

Identification of interdisciplinary research based upon co-cited journals

Marco van Veller

Wageningen University & Research – Library, Wageningen, The Netherlands

Abstract

Purpose – This paper aims to the identification of journal articles that probably report on interdisciplinary research at Wageningen University & Research (WUR).

Design/methodology/approach – For identification of interdisciplinary research, an analysis is performed on journals from which articles have been cited in articles (co-)authored by WUR staff. The journals with cited articles are inventoried from the reference lists of the WUR articles. For each WUR article, a mean dissimilarity is calculated between the journal in which it has been published and the journals inventoried from the reference lists. Dissimilarities are derived from a large matrix with similarity values between journals, calculated from co-occurrence of these journals in the WUR articles' reference lists.

Findings – For 21,191 WUR articles published between 2006 and 2015 in 2,535 journals mean dissimilarities have been calculated. The analysis shows that WUR articles with high mean dissimilarities often are published in multidisciplinary journals. Also, WUR articles with high mean dissimilarities are found in non-multidisciplinary (research field-specific) journals. For these articles (with high mean dissimilarities), this paper shows that citations are often made to more various research fields than for articles with lower mean dissimilarities.

Originality/value – Identification of articles reporting on interdisciplinary research may be important to WUR policy for strategic purposes or for the evaluation of researchers or groups. Also, this analysis enables to identify journals with high mean dissimilarities (due to WUR articles citing more various research fields). Identification of these journals with a more interdisciplinary scope can be important for collection management by the library.

Keywords Journals, Interdisciplinary research, Citations, References, Multidisciplinary journals, Research fields

Paper type Research paper

Introduction

Wageningen University & Research (WUR) is an international research and knowledge center in The Netherlands that consists of a university and various research institutes. The mission of WUR is to explore the potential of nature to improve the quality of life (WUR, 2016). To accomplish this, various research activities are undertaken by 6,500 staff and 10,000 students.

Staff and students present their research in publications of various kinds (e.g. reports, books, book chapters, conference proceedings, journal articles, etc.). Once registered in the Research Information System (RIS) of WUR, these publications are included in WUR's institutional repository (called Staff Publications). Of the publications produced by authors from WUR about one third consists of (refereed) articles in scholarly journals (Staff publications, 2017).

To get an idea of the research topics that are investigated at WUR, 3,020 articles that have been published in 2015 are collected from Web of Science (Clarivate analytics, 2017). To overcome a time lag in registration (in the RIS of WUR) and indexation (in Web of Science) and have a complete overview

of articles, it is chosen to collect the WUR authored articles published in 2015. From the bibliographic data (title and abstract) of these articles a term map based on text data is made with VOSviewer (Eck and Waltman, 2016). Figure 1 shows this term map.

The term map shows that e.g. gene, protein, plant, disease, ecosystem, farmer and climate change are important terms that WUR authors used in their articles. These terms represent a selection of the most important topics that have been investigated at WUR in 2015. Further, the map in Figure 1 shows 24 clusters of term maps. Depending on the settings of VOSviewer while calculating the term map, the number of clusters may vary. Further examination of some of the clusters show that they each indicate a particular research field that is examined at WUR (e.g. the red cluster corresponds with

© Marco van Veller. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

The author thanks wishes to thank Ek'abo Omwana for inspiration on the writing of this paper. Further the author wants to thank Markku A Laitinen for aid in the publication process of this paper.

Received 4 October 2018

Revised 20 February 2019

Accepted 4 March 2019

The current issue and full text archive of this journal is available on Emerald Insight at: www.emeraldinsight.com/2514-9326.htm



Collection and Curation
38/3 (2019) 68–77
Emerald Publishing Limited [ISSN 2514-9326]
[DOI 10.1108/CC-10-2018-0021]

Materials and methods

In an earlier paper, [van Veller \(2013\)](#) describes a methodology to get customized lists of journals that have been cited in WUR articles over a certain period of time. In another paper, [van Veller and Gerritsma \(2015\)](#) make an analysis of the number of times journals have been co-cited in articles published by WUR authors in 2006–2013. Based upon this analysis they propose a measure of similarity between citing journals. The similarity measure that [van Veller and Gerritsma \(2015\)](#) propose is based upon relative abundances of cited journals in the references of the articles published in the (citing) journals for which the similarity is calculated. The more the relative abundances of co-cited journals in the references of articles published in two citing journals are, the higher the similarity between these citing journals is.

This paper applies the similarity measure proposed by [van Veller and Gerritsma \(2015\)](#), but in a modified way. Where in [van Veller and Gerritsma \(2015\)](#) the similarities are calculated for journals with WUR articles (i.e. the citing journals), this paper calculates similarities between the journals included in the references of the WUR articles (i.e. the cited journals). To calculate these similarities a list of cited journals is made. For each of these cited journals, it is inventoried how many times it has been cited in the WUR articles. This inventory is made per combination of a cited and a citing journal, resulting in a list with the number of times a cited journal has been cited by articles published in a citing journal.

