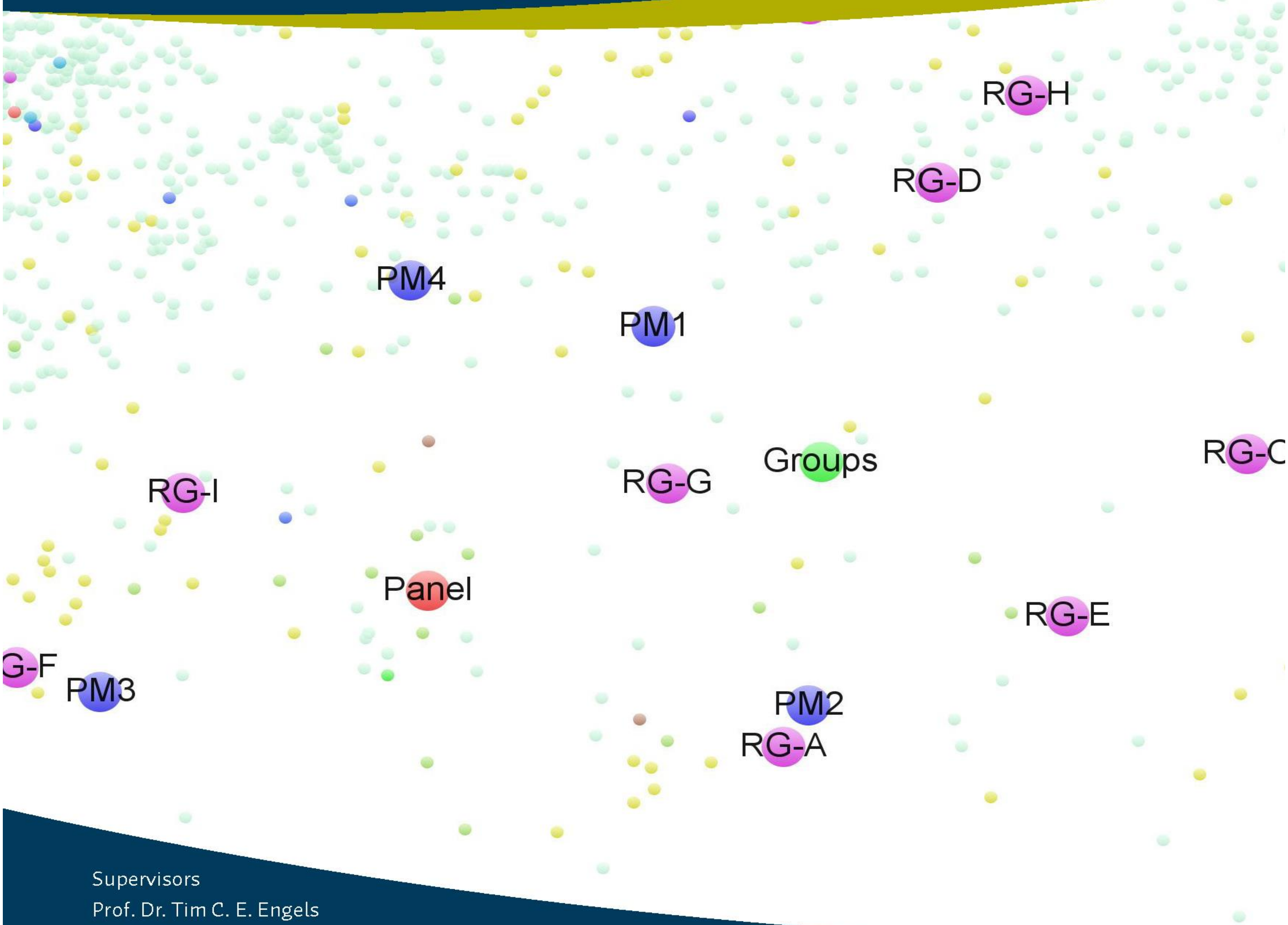
Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: An exploration of informetric methods

Thesis for the degree of doctor in Information and Library Science
at the University of Antwerp to be defended by

A. I. M. Jakaria Rahman



Supervisors
Prof. Dr. Tim C. E. Engels
Dr. Raf Guns

Faculty of Social Sciences
Centre for Research & Developing Monitoring (ECOOM)
Antwerp 2018

Universiteit Antwerpen

Faculty of Social Sciences

Centre for Research & Developing Monitoring (ECOOM)

Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: An exploration of informetric methods

Het bepalen van cognitieve afstand tussen publicatieportfolio's van evaluatoren en geëvalueerden in onderzoeksevaluaties: een verkenning van informetrische methoden

Thesis for the degree of doctor in Information and Library Science at the University of Antwerp to be defended by

A. I. M. JAKARIA RAHMAN

Supervisors:

Prof. Dr. Tim C. E. Engels
Dr. Raf Guns

Antwerp, 2018

The doctoral jury members

Prof. Dr. Sven De Maeyer (chair of the doctoral committee and of the jury)

Prof. Dr. Tim C. E. Engels (supervisor)

Dr. Raf Guns (supervisor)

Dr. Ronald Rousseau

Dr. Ulf Sandström

Prof. Dr. Elías Sanz Casado

Dedication

To my parents

Sabera Khatun and Md. Fazlur Rahman

Acknowledgements

This doctoral thesis presents the research that I have been working for the last four years. It would have not been possible to complete this thesis without support and encouragements from a number of people both at professional and personal level. Therefore, the list of people I would like to thank is probably endless, though I would like to express my gratitude explicitly to:

- Tim Engels and Raf Guns, my supervisors, for their expert guidance, encouragements, and patience. They have been an important source of motivation and inspiration to me. Thank them for their enthusiasm, fresh ideas and their trust in my capabilities.
- Ronald Rousseau, for interesting collaborations and his many insights and suggestions for improving the work.
- Loet Leydesdorff, for interesting collaboration and making the journal similarity data available.
- Nele Dexters and Birgit Houben, for showing me advanced features of Web of Science.
- Truyken Ossenblok, Frederik Verleysen and Linda Sile, my colleagues at the Centre for Research & Development Monitoring (ECOOM)-Antwerp.
- Danny Vindevogel, for English language suggestions.
- all colleagues of the Department of Research Affairs and Innovation (ADOC), for connecting and giving feedback on presentations.
- my conference friends, for inspiration and motivation.
- my parents and brothers, for their endless love and support. I would not have been where I am today.
- Momena Khatun, for her unwavering love, support, and patience at all times.
- Alifa Saberi Rahman, for her arrival in our life just in the middle of the research.
- and finally, last but by no means least, to the Flemish people as this research has been made possible by the financial support of the Flemish government to ECOOM.

English abstract

This doctoral thesis develops informetric methods for determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation. In a discipline specific research evaluation, when an expert panel evaluates research groups, it is an open question how one can determine the extent to which the panel members are in a position to evaluate the research groups. This thesis contributes to the literature by proposing six different informetric approaches to measure the match between evaluators and evaluees using their publications as a representation of their expertise.

An expert panel is specifically appointed for the research evaluation. Experts are typically selected in one of two ways: (1) straightforward selection: the person(s) in charge of the research evaluation has access to a list of acknowledged experts in specific fields, and limits its selection process to ensuring the experts' independence regarding the program under evaluation; and (2) gradual selections: preferred profiles of experts are developed with respect to the specialization under scrutiny in the evaluation. Both ways leave some freedom for an "old boys' network" to appoint someone without properly evaluating their qualifications. There are also other ways for expert selection, for example, inviting open application or the research groups that will be evaluated can propose their choice of experts. In research evaluation, an expert panel usually comprises independent specialists, each of which is recognized in at least one of the fields addressed by the unit under evaluation. The expertise of the panel members should be congruent with the research groups to ensure the quality and trustworthiness of the evaluation. All things being equal, panel members who are credible experts in the field are also most likely to provide valuable, relevant recommendations and suggestions that should lead to improved research quality. However, there was an absence of methods to determine the cognitive distance between evaluators and evaluees in research evaluation when we started working in July 2013.

In this thesis, we develop and test informetric methods to identify the cognitive distances between the (members of) an expert panel on the one hand, and the (whole of the) units of assessment (typically research groups) on the other. More generally, we introduce a number of

methods that allow measuring cognitive distances based on publication portfolios. In academia, publications are considered key indicators of expertise that help to identify qualified or similar experts to assign papers for review, and to form an expert panel. Our main objective is to propose informetric methods to identify panel members who have closely related expertise in the research domain of the research groups based on their publications profile. The main factor that we have taken into account is the cognitive distance between an expert panel and research groups. We consider the publication portfolio of the involved researchers to reflect the position of the unit in cognitive space and, hence, to determine cognitive distance. Expressed in general terms we measure cognitive distance between units based on how often they have published in the same or similar journals. Our investigations lead to the development of new methods of expert panel composition for the research evaluation exercises.

We explore different ways of quantifying the cognitive distance between panel members and research group's publication profiles. We consider all the publications of the research groups (during the eight years preceding their evaluation) and panel members indexed in Web of Science (WoS). We pursue the investigation at two levels of aggregation: WoS subject categories (SCs) and journals. The aggregated citation relations among SCs or journals provide a matrix. From the matrix, one can construct a similarity matrix. From the similarity matrix, one can construct a global SCs or journal map in which similar SCs or journals are located more closely together. The maps can be visualized using a visualization program. During the visualization process, a multi-dimensional space is reduced to a projection in two dimensions. In this process, similar SCs or journals are positioned closer to each other.

We propose three methods, namely the use of barycenters, of similarity-adapted publication vector (SAPV) and of weighted cosine similarity (WCS). We take into account the similarity between WoS SCs and between journals, either by incorporating a similarity matrix (in the case of SAPV and WCS) or a 2-dimensional base map derived from it (in the case of barycenters). We determine the coordinates of barycenters using a 2-dimensional base map based on the publication profiles of research groups and panel members, and calculate the Euclidean distances between the barycenters. We also identify SAPV using the similarity matrix and calculated the Euclidean distances between the SAPVs. Finally, we calculate WCS using the similarity matrix. The SAPV and WCS methods use a square N-dimensional similarity matrix. Here N is equivalent to 224

WoS SCs and 10,675 journals. We used the distance/similarity between panel members and research groups as an indicator of cognitive distance. Small differences in Euclidean distances (both between barycenters and SAPVs) or in cosine similarity values bear little meaning. For this reason, we employ a bootstrapping approach in order to determine a 95% confidence interval (CI) for each distance or similarity value. If two CIs do not overlap, difference between the values is statistically significant at the 0.05 level. Although it is possible for two values to have a statistically significant difference while having overlapping CIs, the difference is less likely to have practical meaning.

Two levels of aggregation and three methods lead to six informetric approaches to quantify the cognitive distance. Our proposed approaches hold advantages over a simple comparison of publication portfolios. Our approaches quantify the cognitive distance between a research group and panel members. We also compare our proposed approaches. We examine which of the approaches best reflects the prior assignment of main assessor to each research group, how much influence the level of aggregation (journals and WoS SCs) plays, and how much the dimensionality matters. The results show that, regardless of the method used, the level of aggregation has only a minor influence, whereas the influence of the number of dimensions is substantial.

The results also show that the number of dimensions plays a major role in the case of identifying shortest cognitive distance. While the SAPV and WCS methods agree at most of cases at both the levels of aggregation the barycenter approaches yield different results. We find that the barycenter approaches score highest at both levels of aggregation to identify the previously assigned main assessor. When it comes to uniquely identifying the main assessor, all methods score better at the journal level than at the WoS SC level. Our approaches, but of course not the numerical result, are independent of the similarity matrix or map used.

All six approaches give the opportunity to assess the composition of the panel in terms of cognitive distance if one or more panel members are replaced and compare the relative contribution of each potential panel member to the panel fit as a whole, by observing the changes to the distance between the panel's and the groups'. In addition, our approaches allow the panel composition authority to see in advance about the panel's fit to the research groups that are going to be evaluated. Therefore, the concerned authority will have the opportunity to replace outliers among the panel members to make the panel fit well with the research groups to be evaluated. For example, the authority can find a best-fitting expert panel by replacing a more distant panel member with a potential panel member located closer to the groups.

Keywords

Barycenter · Bootstrapping · Cognitive distances · Confidence intervals · Expert panel · Journal overlay map · Matching research expertise · Overlay maps · Research evaluation · Similarity matrix · Similarity-adapted publication vector · Web of Science subject categories · overlay map · Weighted cosine similarity.

Nederlandstalig abstract

In dit proefschrift ontwikkelen we informetrische methoden om de cognitieve afstand te bepalen tussen publicatieportfolio's van evaluatoren en geëvalueerden bij onderzoeksevaluatie. In evaluaties op disciplineniveau, waarbij een expertpanel onderzoeksgroepen evalueert, is het een open vraag hoe we kunnen bepalen of de panelleden geschikt zijn om de onderzoeksgroepen te beoordelen. Dit proefschrift stelt zes informetrische benaderingen voor waarmee de inhoudelijke congruentie tussen evaluatoren en geëvalueerden kan worden gemeten; hierbij gebruiken we hun publicaties als een representatie van hun expertise.

Expertenpanels worden specifiek voor de evaluatie aangesteld. Experts worden doorgaans op een van de volgende twee manieren uitgekozen: (1) rechtstreekse selectie: de verantwoordelijken voor de onderzoeksevaluatie gebruiken een lijst van erkende experts in specifieke velden en zorgen er in het selectieproces vooral voor dat de experts onafhankelijk zijn van het te evalueren programma, en (2) graduele selecties: voorkeursprofielen van experts worden ontwikkeld voor de specalisaties die beoordeeld moeten worden. Beide manieren laten echter de ruimte voor een "old boys' network" om iemand aan te stellen zonder dat zijn/haar kwalificaties voldoende onderzocht zijn. Er bestaan ook andere manieren om experts aan te stellen, bijvoorbeeld door spontane kandidaatstellingen aan te moedigen, waarna de te evalueren onderzoeksgroepen hun voorkeur kunnen laten blijken.

Het expertenpanel bestaat gewoonlijk uit onafhankelijke onderzoekers die elk gespecialiseerd zijn in minstens één van de specialismen van het te evalueren programma. Algemeen gesproken doen panelleden die tevens experts zijn in een bepaald veld ook de meest waardevolle en relevante aanbevelingen en suggesties, die tot een verbetering van de kwaliteit van het onderzoek kunnen leiden. Toen we in juli 2013 ons onderzoek aanvatten, waren er echter geen methoden voorhanden waarmee men de cognitieve afstand tussen evaluatoren en geëvalueerden kan bepalen.

We ontwikkelen en testen informetrische methoden om cognitieve afstand te operationaliseren. Deze cognitieve afstanden zijn met name afstanden tussen (de leden van) een expertenpanel enerzijds en (het geheel van) de te evalueren eenheden – doorgaans onderzoeksgroepen – anderzijds. Meer algemeen stellen we methoden voor waarmee men cognitieve afstanden kan bepalen op basis van publicatieportfolio's. In de academische wereld zijn publicaties immers belangrijke indicaties van expertise en kunnen ze bijgevolg helpen om de best geplaatste experts te identificeren. Ons hoofddoel is informetrische methoden voor te stellen waarmee we die panelleden kunnen identificeren die over expertise beschikken die nauw verwant is aan die van de onderzoeksgroepen. De belangrijkste factor waarmee we rekening houden, is de cognitieve afstand tussen expertenpanel en onderzoeksgroepen. We gaan ervan uit dat het publicatieportfolio van de betrokken onderzoekers de positie van de eenheid in de cognitieve ruimte weerspiegelt en aldus kan worden gebruikt om cognitieve aftand te bepalen. In algemene termen gesteld meten we cognitieve afstand tussen eenheden door te bepalen hoe vaak ze in dezelfde of gelijkaardige tijdschriften hebben gepubliceerd. Ons onderzoek leidt tot de ontwikkeling van nieuwe methoden voor de samenstelling van expertenpanels bij onderzoeksevaluatie.

We verkennen verschillende manieren om cognitieve afstand tussen publicatieportfolio's te kwantificeren en maken daarbij gebruik van alle publicaties van de onderzoeksgroepen (uit de acht voorgaande jaren) en panelleden die in Web of Science (WoS) zijn geïndexeerd. De analyse gebeurt op twee aggregatieniveaus: WoS onderwerpscategorieën (*subject categories* of SCs) en tijdschriften. De geaggregeerde citatierelaties tussen SCs of tijdschriften kunnen in een matrix worden samengevat. Vanuit deze citatiematrix kan een similariteitsmatrix worden gemaakt. Op basis van de similariteitsmatrix kunnen we een SC- of tijdschriftenkaart of -map maken waarop gelijkaardige SCs of tijdschriften zich dichter bij elkaar bevinden. Het visualisatieproces houdt in dat de multidimensionale ruimte wordt geprojecteerd naar twee dimensies. In dit proces worden gelijkaardige SCs of tijdschriften dichter bij elkaar geplaatst.

We stellen drie methoden voor, namelijk het gebruik van barycentra, similariteitsgeadapteerde publicatievectoren (SAPV) en gewogen cosinussimilariteit (WCS). We nemen de similariteit tussen SCs of tijdschriften mee, hetzij door een similariteitsmatrix te gebruiken (in het geval van SAPV en WCS) hetzij door een tweedimensionale map te gebruiken die van de matrix is afgeleid (in het geval van barycentra). We bepalen barycentra door een tweedimensionale map te combineren met de publicatieportfolio’s van de onderzoeksgroepen en panelleden, en berekenen de Euclidische afstand tussen de barycentra. SAPVs worden bepaald met behulp van de similariteitsmatrix; vervolgens berekenen we de Euclidische afstand tussen SAPVs. Tot slot berekenen we de WCS met behulp van de similariteitsmatrix. De similariteitsmatrix heeft N dimensies, waarbij in ons geval N gelijk is aan 224 in het geval van SCs en aan 10,675 in het geval van tijdschriften. De afstand of (cosinus)similariteit tussen panelleden en onderzoeksgroepen beschouwen we als een indicator van cognitieve afstand.

Kleine verschillen in afstand of similariteit hebben weinig betekenis. Om deze reden hanteren we een aanpak met *bootstrapping* om een 95% betrouwbaarheidsinterval (*confidence interval* of CI) voor elke afstand/similariteit te bepalen. Indien twee CIs niet overlappen, is het verschil tussen de waarden statistisch significant op het 0,05 niveau. Hoewel het mogelijk is dat het verschil tussen twee waarden met overlappende CIs statistisch significant verschillen, is het minder waarschijnlijk dat het verschil in de praktijk betekenisvol is.

Twee aggregatieniveaus en drie methoden leiden tot zes manieren waarmee cognitieve afstand kan worden gekwantificeerd. Onze voorgestelde methoden hebben duidelijke voordelen ten opzichte van een eenvoudige vergelijking van publicatieportfolio’s, doordat ze de cognitieve afstand tussen groep en panel beter benaderen. We vergelijken de voorgestelde manieren ook onderling. We onderzoeken welke aanpak het best overeenstemt met de toewijzing van hoofdbeoordelaar aan onderzoeksgroep en welke rol zowel aggregatieniveau als dimensionaliteit spelen. De resultaten geven aan dat, ongeacht de gebruikte methode, het aggregatieniveau slechts een beperkte invloed heeft, maar dat met name het aantal dimensies een aanzienlijke invloed uitoefent. De resultaten geven ook aan dat het aantal dimensies een grote rol speelt bij de vraag welke cognitieve afstand de kortste is. Terwijl de SAPV- en WCS-methoden op beide aggregatieniveaus gelijkaardige resultaten opleveren, geeft de aanpak met barycentra vaker andere resultaten. Op beide aggregatieniveaus scoren de barycentra beter dan de andere methoden bij het bepalen van de eerder toegewezen hoofdbeoordelaar. Wanneer de hoofdbeoordelaar uniek moet worden aangeduid, scoren alle methoden beter op het niveau van tijdschriften dan op dat van SCs. Onze manieren zijn onafhankelijk van de gebruikte similariteitsmatrix of map.

Alle zes benaderingen geven de mogelijkheid om de samenstelling van het panel te vergelijken met alternatieve constellaties waarbij een of meer panelleden vervangen worden en om de relatieve bijdrage van ieder potentieel panellid aan het panel als geheel te beoordelen, door wijzigingen in de afstand tussen panel en groep na te gaan. Bovendien laten onze benaderingen toe om reeds op voorhand de overeenstemming van het panel met de onderzoeksgroepen te bekijken. De autoriteit die de evaluaties uitvoert, heeft dan ook de mogelijkheid om eventuele buitenbeentjes in het panel te vervangen en het panel zo goed mogelijk bij de te evalueren onderzoeksgroepen te laten aansluiten. De autoriteit kan er bijvoorbeeld toe besluiten een veraf gelegen panellid te vervangen door een potentieel lid dat zich dichter bij een of meerdere onderzoeksgroepen bevindt.

Content

List of Tables

List of Figures

Chapter I: Introduction

1.1 Introduction

Research evaluation exercises are carried out in many countries and regions across the world including the UK, Norway, Finland, Sweden, Denmark, the Netherlands, Belgium (both Flanders and Wallonia), Italy, Spain, Germany, Czech Republic, Romania, Japan, Hong Kong (China), Australia, New Zealand, and USA (Barker, 2007; Papponetti & Bucchi, 2007; Molas-Gallart, 2012; Simon & Knie, 2013; McKenna, 2015; Milat, Bauman, & Redman, 2015). Discipline-specific research evaluations carried out by panels of peers are a common practice at many universities worldwide. Expert panel review is a standard practice for evaluating research groups (Nedeva, Georghiou, Loveridge, & Cameron, 1996; Butler, 2007; Rons, Bruyn, & Cornelis, 2008; Lawrenz, Thao, & Johnson, 2012; Milat et al., 2015), and for research proposals submitted to research funding organizations (Wessely, 1998; van den Besselaar & Leydesdorff, 2009; Li & Agha, 2015; Pina, Hren, & Marušić, 2015; Wang & Sandström, 2015). In some cases, the evaluations include site visits by an expert panel. The expert panel also files a report. The panels may be required to follow a systematic research evaluation procedure in accordance with an evaluation protocol, for example, the Standard Evaluation Protocol in the Netherlands (VSNU, KNAW, & NWO, 2014). The panel arrives at conclusions and recommendations, preferably through consensus, and provides guidelines for the improvement of research quality and research policy based on the assessments. Depending on the research group or proposal, these recommendations deal with the implementation or have an impact on a program continuation, or part of it. The assessments vary from one to another (Lawrenz et al., 2012) but usually focus on current performance and on future plans and potentials (Hansson, 2010). The principal objective of research groups evaluations is to improve the quality of scientific research groups or departments within a national or regional context (Engels, Goos, Dexters, & Spruyt, 2013). In some cases, in addition to reviewing the quality of research the goal of the evaluation is to allocate research funding on the basis of past performance (Hicks, 2012; Cattaneo, Meoli, & Signori, 2016).

In research evaluation, an expert panel usually comprises independent specialists, each of which is recognized in at least one of the fields addressed by the program under evaluation. The expertise of the panel members should be congruent with the research groups to uphold the quality and trustworthiness of the evaluation. When an expert panel evaluates research groups, it is an open question how one can determine the match between panel members and research groups. This 'match' will be further on refined using the concept of cognitive distance. In this thesis, we aim to develop and test informetric methods to identify the cognitive distances between the (members of) an expert panel on the one hand, and the (whole of the) units of assessment (typically research groups) on the other. More generally, we introduce a number of methods that allow measuring cognitive distances based on publication portfolios. We propose six different informetric approaches according to entities' publication output to measure the match between evaluators and evaluees.

In this chapter, we will introduce the topic of the thesis. First, we briefly discuss the two main concepts in the thesis: peer review and cognitive distance, and how they are used here. In addition, we state the research problem, the purpose of the study, formulate the research questions, and indicate in which chapter they are answered. Moreover, as the data from expert panel evaluations at the University of Antwerp has been used, we briefly discuss this particular evaluation process as an example.

1.1.1 Peer review process

Peer review is an established component of professional practice in research, scholarly communication, the academic reward system, etc. (Lee, Sugimoto, Zhang, & Cronin, 2013). This process is particularly applied to the evaluation of scientific work by one or more persons to maintain standards of quality. In scientific journals, peer review is a key component to determine an article's suitability for publication, scientific accuracy, originality, interest to the journal's readers, contributes to clarity and improvement of the quality of a manuscript, identify any errors or misinterpretations (Nicholas et al., 2015). In the case of article level peer review, often the reviewer is anonymous to the authors of the evaluated publication. When one or more individuals carry out peer review, we refer to it as individual evaluation. Although multiple individuals may evaluate the same thing, they carry out peer review as individuals and without communication with the other reviewers. This kind of peer review is most commonly used for publications.

On the other hand, panel evaluation (Coryn & Scriven, 2008; Abramo & D'Angelo, 2011) refers to a panel of experts working together in their evaluation of, e.g., a research group, an institution or a research grant application (ESF, 2011; Boyack, Chen, & Chacko, 2014). Contrary to individual evaluation, this kind of peer review presupposes frequent contact and communication between the evaluators. It may include site visits by the expert panel members (Borum & Hansen, 2000; Hansson, 2010; Lawrenz et al., 2012). Mixed forms of both types occur frequently. In expert panel evaluation, however, the panel members are visible, and hence the units of assessment themselves can judge the expertise of the panel member and the expert panel in relation to their research domain. In this thesis, we focus on peer review in the form of expert panel evaluations.

