.545%) added an author or editor inappropriately. One listed an author with no initials, and three listed editors as authors.

Title errors

The second largest category was errors in titles of articles, books, or journals. There were 235 (29.672% of all errors) of this type. 29 (3.662% of all errors) of these were missing subtitles. In one case the subtitle appeared but the main title was missing. Errors that occurred

multiple times include words left out, added, reversed, and substituted. In ten cases the citer abbreviated words that were written out in the cited paper.

Two citations listed the wrong journal entirely. For example, an article in the Journal of Geophysical Research was reported to have appeared in Geophysical Research Letters.

Number Errors

The third largest category was page number errors with 90 (11.364% of all errors). 64 (8.081%) of these listed the wrong pages. 3 others (0.379%) failed to include a last page. 23 (2.904%) gave no page numbers at all.

18 (2.782% of all citations with errors) contained wrong volume numbers. Nine citations (1.391%) included incorrect years and one included the wrong month. Although none of the journals required issue numbers, some authors included them, and seven (1.082%) contained errors. Two citations (0.309%) included incorrect AGU numbers.

Publisher or Place Errors

9 (1.391% of all citations with errors) lacked a publisher, a place of publication, or both. Four more (0.618%) listed the correct publisher, but with a mistake in its name.

As mentioned at the beginning of the Results section, 51 apparent flaws in publisher or place were not counted as errors since they seemed to be related to differences between printings of the books. A few citations, however, appeared to contain genuine mistakes in these categories. For example, one book had two publishers, and the citation listed the name of one but gave the address for the other. Five (0.772% of all citations with errors) had such errors in the publisher field, and two (0.309%) had similar errors in place of publication.

Unusual Errors

A few errors were so unusual as to deserve a separate mention:

- * One article was cited twice in the same alphabetical list.
- * In two citations the impersonal article "A" at the beginning of an article title migrated to become an author's middle initial. So "A globally coherent fingerprint..." co-authored by G. Yohe became "Globally coherent fingerprint..." coauthored by G.A. Yohe. These two citations appeared consecutively in the same article.
- * The co-authors of one article made errors in citing three articles they themselves had written.
- * A book with no editor, published by Chapman & Hall/CRC, was cited as edited by C .Hall and published by CRC.
- * One citation conflated co-authors R. Sabatier and J.M. Masson into the non-existent J.M. Sabatier.

Errors in Electronic citations

144 of the citations (5.434%) contained electronic links, either as DOIs or URLs, and 133 of those were checkable. (As stated above, most of the uncheckable ones were considered to be victims of link decay, and were not counted as errors.) Those 133 citations represent 5.019% of the total number of checked citations, and in fact only 26 of them (19.549%) contained flaws, as compared to 24.34% of all the citations.

Journal differences

Of the five journals, Conservation Biology had the lowest percentage of citations with errors (17.138%) and Climatic Change had the highest (38.947%). Conservation Biology contained the only article with no errors, although two of its citations could not be checked. Climatic Change had the article with the highest percentage of citations with errors: 62.857%. In Conservation Biology the largest category of errors was a tie between name and title errors (45% each). In Remote Sensing of Environment name errors was the largest category, with 43.209% of the errors. In each of the other three journals name errors accounted for more than half the mistakes, with the highest being Water Resources Research where names accounted for 56.097% of errors. Title errors was the second largest category in the four journals in which it did not tie for first.

Table 3

The narrowest range of errors within articles was Global Biogeochemical Cycles, which also had the fewest number of articles: 12.698 to 31.944%. Remote Sensing of

Environment had the widest range of error rates within articles, from 10.909 to 55.172%. Figure 2

Water Resources Research and Global Biogeochemical Cycles have the most detailed style instructions, which actually worked against Water Resources Research in the context of this study. For example, both of these journals clearly stated that et al was only to be used if there were more than ten authors, so violations of that were considered errors. There were no such violations in Global Biogeochemical Cycles but there were several in Water Resources Research. The other journals did not specify the number of authors needed, so et al was not treated as an error in their citations as long as there were at least two authors.

Water Resources Research, which contained 36.906% of the searched citations, had 63.91% of the electronic citations. Errors were found in 15.294% of the electronic citations in Water Resources Research, which is lower than the journal's overall rate of 24.131% and the joint rate for electronic citations in all five journals, which was 19.549%.

Practical Results

As stated above, one purpose of citations is to permit the reader to confirm (or refute) the author's claims. Citation errors can significantly affect the scholarship if they make it more difficult to find the cited work.