From the list similarities between cited journals are calculated based upon shares of citations from articles published in the citing journals. There more two journals are cited by comparable shares of citations from articles published in the same citing journals, the higher the similarity between these cited journals is. The calculation of the similarity (S_{gh}) between cited journal g and cited journal h is formalized as follows:

$$S_{gh} = 1 - \frac{\sum_{j=1}^{n_j} \left| \frac{c_{jg}}{\sum_{i=1}^{n_i} c_{jg}} - \frac{c_{jh}}{\sum_{i=1}^{n_i} c_{jh}} \right|}{2}$$

- S_{gh} = similarity between cited journal g and cited journal h.
- c_{jg} = number of citations to journal g in articles published in citing journal j.
- n_j = number of citing journals.
- c_{jh} = number of citations to journal h in articles published in citing journal j.

The calculated similarities between n cited journals are collected in a symmetric $n_g \times n_h$ matrix. At the diagonal, the value 1 represents the complete similarity of these journals with themselves. In this matrix, the similarity S_{gh} between two journals g and h is listed in the cell corresponding with row g and column h.

The similarities between the journals can be visualized with VOSviewer ([Eck and Waltman, 2016](#)). Only journals with minimal 10 citations from WUR articles are included in the visualization. For the cluster analysis and mapping the following parameter settings are applied:

- Mapping attraction: 2
- Mapping repulsion: 1

- Clustering resolution: 1.10
- Minimum cluster size: 10
- Normalization method 1

For all references included in a WUR article, similarities are obtained from the symmetric $n_g \times n_h$ matrix. Hereby, for each reference, the journal in which the article has been published corresponds with row g and the journal to which is cited corresponds with column h. Per WUR article, the mean is calculated over all similarities obtained for its references. When a particular journal has been included in the references of an article more than once, for each occurrence the similarity of this journal with the citing journal is included in the calculation of the mean. The mean similarity of all references in a WUR article is transformed into a dissimilarity value by subtracting it from one. The calculation of the mean dissimilarity (\overline{D}_{gi}) over references j for article i published in journal g is formalized as follows:

$$\overline{D}_{gi} = 1 - \frac{\sum_{j=1}^{n_j} S_{gh_j}}{n_j}$$

- \overline{D}_{gi} = mean dissimilarity of article i published in journal g.
- S_{gh_j} = similarity between journal g (in which the citing article i has been published) and journal h (in which the cited article has been published that is listed in reference j of article i).
- n_j = number of references for article i published in journal g.

The mean dissimilarities for WUR articles published in the same journal g can be used to calculate a mean dissimilarity for journal g (\overline{D}_g). This is formalized as follows:

$$\overline{D}_g = \frac{\sum_{i=1}^{n_i} \overline{D}_{gi}}{n_i}$$

- \overline{D}_g = mean dissimilarity of journal g.
- n_i = number of articles published in journal g.

Results

For the similarity analysis described in this paper, WUR authored articles are collected from the RIS for publication years 2006–2015 and looked up in Web of Science ([Clarivate analytics, 2017](#)). Hereby it is chosen to obtain articles published in 2015 and nine years earlier to prevent a time lag in registration (in the RIS of WUR) and indexation (in Web of Science) and have a complete overview of articles published by WUR over a 10-year time period in journals covered by Web of Science (i.e. ISI journals).

In the period 2006–2015 WUR authors published 21,191 articles in 2,535 ISI journals. From these WUR articles in total 925,809 citations to 13,257 journals have been collected. For each of these cited journals, it is inventoried how many times it has been cited from each of the journals in which the WUR articles were published. Based upon having received citations from the same (citing) journals a similarity value (S_{gh}) was calculated with R ([R Core Team, 2015](#)) for each pairwise combination of cited journals g and h according to the formula listed above. The similarities were

saved in a similarity matrix of $13,257 \times 13,257$ cited journals.

Figure 2 shows a network that represents the similarity between the cited journals. For this network, only journals are selected that received minimal ten citations from the WUR articles in 2006-2015. This selection results in 5,015 of the 13,257 cited journals. Together these more than five thousand journals received 97 per cent of all citations from the WUR articles.

Figure 2 shows that the cited journals fall in six different clusters. Further investigation of these clusters reveals that each broadly corresponds with main scientific disciplines, which are investigated at WUR:

- the red cluster corresponds with journals in agroecosystem sciences and agronomy;
- the green cluster corresponds with journals in chemical sciences and biotechnology;
- the yellow cluster corresponds with journals in environmental sciences;
- the light blue cluster corresponds with journals in plant sciences;
- the dark blue cluster corresponds with journals in health and nutrition; and
- the purple cluster corresponds with journals in animal sciences.

By adjusting the parameter settings for mapping in VOSviewer it is possible to distinguish more clusters that each comprise a smaller set of journals corresponding with a certain scientific discipline.

