

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Germany. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Comments on “Mapping the scientific research on non-point source pollution: a bibliometric analysis” by Yang et al. (2017)

Yuh-Shan Ho¹

Received: 16 May 2017 / Accepted: 2 October 2017 / Published online: 7 October 2017
© Springer-Verlag GmbH Germany 2017

Yang et al. (2017) recently published a paper in this journal entitled “Mapping the scientific research on non-point source pollution: A bibliometric analysis.” Many of the related results presented in the original paper (Yang et al. 2017) are not acceptable because of the use of inappropriate search filters.

Yang et al. (2017) stated in “Materials and methods” that “Non-point source pollution” or “Nonpoint source pollution” or “diffuse pollution” were used to search titles, author keywords, keywords plus, and abstracts of all publications and also in “Results and discussion” that “A total of 2607 publications in ten document types were found in the SCI-Expanded based on the given words during the years 1991–2015.” In fact, a total of 3057 publications in 11 document types were found by the same method as noticed in the original paper (Yang et al. 2017). There is a big difference between 2607 and 3057 publications.

The Science Citation Index Expanded (SCI-EXPANDED) database, originally designed for researchers to find literatures but not bibliometric study. Thus, it is necessary to have a bibliometric treatment when using the Web of Science database. However, Yang et al. (2017) considered keywords contained *KeyWords Plus* which provides search terms extracted from the titles of papers cited in each new article listed in *Current Contents* (Garfield 1990). Those documents that can only be found by *KeyWords Plus* are more likely to be unrelated to the “non-point source pollution.” Ho’s group firstly proposed the “front page” filter (Fu et al. 2012; Fu

and Ho 2014; Ho and Fu 2016)—which covers only documents with searching keywords in their “front page,” including only the title, abstract, and author keywords—might avoid introducing unrelated publications for analysis (Fu et al. 2012). Since any results and discussion depend on the data abstracted by a search filter, an inappropriate filter may lead to inaccurate results and wrong conclusions in the original paper (Yang et al. 2017). In total, 335 searched documents (11% of 3057 documents) unrelated to “non-point source pollution” were also searched out because of using the searching tool in Web of Science without any further data treatment. These include, for example, highly cited articles entitled “SWAT: Model use, calibration, and validation” (Arnold et al. 2012), “Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China” (Yang et al. 2008), and “Asymmetric information and contract design for payments for environmental services” (Ferraro 2008); highly cited reviews entitled “The soil and water assessment tool: Historical development, applications, and future research directions” (Gassman et al. 2007) and “The effects of riparian forest management on the freshwater environment: A literature review of best management practice” (Broadmeadow and Nisbet 2004); and highly cited proceeding papers entitled “Wetland and stream buffer size requirements: A review” (Castelle et al. 1994). None of these papers are related to “non-point source pollution.”

Yang et al. (2017) originally reported that “The characteristics of publication outputs from 1991 to 2015 are shown in Table 1.” Here, these authors have copied the same table and concept from papers published by Ho’s group (Ma et al. 2013).

The authors stated in “Analysis of subject category and journals” that “The 2607 publications from SCI-Expanded database were classified into 59 subject categories identified by WoS. Table 2 represents the top 20 most productive

Responsible editor: Philippe Garrigues

✉ Yuh-Shan Ho
ysho@asia.edu.tw

¹ Trend Research Centre, Asia University, No. 500, Lioufeng Road, Wufeng, Taichung County 41354, Taiwan

subjects.” In fact, there is nothing about “Web of Science categories” but “research areas” only in “Analysis of subject category and journals.” Thus, Table 2 and Fig. 1 in the original paper (Yang et al. 2017) are not correct.

Yang et al. (2017) also did not cite appropriate references for related description. In “Analysis of countries/territories” and “Analysis of institutions,” the authors stated that “The top 30 most productive countries were displayed in Table 4 that contains six indicators: TP, SP, CP, FP, RP, and h-index.” and “The top 30 productive institutions were provided in Table 5 that also includes six indicators: TP, SP, CP, FP, RP, and h-index.” without reporting that Ho’s group (Fu et al. 2010; Wang et al. 2011; Fu et al. 2013; Ho and Fu 2016) had already compared the six indicators such as total articles, independent articles, collaborative articles, first author articles, corresponding author articles, and h-index.

In “Analysis of author keywords,” Yang et al. (2017) noticed that “The top 30 most frequent author keywords during the 5 stages were displayed in Table 6.” It has been pointed out in a comment (Ho 2017) that Ho and his co-workers (Li et al. 2009) have developed a method of combining article titles, author keywords, and *KeyWords Plus* to provide important clues for research hotspots and this was extended in 2010 (Zhang et al. 2010). This analysis, including title words, author keywords, and *KeyWords Plus* together can minimize some limitations, such as the uncompleted meaning of single words in a title, the small sample size for author keywords, and the indirect relationship between *KeyWords Plus* and the research emphases (Fu and Ho 2013). Furthermore, a new method, “word cluster analysis,” has been successfully applied to find the research hotspots in a field (Mao et al. 2010; Fu et al. 2013). Furthermore, similar rebuttals have also been published in *Environmental Earth Sciences* (Ho 2016a), *Scientometrics* (Ho 2016b), and *Journal of Cleaner Production* (Fu and Ho 2017).