Each of the faulty citations to English language journal articles published after 1969 were checked in Web of Science, a database that lists every citation in thousands of scholarly journals. (WEB OF SCIENCE. 2007) Citations were sought using three different strategies:

AJ: Primary author's last name, first initial, and the title of the journal.

Example: ASANO, M*.in Author, Canadian Journal of Anaesthesia in Publication Name.

AT: Primary author's last name, first initial, and the first two words of the article title.

Example: ASANO, M* .in Author, "Improvement of" in Title.

T: First six words of the title.

Example: "Improvement of the accuracy of references" in Title.

Some of the articles were apparently not listed in Web of Science, and they were not considered in the following section. In a few cases, specific searches apparently failed through

eccentricities of the Web of Science database, and they were treated as successes for the sake of this article, since it was not the citation errors that caused the problem.

Of the articles that could be searched in Web of Science, 68 out of 374 (18.18%) failed at least one of the searches. One citation managed to defeat all three search strategies, in spite of the fact that the article was indeed indexed in Web of Science.

Global Biogeochemical Cycles had the lowest percentage of failed searches, with 11.63% of the faulty citations producing at least one failure. Climatic Change was the highest, with a failure rate of 21.15%.

Over all, Author/Title searches were slightly more successful than Author/Journal Name searches. Title searches were the least efficient, being 31.43% more likely to fail than Author/Title.

Table 4

DISCUSSION

It is clear that citation errors remain a serious problem, affecting in the case of this study, almost one citation in four.

The most common type of error involved the author or editor's name. Particularly rampant was the habit of dropping the middle initials of authors and/or editors. This is usually considered to be a minor error, since it does not generally make it difficult to find a publication. (O'CONNOR and KRISTOF 2001) However, it can make it much more problematic to trace the work of an author. To create a hypothetical case, an author named J.Z. Smith might not consider it a minor error to have her article lost in a sea of J. Smiths. BOOTH (2004) also noted that this

means "that the approach that most librarians will utilize instinctively, that is the author search, is the least likely to be successful."

One important lesson of this study is that one cannot assume a citation of an author's middle initial – or lack thereof – is accurate. This makes tracing an author by her or his citations problematic.

One category of error that may cause problems disproportionate to its size is mistakes in the year of publication. As BROADUS (1983) noted, an error in publication date is particularly serious because it can lead to misunderstanding how current or historic a particular study is.

Citations that contain electronic links, either as DOIs or URLs, raise an interesting question. Since many of these links contain information that can be "copy and pasted" it seems reasonable to predict that these citations would be more likely to be accurate than the average citation. In fact, this is exactly what happened: there was almost a 5% difference between the error rate of electronic citations and the overall error rate. Since future articles are likely to contain more electronic citations, this is a hopeful trend.

The question remains: how serious are these errors for a researcher? As it turned out, almost one in five of the studied journal errors had the potential of defeating a search in Web of Science. While a determined searcher who tried all three search techniques would have found all but one of these articles eventually, it is very likely that many would have given up before that point. So, in the case of journal citations, almost twenty percent of the errors could have resulted in a scholar failing to find the source. Another lesson is that in searching Web of Science for a specific article, hunting by author/title is more likely to be successful than other obvious choices.

This article did not attempt to explain the cause of the errors, but three categories of causes come to mind:

1. Inaccurate hand-copying, resulting in the droppings of middle initials, misspelling of title words, and the like.
2. Incomplete copying of electronic citations, resulting in the dropping of the last authors, the last words in titles, etc.
3. “Copy and paste” citing of unread articles that were cited inaccurately in earlier articles. BROADUS discussed the extent of this tendency in 1983.

As electronic articles become a larger percentage of the cited material, category 1 is likely to decrease, but categories 2 and 3 seem even more likely to occur. The results stated above indicate that citations to electronic sources are less likely to have errors, which indicates that category 1 is the most likely culprit.

The broad range of difference in error rate between journals (the highest error rate was more than double that of the lowest) suggests that editors may be able to make a difference in the process. In fact, it is possible that editors might be able to improve the accuracy of the citations in their journals through one small change in their policies. ASANO (1995) reported that one journal was able to reduce error rates by more than half simply by requiring that authors submit a photocopy of the first page of each referenced publication. None of the journals in this study appear to make that requirement. This is an improvement the editors should consider.

CONCLUSION

PLACE (1916) described citation error as an ancient problem when he wrote about it almost a century ago. The situation does not appear to be resolving any time soon.