From the more than 5,000 journals that are presented in the network in Figure 2 only the journals that contain more than

two WUR articles published in 2006-2015 are selected. This selection includes 1,175 journals with 20,362 WUR articles (96 per cent of the articles that are analyzed in this paper).

Based upon the position of each journal in the network presented in Figure 2, a sub-network is presented in Figure 3 with the total number of WUR articles per journal proportional to the size of the bubbles.

From Figure 3, it follows that the top 20 of journals with the highest number of WUR articles published in 2006-2015 can be found back in all six different clusters. Table I lists these 20 journals with information on the number of WUR articles and the cluster they are placed in based upon the similarity analysis described above.

To calculate the mean dissimilarity (\bar{D}_{gi}) over all references for article i published in journal g the similarity matrix of $13,257 \times 13,257$ cited journals is used. Hereby, for each of the 21,191 WUR articles, firstly, the similarity values between the journal in which the article has been published and each of the journals that contains the cited article (included in the references) are obtained from the matrix and averaged over all cited articles in the WUR article. Secondly, the mean similarity value is transformed to a mean dissimilarity by subtraction from 1.

Figure 4 shows the distribution of the mean dissimilarity per WUR article.

The histogram in Figure 4 shows that the mean dissimilarity per WUR article lies between 0 and 1. The most frequent mean dissimilarity lies around 0.70. For a part of the WUR articles the mean dissimilarity is either very low (107 articles with values below 0.02) or high (657 articles with values above 0.98). For the

Figure 2 Network of 5,015 journals based upon a similarity analysis on citations to these journals from research articles published by WUR in 2006-2015

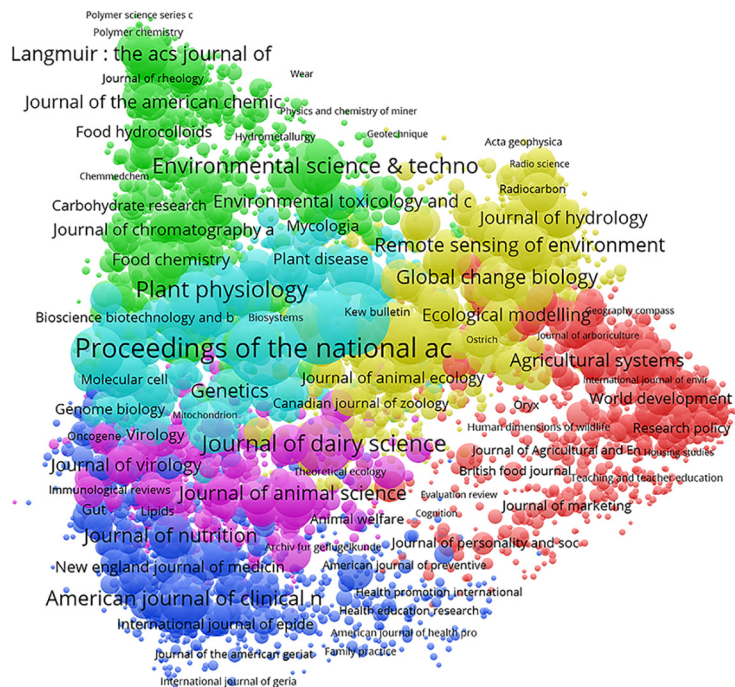
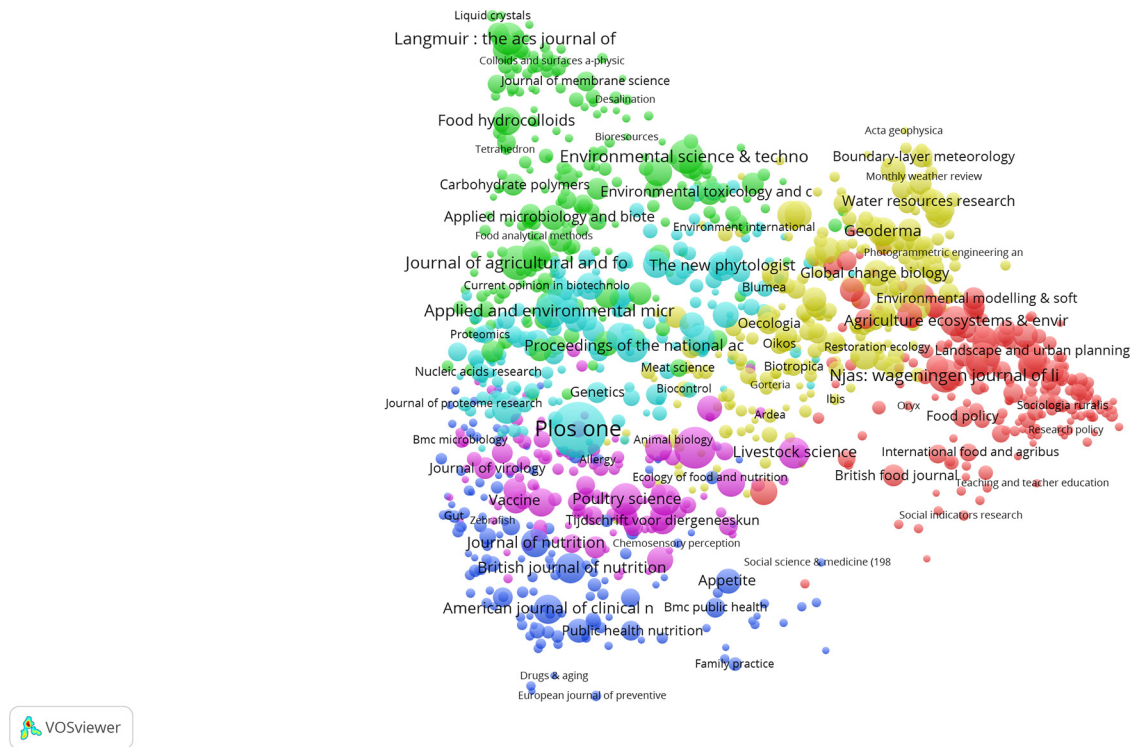