1.1.2 Cognitive distances

The concept of cognitive distance has been developed in the academic literature by Nooteboom and colleagues (Nooteboom, 1999, 2000; Nooteboom, Van Haverbeke, Duysters, Gilsing, & van den Oord, 2007). Cohen & Levinthal (1989, 1990) explained the process by which an individual or organization, by extrapolation can integrate and reuse knowledge from outside sources in research and development, while Nooteboom uses these ideas to define the concept of cognitive distance between individuals and organizations. Nooteboom (2000, p. 73) defines cognitive distance as "a difference in cognitive function. This can be a difference in domain, range, or mapping. People could have a shared domain, but a difference of mapping: two people can make sense of the same phenomena, but do so differently". Thus, cognitive distance describes how two individuals – and, by extension, organizations or groups of individuals – are different, in terms of knowledge, but also in the way they perceive and interpret external phenomena.

The concept of 'cognitive distance' and 'cognitive proximity' is already in use in information science and informetrics literature. Hautala (2013) stated that cognitive proximity is achieved through cooperation and suitable tasks for knowledge creation between international research groups. Science overlay maps have been in use in the informetrics literature to assess the degree of similarity and dissimilarity between research profiles (Boyack, 2009; Rafols, Porter, & Leydesdorff, 2010; Soós & Kampis, 2012). Boyack et al., (2014) used overlay maps to compare the locations of reviewer publications for four expert panels on a base map of science, for the purpose of evaluating a set of grant applications. Molas-Gallart et al. (2015) explored a proximity-based approach in which translational research takes place and what factors may promote collaboration between the different actors who facilitate communication across different communities. A more quantitative way was introduced by Wang & Sandström, (2015), who used bibliographic coupling and topic modelling. We have proposed a number of methods to determine the cognitive distances between publication portfolios of researchers and panel members (Rahman, Guns, Rousseau, & Engels, 2015; Rahman, Guns, Leydesdorff, & Engels, 2016; Rousseau, Guns, Rahman, & Engels, 2017), which are part of this thesis.

1.1.3 Problem statement

The expert panel is specifically appointed for the evaluation. An expert panel usually comprises independent specialists, each of which is recognized in at least one of the fields addressed by the program under evaluation. Experts are typically selected in one of two ways: (1) straightforward selection: the person(s) in charge of the research evaluation has access to a list of acknowledged experts in specific fields, and limits its selection process to ensuring the experts' independence regarding the program under evaluation; and (2) gradual selections: preferred profiles of experts are developed with respect to the specializations under scrutiny in the evaluation. Both these ways leave some freedom for an "old boys' network" to appoint someone without properly evaluating their qualifications (Goldfinch & Yamamoto, 2012). There are also other ways for expert selection, for example, inviting open application or the research groups that will be evaluated can propose their choice of experts. A downside of the peer review process can be the absence of an adequate methodology to find relevant experts (Hofmann, Balog, Bogers, & de Rijke, 2010; Gould, 2013; Berendsen, de Rijke, Balog, Bogers, & Bosch, 2013; Lee et al., 2013; Oleinik, 2014; Buckley, Sciligo, Adair, Case, & Monks, 2014). In research evaluation, the extent to which the expertise of the panel members charged with research assessment is congruent with the research of the units, is crucial to the quality and trustworthiness of the assessment. In addition, the panel members taken together preferably have expertise on the discipline of the research groups; otherwise the trustworthiness of the evaluation is open for discussion. Panel members who are credible experts in the field will be able to provide valuable, relevant recommendations and suggestions that should lead to improved research quality. In practice, however, the degree to which the expertise of the panel members overlaps with the expertise of the unit of assessment has not been quantified (Engels et al., 2013).

The exponential growth of research literature indicates the growth of specialized disciplines (Sobkowicz, 2015) besides the growth of databases themselves. Therefore, an individual panel member may have sufficient expertise in a given field, but collaborative evaluation together with peers is crucial unless and until the individual panel member covers the expertise of the research groups. In this respect, Langfeldt (2004) explored expert panel evaluation and decision making processes, and concluded that overlap of expertise between experts is highly needed in order to foster cooperation among panel members. Tasks are often divided between the panel members according to their field of research. When the panel is heterogeneous, there is a chance that decisions are made in a suboptimal manner and/or not by those panel members that are best suited, i.e. at the closest cognitive distance. Moreover, each research group expects its research interests to be well covered by the expertise of at least one panel member. A sufficiently high degree of congruence between the expertise of the panel members charged with research assessment and the research of the units is a prerequisite for a sound, reliable assessment (Engels et al., 2013). However, to the best of our knowledge, no methods have been established to measure and quantify congruence of expertise or overlap of expertise or cognitive distance between panels and the units of assessment in discipline-specific research evaluation. The main motivation for the work in this thesis is, therefore, the need for a method to find appropriate set of experts and to determine an appropriate expert assignment to research groups in a discipline-specific research evaluation.

1.2 Purpose of the study

We study the problem of composing an expert panel, such that the individual panel members' expertise covers the research domains in the discipline where the units of assessment (in our case: research groups) have publications. In academia, publications are considered key indicators of expertise (Rybak, Balog, & Nørvåg, 2014) that help to identify qualified or similar experts to assign papers for review (Neshati, Beigy, & Hiemstra, 2012), and to form an expert panel (Hashemi, Neshati, & Beigy, 2013). One of the main factors that need to be taken into account is the cognitive distance between an expert panel and research groups (Rahman et al., 2015, 2016; Wang & Sandström, 2015; Rousseau et al., 2017; Rahman, Guns, Rousseau, & Engels, 2017).

In this thesis, we explore informetric approaches to determining cognitive distances between expert panel members and research groups for discipline-specific research evaluation. We specifically focus on the situation where the expert panel needs to evaluate all the research groups of a department. We attempt to gauge cognitive distances between panel and research groups through publishing in the same or similar Web of Science subject categories (WoS SCs) and journals. The goal is therefore to present informetric methodologies to assess the congruence of panel expertise and the research output in the units under assessment.

1.2.1 Objectives of the study

The main objective of this thesis is to propose informetric methods to identify panel members who have closely related expertise in the research domain of the research groups. We aim to develop and test informetric methods to identify the cognitive distances between the (members of) expert panel on the one hand, and between the expert panel and the (whole of the) units of assessment (typically research groups) on the other. More generally, the objective is to develop one or more methods that allow measuring cognitive distances between publication portfolios and understand their properties and interrelations. Our research has led to the development of new methods for determining cognitive distance in the context of expert panel composition.

1.2.2 Research questions

To achieve the objectives of this research the following research questions are formulated:

Research question 1

How can we measure cognitive distance between two entities using publication data, especially between an expert panel and the research groups under evaluation?

As a preliminary exploration, we determine the correlation between the publication output of two expert panel and research groups using Pearson's correlation coefficient, Spearman's rank correlation coefficient, top-down correlation (Iman & Conover, 1987) and cosine similarity (Salton & McGill, 1986). We argue that correlation and similarity measures are insufficient to measure cognitive distance, as they do not consider the relatedness of WoS SCs or journals (discussed in chapter IV). Therefore, in order to answer this first main research question, we formulate the following four sub-questions:

- i) **How can one visualize the expertise of two entities (e.g., a research group and a panel) using publication data?**
- ii) **How can one quantify the cognitive distances (overlap of expertise) between two entities (e.g., a research group and a panel) using the WoS SCs to which their publications belong?**
- iii) **How can one quantify the cognitive distances between two entities using the journals in which they have published?**
- iv) **How can one estimate the uncertainty inherent to these cognitive distances?**

In order to answer the first sub-question, we explore the usefulness of overlay mapping, using two base maps of science, one at the level of WoS SCs and the other at the level of journals as a starting point. We create overlay maps of individual panel members, the whole evaluation panel, individual research groups and the combined research groups to visually compare the cognitive distance between them, thus following earlier work in overlay mapping (Rafols et al., 2010; Leydesdorff, Carley, & Rafols, 2013). The visual comparison based on WoS SCs is discussed in detail in chapter V. The overlay maps based on journals are not discussed explicitly in the body of the thesis but are included in the technical reports that are available online (see section 3.2.8 of chapter III).

To answer the second sub-question, determine an entity's barycenter according to its publications output, we introduce the use of the barycenter method on maps of science (at the level of WoS SCs). The barycenter provides a single point on the map corresponding to an entity's expertise as witnessed by its publication output (see chapter V). The distances between two barycenters were then considered as an operationalization of cognitive distance between the entities involved. In chapter V and the corresponding publication (Rahman et al., 2015), we used the terminology of overlap of expertise. Later, in the PhD project, we came to realize that the term cognitive distance is more appropriate. For the sake of consistency with the published articles, we keep the terminology per chapter as in the articles. This explains the difference in terminology between chapter V and the subsequent chapters.

Further work led to the realization that approaches need not be confined to two dimensions only. In the end, this led to the development of five methods: a benchmark that does not take similarity into account, two methods using barycenters (one in two and one in three dimensions), similarity-adapted publication vectors (SAPV) and weighted cosine similarity (WCS). The benchmark and the last two methods are applied in N-dimensions, where N denotes the total number of WoS SCs. A theoretical comparison between the methods is discussed in chapter VII.

The third sub-question is in fact quite similar to the second one, but at a lower level of aggregation (journals instead of WoS SCs). Hence, our approach to answer this question is also similar. More specifically, we outline three quantitative methods that determine the cognitive distance between evaluators and evaluees, using the journals they have published in. We consider the barycenter, SAPV and WCS method for this purpose. The application of the barycenter and SAPV methods at the aggregation level of journals is discussed in chapter VI, while the result of the WCS method is included in the technical reports (see section 3.2.8 of chapter III).

To answer the fourth sub-question, we use a bootstrapping method, leading to confidence intervals for distances (benchmark, barycenter two-dimensional, barycenter three-dimensional, and SAPV methods) and similarities (WCS method). We discuss the bootstrapping method in chapter VI and chapter VII.

Research question 2

How do the proposed approaches relate?

In order to answer the second main research question, three sub-questions need to be addressed:

i) **What are the correlations between the different approaches? Which aspect (method vs level of aggregation) has the largest influence on the correlation?**

ii) **To what extent do the approaches agree in matching the panel member at the closest cognitive distance from a research group?**

iii) **How accurate are the approaches in matching the main assessor for each research group? How accurate are they to *uniquely* match the main assessor?**

In answer to the first research question, we propose a number of different approaches. We can distinguish between two levels of aggregation – journals and WoS SCs – and three methods – the two-dimensional barycenter method and the N-dimensional SAPV and WCS methods. In total, this leads to six different approaches, all of which are based on the publication profile of research groups and panel members.

To answer the first sub-question, we calculate Spearman's rank-order correlation between the results/values of each pair of the six approaches. We create a heat map with hierarchical clustering based on the correlation results for a visual summary of the results. The hierarchical clustering directly shows which approaches are more closely related.

To answer the second sub-question, we explore whether the approaches agree regarding the first ranked panel member, ignoring the confidence intervals.

To answer the third sub-question, we compare the closest panel member with the main assessor while accounting for overlap between confidence intervals. Here the main assessor refers to the panel members assigned to a research group by the respective panel chair during the research evaluation exercises at the University of Antwerp. We systematically compare how these six approaches fare in predicting the main assessor of each research group. We discuss the above sub-questions in chapter VIII.

1.2.3 Scope of the study

In this thesis, we are especially concerned with the case of expert panel evaluation of research groups. Here the research groups' publications profile covers a period of 8 years and the panel members' publication profile covers the entire publication profile until the year of the evaluation. Therefore, our proposed methods are best suited to the evaluation that covers a longer period with a larger set of publications. The approaches are not tied to any specific map or matrix but can, at least in principle, be applied to any map or similarity matrix.

Our approaches may not be suitable to identify experts who are invited to assess individuals, grants, individual research projects, or to review journal articles. In these cases, the required expertise is more at the topical level than at the discipline level. One can have publications on many topics, but due to the level of granularity both the WoS SC and journal maps and matrices are at the level of disciplines or specializations. Therefore, the proposed approaches cannot identify cognitive distance at the topical level with the data used in this thesis.

1.3 Expert panel evaluation: Example from the University of Antwerp

In this thesis, we particularly look at the research evaluations carried out at the University of Antwerp by its Department of Research Affairs and Innovation (ADOC). The overall annual research output of the University of Antwerp comprises over 2,000 peer-reviewed publications, the large majority of which are included in the WoS (Engels et al., 2013).

In 2007, the University of Antwerp decided to introduce evaluative site visits by expert panels, during which the panel meets the spokesperson of each research group and other relevant stakeholders, and panel members are given every opportunity to enter into a dialogue with academic policymakers and the spokespersons of the research units during on-site interviews. The process facilitated asking additional questions or requesting clarification of specific points described in the self-evaluation report the panel received in advance. This approach ensures maximal involvement of the panel members and guarantees that the panel can reach a balanced decision in a fully interactive environment. The site visits thus guarantee interaction and involvement between experts and research groups.

At the start of a research evaluation, a department – typically encompassing several research groups – is invited to suggest potential panel chairs in addition to those suggested by the ADOC. Preferably, chairs are appointed as full professor, have an excellent publication record, have experience in research evaluations, are editors or board members of important journals, and possess academic management experience. The ADOC verifies whether proposed panel chairs and members have no prior involvement (i.e. no prior joint affiliations, no co-publications, no common projects) with the assessed research groups, and further checks if they are scholars with a prominent publication record in recent years, a proven track record of training young researchers, and sufficient experience in research policy, preferably in academic leadership positions. Furthermore, proposed panel chairs and members are preferably not affiliated with any Flemish institution of higher education and have no formal links to the University of Antwerp. The department that is being evaluated is also allowed to suggest potential panel members, but it should be noted that it is eventually the chair's prerogative to decide on the final composition of the panel.

The combined expertise of all panel members is to cover all subdomains in the discipline that is being evaluated and the panel is preferably balanced in terms of gender and nationality. When a sufficient number of professors have agreed to be on the panel, the university's research council ratifies the panel composition. Furthermore, all research groups belonging to a specific department (e.g., Biomedical Sciences) are to be evaluated by the same panel and the language of communication is English. Following the Dutch Standard Evaluation Protocol (VSNU, 2003; VSNU et al., 2014), the peer panels assess the quality, the productivity, the relevance and the viability of each research group.

To facilitate a critical reflection about the expiring 8-year cycle of site visit-based research assessment exercises and to ensure incorporation of the recommendations and suggestions by the start of the second 8-year cycle, the spokespersons of all research units that have participated so far were invited at the end of 2014 to fill out a short online questionnaire. The questionnaire was sent to 102 (former) spokespersons, i.e., spokespersons of research units belonging to the disciplines for which evaluation reports had already been handed in by the peer review panels. No response bias was found at the level of the scientific fields, while higher response rates were observed for disciplines that had been subjected more recently to assessment.

The vast majority of the respondents concluded that the site visits of the peer panels were useful (73%) and well organized (95%). Eighty-one percent of the respondents claim to have had sufficient time to present their research to the peer panel and 76% consider that also sufficient time was allowed for the subsequent interviews. According to 85% of the respondents, the interview was conducted in an atmosphere of openness. A few spokespersons further propose to devote somewhat more time to the unit's presentation and subsequent interview and three spokespersons advocate the idea that a visit to the research units' labs is incorporated in the site visits.

In general, the spokespersons expressed satisfaction with the external peer review panel charged with the evaluation of their discipline: 85% feel that the panel was (very) well positioned to assess the research in the discipline in general and 77% also believe that the panel's overall assessment was fair. In addition, 82% of the spokespersons are of the opinion that the panel was appropriately composed to evaluate the research conducted by their own research group in particular and 80% feel that their research unit was evaluated in a fair way. Only 8% of the respondents claim that the panel was biased in its judgement of the various research groups under assessment and 14% feel that such was the case for their own research group. Several spokespersons indicate in the accompanying comments that they were satisfied with the fact that the panel covered nearly all research activities of the groups under assessment, despite the diversity of the research that needed to be evaluated for each assessment exercise. The panel members' expertise was also apparent during the interviews, although the extent of preparation of the panel members is questioned by some spokespersons. The switch of panel chair that needed to be carried through for one of the research assessments was unanimously deplored by the spokespersons belonging to this discipline.

1.4 Overview of the chapters

This thesis is structured in nine chapters. In chapter two, we present a literature review of the broader topic and context of the research presented in this thesis.

In chapter three, we provide a detailed description of the data and methods that are used and proposed in this thesis.

Chapter four provides a preliminary exploration, comparing correlations and cosine similarities between the publication profile of evaluators and evaluees. The downside of these correlations and cosine similarity measures is that they do not take into account the relatedness of WoS SCs or journals. This leads to the approaches in the following chapters that do consider similarity.

In chapter five, we introduce the barycenter method at the level of WoS SCs. We explore the usefulness of overlay mapping on a global map of science (with WoS SCs) to gauge cognitive distance and introduce methods to determine an entity's barycenter according to its publication output. The Euclidean distance between barycenters is used as an indicator of cognitive distance. Overlay mapping techniques are used to visualize the barycenters.

In chapter six, we outline two quantitative methods – barycenter and SAPV – that gauge the cognitive distance between evaluators and evaluees, based on the journals they have published in. Both approaches determine an entity's profile based on the journals in which it has published. While the barycenter approach is based on a journal map, the SAPV method is based on the full journal similarity matrix. Subsequently, we determine the Euclidean distance between the barycenters or SAPVs of two entities as an indicator of the cognitive distance between them. Using a bootstrapping approach, we determine confidence intervals for these distances.

In chapter seven, we study the problem of determining the cognitive distance between the publication portfolios of two units from a more general perspective. We provide a systematic overview of five different methods (a benchmark Euclidean distance approach, distance between barycenters in two and in three dimensions, distance between SAPVs, and WCS) to determine cognitive distances on the basis of publication records. The (two- or three-dimensional) barycenter method is based on global maps of science, the SAPV method and WCS method (both in N-dimensions) use a full similarity matrix. We also present a theoretical comparison as well as a small empirical case study.

In chapter eight, we systematically compare how our proposed six approaches are related. We determine the correlation between the approaches, whether the method and level of aggregation has influence on the correlation or not. In addition, we explored, how much the approaches agree in finding the closest panel member, and how much they agree in predicting the main assessor.

Chapter nine contains a summary of the findings, policy recommendations, limitations of the study and suggestions for further research.

Chapter II: Literature review

2.1 Introduction

In this chapter, first we discuss the definition of research and of peer review, and how we have used these terms in this thesis. In section 2.3, we provide an overview of research evaluation focusing on discipline-specific informed peer review evaluation of research groups. We outline the need to determine cognitive distance between expert panel members and research groups (see section 2.4). Furthermore, we discuss the concept of cognitive distance and methods of operationalizing cognitive distance (see section 2.5). In section 2.6, we introduce the topic of bibliometric mapping, how science maps are constructed and visualization techniques. We focus on two specific types of global maps of science – maps based on WoS SCs and journals –, and discuss how these maps can be used to determine cognitive distance between evaluators and evaluees based on their publication portfolios in research evaluation. Finally, in section 2.7, we discuss the research gap in the existing literature and how our research contributes to fills the gap.

2.2 Definition

Before we embark on further discussion, we will discuss the definitions of research and of peer review.

2.2.1 Research

Creswell (2014, p. 4) defined research as a

> *process of steps used to collect and analyze information to increase our understanding of a topic or issue.*

OECD (2015, p. 44) described it as

> *creative and systematic work undertaken in order to increase the stock of knowledge – including knowledge of humankind, culture and society – and to devise new applications of available knowledge.*

According to the UK Research Assessment Exercise, research is defined as

> *... original investigation undertaken in order to gain knowledge and understanding. It includes work of direct relevance to the needs of commerce and industry, as well as to the public and voluntary sectors; scholarship; the invention and generation of ideas, images, performances and artefacts including design, where these lead to new or substantially improved insights; and the use of existing knowledge in experimental development to produce new or substantially improved materials, devices, products and processes, including design and construction.*
>
> *It excludes routine testing and analysis of materials, components and processes, e.g. for the maintenance of national standards, as distinct from the development of new analytical techniques. It also excludes the development of teaching materials that do not embody original research* (RAE2001, 2002, p. 1.12).

For the purpose of the thesis, we will consider the definition of the UK research assessment exercise as this more specified definition applies especially in the context of research evaluation.

2.2.2 Peer Review

Peer review is an umbrella term for expert-based review practices. It is an important mechanism for quality control, assessing scientific works and ensuring trustworthiness of scientific research (Ziman, 2002; Cronin, 2005; Holbrook, 2010; Bornmann, 2011; Lee et al., 2013). The most important forms of formalized peer review are review of research manuscripts, review of funding or grant applications (Guston, 2003; Smith, 2006), job application or career promotion review, review for scientific prizes (like the Nobel Prize) (Hemlin & Rasmussen, 2006), review of research groups and academic institutions (Hemlin, 1996), and national peer review based research assessments (Wouters et al., 2015).

The following possible objectives of peer review are distinguished (Geisler, 2000; Wager, Fiona Godlee, & Jefferson, 2002, as cited in Wouters et al., 2015, pp. 44-45):

1. assess the quality of research results, outcomes, projects and programs;
2. determine the level of performance, either in absolute terms or comparatively, of (parts of) the scientific and innovation system;
3. promote accountability;
4. contribute criteria and evidence for resource allocation;
5. contribute criteria and evidence for science and technology policy making;
6. contribute criteria and evidence for career decisions and human resource policies.

Moed (2005) classified peer review according to the moment it takes place (see also Wouters et al., 2015). They distinguish peer review of:

1. grant proposals in the context of funding decisions;
2. manuscripts in the context of publication decisions by journal or book publishers;
3. scientific data in the context of publication decisions or data repositories;
4. the performance of researchers or research groups in the context of national or international research assessment exercises and awarding scientific or scholarly prizes;
5. the context of foresight exercises and the development of national or international research agendas.

Whitley (2007) characterized peer review as a 'strong research evaluation system' as it is institutionalized and formalized, follows specific procedures, and makes a direct contribution to the concerned authority's purpose.

In the following sections, we will discuss research evaluation focusing on peer review. More specifically, we introduce the problem of expert panel composition, and the two major components of our proposed methods to inform expert panel composition. We will also discuss the research gap concerning the expert panel assignment for a research group in discipline-specific research evaluation and how this thesis contributes to fill the research gap.

2.3 Research evaluation

There has been a paradigm shift in research management when evaluation started to play a more active role in research and innovation governance (Arnold, 2004). Systematic evaluation efforts are perceived as useful instruments for rearranging the science system (Rip & Meulen, 1995; van Steen & Eijffinger, 1998). In many different sectors and fields evaluation of publicly funded research is carried out (Balázs & Arnold, 1998). Evaluation, including research evaluation, has become an institutional part of public sector governance (Dahler-Larsen, 2012; Hammarfelt & de Rijcke, 2015). Evaluation can be divided into two forms: ex-ante (before the event) and ex-post (after the event). The former is conducted prior to, e.g., a research project to assess its potential importance and probability of success, while the latter takes place after the completion of the research to assess its output and impact (Kogan, 1989; Massy, 1996; Suter, 1997; Geuna & Martin, 2003). One can also distinguish between formative and summative evaluation. Summative evaluation is conducted to examine the effects or outcomes of the research unit in comparison to similar units, while formative evaluation has the goal of improving the quality of research. Evaluation results are used as inputs in research management (Geuna & Martin, 2003; McDavid, Huse, & Ingleson, 2012).

Here, we will focus on a more specific research evaluation system, which can be classified as ex-post and carries aspects of both the summative and formative forms of evaluation. At the University of Antwerp, research evaluation is discipline-specific, in that each assessment encompasses all the research groups of that discipline at the same time. The research evaluation results in a report by the external evaluation panel containing both an assessment of the past research performance and suggestions regarding the ongoing research and future plans of each research group and of the department as a whole. Although the panels may make suggestions e.g. to stop or merge some research groups, the main focus rests with the suggestions for the department as a whole and for each of the research groups in the department, i.e. the formative characteristics. We use the terms 'evaluation' and 'assessment' as synonyms.

Since the 1980s a large number of research evaluation programs has emerged in most OECD (Organization for Economic Co-operation and Development) countries, and this on the level of institutions and on national level (OECD, 1997). Many countries have implemented formal policies to assess performance and output of publicly funded research on the national, regional, and institutional level (Whitley, 2007; Hammarfelt & de Rijcke, 2015). To evaluate a research unit's performance, policy makers and funding agencies are using two main approaches: peer review and bibliometrics (Ramos & Sarrico, 2016). Some researchers have compared bibliometric indicators and the outcome of the peer assessment for the same researchers or research groups (Nederhof & van Raan, 1993; Rinia, van Leeuwen, van Vuren, & van Raan, 1998; Aksnes & Taxt, 2004; van Raan, 2006; Lovegrove & Johnson, 2008; Patterson & Harris, 2009; J. Li, Sanderson, Willett, Norris, & Oppenheim, 2010; Franceschet & Costantini, 2011; Wainer & Vieira, 2013; Bertocchi, Gambardella, Jappelli, Nappi, & Peracchi, 2015; Baccini & Nicolao, 2016). There is a positive correlation between peer review score and bibliometric indicators (Oppenheim, 1997; Moed, 2002; Norris & Oppenheim, 2003; van Raan, 2005; Abramo & D'Angelo, 2011).