Specifically, citation errors continue to be a problem in environmental science journals. One hopeful sign is that the percentage of errors was lower for citations that contained electronic links. Since these are likely to increase in the future, error rates may decrease a bit. If the editors of such journals were to adopt ASANO's suggestion mentioned above, this might lead to another improvement. In the meantime, the rule continues to be: let the reader beware.

REFERENCES

- ASANO, M. et al. (1995), Improvement of the accuracy of references in the Canadian Journal of Anaesthesia, *Canadian Journal of Anesthesia*, 42: 370-2.
- BOOTH, A. (2004), A checker's career?, *Health Information and Libraries Journal*, 21: 269-272.
- BROADUS, R.N. (1983), An investigation of the validity of bibliographic citations, *Journal of the American Society for Information Science*, 34: 132-135.
- CENTER FOR JOURNAL RANKING. (2007), *Journal Ranking*, <http://www.journal-ranking.com/ranking>, (accessed 06 September 2007).
- CRONIN, B. (1984), *The citation process: the role and significance of citations in scientific communication*. Taylor Graham, London.
- GOH, D.H., NG, P.K. (2006), Link decay in leading information science journals, *Journal of the American Society for Information Science*, 58: 15-24.
- O'CONNOR, L.G. and KRISTOF, C. (2001), Verify your citations: accuracy of reference citations in twelve business and economics journals, *Journal of Business and Finance Librarianship*, 6: 23-40.
- PLACE, F., Jr. (1916), Verify your references: a word to medical writers, *New York Medical Journal*, 104: 697-699.

TODD, P.A. et al. (2007), Citing practices in ecology: can we believe our own words?,
Oikos, 116: 1599-1601.

WEB OF SCIENCE. *Web of Science Help*.

http://images.isiknowledge.com/WOKRS49B3/help/WOS/h_database.html (accessed 16 June
2010)

Figure 1. Percentage of errors in journal citations by category.

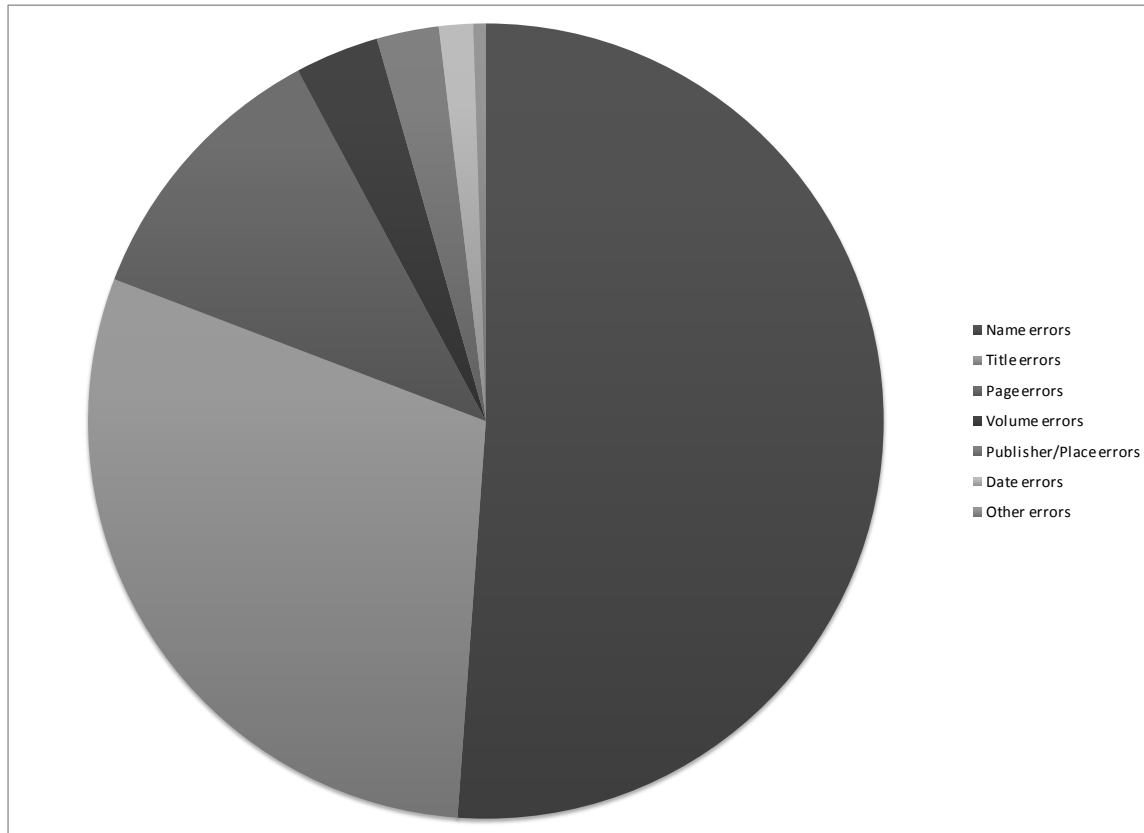


Figure 2. The range of percentage of citations with errors within in each journal. The triangle marks the mean.