Figure 3 Network of 1,175 journals based upon a similarity analysis on citations to these journals from research articles published by WUR in 2006–2015 with mapping of the number of articles published per journal (score proportional to the size of a bubble)



articles with a low mean dissimilarity, further inspection of the data reveals that these articles have only one or a few references. The articles with a high dissimilarity have equal amounts of references as the other WUR articles, but show to have been published in multidisciplinary journals (e.g. *Proceedings of the National Academy of Sciences of the United States*, *Environmental Research Letters*, *Plant Science* or *Current Biology*).

The mean dissimilarity (\overline{D}_{gi}) per article i is averaged over all articles published in journal g resulting in the mean dissimilarity (\overline{D}_g) for journal g . This mean dissimilarity per journal is mapped as score value in the network that represents similarities between journals with more than two WUR articles published in 2006–2015 (Figure 3). Figure 5 shows this mapping.

The network in Figure 5 visualizes the relationships between the selected 1,175 journals based upon similarity in having been cited by the same journals in the WUR articles. The more two journals are placed together, the more similar they are. The size of the bubbles in Figure 5 is proportional to the size of the article output in each of the journals by WUR authors in 2006–2015. The red journals in Figure 5 have high values for the mean dissimilarity per journal.

Journals with high values for the mean dissimilarity per journal (more than 0.75) and a considerable number of WUR articles (more than 100 published between 2006 and 2015) are *Agriculture Ecosystems & Environment*, *The New Phytologist*, *Environmental Science & Technology*, *Proceedings of the National Academy of Sciences of the United States*, *Njas: Wageningen Journal of Life Sciences* and *Plos One*. The mean dissimilarity per journal for the top 20 of journals with the

largest number of WUR articles published in 2006–2015 is included in the last column of table I.

Discussion and conclusions

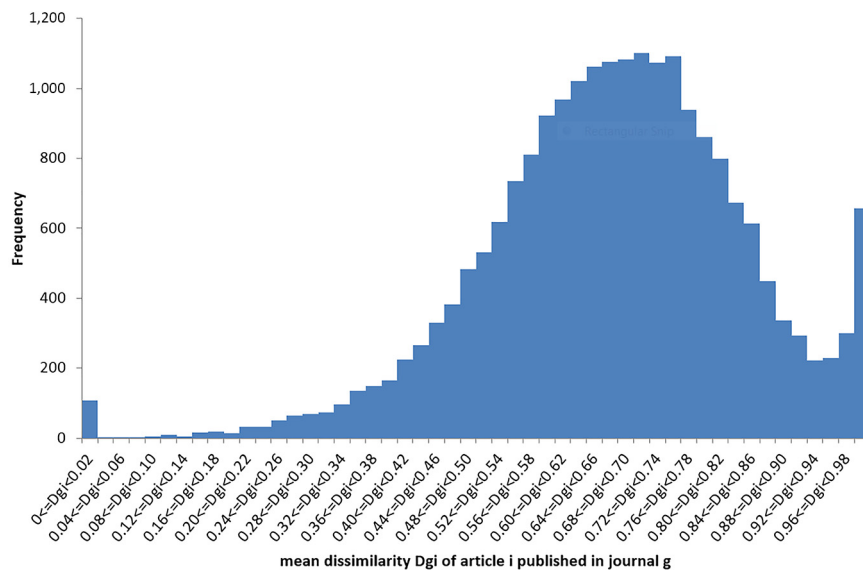
The methodology and analysis described in this paper is based upon WUR articles published in 2006–2015. By analyzing co-citations between journals over this 10-year time period, this paper strives to obtain reliable dissimilarity values between journals that are calculated based upon sufficient data. With derived similarity values, between a journal in which an article has been published and the journals cited in this article, this paper shows how a mean dissimilarity per article can be calculated.