Due to database coverage and varying citation and publication cultures, the correlation between bibliometrics and peer review is weaker in most fields of engineering and computer science (Rahm, 2008), and humanities and social sciences (Abramo & D'Angelo, 2011). Moreover, publication and citation habits vary between different fields (Nederhof, 2006). Research evaluation practices vary according to discipline and country but consultation of peers is normally seen as an 'unavoidable' part of it (Langfeldt, 2004). Warner (2003) argued that expert review is the only system that enjoys both the confidence and the consent of the academic community. Some forms of peer review are vital as practical judgment is required and they cannot be replaced by bibliometric indicators (HEFCE, 2015). However, bibliometric indicators can support the peer review evaluation process (Aksnes & Taxt, 2004; Allen, Jones, Dolby, Lynn, & Walport, 2009; Taylor, 2011, Hicks, Wouters, Waltman, Rijcke, Rafols, 2015). The concept of informed peer review, in which bibliometric techniques and other information support the peer review evaluation process, has been proposed (van Raan, 1996; Butler, 2007; Moed, 2007). An example of such informed peer review is the United Kingdom's Research Excellence

Framework (REF) system for assessing the quality of research in UK higher education institutions evaluation (REF2014, 2014).

2.4 Problems with expert panel composition

Eisenhart (2002) stated that active researchers are the best-suited persons to do peer review and these researchers need to be from the same or a broader field as the research to be evaluated. In research evaluations, a group of reviewers evaluates the research outputs. These reviewers are termed *peers*. These peers are well-informed experts in the sense that they are aware of the discipline's research literature, as well as the state of the art, the challenges, and the research frontier in their field (Nederhof & van Raan, 1993; Langfeldt, 2004; Fedderke, 2013). In addition, they are typically active researchers who have the ability to evaluate fellowship or grant applications, manuscripts, published research, and the like. Their task is to make explicit recommendations on whether certain quality standards have been met, whether the research contributes to the knowledge base, which research proposal is expected to have a greater impact etc. (Eisenhart, 2002; Bornmann, 2011; Hammarfelt & de Rijcke, 2015). A downside of the peer review process can be the absence of an adequate methodology to find relevant experts (Hofmann et al., 2010; Gould, 2013; Berendsen et al., 2013; Lee et al., 2013; Oleinik, 2014; Buckley et al., 2014).

Expert panel review is a standard practice for evaluating research groups (Nedeva et al., 1996; Butler, 2007; Rons et al., 2008; Lawrenz et al., 2012; Milat et al., 2015) and for research proposals submitted to research funding organizations (Wessely, 1998; van den Besselaar & Leydesdorff, 2009; Li & Agha, 2015; Pina, Hren, & Marušić, 2015; Wang & Sandström, 2015). In many peer review processes the referees are anonymous to the researchers whose work is under assessment; in expert panel evaluation, however, the panel members are visible, and hence the units of assessment themselves can judge the expertise of the panel members and the expert panel in relation to their research domain.

The exponential growth of the research literature indicates the growth of specialized disciplines (Sobkowicz, 2015). Therefore, an individual panel member may have ample expertise in a subfield, but collaborative evaluation together with peers is still needed unless and until the individual panel member covers the expertise of the research groups. It has been argued that peer review can be biased and unreliable (Cicchetti, 1991; Bazeley, 1998; Wessely, 1998; Langfeldt, 2004; Bornmann & Daniel, 2005; Bornmann, 2011; Lee et al., 2013). Sometimes the evaluees are not pleased with the evaluation because 'reviewers or panelists are not expert in the field, poorly chosen, or poorly qualified' (McCullough, 1989, p. 82). The reliability and validity of peer review are not a given (Cicchetti, 1991; Bazeley, 1998; Langfeldt, 2004; Bornmann & Daniel, 2005; Benda & Engels, 2011; Bornmann, 2011; Lee et al., 2013). There may be controversy on a panel's composition: the expertise of the reviewers is frequently questioned (Over, 1996; Bornmann & Daniel, 2006; Daniel, Mittag, & Bornman, 2007). Evaluation done by people perceived as being non-experts raises questions about its credibility (Langfeldt, 2004). One way the credibility of peer review could be supported, is through measurement of the match between the expertise of the panel member and the research interests of the research groups. Such methods should be able to quantify the cognitive distance between the expert panel and the research groups. Engels et al., (2013) argue that a methodology is required to measure and quantify congruence of expertise or cognitive distance between panels and research groups in discipline-specific research evaluation. In this thesis, we focus on approaches that do exactly that. In the following section, we discuss cognitive distance and proximity.

2.5 Cognitive distance/proximity

Nooteboom (2000, p. 73) defines cognitive distance as "a difference in cognitive function". He explains this as follows: "This can be a difference in domain, range, or mapping. People could have a shared domain but a difference of mapping: two people can make sense of the same phenomena, but do so differently". Cohen & Levinthal (1989, 1990) explained the process by which an individual or organization by extrapolation can integrate and reuse knowledge from outside sources in research and development, while Nooteboom uses these ideas to define the concept of cognitive distance between individuals and organizations. Broström & McKelvey (2016, p. 6) define cognitive distance as "an inverse characterization of the degree of overlap between two people in terms of knowledge bases, values, norms and the heuristics of attribution and decision making", while exploring the interaction between experts and policy makers for generating policy learning and implementation. Cognitive distance may lead to misunderstanding or disagreement, but it also has positive aspects, e.g. in mutual learning. Wuyts, Colombo, Dutta, & Nooteboom (2005) note that the distinctiveness value (standing out as being better due to a specific feature) of a relation between two firms increases with cognitive distance and decreases with mutual understanding. Organizations need to aim at avoiding too great cognitive distance between its members to understand each other and achieve organizational goals; however, a certain amount of cognitive distance between the members is necessary (Granovetter, 1973). In a densely connected organization (strong ties, small cognitive distance) organizational knowledge is implicit and taken for granted, while a loosely connected organization has more weak ties that typically represent a greater cognitive distance between the involved actors. Consequently, the latter situation represents more opportunities for mutual learning (Granovetter, 1983). The capacity of sharing knowledge between communities depends on the degree of similarity between communities while the capacity of learning new things depends on the degree of dissimilarity (Grabher & Ibert, 2013).

If there is a large cognitive distance, it is hard to have a meaningful exchange between experts and policy makers (Howlett, 2009). Cognitive distance is also a factor in the collaboration between universities and industry for knowledge transfer activities. Different values, norms and mindsets from two different entities can increase the cognitive distance in this context (Muscio & Pozzali, 2013). Boschma (2005) explained that cognitive proximity is one of the factors that facilitate effective collaboration between the different actors in translational research. However, an excessive amount of or too minimal cognitive proximity might hinder learning and innovation (Boschma, 2005; Nooteboom et al., 2007; Hautala, 2013; Molas-Gallart, D'Este, Llopis, & Rafols, 2015).

In sum, the term cognitive distance has been used in different fields. In the literature, cognitive distance has been studied especially in the social and behavioral sciences (Golledge, 1987; Montello, 1991). For example, in the tourism literature, cognitive distance has been identified as a mental representation of actual geographical distance based on an individual's social, cultural and life experiences (Harrison-Hill, 2001). Cognitive distance also occurs in individuals depending on cognitive and cultural differences. Tourists take decisions based on their own perception of distance (Ankomah & Crompton, 1992; Ankomah, Crompton, & Baker, 1996).

In summary, the term ‘cognitive distance’ refers to the way in which two persons, and by extension, two organizations or groups of persons, are different, not only in terms of knowledge, but also in the way they perceive and interpret external phenomena. Like many other notions used in the social sciences – the notions of the impact, inequality, visibility come to mind –, the notion of cognitive distance must be operationalized. This operationalization can be done in many different ways.

In this thesis we consider the publication portfolio of the involved researchers to reflect the position of the unit in cognitive space and, hence, to determine cognitive distance. Expressed in general terms we measure cognitive distance between units based on how often they have published in the same or similar journals. One can think of other informetric ways to determine cognitive distance between scientists. Wang & Sandström (2015) for example use bibliographic coupling and topic modelling to determine cognitive distance between publication portfolios. Besides using publication portfolios, one could also measure cognitive distance between patent portfolios, in terms of conference participation, in terms of diplomas, and so on. Moreover, cognitive distance is relevant in many other social and political contexts as well, e.g. when hiring employees, when comparing the programs of political parties, or to understand cultural differences.

2.6 Bibliometric mapping

Bibliometrics is the scientific field that uses statistical analyses of the research literature and covers a wide range of laws and methodologies (Godin, 2006). A well-known application of bibliometrics is the comparative evaluation of countries, universities, research organizations, individual researchers etc. on the basis of their publication profiles. Bibliometrics is also used to investigate the structure of a research field or to determine the growth of research topics (Borgman & Furner, 2002). Bibliometric mapping is one of the major research topics in the field (Börner, Chen, & Boyack, 2003). For bibliometric mapping, the first requirement is to build a bibliometric network, often using a normalization process of the relation (edges) between its nodes by using similarity measures. “Maps are built on the basis of a matrix of similarity measures computed from correlation functions among information items present in different elements (e.g., co-occurrence of the same author in various articles)” (Rafols, et al., 2010).

The nodes can be different entities, e.g., publications, journals, researchers, subject categories, or keywords occurring in research papers. The edges refer to the relation between pairs of nodes, for example, – in a co-authorship network – which authors co-author papers or – in a citation network – who cites whom. Bibliometric networks are usually weighted networks. Therefore, edges indicate the relation and strength of the relation between two nodes (van Eck & Waltman, 2014). Since the beginning of bibliometric research, visualization has received much attention. The visualization of bibliometric networks is referred to as 'science mapping' or 'bibliometric mapping' (van Eck & Waltman, 2009; Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011). In bibliometric mapping, similar nodes are placed closer to each other and dissimilar nodes are more distant. Garfield, Sher, & Torpie (1964) manually constructed the first bibliometric maps of citation networks. Many different approaches have been developed to extract networks using, e.g., citation, co-citation, bibliographic coupling, keyword co-occurrences, and co-authorship. It is important in this process to choose an adequate unit of analysis (authors, documents, journals, terms etc.). For details on methods of bibliometric mapping we refer to the literature (Leydesdorff, 1987; Small, 1999; Noyons, 2001, 2004; Börner et al., 2003; Boyack, Klavans, & Börner, 2005; van Eck, 2011; Boyack & Klavans, 2014a, 2014b).

Bibliometric mapping thus refers to a set of quantitative methods to visually represent some aspects of the research literature by visualizing the relations between entities. Maps of science provide a visual overview of a knowledge field and can be a useful tool for evaluative bibliometric studies (Buter, Noyons, & Van Raan, 2004). Most bibliometric maps are using proximity-based approaches where proximity between two nodes indicates the relatedness of the nodes – the closer they are, the more they are related. Multidimensional scaling (MDS) (Borg & Groenen, 2005) is a common proximity-based technique for determining the location of nodes. Alternatives to MDS include VxOrd (Boyack et al., 2005; Klavans & Boyack, 2006) and the VOS technique (van Eck & Waltman, 2007). VOS, which stands for Visualization of Similarities, and MDS are mathematically closely related to each other. Like most mapping techniques, VOS and MDS locate items in a low-dimensional space and reflect the similarity or relatedness. VOS provides distance-based visualizations of bibliometric networks and display the edges between the nodes. VOSviewer is especially suitable for visualizing larger networks (van

Eck & Waltman, 2014). For details about the VOS mapping technique we refer to van Eck & Waltman (2010).

Graph-based maps determine distances between nodes based on the occurrence and weight of links. Typically, the links are drawn as well, which is less common in distance-based approaches. The Kamada and Kawai (1989) algorithm is a commonly used technique for creating graph-based visualizations. Kamada-Kawai is a spring-based layout algorithm for connected undirected graphs. It calculates the total balance of the graph, as the square summation of the differences between the ideal distance and the actual distance for all vertices and produces layouts with small amounts of edge crossings (Kamada & Kawai, 1989). Pajek is a professional software package for performing network analysis that implements the Kamada-Kawai algorithm (de Nooy, Mrvar, & Batagelj, 2012). For details about the Kamada-Kawai algorithm we refer to Kamada & Kawai (1989). Both in the graph-based approach and the distance-based approach nodes are positioned in a two-dimensional space. Among other approaches, Garfield, Pudovkin, & Istomin (2003) and Chen (2006a) used a timeline-based approach where the position of each node is constrained by publication year. In this thesis, we have used the VOS and Kamada-Kawai techniques for visualization purposes.

Bibliometric mapping simplifies the job of preparing bibliometric data for display and facilitates the reading of bibliometric information by non-experts (Rafols, et al., 2010). Its initial purpose was to provide suggestions for policy-related decision making by governments, funding agencies, and universities (Healey, Rothman, & Hoch, 1986; Franklin & Johnston, 1988; Noyons, 2001, 2004). Moreover, bibliometric maps can be used to get an overview of the scientific literature in a certain domain or on a certain topic (van Eck, 2011). Henry Small has pioneered several bibliometric techniques for mapping science (Griffith, Small, Stonehill, & Dey, 1974; Small & Garfield, 1985; Small, 1999). We refer to several articles for overviews of the bibliometric mapping literature regarding visualization of knowledge domains (Börner et al., 2003; Shiffrin & Börner, 2004), contemporary scientific data visualization (Börner, 2010), information visualization with a specific focus on science mapping (Chen, 2006b, 2013), characterizing individual bibliographic entities (Morris & Martens, 2008). Due to technological advancements, maps at the level of individual papers, containing millions of nodes, have also become possible in recent years (Boyack & Klavans, 2014a, 2014b). Bibliometricians, research institutions, funding agencies, and publishers have shown interest in bibliometric network visualizations for their respective purposes. Nowadays, a number of software tools are available. For example Pajek and Gephi are general tools for the analysis and visualization of networks, while CitNetExplorer, CiteSpace, HistCite, Sci^2, and VOSviewer are more specialized tools that focus on bibliometric networks.

Bibliographic databases are one of the main sources to create science maps. When a map of a specific discipline is created, it is called a ‘local science map’ while a ‘global science map’ is based on an entire bibliographical database and hence, in principle, covers all fields of research. In this thesis, we will work with global maps of science. Examples of general bibliographic databases include multidisciplinary citation indexes, like WoS by Clarivate Analytics (formerly Thomson Reuters) and Scopus by Elsevier. At present, WoS and Scopus are the main sources for citation data. WoS contains publications from the year 1900 to the present (Clarivate Analytics, 2016) while the Scopus database initially covered 1996 to the present. Its coverage nowadays dates back to 1970 (Elsevier, 2016). Scopus has greater coverage of some disciplines, including computer science, engineering, clinical medicine and biochemistry, literature from Asia and the Far east (Klavans & Boyack, 2007), but its coverage of pre-1996 publications and citations is lower than that of the WoS (Harzing & Alakangas, 2016; Mongeon & Paul-Hus, 2016).

WoS and Scopus contain citations and references of publications, while this information is lacking in most disciplinary databases like Chemical Abstracts for chemistry, MEDLINE in medicine, etc. Usually, the multidisciplinary databases have advantages for bibliometric mapping purposes as they contain cited references of a document (van Eck, 2011). A bibliometric map can be constructed, if one has access to a bibliographic database of the domain.

Now we will focus on two types of global science maps – maps whose nodes are WoS SCs and maps whose nodes are journals. In WoS, the Journal Citation Reports (JCR) are one of the main sources of citation data on journals that are indexed in the Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI). The JCR help to measure research influence and impact at the journal level, and show the relationship between citing and cited journals of SCIE and SSCI.

The aggregated citation relations between journals contain information about of the structure of and relations between disciplines and specialties (Leydesdorff, 2006). The aggregated journal-journal citation data is however not publicly available and must be purchased from Clarivate Analytics or a similar company. The aggregated citation relations among journals can be used to construct a journal similarity matrix (Leydesdorff, 2004, 2006). The journal similarity matrix can be considered as an adjacency matrix, and thus is equivalent to a weighted network where similar journals are linked and link weights increase with similarity strength. Leydesdorff & Rafols, (2012) generated global maps of science at the journal level from the aggregated journal–journal citation data of the SCIE and SSCI. This global map of science based on journal similarity is available at http://www.leydesdorff.net/journals12. 'Using a global map of journals, one can assess the portfolio in terms of the spread across journals and journal categories, and also measure "interdisciplinarity" in terms of the journal coverage of the set(s) under study' (Leydesdorff, Moya-Anegón, & Guerrero-Bote, 2015, p. 1001).

A global map of science has been constructed based on aggregated journal–journal citations using Scopus data too (using the entire set of 1996–2012 journals) (Leydesdorff, de Moya-Anegón, & Guerrero-Bote, 2010; Leydesdorff, Moya-Anegón, et al., 2015). The Scopus maps have greater coverage than WoS data of, for example, the humanities (Leydesdorff et al., 2010).

WoS also contains a classification of journals into so-called Subject Categories (Clarivate Analytics, 2017). Clarivate Analytics (formerly Thomson Reuters) has added one or more subject categories (SCs) to WoS-indexed journals based on 'subjective, heuristic methods', and has received criticism for being crude for some research areas (Pudovkin & Garfield, 2002). However, the WoS SCs are used for the evaluation of scientific performance (van Leeuwen, Visser, Moed, Nederhof, & Van Raan, 2003), schematic visualizations of scientific domains (Moya-Anegón et al., 2004), evaluation of the quality of research work in different subject categories (Sombatsompop & Markpin, 2005), analysis of citation or publishing patterns (Guerrero-Bote, Zapico-Alonso, Espinosa-Calvo, Gómez-Crisóstomo, & Moya-Anegón, 2007; Lancho-Barrantes, Guerrero-Bote, & Moya-Anegón, 2010a) etc. In addition, WoS SCs cover all disciplines and are generally accepted and used by bibliometric practitioners (Rehn, Kronman, Gornitzki, Larsson, & Wadskog, 2014; Leydesdorff & Bornmann, 2015).

A global map of science using the aggregated WoS SC-SC citation matrix has been developed by Leydesdorff, Rafols and colleagues (Leydesdorff & Rafols, 2009; Leydesdorff, Carley, et al., 2013). Based on JCR 2011 (which contains journals from the SCIE and SSCI), a matrix of citing to cited SCs has been created that has been cosine normalized in the citing direction. The map based on WoS SC data is available at http://www.leydesdorff.net/overlaytoolkit/map10.paj (Leydesdorff, Carley, et al., 2013). The file 'map10.paj' contains a weighted network (where edge weights represent similarity between the nodes involved) of WoS SCs. WoS has 250+ SCs while Scopus distinguishes about 330 so-called minor subject areas (based on all science journal classification' (ASJC) codes, equivalent to WoS SCs), but there is no map for the latter.

Overlay maps are a kind of visualization, where a similarity matrix is the source of a base map. The base map is used as a template on which the result of an analysis can be projected. The overlay map is well known from Google Earth, Google Maps, and/or network visualization programs such as Pajek (Boyack et al., 2005; Bornmann & Leydesdorff, 2011; Leydesdorff & Persson, 2010; Leydesdorff & Bornmann, 2012; Leydesdorff, Kushnir, & Rafols, 2014) and VOSviewer (van Eck & Waltman, 2007). Rafols et al., (2010) first used WoS SCs for developing interactive (responding to a user's input) overlays on base maps. Subsequently, interactive overlays on a journal base map using VOSviewer for the visualization of journals (Leydesdorff & Rafols, 2012; Leydesdorff, Rafols, & Chen, 2013), patent categories (Leydesdorff et al., 2014; Kay, Newman, Youtie, Porter, & Rafols, 2014) and clusters of papers (Boyack & Klavans, 2014b) have been developed. Overlay maps based on disciplines (such as WoS SCs) or journals can be used for research policy questions (Rafols, et al., 2010). However, document-level maps of science are required for questions related to research planning, for example the article-level map of science covering 16 years and nearly 20 million articles using co-citation-based techniques created by Boyack & Klavans (2014a).

2.7 Research gap

The literature addressing the issue of reviewer assignments focuses on submissions to academic journals or conferences (Campanario, 1998a, 1998b; Bornmann, 2011; Gould, 2013), grant applications (Wood & Wesseley, 2003; Demicheli & Pietrantonj, 2007), and fellowship applications (Bornmann & Daniel, 2005, Reinhart, 2009). In each of these contexts the question of optimization of the assignment of the 'object' under evaluation to reviewers arises. This kind of assignment is handled by the concerned authority, which relies on field knowledge that is relevant to the item of assessment (Wang, Zhou, & Shi, 2013). Such processes are a potential source of bias. To have an unbiased process of selection of reviewers several studies have been carried out, mainly with two approaches: a modeling or algorithmic approach, and an information retrieval approach. Based on one or both approaches reviewer assignment systems have been developed. For example, Dumais & Nielsen (1992) proposed an automated method for assigning manuscripts to reviewers "based on information retrieval principles and Latent Semantic Indexing." The abstracts of the reviewers' publications were considered as a description of their sub-areas of interest and expertise. Latent Semantic Indexing was then used to compare the match between the reviewers' expertise and the abstracts of the submitted papers.

Tian, Ma, & Liu (2002) proposed a decision support system for research and development project selection. In this model, each of the proposals and external reviewers have to give two keywords to describe the discipline areas they belong to. Through a combination of keyword matching and knowledge rules set by the funding organizations, external reviewers are matched with selected proposals. Janak, Taylor, Floudas, Burka, & Mountziaris (2006) have introduced a mathematical framework to address panel-assignment issues. This approach takes into account the number of proposals to consider, the number of reviewers to consider, and the number of reviewers needed per proposal. Rodriguez & Bollen (2008) used a relative-rank particle-swarm algorithm on a co-authorship network to determine the most appropriate reviewers for a manuscript. This approach used the reference list of a manuscript to represent the authors of its subject domain. From that author list, this approach identifies related authors in a co-authorship network for potential reviewers of the submitted manuscript. Wang, Zhou, & Shi (2013) proposed a two-phase stochastic-based 'greedy algorithm' for a group-to-group reviewer assignment, where groups of reviewers are assigned to groups of manuscripts. All reviewers in the same group will review each manuscript in the assigned manuscript group. The reviewers need to select a number of keywords that describe their expertise. Similarly, a set of keywords is used to represent the manuscript. All the keywords are selected from the subject classification tree used by the National Science Foundation of China. In the first phase, the manuscripts are clustered into groups and in the second phase reviewers are assigned to the manuscript groups.

There are some systems available that support the concerned authority to decide which submissions are to be assigned to which reviewer, for example, CMT: Microsoft's Academic Conference Management Service (https://cmt.research.microsoft.com/cmt), Easychair (http://www.easychair.org), HotCRP (http://www.read.seas.harvard.edu/~kohler/hotcrp), Linklings (http://www.linklings.com), Softconf (http://www.softconf.com), or Web Submission and Review Software (http://people.csail.mit.edu/shaih/websubrev). In these systems, the reviewers need to declare any potential conflict of interest and state their preferences for certain papers (paper bidding). Based on this information the system selects potential reviewers.

Elsevier provides a back-end software system named Fingerprint Engine (https://www.elsevier.com/solutions/elsevier-fingerprint-engine). This software mines text from publication abstracts, funding announcements and awards, project summaries, patents, proposals/applications and creates an index of collections of weighted key concepts. This software maps text to semantic 'fingerprints'. By using this technology, Elsevier provides services like creating expertise profiles to enable collaboration (Pure, https://www.elsevier.com/solutions/pure) and comparing Fingerprints to find reviewers (Expert Lookup, https://www.elsevier.com/solutions/expert-lookup). The existing approaches discussed in this section look for reviewers for individual items (e.g. a manuscript, research proposal, or grant application) or explore connections among researchers in view of potential collaboration. We have not found any academic literature that discusses the algorithms or techniques behind the software in greater depth.

The existing approaches do not seem to address the issue of expert panel composition. Moreover, as far as we know there are no studies focusing on issues related to expert panel composition and the assignment of panel members to research groups in discipline-specific research evaluation. In this thesis, we aim to fill this research gap. In our case, expert panels are composed and the panel members are assigned to evaluate one or more research groups of a discipline. These expert panel members have a wide range of experience with a large publication profile. In this thesis, the publications of a panel member and research group are considered as indicative of their respective research expertise. For the purpose of research evaluation, each of the expert panel members should have expertise that is relevant to one or more research groups that will be evaluated. In this thesis, we propose six different informetric approaches to measure this match between evaluators and evaluees using their publications as a representation of their expertise.

2.8 Conclusion

In discipline-specific research evaluation, peer review is widely accepted by the scientific community. As the literature indicates, the reviewer assignment problem has been discussed in the scenario of academic journals, conferences, and grant and fellowship selection. In order to improve the trustworthiness of peer review, several mathematical models and algorithmic approaches have been proposed and some software tools are in use. Discipline-specific research evaluation, however, is peer review in a different context. In this case, the entire panel is responsible to evaluate a number of research groups in a department or discipline. Each of the expert panel members have expertise that is relevant to one or more research groups that will be evaluated. Here we consider a problem, to the best of our knowledge, that has not been addressed until now: how to assign reviewers (expert panel members) to a research group in the context of discipline-specific research evaluation. In this thesis, to fill this research gap, we propose six informetric approaches to determine cognitive distances between the publications of panel members and those of research groups.