The Web of Science database was originally designed not for bibliometric study. An appropriate data treatment is needed for bibliometric study. The inappropriate use of search filters can have enormous effects on the results obtained, and thus great attention should be attached to the search filters chosen. Citing an original paper not only respects those authors who presented a novel idea but it also directs readers to the details of the original work (Ho 2010). In my view, Yang et al. should have cited the original papers for all the indicators and concepts they discuss, thereby providing greater accuracy and detailed information about the bibliometric concepts that they employed.

References

- Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, Van Liew MW, Kannan N, Jha MK (2012) SWAT: model use, calibration, and validation. *Trans ASABE* 55(4):1491–1508
- Broadmeadow S, Nisbet TR (2004) The effects of riparian forest management on the freshwater environment: a literature review of best management practice. *Hydrol Earth Syst Sci* 8(3):286–305
- Castelle AJ, Johnsn AW, Conolly C (1994) Wetland and stream buffer size requirements: a review. *J Environ Qual* 23(5):878–882
- Ferraro PJ (2008) Asymmetric information and contract design for payments for environmental services. *Ecol Econ* 65(4):810–821
- Fu HZ, Ho YS (2013) Independent research of China in Science Citation Index Expanded during 1980–2011. *J Informetrics* 7(1):210–222
- Fu HZ, Ho YS (2014) Top cited articles in adsorption research using Y-index. *Res Eval* 23(1):12–20
- Fu HZ, Ho YS (2017) Comments on “State of the art on food waste research: a bibliometrics study from 1997 to 2014”. *J Clean Prod* 157:167
- Fu HZ, Ho YS, Sui YM, Li ZS (2010) A bibliometric analysis of solid waste research during the period 1993–2008. *Waste Manag* 30(12):2410–2417
- Fu HZ, Wang MH, Ho YS (2012) The most frequently cited adsorption research articles in the Science Citation Index (Expanded). *J Colloid Interface Sci* 379(1):148–156
- Fu HZ, Wang MH, Ho YS (2013) Mapping of drinking water research: a bibliometric analysis of research output during 1992–2011. *Sci Total Environ* 443:757–765
- Garfield E (1990) *KeyWords Plus™*—ISIS breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette. *Curr Contents* 32:5–9
- Gassman PW, Reyes MR, Green CH, Arnold JG (2007) The soil and water assessment tool: historical development, applications, and future research directions. *Trans ASABE* 50(4):1211–1250
- Ho YS (2010) Letters to the editor. *Adsorpt Sci Technol* 28(5):465
- Ho YS (2016a) Rebuttal to: “a bibliometric review on carbon cycling research during 1993–2013” by Zhi et al. (*environ earth Sci* 2015, 74(7): 6065–6075). *Environ Earth Sci* 75 (9), Article Number: 819
- Ho YS (2016b) Rebuttal to: Liu et al. “Progress in global parallel computing research: a bibliometric approach”, vol. 95, pp 967–983. *Scientometrics* 108(3):1693–1694
- Ho YS (2017) Comments on “Past, current and future of biomass energy research: a bibliometric analysis” by Mao et al. (2015). *Renew Sustain Energy Rev*, (in press)
- Ho YS, Fu HZ (2016) Mapping of metal-organic frameworks publications: a bibliometric analysis. *Inorg Chem Commun* 73:174–182
- Li LL, Ding GH, Feng N, Wang MH, Ho YS (2009) Global stem cell research trend: bibliometric analysis as a tool for mapping of trends from 1991 to 2006. *Scientometrics* 80(1):39–58
- Ma JP, Fu HZ, Ho YS (2013) The top-cited wetland articles in Science Citation Index Expanded: characteristics and hotspots. *Environ Earth Sci* 70(3):1039–1046
- Mao N, Wang MH, Ho YS (2010) A bibliometric study of the trend in articles related to risk assessment published in Science Citation Index. *Hum Ecol Risk Assess* 16(4):801–824
- Wang MH, Li JF, Ho YS (2011) Research articles published in water resources journals: a bibliometric analysis. *Desalin Water Treat* 28(1–3):353–365
- Yang J, Reichert P, Abbaspour KC, Xia J, Yang H (2008) Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J Hydrol* 358(1–2):1–23
- Yang BB, Huang K, Sun DZ, Zhang Y (2017) Mapping the scientific research on non-point source pollution: a bibliometric analysis. *Environ Sci Pollut Res* 24(5):4352–4366
- Zhang GF, Xie SD, Ho YS (2010) A bibliometric analysis of world volatile organic compounds research trends. *Scientometrics* 83(2):477–492