The distribution of the mean dissimilarity per article shows there are 107 WUR articles with mean dissimilarities below 0.02. Further inspection of these articles reveals they each have a very short reference list (on average three references) in which only articles are included that are published in the same journal. Contrary to these WUR articles with very low mean dissimilarities there are 657 articles with a mean dissimilarity above 0.98. When these articles with high dissimilarities are further examined, it is found that their reference lists on average contain 39 references per article. This is close to the average number of references per article (38) for all WUR articles that have been analyzed in this paper. Of the 657 articles with high dissimilarities, there are 19 WUR articles with a dissimilarity of one. Further inspection of these articles reveals that they each have a short reference list (on average seven references) in

Table I Top 20 of journals with the largest number of WUR articles published in 2006–2015 with information on the number of articles, clustering and mean dissimilarity

Journal	No. of WUR articles in 2006–2015	Clustering according to network presented in Figure 2	Mean dissimilarity with journals cited from articles
<i>Plos One</i>	611	Plant sciences (light blue)	0.76
<i>Journal of Dairy Science</i>	282	Animal sciences (purple)	0.45
<i>Journal of Agricultural and Food Chemistry</i>	173	Chemical sciences and biotechnology (green)	0.64
<i>Environmental Science & Technology</i>	163	Chemical sciences and biotechnology (green)	0.83
<i>Applied and Environmental Microbiology</i>	160	Plant sciences (light blue)	0.56
<i>NJAS: Wageningen Journal of Life Sciences</i>	154	Agroecosystem sciences and agronomy (red)	0.81
<i>Agricultural Systems</i>	148	Agroecosystem sciences and agronomy (red)	0.61
<i>Langmuir : The ACS Journal of Surfaces and Colloids</i>	144	Chemical sciences and biotechnology (green)	0.53
<i>Proceedings of the National Academy of Science of the United States</i>	140	Plant sciences (light blue)	0.85
<i>Poultry Science</i>	136	Animal sciences (purple)	0.49
<i>Journal of Animal Science</i>	128	Animal sciences (purple)	0.55
<i>Livestock Science</i>	127	Animal sciences (purple)	0.64
<i>British Journal of Nutrition</i>	124	Health and nutrition (dark blue)	0.63
<i>BMC Genomics</i>	121	Plant sciences (light blue)	0.65
<i>Agriculture Ecosystems & Environment</i>	120	Agroecosystem sciences and agronomy (red)	0.89
<i>American Journal of Clinical Nutrition</i>	107	Health and nutrition (dark blue)	0.55
<i>Geoderma</i>	105	Environmental sciences (yellow)	0.57
<i>The New Phytologist</i>	105	Plant sciences (light blue)	0.85
<i>ICES Journal of Marine Science</i>	103	Environmental sciences (yellow)	0.47
<i>Animal : An International Journal of Animal Bioscience</i>	102	Animal sciences (purple)	0.61

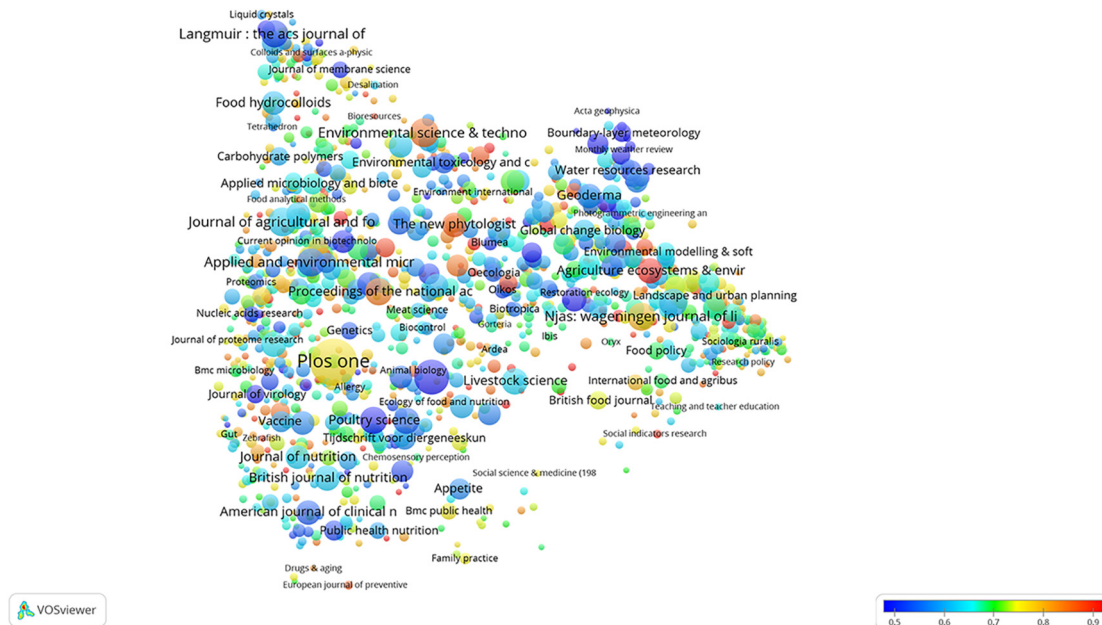
Figure 4 Distribution of the mean dissimilarity of 21,191 WUR articles calculated from dissimilarities between the journal in which each article is published and the journals from which articles have been cited in each WUR article