Chapter III: Data and Methodology

Imagine that person A publishes all his papers in the *Journal of Documentation* and person B publishes all his papers in *Journal of Information Science*; this would yield a cosine (or any other similarity) measure between A and B equal to zero. However, since the journals are very similar in topical scope, that score does not accurately reflect the ‘real’ similarity of their publication portfolios. This is in fact clearer when looking at a map of science, where the two journals would likely be positioned very close to each other. If we want to quantify that, we need to find something like an 'average' location of the group and panel and look at the distance between them.

The overall aim of the thesis is to develop methods to solve the kind of problem illustrated by the example above. In other words, we have developed a number of methods that not only consider the number of publications in each WoS SC or journal, but also the similarity of WoS SCs and journals. In this chapter, we explain how these methods are constructed. We describe how the individual research groups’ and panel members’ publications data are collected, explain our use of similarity matrices (specifically the WoS SC similarity matrix and the journal similarity matrix) and maps of science that are derived from these matrices, as well as our proposed methods in detail.

3.1 Data

The data in this thesis stem from the research assessments from 2009–2014 of six departments belonging to the University of Antwerp through site visits by expert panel members. We use the data collected in the framework of these completed research evaluations. Altogether, there are 58 research groups and 6 expert panels that are dedicated to one of the departments. We specifically focus on the situation where the expert panel needs to evaluate all the research groups of a department. These evaluations consider the entire research groups’ scientific activity for a specific period, typically eight years preceding the year of evaluation. All articles, letters, notes, proceeding papers, and reviews by the research groups published during the reference period were considered in the evaluation. For the purpose of the thesis, we consider only the publications that are indexed in SCIE and SSCI of WoS.

3.1.1 Data collection

3.1.1.1 Research groups data collection

Research groups at the University of Antwerp consist of professors (of all ranks), research and teaching assistants, and researchers (PhD students and postdocs). A research group consists either of one professor assisted by junior and/or senior researchers, or of a group of professors and a number of researchers linked to them.

First, we received all the WoS accession numbers of the publications of each research group from the ADOC of the University of Antwerp. Research group names have been standardized using the first four letters of the corresponding department, for example, BIOL-A for Biology research group A, PHAR-B for Pharmaceuticals sciences research group B, etc. We do a basic search in WoS with the accession numbers of each research group, keeping the time span to all years and searching SCIE and SSCI. Subsequently, we analyze the search result with the 'Analyze Results' option in the WoS according to 'WoS SCs' and the 'Source titles' (here after: journal titles). We repeat this procedure for each of the research groups. For both cases, we combine the search sets for each research group and get the data for the publications of the research groups as a whole, i.e. the department. In this way, any publication that has been co-authored by members of two or more research groups is counted only once.

Table 1 lists the publication statistics of the nine Biology research groups that generated 1158 publications in 372 journals. In total, their publications are distributed over 90 WoS SCs.

Table 2 lists the publication statistics of the fifteen Biomedical Sciences research groups that generated 1234 publications in 476 journals. In total, their publications are distributed over 103 WoS SCs.

Table 1: Publication statistics of Biology research groups (2004-2010)

Group code	Number of publications	Number of journals	Number of WoS SCs
BIOL-A	168	53	26
BIOL-B	58	33	13
BIOL-C	212	75	36
BIOL-D	176	68	26
BIOL-E	169	69	28
BIOL-F	58	35	18
BIOL-G	280	139	55
BIOL-H	67	42	25
BIOL-I	86	52	24
All groups	**1158**	**372**	**90**

Table 2: Publication statistics of Biomedical Sciences research groups (2006-2013)

Group code	Number of publications	Number of journals	Number of WoS SCs
BIOM-A	96	55	42
BIOM-B	43	27	16
BIOM-C	107	47	24
BIOM-D	201	95	43
BIOM-E	70	34	15
BIOM-F	27	17	12
BIOM-G	241	115	45
BIOM-H	50	29	17
BIOM-I	89	55	27
BIOM-J	47	27	21
BIOM-K	74	43	28
BIOM-L	12	11	7
BIOM-M	164	67	22
BIOM-N	114	43	12
BIOM-O	60	32	13
All groups	**1234**	**476**	**103**

Table 3 lists the publication statistics of twelve Chemistry research groups that generated 920 publications in 300 journals. In total, their publications are distributed over 94 WoS SCs. Table 4 lists the publication statistics of ten Pharmaceutical Sciences research groups that generated 376 publications in 180 journals. In total, their publications are distributed over 67 WoS SCs. Table 5 lists the publication statistics of nine physics research groups that generated 1739 publications in 353 journals. In total, their publications are distributed over 108 WoS SCs.

Table 3: Publication statistics of Chemistry research groups (2001-2008)

Group code	Number of publications	Number of journals	Number of WoS SCs
CHEM-A	129	47	27
CHEM-B	65	24	17
CHEM-C	156	52	26
CHEM-D	32	17	13
CHEM-E	70	39	23
CHEM-F	21	17	8
CHEM-G	161	47	42
CHEM-H	62	33	28
CHEM-I	51	24	19
CHEM-J	27	11	15
CHEM-K	97	66	48
CHEM-L	92	42	24
All groups	**920**	**300**	**94**

Table 4: Publication statistics of Pharmaceutical Sciences research groups (2001-2008)

Group code	Number of publications	Number of journals	Number of WoS SCs
PHAR-A	40	22	19
PHAR-B	62	32	21
PHAR-C	61	35	25
PHAR-D	32	17	13
PHAR-E	64	42	31
PHAR-F	34	21	8
PHAR-G	67	31	14
PHAR-H	39	27	21
PHAR-I	29	10	6
PHAR-J	11	9	10
All groups	**376**	**180**	**67**

Table 6 lists the publication statistics of three Veterinary Sciences research groups that generated 231 publications in 146 journals. In total, their publications are distributed over 61 WoS SCs. Table 7 shows how many research groups collaborated and how often this has happened.

Table 5: Publication statistics of Physics research groups (2002-2009)

Group code	Number of publications	Number of journals	Number of WoS SCs
PHYS-A	125	53	44
PHYS-B	486	66	25
PHYS-C	525	147	46
PHYS-D	269	17	7
PHYS-E	159	55	28
PHYS-F	42	23	13
PHYS-G	43	26	12
PHYS-H	132	31	12
PHYS-I	115	63	49
All groups	**1739**	**353**	**108**

Table 6: Publication statistics of Veterinary Sciences research groups (2006-2013)

Group code	Number of publications	Number of journals	Number of WoS SCs
VETE-A	143	102	55
VETE-B	41	33	25
VETE-C	52	21	16
All groups	**231**	**146**	**61**

Table 7: Number of publications co-authored by research groups

Department	Two groups	Three groups	Four groups	Five groups
Biology	113	3	-	-
Biomedical Sciences	142	16	2	1
Chemistry	43	-	-	-
Pharmaceutical Sciences	59	4	-	-
Physics	150	7	-	-
Veterinary Sciences	5	-	-	-

3.1.1.2 Panel members data collection

We have obtained the names and curricula vitae of the panel members from the ADOC. The panel member names are standardized as PM1, PM2 etc. We perform an advanced search for each panel member in WoS through checking the SCIE and SSCI. All the publications of the individual panel members up to the year of the research assessment at University of Antwerp were taken into account. We analyze the search result with the 'Analyze Results' option in the

WoS according to ‘WoS SCs’ and then ‘Source titles’ (here after: journal titles). We repeat this procedure for each of the panel members. For both the cases, we combine the search sets for each panel member and get the data for the publications of the panel as a whole. In this way, any publication that has been co-authored by two or more panel members is counted only once. Co-authorship between panel members only occurs in the case of Chemistry.

The Biology panel was composed of five panel members (including the chair). Table 8 lists the publication statistics of the Biology panel members. The combined publication output of the Biology panel members consists of 786 publications. The number of publications per panel member ranges from 76 to 262. In total, these publications appeared in 217 different journals and are assigned to 54 different WoS SCs.

Table 8: Publication statistics of Biology panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	146	48	20
PM2	117	49	24
PM3	76	35	15
PM4	185	49	13
PM5	262	76	28
Panel	**786**	**217**	**54**

Table 9: Publication statistics of Biomedical Sciences panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	153	78	30
PM2	201	81	26
PM3	261	79	22
PM4	240	86	39
PM5	74	37	18
PM6	109	35	23
PM7	194	68	21
PM8	101	32	23
Panel	**1333**	**395**	**80**

Table 10: Publication statistics of Chemistry panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	694	72	11
PM2	221	60	25
PM3	152	42	16
PM4	254	55	34
PM5	206	58	34
PM6	113	32	14
PM7	512	69	12
Panel	**2150**	**248**	**66**

Table 11: Publication statistics of Pharmaceutical Sciences panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	122	39	17
PM2	351	93	36
PM3	259	91	33
PM4	124	67	31
PM5	180	86	33
Panel	**1036**	**300**	**68**

Table 12: Publication statistics of Physics panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	117	7	3
PM2	168	15	4
PM3	124	49	10
PM4	166	40	10
PM5	247	87	10
PM6	282	54	10
Panel	**1104**	**204**	**46**

The Biomedical Sciences panel was composed of eight panel members (including the chair). Table 9 lists the publication statistics of the Biomedical Sciences panel members. The combined publication output of the Biomedical Sciences panel members consists of 1333 publications. The number of publications per panel member ranges from 74 to 261. In total, these publications appeared in 395 different journals and are assigned to 80 different WoS SCs.

Table 13: Publication statistics of Veterinary Sciences panel members

Panel member code	No. of publications	No. of journals	No. of WoS SCs
PM1	313	50	21
PM2	121	66	31
PM3	272	46	8
PM4	131	53	19
Panel	**837**	**200**	**55**

Table 10 lists the publication statistics of the Chemistry panel members. The Chemistry panel was composed of seven panel members (including the chair). The combined publication output of the Chemistry panel members consists of 2150 publications, two of which are co-authored publications between two panel members. The number of publications per panel member ranges from 113 to 694. In total, these publications appeared in 248 different journals and are assigned to 66 different WoS SCs.

The Pharmaceutical Sciences panel was composed of five panel members (including the chair). Table 11 lists the publication statistics of the Pharmaceutical Sciences panel members. The combined publication output of the Pharmaceutical Sciences panel members consists of 1036 publications. The number of publications per panel member ranges from 122 to 351. In total, these publications appeared in 300 different journals and are assigned to 68 different WoS SCs.

The Physics panel was composed of six panel members (including the chair). Table 12 lists the publication statistics of the Physics panel members. The combined publication output of the Physics panel members consists of 1104 publications. The number of publications per panel member ranges from 117 to 282. In total, these publications appeared in 204 different journals and are assigned to 46 different WoS SCs.

The Veterinary Sciences panel was composed of four panel members (including the chair). Table 13 lists the publication statistics of the Veterinary Sciences panel members. The combined publication output of the Veterinary Sciences panel members consists of 837 publications. The number of publications per panel member ranges from 121 to 313. In total, these publications appeared in 200 different journals and are assigned to 55 different WoS SCs.

Table 14: Publication statistics of the research groups and panels

Name of the Department	Research groups					Panel			
	Assessment year	No. of research groups	No. of journals	No. of publications	No. of WoS SCs	No. of panel members	No. of journals	No. of publications	No. of WoS SCs
Biology	2011	9	372	1158	90	5	217	786	54
Biomedical Sciences	2014	15	476	1234	103	8	395	1333	80
Chemistry	2009	12	300	920	94	7	248	2150	66
Pharmaceutical Sciences	2009	10	180	376	67	5	300	1036	68
Physics	2010	9	353	1739	108	6	204	1104	46
Veterinary Sciences	2014	3	146	231	61	4	200	837	55

Table 14 lists the number of publications of the research groups during the eight years preceding their evaluation, and the entire publication profile of the panel (members) up to the year of assessment. Altogether, there are 58 research groups in six departments. The number of publications per department ranges from 231 to 1739. In total, these publications appeared in 146 to 476 different journals, and publications are distributed over 61 to 108 WoS SCs. In all cases, two or more research groups from the same department co-authored some publications.

Table 14 also shows that in total, there were 35 panel members involved in the research evaluations of the six University of Antwerp departments. The number of panel members ranges from 4 to 8 for each department. The number of publications per panel ranges from 786 to 2150. In total, these publications appeared in 200 to 395 different journals, and were distributed over 54 to 80 WoS SCs. There is no shared authorship between panel members and research groups in any of the cases. None of the panels have any co-authored publications among the respective panel members, except for two Chemistry panel members with two co-authored publications.

3.2 Methods

Our methods are based on the assumptions that for the evaluation of a research group by a panel, the shorter the cognitive distances or the higher the similarity between a research group and panel (members) the better the fit of the expert panel. We use the global map of sciences based on WoS data, with the subject categories or the journals as nodes.

WoS SCs cover all disciplines and may comprise a wide array of different subfields and topics (Bornmann, Mutz, Marx, Schier, & Daniel, 2011). In addition, most journals cover closely related subfields and topics (Tseng & Tsay, 2013). The methods take into account the similarity between WoS SCs and between journals: If the publications of a panel member and a research group appear in similar or closely related journals, they may still cover the same or similar subfield or topic.

The global map of science is based on a matrix of similarity measures computed from correlation functions among the SCs. The similarity matrix contains two structures: a cited and a citing one. Loet Leydesdorff, Ismael Rafols and colleagues (Leydesdorff & Rafols, 2009; Rafols et al., 2010; Leydesdorff, Carley, et al., 2013) created a matrix of citing to cited WoS SCs based on the SCIE and the SSCI, which was subsequently normalized in the citing direction with a threshold cutoff at a *cosine* similarity > 0.15 between two SCs. The result is a symmetric $N \times N$ similarity matrix (here, N=224). If we interpret it as an adjacency matrix, we see that it is equivalent to a weighted network, in which similar categories are linked (the higher the link weight, the stronger the similarity). The file 'map10.paj' contains this weighted network of WoS SCs (available at http://www.leydesdorff.net/overlaytoolkit/map10.paj). The information in the network file can be visualized. The subfield of bibliometric mapping is dedicated to the visualization, clustering, and interpretation of similarity matrices or networks like the one we use. Many different algorithms or layout techniques have been developed for this purpose, for example, Kamada-Kawai (Kamada & Kawai, 1989) or VOS (visualization of similarities) (van Eck & Waltman, 2007).

We also use a global map of science based on journal similarity. We have received the data of the underlying similarity matrix from Loet Leydesdorff in the context of a joint paper (Rahman, Guns, Leydesdorff and Engels, 2016). While we did not construct this similarity matrix ourselves, we briefly outline the main steps that were taken to create it. The data was harvested from Thomson Reuters' (currently Clarivate Analytics) JCR of the Science and Social Science Editions 2011.

An aggregated journal-journal citation matrix of 10,675 journals[1] was constructed with a grand total of 35,295,459 citations over the entire matrix, which was subsequently normalized in the citing direction. The similarities between journals are calculated using the cosine similarity between their citing distributions respectively (see Leydesdorff, Rafols, et al., 2013 for details). The resulting journal similarity matrix can again be considered as an adjacency matrix, and thus is equivalent to a weighted network where similar journals are linked and link weights increase with similarity strength. However, as some of the journal names underwent a name change over time, we had to find a way to handle these changes in a uniform way. For the detailed guidelines, we refer to Chapter VI of this thesis.

Leydesdorff & Rafols (2012) used the JCR 2009 of the SCIE and SSCI, containing 9162 journals in total, to generate global maps of science at the journal level from the aggregated journal–journal citation. Based on the data the authors created two base maps – a citing map (8860 journals) and a cited map (9162 journals). The data was harvested from JCR 2009 and an aggregated journal–journal citation matrix was constructed. The matrix was transformed into a cosine-normalized matrix both in the cited and in the citing dimensions. Simultaneously, to create citing and cited overlay maps two corresponding computer programs (citing.exe and cited.exe) are made available. These programs can process data downloaded from WoS to generate overlay maps (see http://www.leydesdorff.net/journalmaps for details). Updated maps (both citing and cited) were created based on the JCR 2011 (SCIE and SSCI). This edition of the JCR contains 10,675 journals. This time, a third program (analyze.exe) was made available, which uses the results of the option 'Analyze Results' in WoS directly. This program no longer requires downloading the data sets and is hence more efficient (Leydesdorff, Rafols, & Chen, 2013; see http://www.leydesdorff.net/journals11 for details). Later, Leydesdorff and collaborators created an update that was based on the JCR 2012 (SCIE and SSCI), containing 10,936 journals in total. This time a citing map (10,546 journals) is offered with two programs (citing.exe and crciting.exe) for creating overlay maps. They also made available the third

[1] The Science and Social Science Editions 2011 contain 8281 and 2943 journals respectively. Of these journals, 549 are contained in both databases.

program (analyze.exe, see http://www.leydesdorff.net/journals12 for details) (Leydesdorff, de Moya-Anegón, & de Nooy, 2016) for the same purpose.

The citing and cited maps based on the JCR 2009 require downloading all records, while the maps based on the JCR 2011 do not require to download all records. The citing base map based on the JCR 2011 includes more journals (10,673 journals) than the cited base map (10,256 journals). Moreover, the majority of our research groups' publications data are close to the JCR 2011. For these reasons, we use the 2-dimensional citing journal base map based on JCR 2011 and that is available at http://www.leydesdorff.net/journals11/citing_all.txt (see http://www.leydesdorff.net/journals11).

3.2.1 Overlay maps

We created overlay maps based on a base map of science derived from WoS SCs (Leydesdorff & Rafols, 2009; Rafols et al., 2010) and journals (Leydesdorff & Rafols, 2012; Leydesdorff, Rafols, et al., 2013). Combining the base maps described in the previous section with publication data (how many publications in which SCs?), one can create overlay maps as the visual representation of the expertise of a research unit.

In an overlay map, the original map – referred to as the *base map* – provides the location (and sometimes cluster) of each SC/journal, whereas publication data is used to visualize the unit's publication intensity for each SC/journal. Typically, this is done by scaling the size of each node according to the number of publications. Hence, overlay maps can also be used for visual comparison and estimation of the degree of overlap of two or more entities in exploratory analysis.

During WoS SCs data collection for each entity (individual research groups, panel members, research groups together and panel) the resulting files are downloaded using the default name 'analyze.txt' for all research groups and panel members. We download the 'WC10.exe' program from http://www.leydesdorff.net/overlay toolkit.

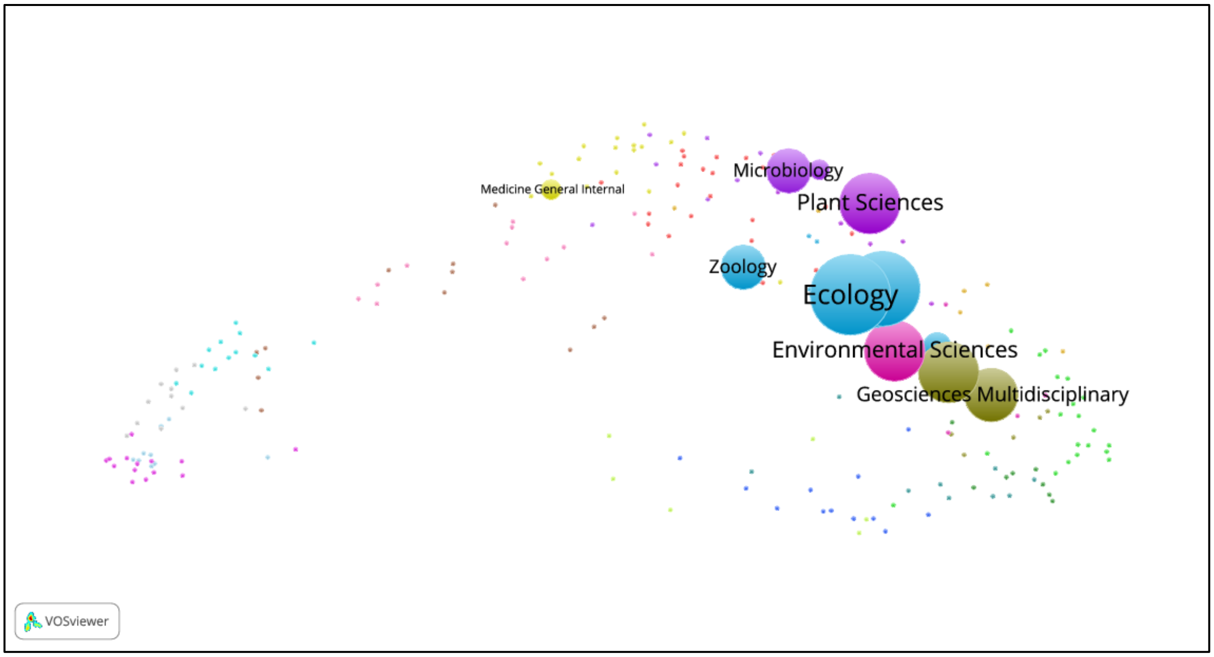


Figure 1: BIOL-B research group's publication overlay map in WoS SCs

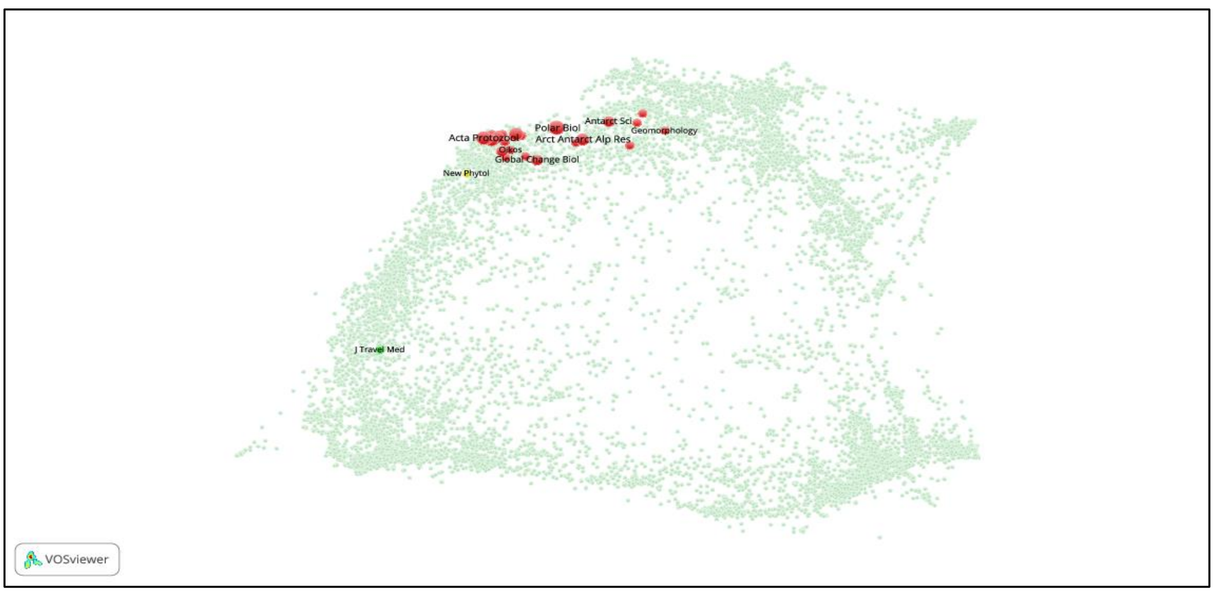


Figure 2: Journal overlay map of the BIOL-B research group

This file 'analyze.txt' is transformed by the mini-program 'WC10.exe' to 'WC10.vec' for upload into Pajek as a vector, and generates files like 'vos4.csv', 'vos6.csv', and 'vos19.csv' for use in VOSviewer (with 4, 6 or 19 base colors for the clusters, respectively) following the instruction given in (Riopelle, Leydesdorff, & Jie, 2014). We use the VOSviewer software for visualization of overlay maps. For example, Figure 1 shows BIOL-B research group's publications overlay map in WoS SCs.

Subsequently, the journals data for each entity's resulting files were downloaded from WoS. We followed the instruction of creating journal overlay maps available at http://www.leydesdorff.net/journals12. We obtain journal overlay maps by using VOSviewer. For example, Figure 2 shows the journal overlay map of the BIOL-B research group.

We prepare both the WoS SCs and journal overlay maps for each research group, each panel member, research groups together and panel. The maps are reported in the online technical reports (The full reference is mentioned at the end of this chapter) of each department.

3.2.2 The benchmark

The panel members and the research groups are represented as N-dimensional publication vectors. As a start (benchmark) we just calculate the Euclidean distance between the L_1-normalized (Golub & Van Loan, 1996) (see section 3.2.4) arrays of each panel member and each research group. The Euclidean distance between two vectors $a = (a_n)_{n=1,\ldots,k}$ and $b = (b_n)_{n=1,\ldots k}$ in $\mathbf{R}^k$, for any strictly positive integer k, is given as:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_k - b_k)^2} \tag{1}$$

In this thesis, we will use formula (1) for k = 2, k=3 and k = N. It is mentionable that the benchmark does not consider the similarity of WoS SCs or journals, while our proposed methods consider the similarity. We have used and present results of this benchmark method in the chapter VII, where we compare each of the methods we have developed.