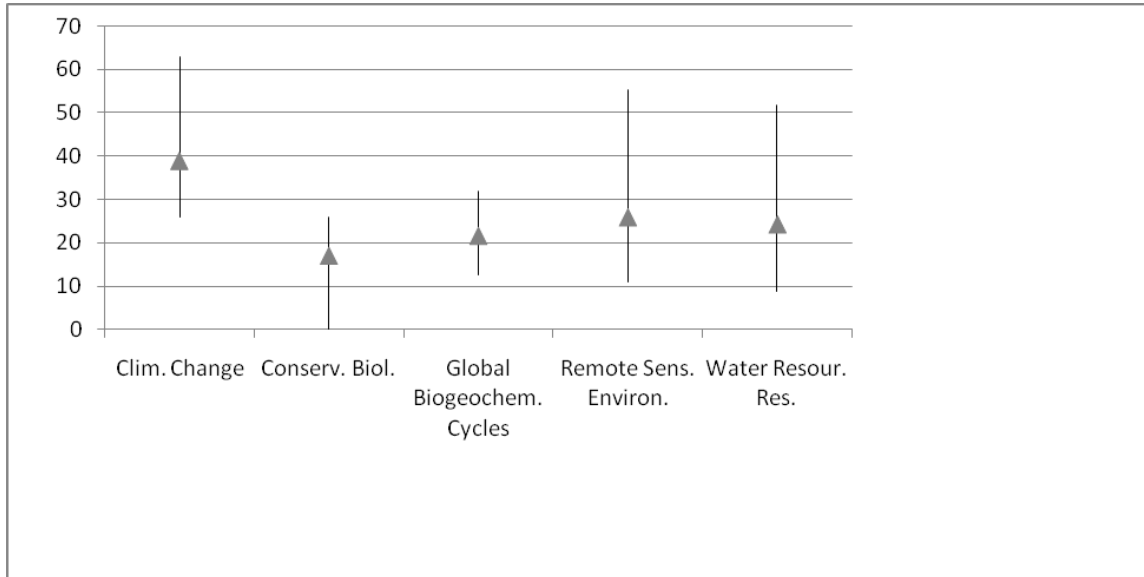


Table 1. Articles and citations checked.

JOURNAL	ARTICLES CHECKED	TOTAL CITATIONS	CITATIONS CHECKED
Climatic Change	6	573	566
Conservation Biology	10	307	285
Global Biogeochemical Cycles	5	296	290
Remote Sensing of Environment	12	551	531
Water Resource Research	20	997	978
TOTAL	53	2724	2650

Table 2. Number of errors per citation

JOURNAL	No errors	1 error	2 errors	3 errors	flawed citations	total errors
Climatic Change	469	76	19	2	97	120
Conservation Biology	174	79	25	7	111	150
Global Biogeochemical Cycles	227	54	8	1	63	73
Remote Sensing of Environment	393	117	18	3	138	162
Water Resource Research	740	197	33	8	236	287
TOTAL	2005	523	103	21	647	792

Table 3. Error types by percentage and journal

	Name	Title	Page	Date	Volume	Pub/Place	Other
Climatic Change	54.305	25.828	12.583	0.662	1.324	5.298	0
Conservation Biology	45	45	5	1.667	1.667	1.667	0
Global Biogeochemical Cycles	52.778	31.944	6.944	2.778	2.778	2.778	0
Remote Sensing of Environment	43.478	27.329	20.497	1.242	5.59	1.242	0.621
Water Resources Research	56.097	25.784	9.408	1.394	4.181	2.09	1.045

Table 4. Citations with Errors Checked in Web of Science (1970-2010)

	Total	AJ	AT	T	SC
Climatic Change	52	7	5	9	11 (21.15%)
Conservation Biology	43	2	3	7	9 (20.93%)
Global Biogeochemical Cycles	43	2	3	3	5 (11.63%)
Remote Sensing of Environment	90	6	6	9	15 (16.67%)
Water Resources Research	146	20	18	18	28 (19.18%)
TOTAL	374	37	35	46	68 (18.18%)

SC= Searched citations with failed searches