which only articles are included that are published in other journals than the journal in which the WUR article is published. The results described above show that the calculation of the mean dissimilarity per article is sensitive for the number

of cited articles included in the reference list. For articles with short reference lists extreme values of dissimilarities can be found because the calculation of these dissimilarities is based upon only a few cited articles that all are either

Figure 5 Network of 1,175 journals that contain more than two WUR articles published in 2006–2015 with mapping of the mean dissimilarity per journal (score corresponds with the color of a bubble) and the number of articles published per journal (score proportional to the size of a bubble)



published in the same journal or different journals as the journal in which the citing article has been published. To correct for this anomaly, a selection of 19,789 WUR articles that each contain at least ten references is made. These selected articles together contain 99 per cent of all citations to journal articles that are inventoried from the WUR articles that are published in 2006–2015 and analyzed in this paper.

To investigate whether high mean dissimilarities for articles co-occur with publication of these articles in multidisciplinary journals, the selected articles have been divided in ten classes based upon the deciles of their dissimilarities. Per class, the percentage of articles published in multidisciplinary journals have been calculated. For the identification of multidisciplinary journals, the category “Multidisciplinary sciences” with 63 journals from the InCites Journal Citations Reports (2017) have been used. Figure 6 shows the percentages of selected articles published in multidisciplinary journals per class (decile) of dissimilarities.

From Figure 6 it follows that the classes with articles that have low dissimilarities also have low percentages of articles published in multidisciplinary journals. For the class of articles with the highest dissimilarity (above 0.90), 11 per cent of the articles are published in multidisciplinary journals. These findings indicate that high values of dissimilarities for articles can be an indication that these articles have been published in multidisciplinary journals.

However, it also may be that an article describes interdisciplinary research, but is published in a journal that only represents a particular research field. To illustrate this, two articles from the journal *Ecology* (ISSN: 0012-9658) are selected and represented in Figure 7. In the Journal Citation Reports (InCites, 2017), the journal *Ecology* is classified in

the category “Ecology.” Also, the website of the journal ([Journal Ecology, 2017](#)) indicates that this journal publishes articles that present ecological research and ecological phenomena in particular.

From the journal *Ecology* two WUR articles are selected; one with a low dissimilarity (0.38) and the other with a high dissimilarity (0.99). For both articles, it is inventoried how many articles from which journals have been cited and to which categories in the Journal Citation Reports (InCites, 2017) these journals (with cited articles) belong. Figure 7 shows the shares and variety in the categories (research fields) to which these citations have been made from each article.

From Figure 7 it follows that although published in the same journals, the article with the high dissimilarity (right pie diagram) cites articles from more different categories or research fields (22 categories for the article with the high dissimilarity as opposed to four categories for the article with the low dissimilarity (left pie diagram)). For the article with the low dissimilarity 72 per cent of the citations are made to articles in the same category (research field) as the category (research field) in which the article has been published (i.e. Ecology). For the article with the high dissimilarity 72 per cent of the citations are made in 11 categories, thereby covering a larger variety of research fields.

It is also investigated whether in general from articles with higher dissimilarities more different categories (research fields) are represented by the journals from which articles have been cited by WUR authors. This inventory of categories have been made on 18,853 WUR articles published in non-multidisciplinary journals with each at least ten citations to journal articles. Figure 8 shows the mean number of different

Figure 6 Percentages of WUR articles (each with at least ten cited articles) published in multidisciplinary journals for ten classes of articles based upon deciles of the mean dissimilarity of each article