3.2.3 Barycenter method

A barycenter is an entity's weighted average location on a map. The barycenter method - the center of publication or publication barycenter is introduced by (Rousseau, 1989a, 1989b). Jin & Rousseau (2001) applied the method in a practical situation to identify the mean center of

Chinese publications. Later, using barycenter places of publication of monographs, edited books and book chapters, Verleysen & Engels (2014b) measured internationalization of peer reviewed and non-peer reviewed book publishing in the Social Sciences and Humanities (SSH) as practiced at universities in Flanders. Further, they used barycenters to compare the geographic center of weight of book publishing between the SSH (Verleysen & Engels, 2014a)

In our case, an entity's barycenter is the center of weight (Rousseau, 1989a, 1989b, 2008; Jin & Rousseau, 2001) of the WoS SCs or journals in which it has publications. The weight is the number of publications of the panel member or the research group in the WoS SCs or in the journals in which it has publications. In the 2-dimensional WoS SCs map or journal map each SC/journal has a place on this map, characterized by the corresponding coordinates, denoted as $(L_{j,1}, L_{j,2})$, $j = 1, \ldots, N$. For each panel member and for each research group a barycenter is calculated and Euclidean distances between barycenters can be determined. Coordinates of these barycenters (in 2-dimension) are given as the point $C = (C_1, C_2)$, where

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \;\; ; \; C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} \qquad (2)$$

In the case of WoS SCs, m_j is the number of publications of the unit under investigation (panel member, research group) belonging to category j; this category j has coordinates ($L_{j,1}$, $L_{j,2}$) in the base map and $T = \sum_{j=1}^{N} m_j$. Note that T is larger than the total number of publications as we use full counting of WoS SCs: if a publication appears in a journal belonging to two categories, it will be counted twice.

In the case of journals, $L_{j,1}$ and $L_{j,2}$ are the horizontal and vertical coordinates of journal j on the map, m_j is the number of publications in journal j of the unit under investigation (panel member, research group), and $T = \sum_{j=1}^{N} m_j$ is the total number of publications of the entity.

We point out that the term 'barycenter' taken on its own, has no meaning. Any point can be the barycenter of infinitely many sets of points, possibly using sets of weights. We refer to a formal description of the notion of a barycenter: A barycenter is the result of an operation performed on a set of vectors. Let $X = (X_n)_{n=1,\ldots,k}$ be a set of vectors in m-dimensional space, $\mathbf{R}^m$. Then its barycenter B_X is the result of the following mapping:

$$B:(\boldsymbol{R}^m)^k \rightarrow R^m:(X_n)_{n=1,\dots,k} \rightarrow B_X = \frac{1}{k}\sum_{n=1}^{k} X_n \tag{3}$$

An example: let m = 2, k = 4 and $X_1 = \binom{0}{0}, X_2 = \binom{0}{1}, X_3 = \binom{2}{1}$ and $X_4 = \binom{2}{0}$. Then the barycenter of this set of four vectors is: $\frac{1}{4}\left(\binom{0}{0} + \binom{0}{1} + \binom{2}{1} + \binom{2}{0}\right) = \frac{1}{4}\binom{4}{2} = \binom{1}{0.5}$. This is the standard barycenter of the set of vertices X_1, X_2, X_3 and X_4 of a rectangle in the plane. More generally, one may assign a positive weight to each vector. If m_n is the weight assigned to vector X_n then the (generalized) barycenter (or center of gravity) is the result of the following mapping:

$$B:(\mathrm{R}^+, R^m)^k \rightarrow R^m:(m_n, X_n)_{n=1,\dots,k} \rightarrow B_x = \frac{1}{T}\sum_{n=1}^{k} m_n X_n \tag{4}$$

where $T = \sum_{n=1}^{k} m_n$. If all weights are set equal to 1 then one recovers formula (3).

Clearly, any vector can be the barycenter of infinitely many sets of vectors and weights. This is the main reason why the term 'barycenter' has no meaning on its own. In an extremely formal way, one may even say that any vector X_0 is the barycenter of itself, by taking the set of vectors equal to the singleton set {X} and weight equal to 1.

We further note that in order to obtain meaningful distances these values must be scale-invariant. This means that the distance between points *P* and *Q* must be the same as the distance between the points *P* and *cQ*, where *c* is a strictly positive number. Indeed: the total output of a research group can be several orders of magnitude larger than that of one expert. This difference must not play a role in determining cognitive distances. The barycenter method explained above and in particular formulae (2) satisfy this requirement as multiplying all m_j*s* with the same strictly positive factor leads to the same barycenter.

Although it is convenient to perform visualization and to determine cognitive distance in the plane, there is no theoretical reason to perform these acts in two dimensions. Likewise, there are no strong reasons to do both in the same dimension. The barycenter method can, at least in theory, be applied in any strictly positive dimension smaller than or equal to N. In Chapter VII, we compare barycenters in two and three dimensions; in the remainder of the thesis we limit ourselves to barycenters in two dimensions.

For three dimensions, we use the VOS algorithm, but resulting in a three-dimensional base map. This map was based on the network in http://www.leydesdorff.net/overlaytoolkit /map10.paj and obtained using Pajek, which implements the VOS algorithm both in two and three dimensions. Again, each SC has a place on this map, characterized by corresponding coordinates, denoted as $(L_{j,1}, L_{j,2}, L_{j,3}),\ j = 1, \ldots, N$, and for each panel member and for each research group a barycenter derived from their publication profiles is calculated. Coordinates in 3-dimensions are given as

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \;;\; C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T};\; C_3 = \frac{\sum_{j=1}^{N} m_j L_{j,3}}{T} \tag{5}$$

The meaning of the symbols T and m_j in formula (5) is the same as in formula (2). The Euclidean distance between barycenters is calculated with the formula (1). We have used the barycenter method in chapters V to VIII. The 3-dimensional variant of the barycenter method is used in chapter VII.

3.2.4 Similarity-adapted publication vector method

A regular publication vector simply counts publications per journal (or subject category), whereas in a SAPV these counts are adapted to account for similarity between WoS SCs or journals. We use normalized SAPVs, such that there is scale invariance and publication vectors of entities of varying size can be meaningfully compared.

We calculate SAPVs for each entity, starting from the original journal similarity matrix, (where N = 10,675) and original WoS SCs similarity matrix (where N = 224) is the number of rows or columns in the matrix. Based on their respective SAPVs, the distance can be calculated between the expert panel, panel members, groups, and separate groups.

A SAPV is determined as the vector $\mathrm{C} = (C_1, C_2, \ldots, C_N)$, where*:*

$$C_k = \frac{\sum_{j=1}^{N} s_{kj} m_j}{\sum_{i=1}^{N}\sum_{j=1}^{N} s_{ij} m_j} = \frac{(S * M)_k}{\|S * M\|_1} \tag{6}$$

In case of WoS SCs, $s_{j,k}$ denotes the similarity value between the k-th and the j-th WoS SC, and m_j is the number of publications in WoS SC j. The numerator of Equation (6) is equal to the k-th element of $S * M$, the multiplication of the WoS SCs similarity matrix S and the column matrix of publications $M = \left(m_j\right)_j$. The denominator is the L_1-norm of the unnormalized vector. We observe that the L_1-norm of the normalized vector C is indeed equal to 1.

In case of journals, $s_{j,k}$ denotes the k-th coordinate of journal j and m_j is the number of publications in journal j. The numerator of Equation (6) is equal to the k-th element of $S * M$, the multiplication of the journals similarity matrix S and the column matrix of publications $M = \left(m_j\right)_j$.

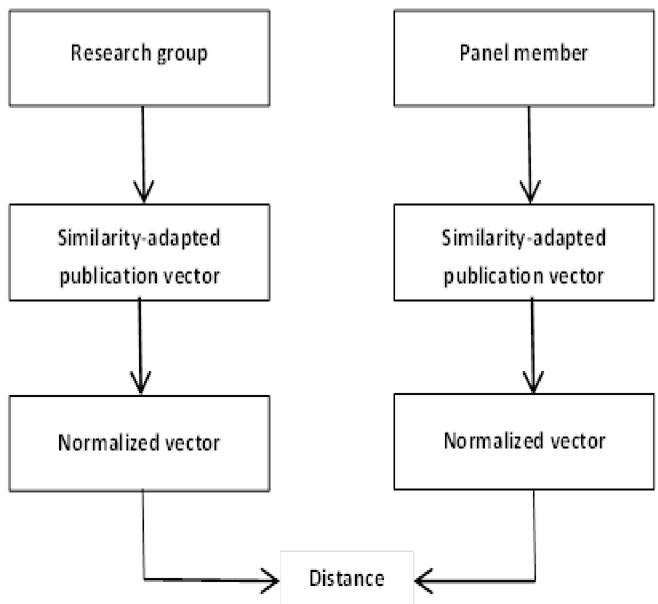


Figure 3: Workflow for determining distances between SAPVs

The multiplication $S * M$, i.e. applying the linear map with matrix representation S to the publication vector M leads to a new vector that we termed a SAPV. If we ignore similarity then S is the identity matrix and publication columns stay unchanged. We consider the SAPV method to be quite interesting as it provides a solution to the problem that WoS SCs overlap and are sometimes poorly defined, the SC *Information Science & Library Science* being a well-known example.

Figure 3 shows the workflow for determining distances between SAPVs. In particular, scale-invariance can be obtained through normalization as illustrated (for 3-dimensions) in Figure 4. All points situated on the straight line through the origin are represented by the point in the plane with equation x + y + z = 1.

This is so-called L_1-normalization: by dividing each coordinate by the sum of all coordinates, one obtains a new array for which the sum of all coordinates is one (taking into account that no coordinate is negative). One could equally well divide by an array's Euclidean length (so-called L_2-normalization (Golub & Van Loan, 1996) but as we do not see an advantage for any of the two approaches, we applied L_1-normalization as is done in diversity studies.

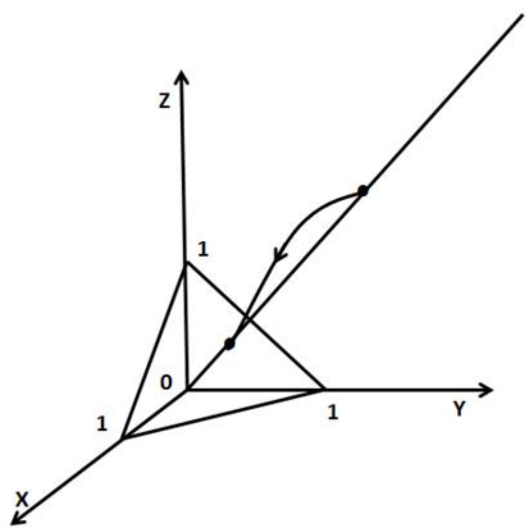


Figure 4: Normalization, leading to a scale invariant approach

Subsequently, we determine the Euclidean distance between the SAPVs of panel members, individual research groups, panel and research groups together using formula (1). Although the matrix and vectors are large, the calculation of SAPV and distances is relatively fast, due to the use of efficient matrix procedures implemented in NumPy and SciPy.[2]

[2] http://www.numpy.org and http://scipy.org

The distances thus obtained through barycenter and SAPV methods should be interpreted as having arbitrary units on a ratio scale (Egghe & Rousseau, 1990). This means there is a fixed meaningful zero (distance zero in the map), and distances can be compared in terms of percentage or fraction (e.g. the distance between A and B is 1.5 times larger than the distance between C and D).

We have used the SAPV method in chapters VI, VII, and VIII.

3.2.5 Weighted cosine similarity method

We also consider a weighted similarity method (generalized cosine similarity). The weighted cosine similarity (WCS) between panel member (PM) k and research group m, according to Zhou et al. (2012) is:

$$\frac{\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)}{\sqrt{\left(\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} M_j^k s_{ji}\right)\right) \cdot \left(\sum_{i=1}^{N} R_i^m \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)\right)}}$$

$$= \frac{\left(M^k\right)^t * S * R^m}{\sqrt{\left(M^k\right)^t * S * M^k} \cdot \sqrt{\left(R^m\right)^t * S * R^m}} \tag{7}$$

The numerator is the matrix multiplication: $\left(M^k\right)^t * S * R^m$, where ^{t} denotes matrix transposition, *S* is the journal/WoS SCs similarity matrix, M^k denotes the column matrix of publications of panel member k and R^m denotes the column matrix of publications of research group m. Similarly, the two products under the square root in the denominator are: $\left(M^k\right)^t * S * M^k$ and $\left(R^m\right)^t * S * R^m$. The result is the WCS value between panel member k and research group m. This value is calculated for each panel member and each research group. The larger the value the more similar they are.

Formula (7) is scale-invariant: multiplying M^k or R^m with a fixed constant does not change the result. Note that if S is the identity matrix (similarity is not taken into account), formula (7) reduces to regular cosine similarity (see chapter IV section 4.2.3). A similarity or proximity can be considered as the opposite of a distance: the higher the similarity the better the match – the closer the distance – between a panel member and a research group. This value too is calculated for each panel member and each research group.

We note that this method may lead to mathematical problems when applied in general vector spaces, but that these problems do not occur in the particular framework used in this thesis (in mathematical terms: we work in the positive cone $(\mathbf{R}^{+})^{N}$, where **R** denotes the real numbers). In the following example, we show that WCS cannot be used with any similarity matrix but that the problem does not occur for the similarity matrices (WoS SCs and journals) used by us. We illustrate this with the regular cosine similarity (the numerator of formula 7).

In a general (real or complex) vector space it is possible that if expressions of the form $\left(M^{k}\right)^{t} * S * R^{m}$, with S a symmetric matrix, are used as similarity measures, some non-null vectors have similarity zero to themselves. This excludes this type of construction as a general method for calculating similarities.

We consider the symmetric matrix $S = \begin{pmatrix} 1 & 0.8 & 0.9 \\ 0.8 & 1 & 0 \\ 0.9 & 0 & 1 \end{pmatrix}$, see (Zhou et al., 2012), and want to find a vector X = $(u,v,w)^{t}$, (u,v,w: real numbers) such that $(X)^{t} * S * X = 0$. Replacing X by $(u,v,w)^{t}$ leads to the requirement: $u^2 + 1.6uv + 1.8uw + v^2 + w^2 = 0$. Taking u = 1, v ≈ -1.44031 and w = -1.1 provides an (approximate) solution. In fact, this is just one solution among infinitely many.

If u = 1 and w = K then $v_1 = -\left(\sqrt{-K^2 - 1.8 * K - 0.36} + 0.8\right)$ and $v_2 = \left(\sqrt{-K^2 - 1.8 * K - 0.36} - 0.8\right)$ always provide solutions (some of which may be complex numbers). The one given above is v_1 with K = -1.1. This solution was obtained using TI-*n*spire software.

We check now that v_1 and v_2 as given above, indeed lead to the perfect null solution. Writing $\sqrt{-K^2 - 1.8 * K - 0.36}$ as R and using v_1 we find:

$u^2 + 1.6\,uv + 1.8uw + v^2 + w^2$

$= 1 - 1.6R – 1.28 + 1.8K + (- K^2 - 1.8K - 0.36) + 1.6R + 0.64 + K^2$

$= (1 - 1.28 - 0.36 + 0.64) + (1.8 - 1.8)K + (- K^2 + K^2) + R (- 1.6 + 1.6) = 0$

Similarly, with v_2 we obtain:

$u^2 + 1.6uv + 1.8uw + v^2 + w^2$

$= 1 +1.6 R – 1.28 + 1.8 K + (- K^2 – 1.8K - 0.36) – 1.6 R + 0.64 + K^2$

$= (1 - 1.28 - 0.36 + 0.64) + (1.8 - 1.8) K + (K^2- K^2) + R (1.6 - 1.6) = 0.$

However, this problem cannot occur when the matrix S has non-negative values and when, moreover, the vector X has only non-negative values, which is precisely the context in which we work. Indeed: under these circumstances the expression $(X)^t * S * X$ is always non-negative and only zero when X = 0 (the zero-vector) and this in any dimension. Note that the example presented above led to a vector X with two negative coordinates.

We have used the WCS method in chapters VII and VIII.

3.2.6 Bootstrapping and confidence intervals

The barycenter and SAPV methods determine cognitive distance, on the basis of the WoS SCs/journals in which the groups and panel members have published. In the same way, the WCS method determines similarity on the basis of the WoS SCs/journals in which the groups and panel members have published. However, such information is not entirely deterministic; it is, for instance, dependent on the database used as well as environmental factors like the speed with which a journal processes a submission. It logically follows that small differences in Euclidean distances or similarity bear little meaning.

To study this problem in a more systematic way, we employ a bootstrapping approach in order to determine 95 % confidence intervals (CIs) to each Euclidean distance (both between barycenters and SAPVs) and similarity. If two CIs do not overlap, the difference between the distances is statistically significant at the 0.05 level. Although it is possible for overlapping CIs to have a statistically significant difference between the corresponding distances, the difference between the distances is less likely to have practical meaning. If the confidence interval of the panel member who is closest to a given research group overlaps with that of the panel member who ranks second (and maybe even with the panel members ranking third or fourth) we say that there is no (statistical) difference in cognitive distance.

Bootstrapping (Efron & Tibshirani, 1998) is a simulation-based method for estimating standard error and confidence intervals. Bootstrapping depends on the notion of a bootstrap sample. To determine a bootstrap sample for a panel member or research group with N publications, we randomly sample with replacement N publications from its set of publications. In other words, the same publication can be chosen multiple times. Some publications in the original data set will not occur in the bootstrap data set, whereas others will occur once, twice or even more times. From the bootstrap sample, one can calculate a bootstrap replication, in our case a benchmark using formula (1), barycenter 2D using formula (2), barycenter 3D using formula (5), SAPV using formula (6), and WCS using formula (7).

By generating a large amount of independent bootstrap samples (in our case 1000) and each time calculating the bootstrap replication, we can approximate the variability within the data set. Since we have a two-sample problem (distance between two entities; Efron & Tibshirani, 1998, Ch. 8), we calculate the distances between pairs of bootstrap replications, from which we obtain a CI using a bootstrap percentile approach (Efron & Tibshirani, 1998, Ch. 13). In the case of WCS, we generate 1000 independent bootstrap samples for both entities and calculate the similarity between them using formula (7). A more detailed explanation can be found online (Guns, 2016a, 2016b).

The bootstrap replications of barycenters are used to add a 95% confidence region for each barycenter to the maps. For each barycenter, we have a cloud of 1000 points (bootstrapped barycenters) surrounding it. The confidence region is an ellipse that covers 95% of the bootstrapped barycenters. The larger the confidence region, the less stable the barycenter is. Although the CI of the distance between two barycenters and their confidence regions are related, the two should not be conflated. In particular, we stress that overlapping confidence regions do not correspond to overlap between CIs for distances.

The maps were plotted using Matplotlib (http://matplotlib.org). First, the base map was plotted using the pre-existing coordinates. Next, the barycenters were added as slightly larger red or green points. Finally, a partially translucent confidence region (ellipse) was calculated and superimposed on the map. Calculation of the confidence region was done using an implementation by Kington (2014). We briefly outline what elements determine the location and placement of such a confidence ellipse. The center of the ellipse is simply the mean of all bootstrapped barycenters. The width and height of the ellipse (or its axes) depend on the variance in the cloud of points. Finally, the orientation of the ellipse is obtained from the largest eigenvector.

3.2.7 Methods at a glance

In total, we introduced five methods: a benchmark, two methods using barycenters (one in two and one in three dimensions), SAPV and WCS. The benchmark and the SAPV and WCS calculations are applied in N dimensions, where N denotes the total number of WoS SCs/journals. One of the main components of the proposed methods is that they take into account the similarity matrix of WoS SCs and journals. The benchmark method, however, does not take this relatedness of WoS SCs or journals into account. The 3-dimensional barycenter is a variant of the 2-dimensional barycenter method. Therefore, we consider that we have introduced three main methods (barycenter in two dimensions, SAPV and WCS).

For each of these three methods we have two levels of aggregation – WoS SCs and journals. For each level, there is a similarity matrix (N-dimensions, with N the number of WoS SCs or journals) and a 2-dimension base map derived from the respective similarity matrix.

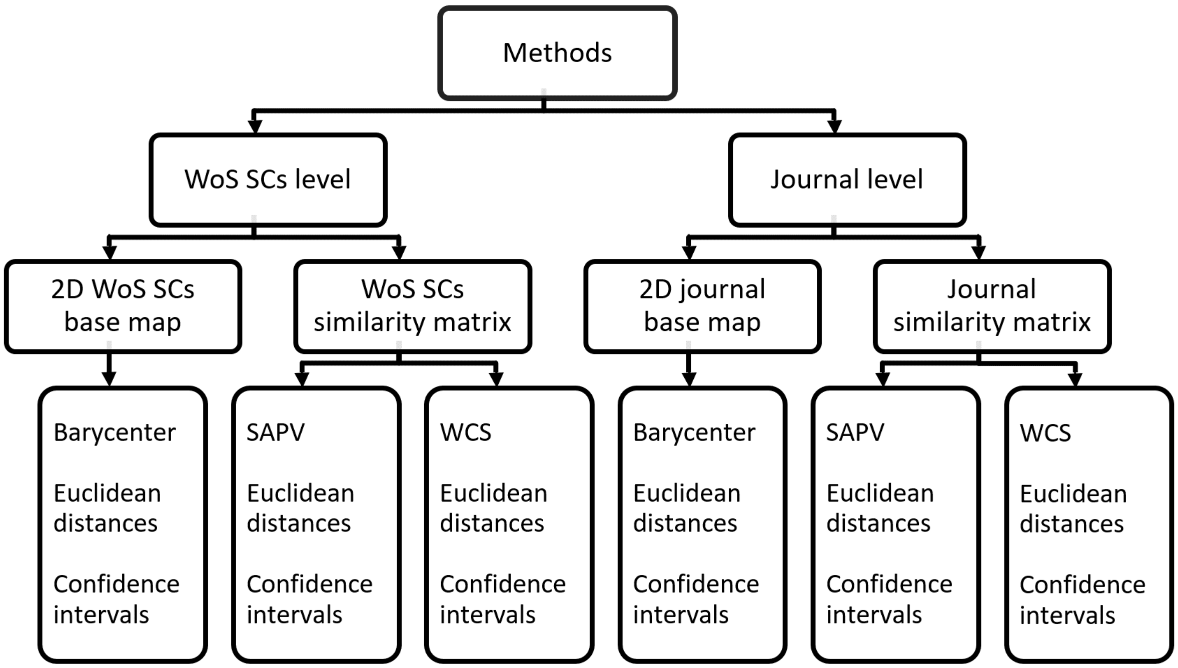


Figure 5: Main components of the six approaches at a glance

The SAPV and WCS methods operate at the level of N-dimensions, whereas the barycenter method uses the 2-dimensional base map. For the case of SAPV and barycenter, first we determine the location of an entity in the similarity matrix or in the 2-dimensional base map based on entities publication profile. Later we calculate the Euclidian distances between the locations (SAPV or barycenter) of the entities.

Further, with the bootstrapping approach we determine the confidence intervals of the distance between two entities. Subsequently, for the case of WCS, we calculate the similarity between the entities and, using bootstrapping approach, we determine the confidence intervals of the similarities. All together, these leads to six informetric approaches. Figure 5 illustrates the main components of the six approaches at a glance.

3.2.8 Technical reports

In the frame of the journal articles included as chapter V to VIII, it was not possible to include data from all six departments both due to time constraints and to space limitations. For the entire thesis, we have done all calculations for the six departments included in the thesis. Therefore, we have prepared technical reports on each of the six departments. In the technical reports, we described step by step the data collection, correlation calculations, barycenter calculation, SAPV calculation, WCS calculation, Euclidean distance calculation, WoS SCs and journal overlay map creation and programming codes. These technical reports are self-sufficient to understand the technical details from data collection to calculation. The reports are available online as depicted below:

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Biology department (Technical report) (p. x, 77). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431560151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Biomedical Sciences department (Technical report) (p. xii, 93). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431570151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Chemistry department (Technical report) (p. x, 87). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431580151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Pharmaceutical Sciences department (Technical report) (p. x, 79). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431590151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Physics department (Technical report) (p. xii, 80). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431600151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: a case study of Veterinary Sciences department (Technical report) (p. x, 68). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431610151162165141

Chapter IV: Correlation and similarity between publication profiles of panels and research groups

4.1 Introduction

Across all disciplines, we find that both the panel and the research groups have publications in the same or similar WoS SCs or journals. As a preliminary exploration, in this chapter we focus on the correlation to measure the strength and direction (positive/negative) of the association between the publication profile of the research groups and the panel. With respect to a specific journal or WoS SC, three situations can occur: (i) both the panel and the research groups have publications in it; (ii) either the panel or the research groups (but not both) have publications in it; or (iii) neither the panel nor the research groups have publications in it. For the first two cases, we determine the correlation between the publication output of two entities using Pearson's correlation coefficient and Spearman's rank correlation coefficient. Since situation (iii) is usually the most common one, we adopt top-down correlation (Iman & Conover, 1987), a correlation coefficient that is better able to handle zero-inflated situations. Salton's cosine measure (Salton & McGill, 1986) has also been shown to be insensitive to common zeros (Ahlgren, Jarneving, & Rousseau, 2003). Its use to determine the similarity between publication profiles is explored in sections 4.2.3 and 4.3.3.

4.2 Comparison of panel and research group profiles at the level of subject categories

Table 15 shows that together, the Chemistry panel and groups have published in 108 WoS SCs; when considered separately, panel publications appear in 66 categories, and group publications in 94 categories. Similarly, the Physics panel and group publications are found in 112 WoS SCs, with the Physics panel publications appearing in 46 categories, and the groups' publications in 108 categories.

Table 15: Distribution of the panels' and research groups' publications in WoS SCs

Department	Panel and groups total	Panel and groups common	Groups	Panel
Biology	101	43	90	54
Biomedical Sciences	112	71	103	80
Chemistry	108	51	94	66
Pharmaceutical Sciences	86	49	67	68
Physics	112	42	108	46
Veterinary Sciences	71	45	61	55

The Biology and Biomedical panels have respectively 43 and 71 WoS SCs in common with the research groups in these disciplines. For five of the six disciplines, the research groups' publications fall in more WoS SCs than the panels' publications; Pharmaceutical Sciences being the exception.

Table 15 shows that there are common WoS SCs where both the panel and research groups from the same department have publications. Simultaneously, there are some WoS SCs where either the panel or the research groups have publications. The WoS SCs where neither the panel nor the research groups have publications are excluded.

We ranked the WoS SCs in decreasing order of the number of records for each department and each panel. Table 16 presents a direct comparison of the top five WoS SCs for each panel and the research groups (i.e., all research groups of a department taken together). There are two common WoS SCs in the top five between the panel and the department in Chemistry, while there are three common WoS SCs between the panel and the department in the Physics, Biology, Biomedical Sciences, Veterinary Sciences, and Pharmaceutical Sciences.

We determine the correlation between the rankings of WoS SCs in two publication portfolios using Pearson's correlation coefficient (r) and Spearman's rank order correlation coefficient (ρ). To calculate correlation, the value zero was kept on the corresponding categories in which either the panel or the groups had no publications (but not both).

Table 16: Comparison of top five WoS SCs for panels' and research groups' per department

Panel publications		Research groups publications	
WoS SCs	**Records**	**WoS SCs**	**Records**
Biology department			
Ecology	187	*Ecology*	256
Plant sciences	102	*Environmental sciences*	215
Biochemistry molecular biology	99	*Zoology*	170
Zoology	95	*Plant sciences*	136
Entomology	89	*Evolutionary biology*	113
Biomedical Sciences department			
Biochemistry molecular biology	301	*Neurosciences*	264
Cell biology	165	*Genetics heredity*	227
Neurosciences	155	*Clinical neurology*	210
Genetics heredity	131	*Biochemistry molecular biology*	142
Biophysics	91	*Pharmacology Pharmacy*	82
Chemistry department			
Chemistry inorganic & nuclear	798	*Chemistry physical*	198
Chemistry organic	458	*Chemistry analytical*	194
Chemistry analytical	350	*Spectroscopy*	164
Chemistry multidisciplinary	324	*Physics atomic molecular & chemical*	100
Chemistry physical	177	*Physics applied*	77
Pharmaceutical Sciences department			
Pharmacology pharmacy	509	*Pharmacology pharmacy*	91
Chemistry multidisciplinary	260	*Chemistry medicinal*	53
Biochemistry molecular biology	99	*Environmental sciences*	48
Chemistry analytical	98	*Chemistry analytical*	36
Chemistry medicinal	95	*Plant sciences*	32
Physics department			
Physics condensed matter	410	*Physics condensed matter*	515
Physics multidisciplinary	188	*Physics applied*	252
Chemistry physical	182	*Physics multidisciplinary*	231
Physics applied	159	*Materials science multidisciplinary*	226
Optics	124	*Chemistry physical*	193
Veterinary Sciences			
Veterinary sciences	455	*Veterinary sciences*	42
Reproductive biology	202	*Reproductive biology*	29
Agriculture dairy animal science	107	*Cell biology*	28
Anatomy morphology	67	*Neurosciences*	27
Immunology	62	*Agriculture dairy animal science*	20

4.2.1 Correlation coefficient

An obvious improvement to this rather crude approach consists in taking more than just the top five (or some other threshold) into account.

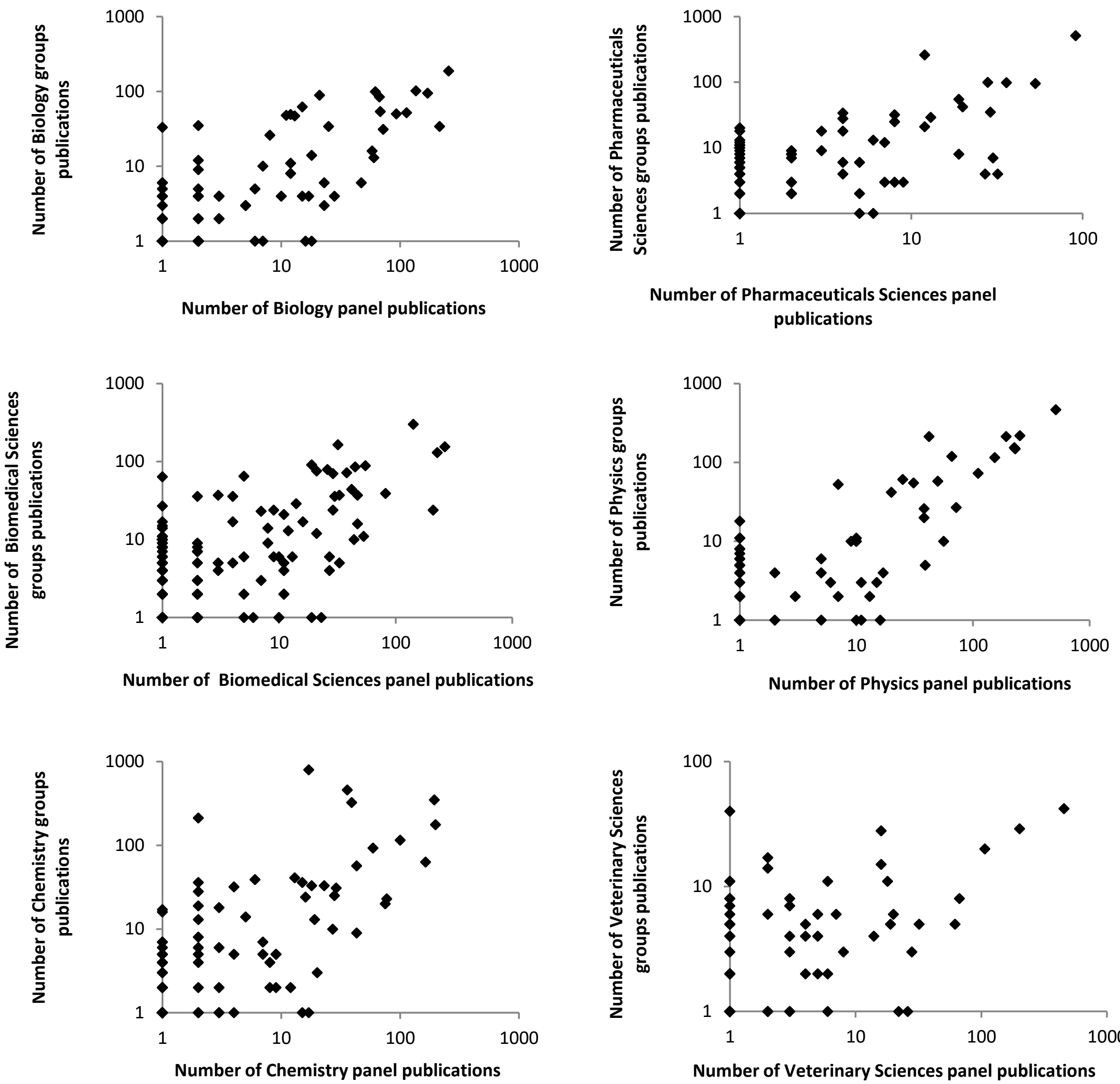


Figure 6: Log-log plots of the number of publications (log-log scale) per WoS SC for the panel (horizontal axis) and research groups together (vertical axis) for all the six departments

Figure 6 presents log-log plots of the correlation between panel and corresponding department's publications occurrence in the WoS SCS. The correlation coefficient between the panel and the department in Biology is ($r = 0.78$, $\rho = 0.53$), Biomedical Sciences ($r = 0.63$, $\rho = 0.57$), Chemistry ($r = 0.39$, $\rho = 0.37$), Pharmaceutical Sciences ($r = 0.70$, $\rho = 0.29$), Physics ($r = 0.92$, $\rho = 0.52$), Veterinary Sciences ($r = 0.70$, $\rho = 0.30$). Pearson correlation is strong to moderate while the Spearman's rank order correlation coefficient is moderate to low. An overview of the correlations is shown in the Table 19.

4.2.2 Top-down correlation coefficient

In our data, there is a large number of zeros. For example, in total there are N WoS SCs. If x is the number of WoS SCs in which either research group or panel (member) has published, then typically $x \ll N$. Since traditional correlation coefficients like Pearson and Spearman are not well-adapted to zero-inflated data (i.e., data with a large amounts of zeroes; Tu, 2006), we adopt the top-down correlation coefficient (Iman & Conover, 1987). For this case, all N WoS SCs are considered, including the ones with common zeros. The top-down coefficient places emphasis on the higher ranked data by computing the correlation using Savage scores derived from the ranked data.

Savage scores are calculated as follows:

$$S_i = \sum_{j=i}^{n} 1/j \tag{8}$$

where i is an item's rank among a set of n items. For instance, if $n = 3$, the three Savage scores are $S_1 = 1 + \frac{1}{2} + \frac{1}{3}$, $S_2 = \frac{1}{2} + \frac{1}{3}$, and $S_3 = \frac{1}{3}$. The top-down correlation coefficient is calculated as:

$$r_{td} = (\sum_{i=1}^{n} S_{R_i}\ S_{Q_i} - n)/(n - S_1) \tag{9}$$

where S is the Savage score, R_i and Q_i are the ranks of the data in the two samples, and n is the sample size. In case of ties, we use the average Savage score. This correlation coefficient was found to be an adequate rank correlation coefficient for zero-inflated data (Huson, 2007). For a full description of the top-down correlation coefficient we refer to Iman and Conover (1987).

We find positive, moderate correlations between the panel and the department of Chemistry (0.54), Pharmaceutical Sciences (0.53), Veterinary Sciences (0.52) and rather low correlations in Physics (0.39), Biomedical Sciences (0.30), and Biology (0.23).

4.2.3 Cosine similarity

We measure cosine similarity (Salton & McGill, 1986) between the publication vectors of panel and research groups. Cosine similarity is a measure of similarity between two non-zero vectors. Common zeros do not affect the cosine similarity value (Ahlgren et al., 2003). In our case, we represent panel and research group by vectors in which each item represents the number of publications in a particular WoS SC (or journal). We use the following formula:

$$\text{Cosine similarity (A, B)} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \qquad (10)$$

where A and B are the vectors for group and panel. In cosine similarity, the value ranges from -1 to 1, therefore we have three classes of similarities (1: the same, 0: dissimilar, -1: opposite in nature). It is mentionable that in our case, the publication vector has only non-negative values. Therefore, the cosine similarity value ranges from 0 to 1.

We find high similarity between the panel and the research groups of Chemistry (0.94), moderate similarity for Pharmaceutical Sciences (0.58) and Veterinary Sciences (0.45), and low similarity for Biology (0.32), Biomedical Sciences (0.17) and Physics (0.14). We have done the same exercise at the journal level as explained in the next section.

A weighted generalization of cosine similarity (Zhou et al., 2012) is discussed in chapter III, section 3.2.5.

4.4.3 Comparison of panel and research group profiles at the level of journals

Table 17 shows that together, the Biomedical Science panel and groups have published in 744 journals, when considered separately, panel publications appear in 395 journals, and group publications in 476 journals. Chemistry and Physics panel and groups have 79 and 98 journals in common respectively.

Table 17: Distribution of the panels' and groups' publications in journals

Department	Panel and groups total	Panel and groups common	Groups	Panel
Biology	496	93	372	217
Biomedical Sciences	744	127	476	395
Chemistry	506	79	293	292
Pharmaceutical Sciences	419	61	180	299
Physics	514	98	355	257
Veterinary Sciences	318	28	146	200

In Biology, 124 journals contained panel publications, but no group publications, and 279 journals have group publications but no panel publications. The Pharmaceutical Sciences and Veterinary Sciences panels have published in more journals than the research groups, while for Chemistry the number of journals used is almost equal. For the remainder of the disciplines it appears that the groups have published in more journals than their respective evaluation panels.

We ranked the journals in decreasing order of the number of records for each department and each panel. Table 18 presents a direct comparison of the top five journals for each panel and the research groups (i.e., all research groups of a department taken together).

There are no common journals in the top five journals between the panels and departments in the Chemistry, Biomedical Sciences, and Pharmaceutical Sciences. For Biology, Physics and Veterinary Sciences, Table 18 shows that there are journals in the top 5 where both the panel and the department have publications.

Table 18: Comparison of top five journals for panel and research groups per department

Panel publications Journal Title	Records	Research groups publications Journal Title	Records
Biology			
Experimental and Applied Acarology	35	*Environmental Pollution*	40
General and Comparative Endocrinology	33	*Biological Journal of the Linnean Society*	33
Journal of Experimental Biology	30	*Journal of Experimental Biology*	26
Proceedings of the Royal Society B: Biological Sciences	22	*Aquatic Toxicology*	23
New Phytologist	22	*Environmental Science Technology*	22
Biomedical Sciences			
Magnetic Resonance in Medicine	36	*PLoS ONE*	33
Journal of Virology	27	*Human Mutation*	32
Journal of Biological Chemistry	26	*Neurology*	28
Science	24	*Neurobiology of Aging*	20
Acta crystallographica Section d Biological Crystallography	22	*American Journal of Human Genetics*	20
Chemistry department			
Inorganic Chemistry	213	*Spectrochimica Acta Part B Atomic Spectroscopy*	37
Organometallics	174	*Journal of Physical Chemistry A*	37
Journal of Organometallic Chemistry	106	*Journal of Analytical Atomic Spectrometry*	35
Analytical Chemistry	94	*X ray Spectrometry*	27
Journal of the American Chemical Society	86	*Analytical Chemistry*	26
Pharmaceutical Sciences			
Pharmaceutical Research	52	*Kidney International*	13
British Journal of Clinical Pharmacology	35	*Planta Medica*	11
Archiv der Pharmazie	35	*Environmental Science Technology*	8
Clinical Pharmacology Therapeutics	27	*Journal of Mass Spectrometry*	7
Monatshefte Fur Chemie	23	*Chemosphere*	7
Physics department			
Physical Review B	246	*Physical Review B*	291
Physical Review Letters	98	*European Physical Journal C*	108
Surface Science	56	*Physics Letters B*	72
Applied Physics Letters	54	*Physical Review Letters*	60
Physics Letters B	51	*Physica C Superconductivity and its Applications*	43
Veterinary Sciences			
Theriogenology	71	*Theriogenology*	11
Reproduction in Domestic Animals	47	*Reproduction in Domestic Animals*	10
Animal reproduction science	40	*Neurogastroenterology and Motility*	7
Berliner und Munchener Tierarztliche Wochenschrift	38	*Histochemistry and Cell Biology*	5
Kleintierpraxis	30	*Vlaams Diergeneeskundig Tijdschrift*	4

More precisely Table 18 shows that there is one common journal – *Journal of Experimental Biology* – between the Biology panel and department, two common journals – *Theriogenology* and *Reproduction in Domestic Animals* – between the Veterinary panel and department, and three common journals – *Physical Review B, Physical Review Letters, Physics Letters B* – in the case of Physics. The Physics, Biology and Veterinary Sciences departments fare somewhat better than the other three disciplines. We conclude that, overall, the overlap between the top five of journals used by the department and the panel is quite low.

4.3.1 Correlation coefficient

We determine the correlation between the rankings of journals in two publication portfolios using Pearson's correlation coefficient and Spearman's rank order correlation coefficient. Scatter plots of correlation between panel and corresponding department's publications occurrence in the journals are shown in Figure 7.

Values for the Pearson's correlation coefficient vary from high to low – Physics ($r = 0.85$), Veterinary Sciences ($r = 0.29$), Biology ($r = 0.26$), Chemistry ($r = 0.16$), Biomedical Sciences ($r = 0.13$), except Pharmaceutical Sciences where the correlation is negative ($r = -0.05$). The Spearman's rank correlation coefficient between the panel and the department is negative in all cases: Biology ($\rho = -0.28$), Chemistry ($\rho = -0.36$), Physics ($\rho = -0.24$), Biomedical Sciences ($\rho = -0.43$), Pharmaceutical Sciences ($\rho = -0.47$) and Veterinary Sciences ($\rho = -0.66$).

We find that the two correlation coefficients are very different in this case. For the same data, the Pearson's correlation coefficient is positive in five cases, while Spearman's rank correlation coefficient is negative in all the cases. The difference between the correlations is large for the case of Physics due to an outlier – *Physical Review B* – where both the sides have over 250 publications. If we exclude that outlier the Pearson's correlation coefficient drops ($r = 0.51$) while Spearman's rank correlation coefficient remains the same.

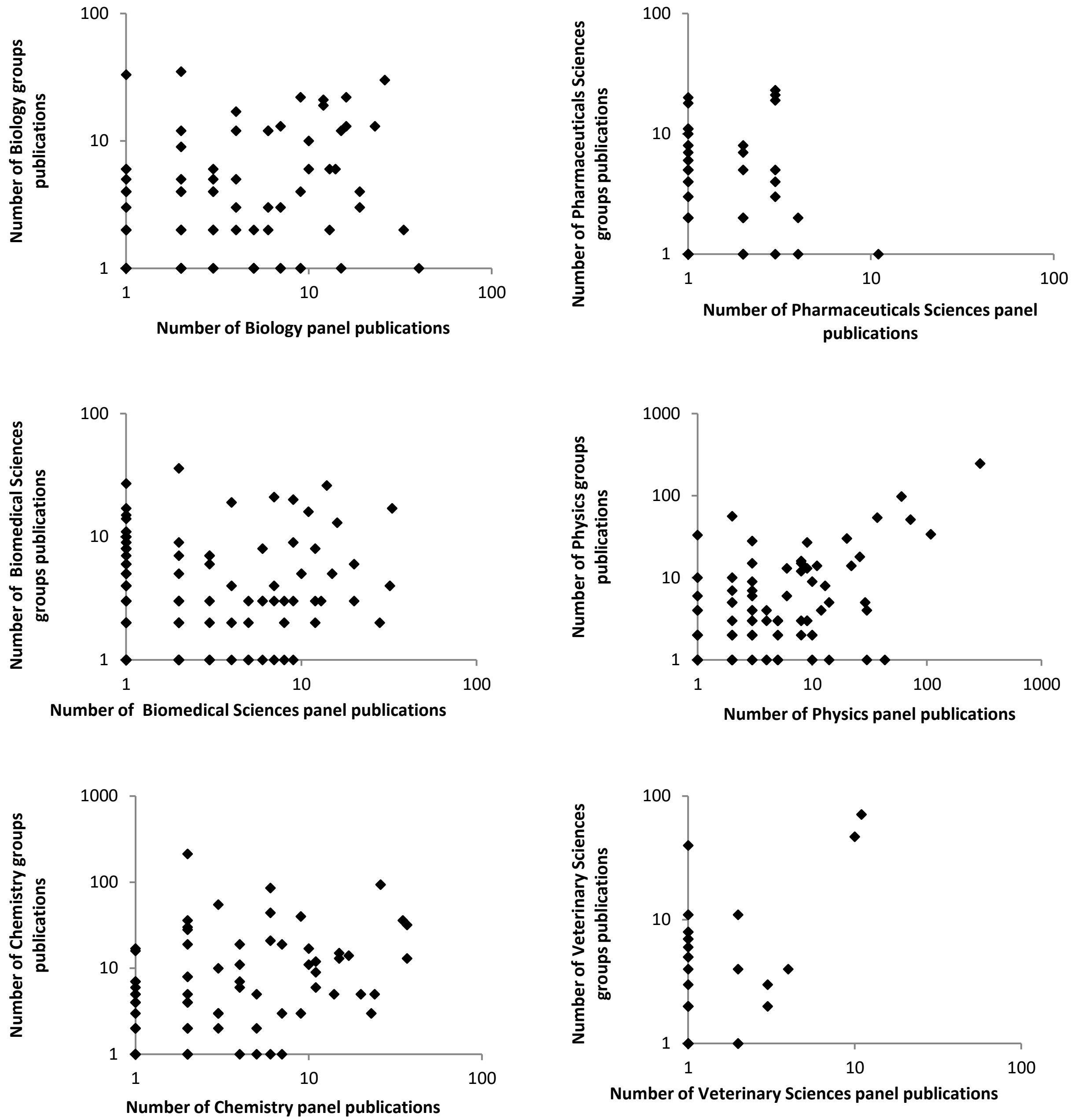


Figure 7: Scatter plot of the number of publications (log-log scale) per journals for the panel (horizontal axis) and research groups together (vertical axis) for all the six departments

4.3.2 Top-down correlation coefficient

We also determine the correlation between the rankings of journals using the top-down correlation. As explained earlier, we employ the top-down correlation to account for the large number of zeroes in our data (Huson, 2007). We find a positive but very low correlation in all cases: Chemistry (0.14), Biology (0.15), Physics (0.16), Biomedical Sciences (0.16). Pharmaceutical Sciences (0.09), Veterinary Sciences (0.06).

4.3.3 Cosine similarity

Using formula (10), we measure the cosine similarity at the journal level. For all the cases, we find low similarity values: Biology (0.19), Biomedical Sciences (0.16), Pharmaceutical Sciences (0.14), Chemistry (0.10), Veterinary Sciences (0.04) and Physics (0.03).

4.4 Conclusion

We have determined the association between the publication output of two entities using Pearson's correlation coefficient and Spearman's rank correlation coefficient, the top-down correlation coefficient, and cosine similarity.

Table 19 summarizes the obtained correlation and similarity values between each department and corresponding panel. At the level of WoS SCs, both the Spearman and top-down correlations are moderate to low. The Pearson's correlation coefficients are strong to moderate. The cosine similarity for Chemistry is remarkably high, while the others have moderate to low similarity values. At the level of journals, all the values are high to low for Pearson's correlation except Pharmaceutical Sciences; however, values are negative in all cases for Spearman correlation. Top-down correlation and cosine similarity are low in all the cases at the level of journals.

The correlation coefficients and cosine similarity in WoS SCs are strong to moderate while in journals it is low to negative. This is likely due to the fact that the total number of WoS SCs (224 SCs) is lower than the total number of journals (10,675 journals). Hence, two publications are far more likely to be in the same WoS SC than in the same journal.

Table 19: Correlation and similarity in the six departments

Name of the department	WoS SCs				Journals			
	Pearson	Spearman	Top-Down	Cosine Similarity	Pearson	Spearman	Top-Down	Cosine Similarity
Biology	0.78	0.53	0.23	0.32	0.26	- 0.28	0.15	0.19
Biomedical Sciences	0.63	0.57	0.30	0.17	0.13	- 0.43	0.16	0.16
Chemistry	0.39	0.37	0.54	0.94	0.16	- 0.36	0.14	0.10
Pharmaceutical Sciences	0.70	0.29	0.53	0.58	-0.05	- 0.47	0.09	0.14
Physics	0.92	0.52	0.39	0.14	0.85	- 0.24	0.16	0.03
Veterinary Sciences	0.70	0.30	0.52	0.45	0.29	- 0.66	0.06	0.04

We argue that such correlation and similarity measures are insufficient, since they do not take into account the relatedness of WoS SCs or journals. One can intuitively understand that some categories are much more closely related than others. If a panel member has many publications in closely related WoS SCs or journals, she may still have relevant expertise, even if she has no publications in the exact same category or journal as the group to be evaluated. This is reminiscent of the way diversity is sometimes studied using only the dimensions of variety and balance. As discussed by Stirling (2007), the additional dimension of disparity – the opposite concept of similarity – is needed to provide a complete picture. Likewise, a comparison of journal publication profiles that does not take WoS SC similarity or journal similarity into account might yield distorted results.

Chapter V: Is the expertise of evaluation panels congruent with the research interests of the research groups: A quantitative approach based on barycenters[3]

5.1 Introduction

Using data collected in the framework of two completed research evaluations, this chapter focuses on the expertise overlap between expert panels and the research groups involved in the evaluation. An expert panel typically consists of independent specialists, and is multidisciplinary and/or interdisciplinary in its composition; each of the members are recognized experts in at least one of the fields addressed by the department under evaluation. Surprisingly, the degree to which the expertise of the panel (members) overlaps with the expertise of the research groups has not been quantified to date. The goal of this chapter is therefore to present a bibliometric methodology to assess the congruence of panel expertise and research interests in the units under assessment. As such, we present a bibliometric analysis of the overlap of expertise between research groups in the Departments of Chemistry and Physics and the respective expert panels based on two research evaluations carried out at the University of Antwerp. We focus on the following research questions:

v) How can one visualize the expertise of two entities (e.g., a research group and a panel) using publication data?

vi) How can one quantify the cognitive distances (overlap of expertise) between two entities (e.g., a research group and a panel) using the WoS SCs to which their publications belong?

[3] This chapter is based on Rahman, Guns, Rousseau & Engels (2015) and takes into account the correction published in Rahman, Guns, Rousseau & Engels (2016).

We address these questions in the context of expert panel reviews. Specifically, we focus on comparing:

- panel and individual research group;
- panel member and individual research group (even if the panel does not cover a group's expertise well, it may suffice that one panel member does);
- panel and all reviewed research groups (e.g., all physics research groups).

5.2 Data

The data in this chapter stem from the 2009 assessment of the twelve research groups (referred to as CHEM-A, CHEM-B and so on) belonging to the Department of Chemistry, and the 2010 assessment of the nine research groups (referred to as PHYS-A, PHYS-B and so on) belonging to the Department of Physics, University of Antwerp.