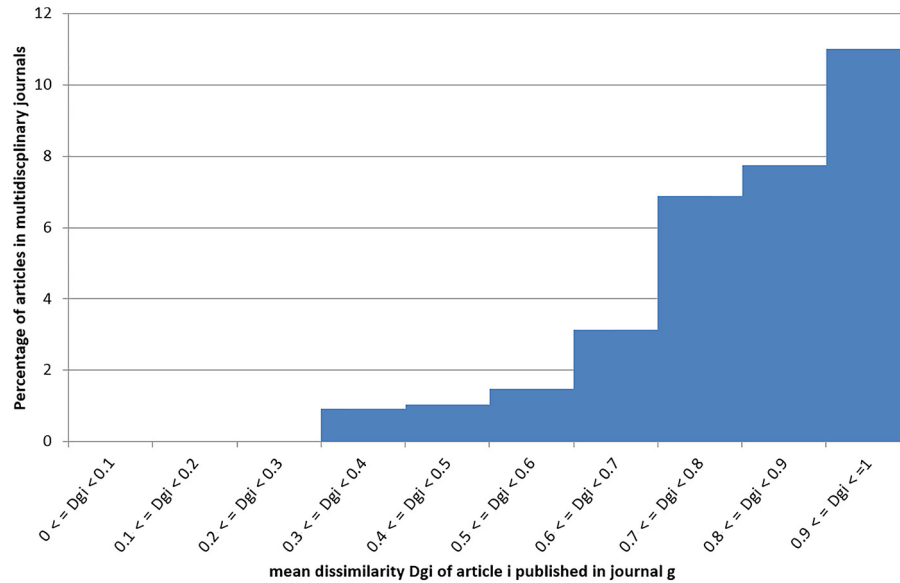


Figure 7 Shares of the categories (research fields) to which the journals (according to the Journal Citation Reports [InCites, 2017]) belong that contain the articles cited in a WUR article with a low dissimilarity (left) and a WUR article with a high dissimilarity (right). Both WUR articles have been published in the journal Ecology and respectively cite 31 and 35 articles in journals for which category information can be obtained from the Journal Citation Reports (InCites, 2017)

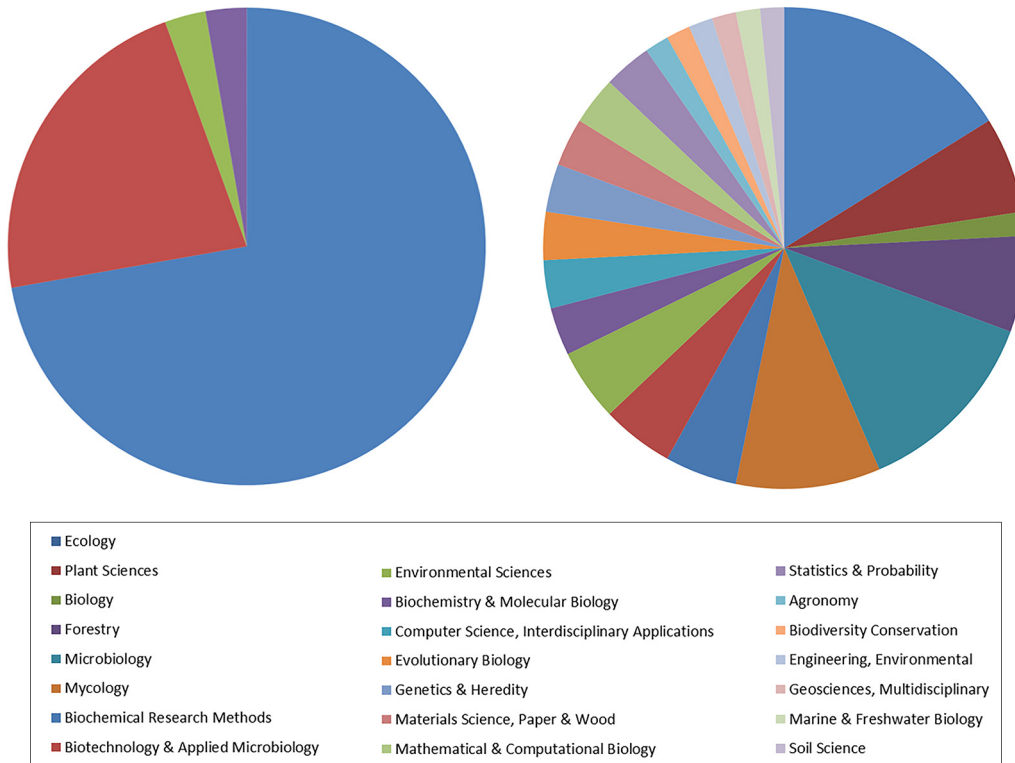
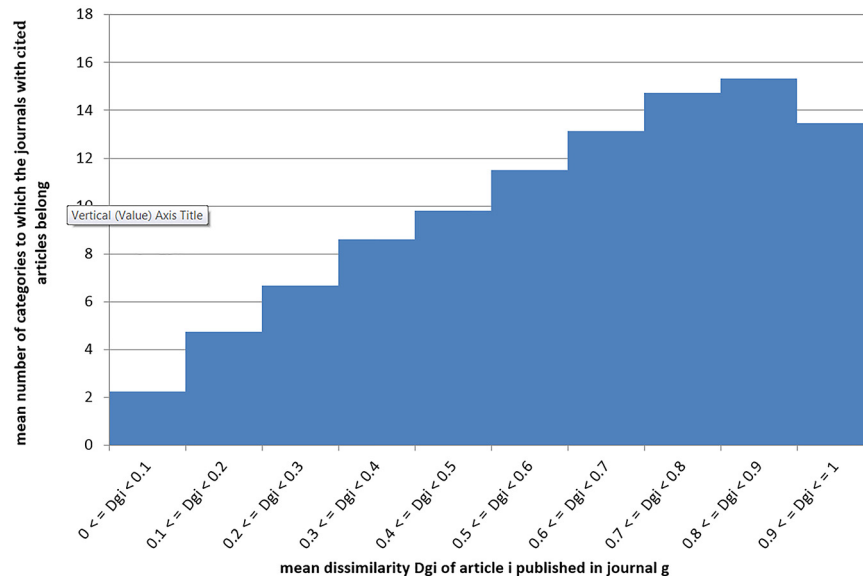