The reference period encompasses eight years preceding the evaluation. In principle all articles, letters, notes, proceedings papers, and reviews by the research groups published during the reference period were considered in the evaluation. In this thesis, we only consider the publications that are indexed in SCIE and SSCI of WoS.

Table 20 lists the number of publications of the research groups during the eight years preceding their evaluation. The Chemistry research groups published 920 publications in 300 journals, including 43 joint publications between two Chemistry research groups. In total, their publications are distributed over 94 WoS SCs. While the Physics research groups generated 1739 publications in 353 journals, with 150 publications co-authored by members of two and seven publications co-authored by members of three research groups. In total, their publications are distributed over 108 WoS SCs.

The Chemistry and Physics panels were composed of seven and six members (both including the chair), respectively. All the publications of the individual panel members up to the year of assessment were taken into account.

Table 20: Publication profile of the Chemistry and Physics research groups

Group code	Number of Publications	Number of Journals	Number of WoS SCs
Chemistry research groups (2001-2008)			
CHEM-A	129	47	27
CHEM-B	65	24	17
CHEM-C	156	52	26
CHEM-D	32	17	13
CHEM-E	70	39	23
CHEM-F	21	17	8
CHEM-G	161	47	42
CHEM-H	62	33	28
CHEM-I	51	24	19
CHEM-J	27	11	15
CHEM-K	97	66	48
CHEM-L	92	42	24
Total	**920**	**300**	**94**
Physics research groups (2002-2009)			
PHYS-A	125	53	44
PHYS-B	486	66	25
PHYS-C	525	147	46
PHYS-D	269	17	7
PHYS-E	159	55	28
PHYS-F	42	23	13
PHYS-G	43	26	12
PHYS-H	132	31	12
PHYS-I	115	63	49
Total	**1739**	**353**	**108**

The Chemistry panel members' publication output amounts to 2150 publications in 248 different journals and are assigned to 66 different WoS SCs. The number of publications per panel member ranges from 113 to 694. Panel members one and seven have two joint publications. The combined publication output of the Physics panel members is 1104 publications, none of which are co-authored publications between panel members. The number of publications per panel member ranges from 117 to 282. In total, these publications appeared in 204 different journals and are assigned to 46 different WoS SCs.

5.3 Methods

5.3.1 Subject category similarity matrix and maps

Each journal in WoS is assigned to one or more WoS SCs. Our method is based on the assumption that entities with more publications in the same or similar WOS SCs have greater expertise overlap. While WoS SCs have been criticized for being crude (Leydesdorff & Rafols, 2009; Leydesdorff & Bornmann, 2016) they are considered sufficient for evaluation of a given discipline (van Leeuwen & Medina, 2012), and are widely accepted and used by bibliometric practitioners. Moreover, the categories cover all disciplines (Rehn, et. al., 2014; Leydesdorff & Bornmann, 2015).

To operationalize the relatedness or similarity of WoS SCs, we draw upon data made available by Rafols, et al. (2010) at http://www.leydesdorff.net/overlaytoolkit/map10.paj. These authors created a matrix of citing to cited SCs based on the SCIE and SSCI, which was subsequently normalized in the citing direction. Only cosine values > 0.15 were retained. The result is a symmetric $N \times N$ similarity matrix (here, N=224). If we interpret it as an adjacency matrix, we see that it is equivalent to a weighted network, in which similar categories are linked (the higher the link weight, the stronger the similarity). The two most similar SCs are *Nanoscience & Nanotechnology* and *Materials Science, Multidisciplinary*, which have a cosine similarity of 0.978.

The information in the similarity matrix can be visualized. The subfield of bibliometric mapping is dedicated to the visualization, clustering and interpretation of similarity matrices or networks like the one we use. Many different algorithms or layout techniques have been developed for this purpose. In this chapter, we use two:

- Kamada-Kawai (Kamada & Kawai, 1989) is a spring-based layout algorithm for networks, which is implemented in, among others, Pajek (de Nooy et al., 2012). Kamada-Kawai is the algorithm used by (Rafols et al., 2010).

- VOS (van Eck & Waltman, 2007) stands for ‘visualization of similarities’ and is a variant of multidimensional scaling (Borg & Groenen, 2005; van Eck, Waltman, Dekker, & van den Berg, 2010). It is implemented in VOSviewer and in recent versions of Pajek.

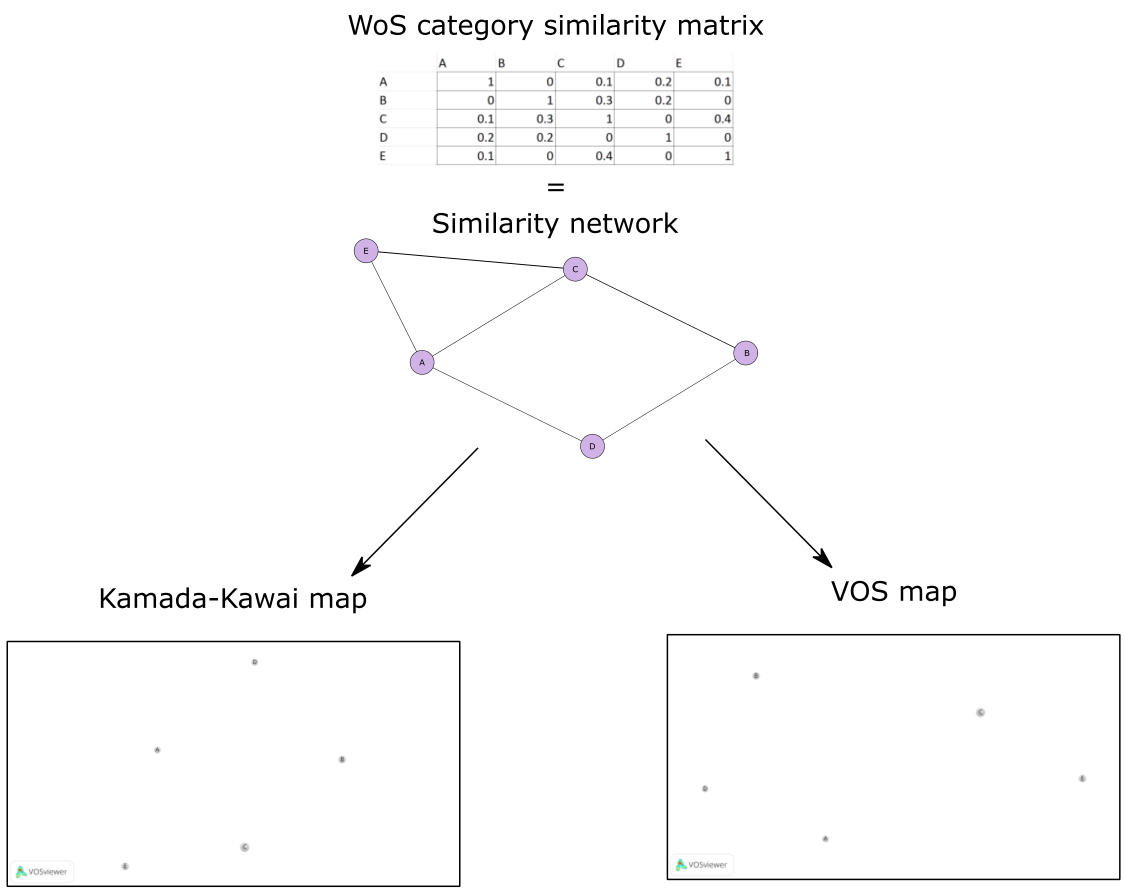


Figure 8: Overview of similarity matrix and maps

Figure 8 provides an overview of the relations between similarity matrix, network and the two maps. Since the source data include all research fields included in the SCIE and SSCI, the resulting maps are global maps of science (as opposed to local maps of science, which focus on one or a few disciplines).

5.3.2 Overlay maps

Combining the maps described in the previous section with publication data (how many publications in which SCs?), one can create overlay maps as visual representation of the expertise of a research unit (Leydesdorff & Rafols, 2009; Leydesdorff, Carley, et al., 2013; Rafols et al., 2010). To answer the first research question of this chapter, we created overlay maps based on a base map of science. In an overlay map, the original map – referred to as the *base map* – provides the location (and sometimes cluster) of each SC, whereas publication data is used to visualize the unit's publication intensity for each SC. Typically, this is done by scaling the size of each node according to the number of publications. Hence, overlay maps can also be used for visual comparison and estimation of the degree of overlap of two or more entities in exploratory analysis. These overlay maps provide an answer to the first research question.

In the 'Results' section, we present several overlay maps. Some of these are zoomed in to better highlight places of interest. All distances presented are based on the coordinates in the original maps and hence independent of whether the figures are enlarged.

For our purposes, however, overlay maps have an important limitation. Despite their value in an exploratory analysis, overlay maps are hard to compare. It is not always obvious, for instance, which of several candidate panel members has better overlap of expertise with a given group or department. This is especially the case if the entities publish in many different categories or in categories that are quite close to one another.

In order to answer the second research question of this chapter, we use the barycenter approach to estimate an entity's 'average' or 'overall' position. Consequently, one can determine and compare the cognitive distance between entities, thus adding a measure to the qualitative visual comparison facilitated by overlay maps.

5.3.3 Barycenter and distance calculation

An entity's barycenter is the center of weight (Rousseau, 1989a, 1989b; Jin & Rousseau, 2001; Rousseau, 2008) of the SCs in which it has published, where a SC's weight is the entity's number of publications therein. Now for each panel member and for each research group a barycenter is calculated and Euclidean distances between barycenters can be calculated. The coordinates of these barycenters on a two-dimensional base map is defined as the point $C = (C_1, C_2)$ where

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \;\; ; \; C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} \tag{11}$$

where m_j is the number of publications of the unit under investigation (panel member, research group) belonging to category j; this category j has coordinates ($L_{j,1}$, $L_{j,2}$) in the base map, and $T = \sum_{j=1}^{N} m_j$. Note that T is larger than the total number of publications as we use full counting of WoS SCs: if a publication appears in a journal belonging to two categories, it will be counted twice.

Having obtained barycenters for each entity, we can determine the distance between (the barycenters of) the expert panel as a whole, individual panel members, the combined group, and individual groups. The Euclidean distance between points $C = (C_1, C_2)\ and\ D = (D_1, D_2)$ is calculated with the formula:

$$d = \sqrt{(C_1 - D_1)^2 + (C_2 - D_2)^2}. \tag{12}$$

The distances thus obtained should be interpreted as having arbitrary units on a ratio scale (Egghe & Rousseau, 1990). This means there is a fixed meaningful zero (distance zero in the map), and distances can be compared in terms of percentage or fraction (e.g. the distance between A and B is 1.5 times larger than the distance between C and D).

This two-dimensional approach allows for easy visualization of barycenters: C_1 and C_2 are, respectively, horizontal and vertical coordinates.

5.4 Results

We start by calculating barycenters and Euclidean distances between the barycenters to gauge the differences between the Kamada-Kawai and VOS mapping techniques for both case studies. We can use both the VOS map and Kamada-Kawai map as a basis for visual exploration and barycenter calculation and comparison. We use a Kamada-Kawai map for the global map of science as introduced by Leydesdorff and Rafols (2009) and VOS map that is readily available (http://www.leydesdorff.net/overlaytoolkit) and applied for creating overlay maps (Leydesdorff, Carley, et al., 2013; Rafols et al., 2010). In the second part of this section, we focus on the two cases of research assessment at the university of Antwerp.

5.4.1 Euclidean distances between barycenters

Table 21 and Table 23 provide the distances based on the VOS map for the cases of Chemistry and Physics, respectively, while the distances based on the Kamada-Kawai map are provided in Table 22 and Table 24.

Comparing the results for the two approaches, we find that 5 out of 12 Chemistry groups are most closely located to the same panel member in the VOS-map and Kamada-Kawai approach. Likewise, 6 out of 9 Physics groups are most closely located to the same panel member.

Table 21: Euclidean distances between barycenters of Chemistry research groups, panel members, panel and research groups together using the VOS map

	All groups	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
Panel	0.105	0.166	0.141	0.202	0.123	0.275	0.284	0.108	0.107	0.044	0.326	0.384	0.141
PM 1	0.168	0.167	0.129	0.217	0.165	0.329	0.337	0.179	0.165	0.111	0.394	0.454	0.127
PM 2	0.200	0.350	0.342	0.362	0.129	**<u>0.079</u>**	**<u>0.090</u>**	0.145	0.215	0.199	0.259	**<u>0.228</u>**	0.342
PM 3	0.054	0.171	0.161	0.192	0.129	0.252	0.263	**<u>0.053</u>**	**<u>0.061</u>**	**<u>0.020</u>**	0.269	0.330	0.161
PM 4	0.119	0.269	0.262	0.280	0.108	0.158	0.170	0.063	0.134	0.121	**<u>0.232</u>**	0.250	0.263
PM 5	0.106	**<u>0.056</u>**	**<u>0.055</u>**	**<u>0.091</u>**	0.232	0.367	0.378	0.154	0.093	0.099	0.315	0.411	**<u>0.057</u>**
PM 6	0.200	0.302	0.276	0.335	**<u>0.027</u>**	0.175	0.181	0.161	0.210	0.156	0.366	0.370	0.275
PM 7	0.186	0.116	0.072	0.172	0.235	0.395	0.404	0.216	0.178	0.144	0.410	0.491	0.070

For each research group, we determined the panel member at the shortest distance. Average of shortest distances is 0.087 (SD 0.070). The number in the row of this panel member is indicated in bold and underlined.

Table 22: Euclidean distances between barycenters of Chemistry research groups, panel members, panel and research groups together using the Kamada-Kawai map

	All groups	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
Panel	0.047	0.055	0.039	0.093	0.025	0.044	0.055	0.013	0.058	0.056	0.125	0.109	0.036
PM 1	0.081	0.080	0.063	0.125	0.027	0.078	0.091	0.048	0.090	0.092	0.160	0.146	0.054
PM 2	0.050	0.080	0.079	0.077	0.082	**<u>0.028</u>**	**<u>0.016</u>**	0.062	0.062	0.033	**<u>0.084</u>**	**<u>0.045</u>**	0.085
PM 3	0.011	**<u>0.022</u>**	**<u>0.021</u>**	**<u>0.052</u>**	0.066	0.049	0.044	0.029	**<u>0.017</u>**	**<u>0.029</u>**	0.087	0.085	0.030
PM 4	0.049	0.058	0.043	0.095	**<u>0.022</u>**	0.042	0.054	0.015	0.060	0.056	0.126	0.109	0.040
PM 5	0.024	0.038	0.027	0.071	0.047	0.037	0.041	**<u>0.010</u>**	0.036	0.035	0.102	0.091	0.030
PM 6	0.089	0.090	0.073	0.134	0.024	0.078	0.093	0.054	0.099	0.097	0.166	0.148	0.066
PM 7	0.054	0.050	0.033	0.095	0.035	0.063	0.071	0.025	0.061	0.067	0.132	0.124	**<u>0.025</u>**

For each research group, we determined the panel member at the shortest distance. Average of shortest distances is 0.031 (SD 0.020). The number in the row of this panel member is indicated in bold and underlined.

Table 23: Euclidean distances between barycenters of Physics research groups, panel members, panel and research groups together using the VOS map

	All groups	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
Panel	0.135	1.115	0.025	0.078	0.125	0.033	0.239	0.383	0.040	0.607
PM 1	0.230	1.173	0.123	0.215	**<u>0.017</u>**	0.145	0.208	0.495	0.120	0.664
PM 2	0.214	1.195	0.067	0.109	0.158	0.118	0.316	0.443	0.056	0.688
PM 3	0.131	1.041	0.146	0.194	0.116	0.113	**<u>0.104</u>**	0.387	0.157	0.532
PM 4	0.100	**<u>1.020</u>**	0.168	0.085	0.263	0.132	0.295	**<u>0.249</u>**	0.179	**<u>0.522</u>**
PM 5	0.156	1.136	0.046	**<u>0.055</u>**	0.159	**<u>0.069</u>**	0.281	0.385	0.050	0.629
PM 6	0.175	1.157	**<u>0.031</u>**	0.084	0.138	0.078	0.280	0.412	**<u>0.026</u>**	0.649

For each research group, we determined the panel member at the shortest distance. Average of shortest distances is 0.232 (SD 0.337). The number in the row of this panel member is indicated in bold and underlined.

Table 24: Euclidean distances between barycenters of Physics research groups, panel members, panel and research groups together using the Kamada-Kawai map

	All groups	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
Panel	0.050	0.349	0.018	0.047	0.154	0.009	0.118	0.093	0.015	0.127
PM 1	0.246	0.537	0.206	0.244	**<u>0.045</u>**	0.207	0.128	0.291	0.197	0.315
PM 2	0.023	0.322	0.022	0.027	0.180	0.020	0.136	0.070	0.029	**<u>0.101</u>**
PM 3	0.130	0.423	0.090	0.130	0.072	0.092	**<u>0.064</u>**	0.176	0.081	0.200
PM 4	0.044	**<u>0.309</u>**	0.065	**<u>0.017</u>**	0.214	0.053	0.179	**<u>0.034</u>**	0.071	0.103
PM 5	0.026	0.323	0.028	0.020	0.183	0.020	0.142	0.065	0.034	0.104
PM 6	0.047	0.346	**<u>0.017</u>**	0.044	0.158	**<u>0.006</u>**	0.120	0.090	**<u>0.015</u>**	0.124

For each research group, we determined the panel member at the shortest distance. Average of shortest distances is 0.067 (SD 0.095). The number in the row of this panel member is indicated in bold and underlined.

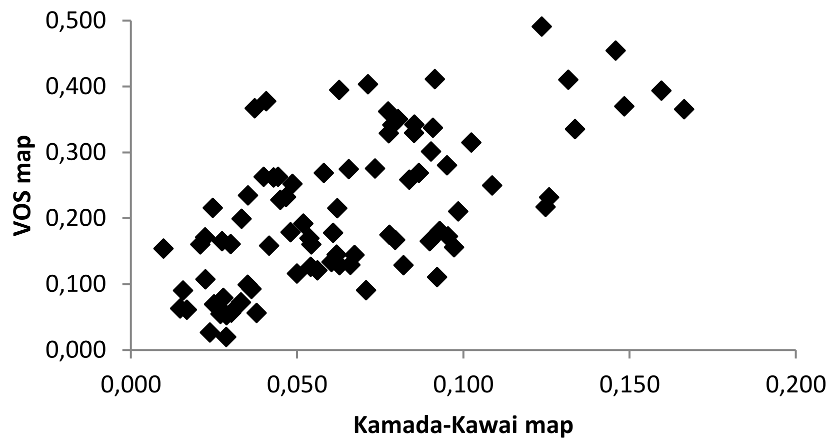


Figure 9: Scatter plots of the Euclidean distances between barycenters of research groups and individual panel members between VOS map and Kamada-Kawai map in the Chemistry department

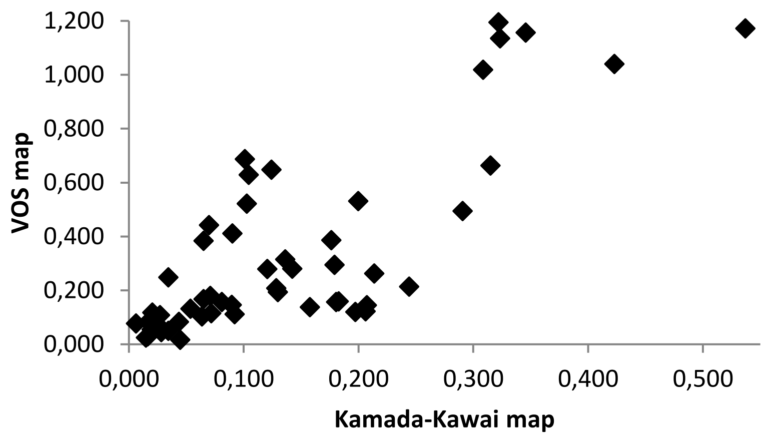


Figure 10: Scatter plots of the Euclidean distances between barycenters of research groups and individual panel members between VOS map and Kamada-Kawai map in the Physics department

Spearman's rank correlation coefficients (ρ) were calculated between the two techniques. For this, we take the Euclidean distances between barycenters of research groups and panel members only. Although there are co-publications between groups, the Euclidean distance between barycenters of research groups and individual panel members can be (or at least are) considered independent, and have been included in the correlation calculation.

Correlations for the Chemistry department (see Figure 9) are moderately strong between the VOS map and the Kamada-Kawai map (ρ = 0.64), and strong in the Physics department (ρ = 0.80, see Figure 10). In summary, the barycenter distances between the VOS-map and the Kamada-Kawai map are fairly strongly correlated. In the remainder of this chapter, calculations of barycenters, Euclidean distances, comparisons, and visual explorations are based on the VOS map.

5.4.2 Case studies of University of Antwerp research assessments

5.4.2.1 Chemistry assessment

5.4.2.1.1 Panel profile versus groups profile

The overlay maps for the Chemistry panel (Figure 11) and the combined groups (Figure 12) clearly show that the publication scope of the combined chemistry groups is wider than that of the panel.

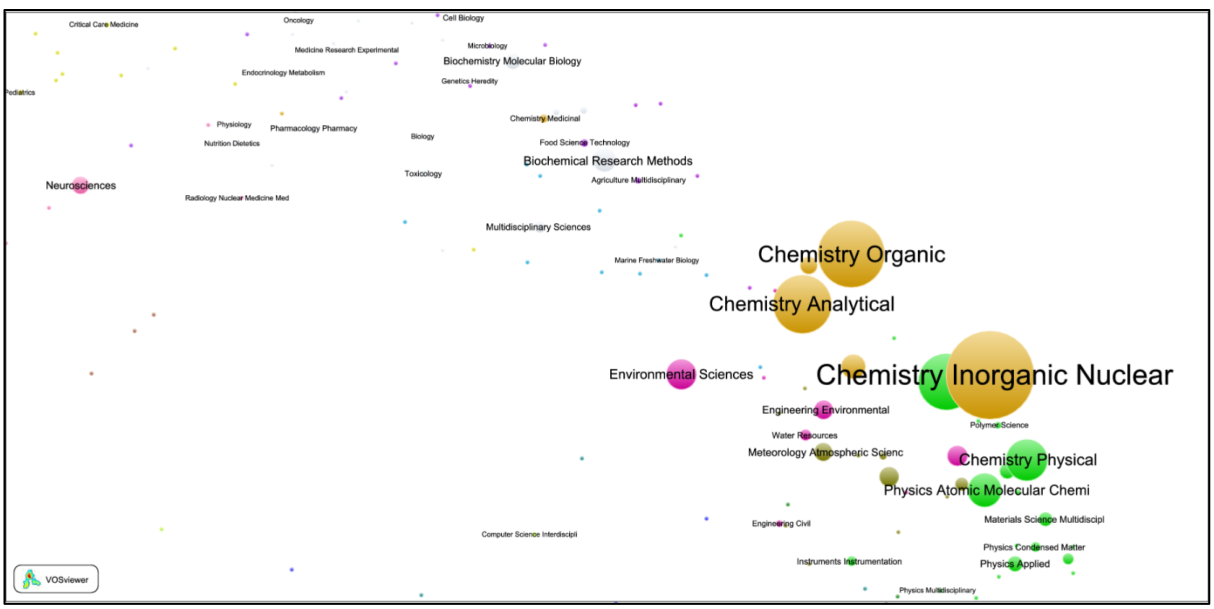


Figure 11: Chemistry panel members' publication overlay map

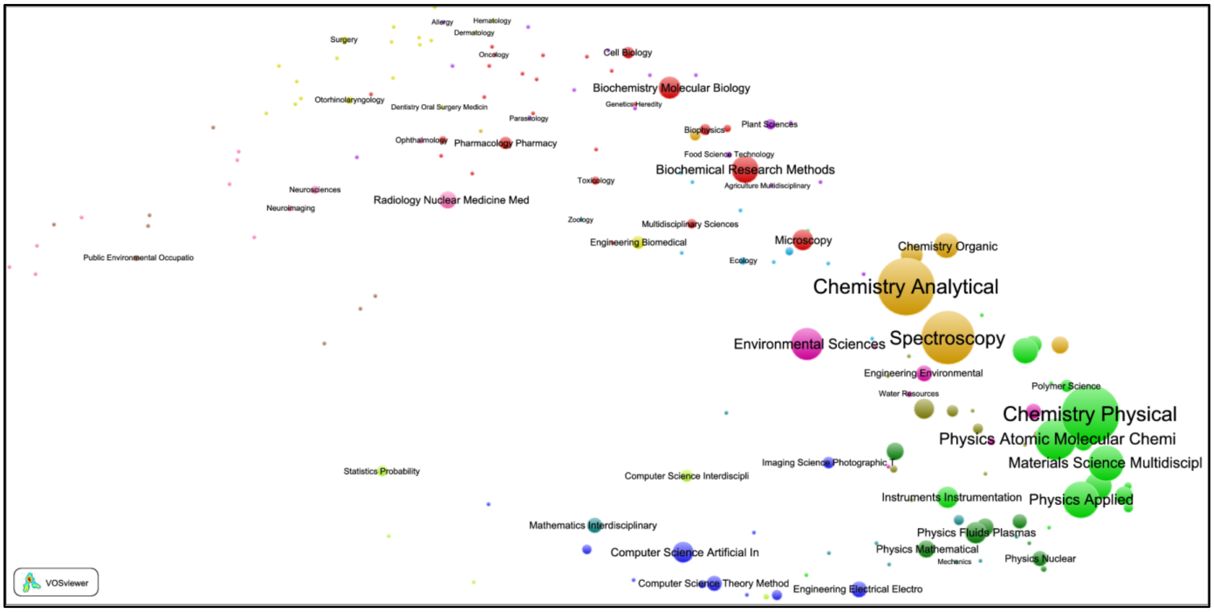


Figure 12: Chemistry groups' publication overlay map

The panel publications are strongly (74.67%) represented in the WoS SCs of *Chemistry inorganic & nuclear*, *Chemistry organic*, and *Chemistry analytical*, whereas the research group publications are predominantly clustered (60.43%) in *Chemistry physical*, *Chemistry analytical*, and *Spectroscopy*.