Figure 8 Mean number of categories (research fields) to which the journals with cited articles belong for ten classes of WUR articles from which the citations have been made. The WUR articles have been published in non-multidisciplinary journals and each cite at least ten journals articles. The classes are based upon deciles of the mean dissimilarities of the WUR articles



categories to which the journals with the cited articles belong for the 10 deciles (that also have been used in Figure 6) for the mean dissimilarities of WUR articles.

In accordance with the findings for the two WUR articles published in the journal *Ecology*, Figure 8 also shows that articles in journals that belong to more different categories (research fields) tend to be cited from WUR articles with higher dissimilarities. These findings support that for WUR articles with high dissimilarities the indications are stronger that they report on interdisciplinary research; either by publication of research in a multidisciplinary journal (Figure 6) or by publication of an interdisciplinary research in a non-multidisciplinary journal (Figures 7 and 8).

The results described in this paper show that the dissimilarities that are calculated for articles can be aggregated to calculate mean dissimilarities for journals. The mean dissimilarity per journal can be used to identify journals that have relatively larger shares of published articles that describe interdisciplinary research. The networks presented in Figures 3 and 5 show that for WUR the journals with higher shares of “interdisciplinary articles” especially can be found in the scientific disciplines of chemical sciences and biotechnology, plant sciences and agroecosystem sciences and agronomy. It may be interesting to identify journals that are important to authors for publishing interdisciplinary research findings in particular fields. When these fields are important for the research institution collection management of the library can be adapted to these findings.

Based upon the study described in this paper it is possible to identify WUR articles that to a higher extent report on the findings of interdisciplinary research. Hereby, calculated mean dissimilarities of the WUR articles can be used as a proxy for interdisciplinarity. Once more interdisciplinary

WUR articles have been identified, the RIS of WUR can be used to identify authors and research groups that have published these articles. Hereby, authors or research groups can be ranked, and additionally evaluated, on the level of interdisciplinarity of their research described in the WUR articles. Besides ranking also identification of authors or research groups that often use an interdisciplinary approach in their research might be interesting in the WUR policy for research strategies.

References

- Clarivate analytics (2017), “Web of science”, available at: <https://apps- webofknowledge-com> (accessed 28 February 2017).
- Eck, N.J.v. and Waltman, L. (2016), *VOSviewer Version 1.6.5. Center for Science and Technology Studies, Leiden University, Leiden*, available at: www.vosviewer.com/
- InCites (2017), “Journal citation reports”, available at: <http://jcr.incites.thomsonreuters.com> (accessed 10 April 2017).
- Journal Ecology (2017), available at: <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-ECY.html> (accessed 13 April 2017).
- Leydesdorff, L., Wagner, C.S. and Bornmann, L. (2017), “Betweenness and diversity in journal citation networks as measures of interdisciplinarity – a tribute to eugene garfield”, *Scientometrics*, Vol. 114 No. 2, pp. 567-592.
- R Core Team (2015), “R: a language and environment for statistical computing”, R Foundation for statistical Computing, Vienna, Austria, available at: www.R-project.org/
- Staff publications (2017), available at: <http://library.wur.nl/WebQuery/wurpubs/show> (accessed 13 April 2017).

- van Veller, M.G.P. (2013), "Analysis of journal usage by WUR staff members via article references", *Qualitative and Quantitative Methods in Libraries*, Vol. 2 No. 3, pp. 231-244.
- van Veller, M.G.P. and Gerritsma, W. (2015), "Development of a journal recommendation tool based upon co-citation analysis of journals cited in WUR research articles", *Qualitative and Quantitative Methods in Libraries*, Vol. 4 No. 2, pp. 233-257.
- WUR (2016), *Annual Report WUR 2015*, WUR, Wageningen, p. 122.

- Zhang, L., Rousseau, R. and Glänzel, W. (2016), "Diversity of references as an indicator for interdisciplinarity of journals: taking similarity between subject fields into account", *Journal of the Association for Information Science and Technology*, Vol. 67 No. 5, pp. 1257-1265.

Corresponding author

Marco van Veller can be contacted at: marco.vanveller@wur.nl

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com