5.4.2.1.2 Panel profile versus individual group profile

Overlay maps of the publications of the individual groups were created, and subsequently compared with the panel overlay map (see Figure 11). We present the data for CHEM-A as an example. Figure 13 shows the corresponding overlay map. The majority of the publications of the CHEM-A group fall in *Chemistry physical* (48.06%) and *Physics atomic molecular & chemical* (34.88%) WoS SCs. We have found that the research output of six (CHEM-A, CHEM-B, CHEM-D, CHEM-G, CHEM-I, and CHEM-L) of the twelve research groups, are thematically well covered by the panel's expertise, i.e., the majority of the panel's publications can be classified in WoS SCs, where also the majority of the corresponding group publications is found.

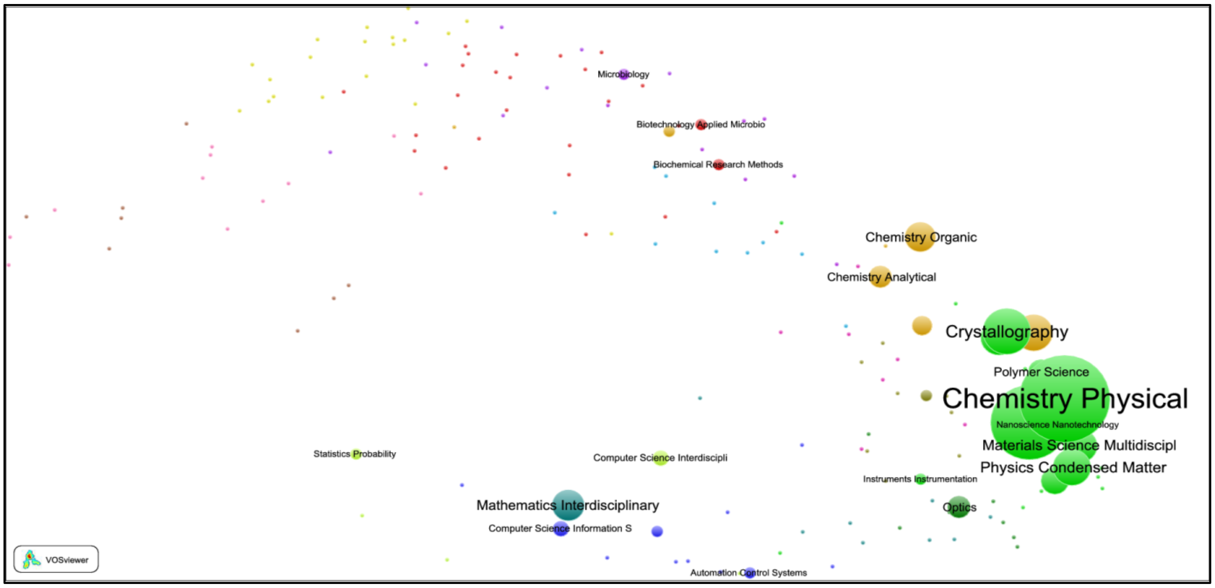


Figure 13: Overlay map of CHEM-A research group's publications

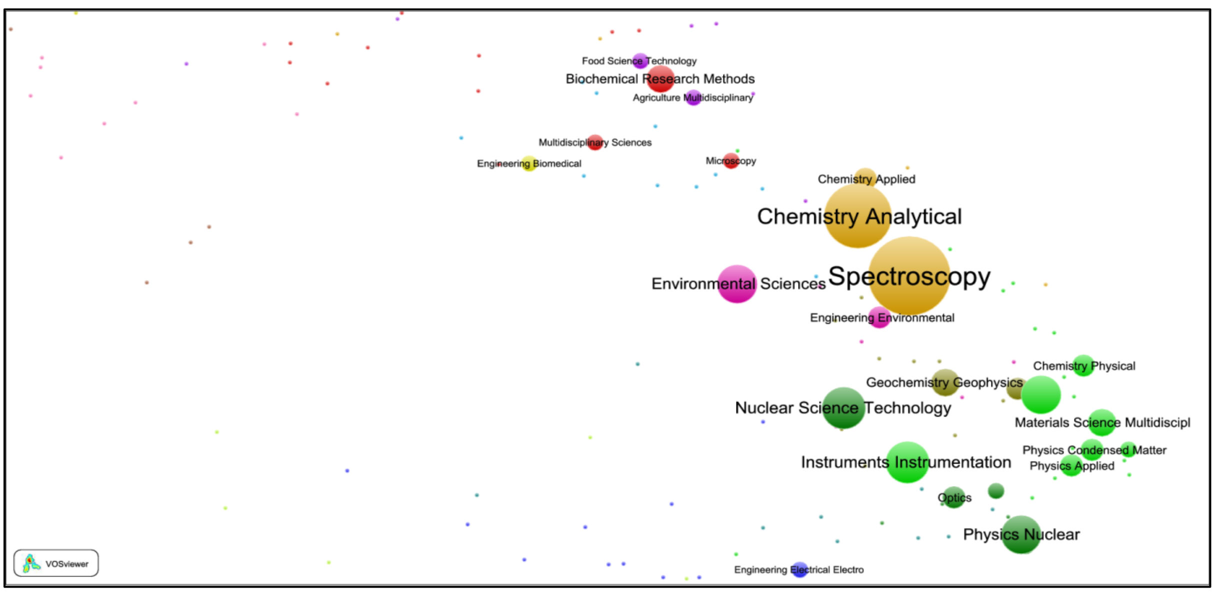


Figure 14: Overlay map of CHEM-H research group's publications

Furthermore, the majority of the CHEM-C group publications falls in *Physics applied* (35.25%) and *Spectroscopy* (23.71%); for CHEM-H the dominant SCs are *Spectroscopy* (40.32%) and *Chemistry analytical* (27.41%; Figure 14). These two research groups have a large number of publications in WoS SCs where the publications output of the panel tends to be limited.

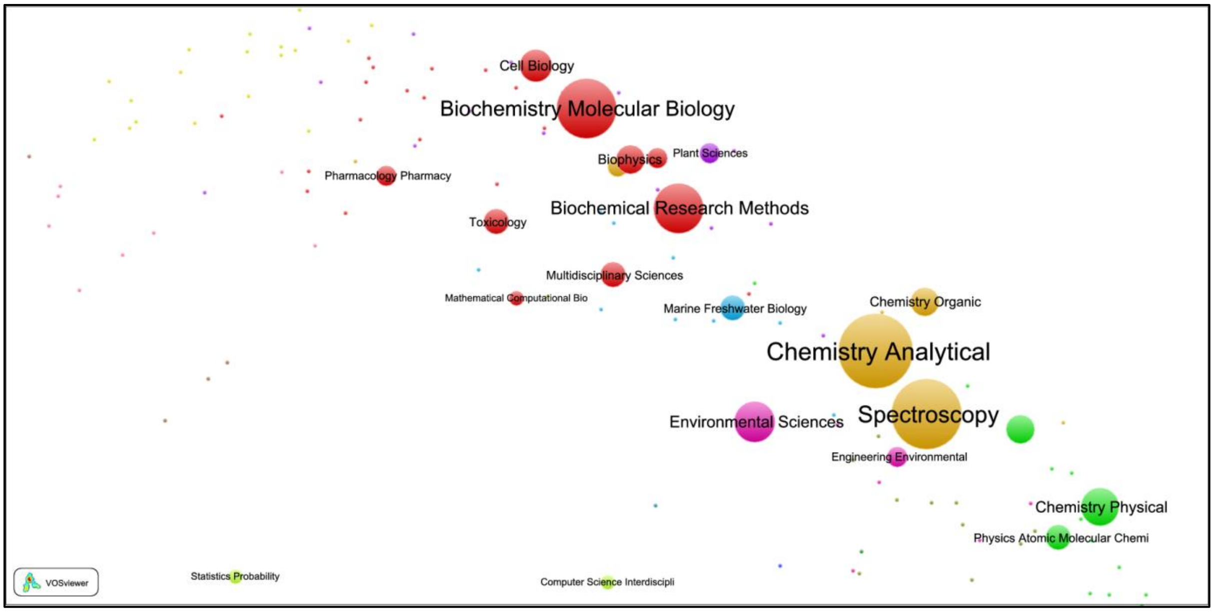


Figure 15: Overlay map of CHEM-E research group's publications

Therefore, the research output of these two research groups is only partially covered by the panel's expertise.

Likewise, the majority of the publications of CHEM-E group (Figure 15) fall in *Chemistry Analytical* (38.57%) and *Spectroscopy* (34.28%); CHEM-F group: *Chemistry analytical* (66.66%) and *Biochemical research methods* (23.81%); CHEM-J group: *Chemistry analytical* (48.14%) and *Instruments Instrumentation* (33.33%); CHEM-K group: *Microscopy* (81.48%) and C*omputer science artificial intelligence* (70.37%) WoS SCs. Therefore, these four research groups hardly have any overlap in terms of the share of their publications in WoS SCs where the evaluation panel has publications.

In summary, of the twelve Chemistry groups, six groups are well covered, two groups are partially covered, and the remaining four groups seem poorly covered by the Chemistry panel's expertise as far as publication output is described via WoS SCs.

5.4.2.1.3 Distances between barycenters

Figure 16 and Table 21 provide data on the distances between the Chemistry panel's barycenter/coordinates and those of the individual Chemistry groups (panel members are indicated by the symbol PM). The CHEM-I group is very close to the panel while CHEM-K group is almost 2.3 times farther away than CHEM-A group, and CHEM-F group is 2.6 times further away than CHEM-G group. CHEM-B (0.141), CHEM-D (0.123), CHEM-G (0.108), CHEM-H (0.107), and CHEM-L (0.141) are situated comparatively close to the panel's coordinates. CHEM-F (0.284), CHEM-J (0.326), and CHEM-K (0.384) are located farther away, and CHEM-A (0.166) and CHEM-C (0.202) too are found at a considerable distance from the panel's barycenter.

A further comparison of the distances between the Chemistry groups and individual Chemistry panel members as presented in Table 21 reveals that the partially covered CHEM-C and CHEM-H groups, while located moderately far away from the panel, are relatively close to PM5 (0.091) and PM3 (0.061), respectively.

Figure 16: Barycenter overlay map of Chemistry panel, panel members (PM) and research groups

Similarly, the less covered groups CHEM-E and CHEM-F are found relatively close to PM2, CHEM-J close to PM4, and CHEM-K situated at a remote distance from the panel's coordinates. In Table 21, the shortest distance between the Chemistry groups and a panel member is printed in bold and underlined. The average of these distances is 0.087 (SD 0.070) and can be used as a measure of the fit between the expertise of the Chemistry panel and the research interests of the Chemistry research groups.

5.4.2.2 Physics assessment

5.4.2.2.1 Panel profile versus group profile

The overlay maps for physics similarly revealed a wider publication scope for the combined research groups (Figure 17) compared to the Physics panel (Figure 18).

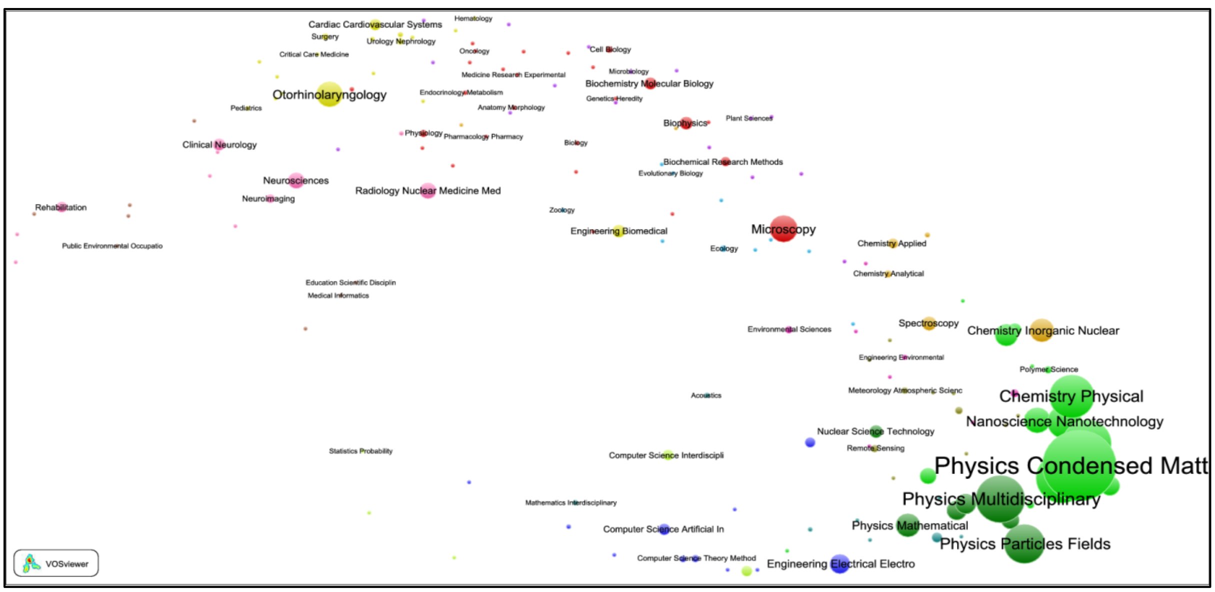


Figure 17: Physics groups' publication overlay map

The panel's publications are strong (68.75%) in *Physics condensed matter*, *Physics multidisciplinary*, *Chemistry physical*, and *Physics applied*, whereas the groups' publications tend to be mainly clustered (57.62%) in *Physics condensed matter*, *Physics multidisciplinary*, *Physics applied*, and *Materials science multidisciplinary*.

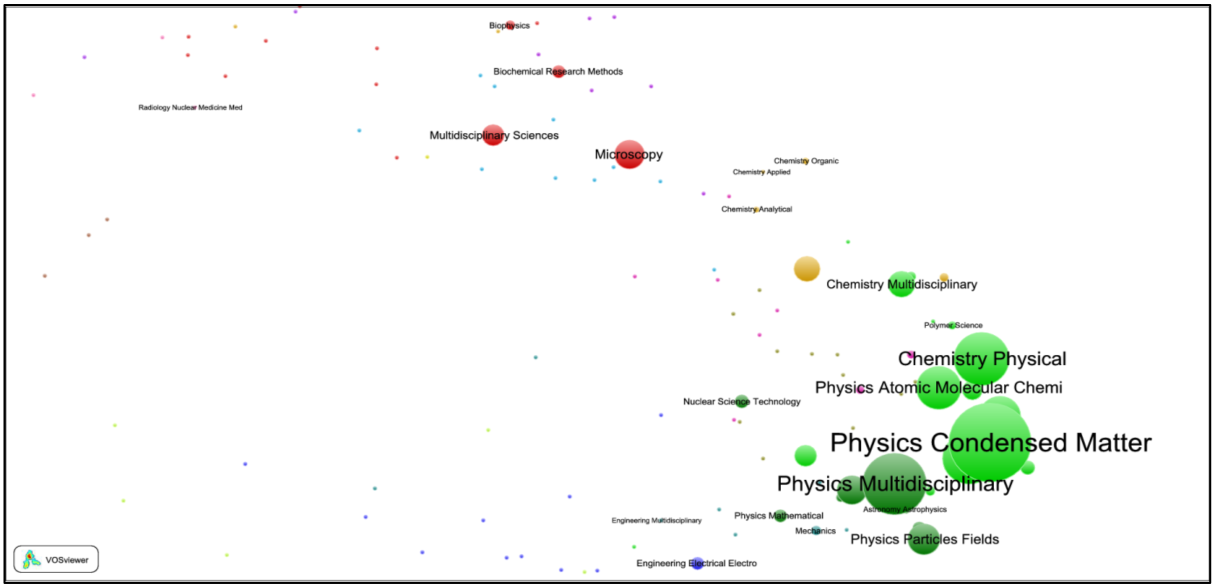


Figure 18: Physics panel members' publication overlay map

5.4.2.2.2 Panel profile versus individual group profile

Overlay maps of the publications of the individual groups were created, and subsequently compared with the panel overlay maps (see Figure 18).

PHYS-B group: *Physics condensed matter* (59.67%) and *Physics applied* (19.34%; Figure 19); PHYS-C: *Materials science multidisciplinary* (35.61%) and *Chemistry physical* (29.9%); PHYS-D: *Physics particles fields* (56.87%) and *Physics multidisciplinary* (40.89%); PHYS- E: *Physics multidisciplinary* (25.15%), *Physics particles fields* (24.52%), and *Physics condensed matter* (20.75%); PHYS-H: *Physics condensed matter* (61.06%) and *Physics applied* (19.08%).

These data show that five of the nine Physics groups (PHYS-B, PHYS-C, PHYS-E, PHYS-F, and PHYS-H) are thematically well covered by the panels' expertise as the majority of the groups publications are found in WoS SCs where the majority of the panels' publications have been classified.

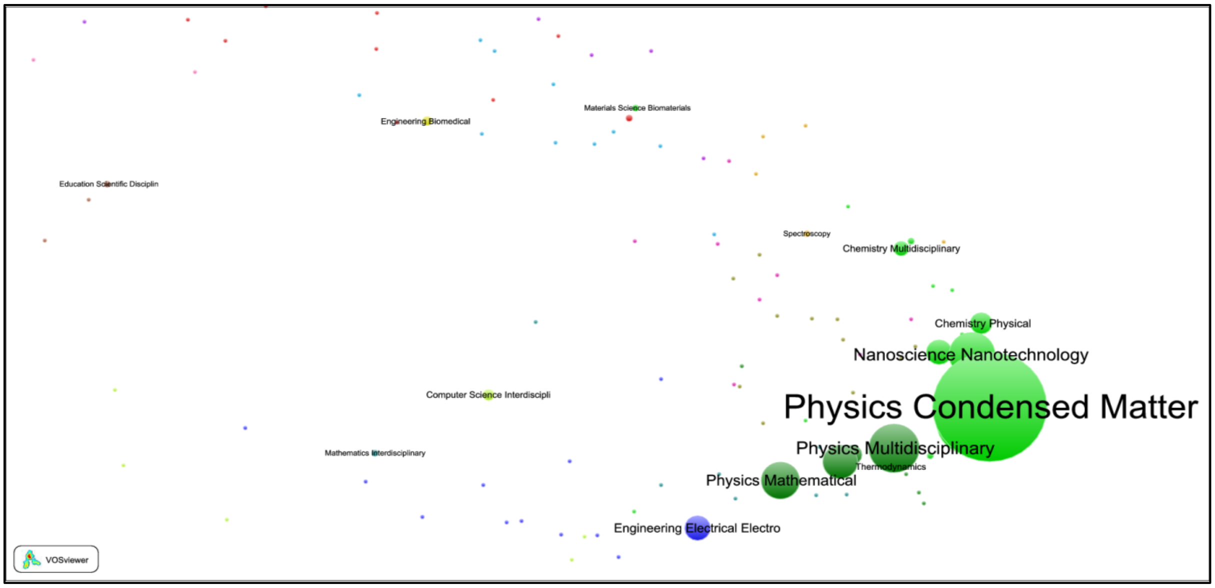


Figure 19: Overlay map of PHYS-B group's publications

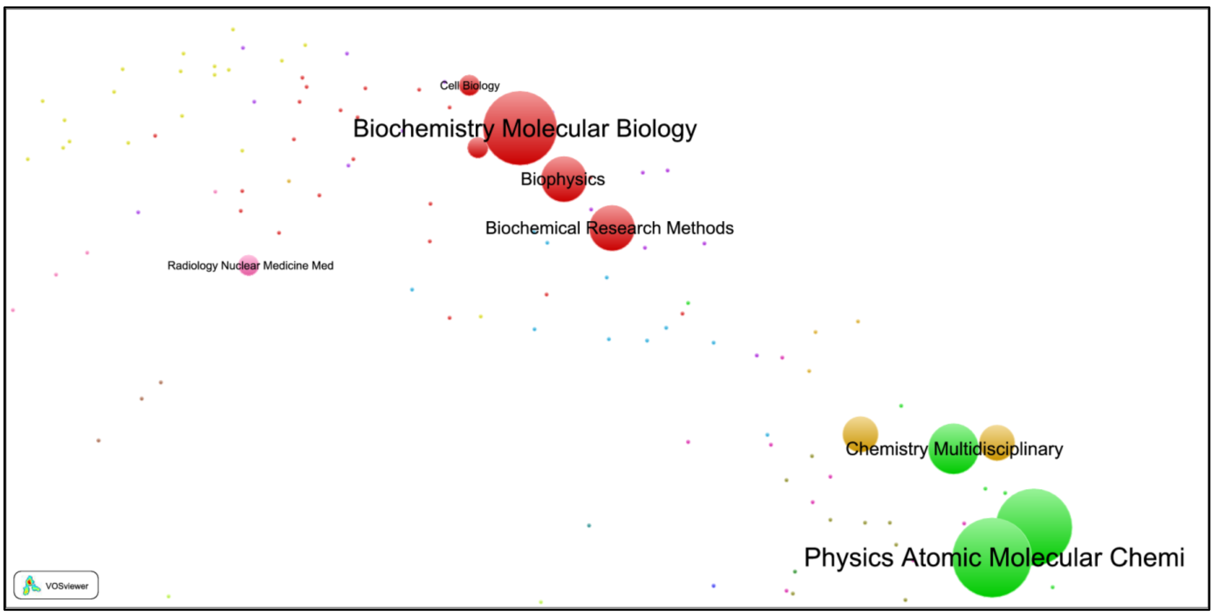


Figure 20: Overlay map of PHYS-G research group's publications

PHYS-F: *Physics multidisciplinary* (59.52%) and *Physics mathematical* (42.85%); PHYS-G: *Physics atomic molecular chemical* (34.88%) and *Chemistry physical* (32.55%; Figure 20). Two physics groups (PHYS-F and PHYS-G) have a large number of publications in WoS SCs where the publication output of their respective panels tends to be limited. The research output of these four groups is therefore only partially covered by the respective panels' expertise.

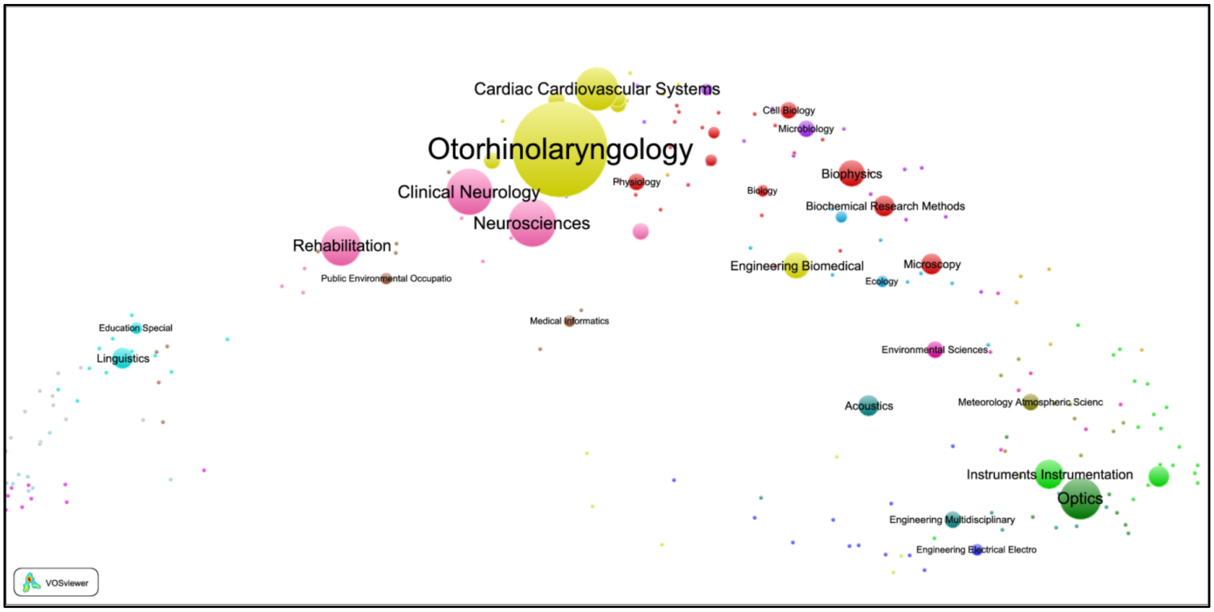


Figure 21: Overlay map of PHYS-A research group's publications

PHYS-A: Otorhinolaryngology (51.2%) and Audiology speech language pathology (14.4%; Figure 21); PHYS-I: Microscopy (26.08%) and Radiology nuclear medicine medical imaging (20.87%) WoS SCs. There was hardly any overlap in terms of the share of their publications in WoS SCs between these groups and the evaluation panel.

In summary, of the nine Physics groups, five groups are well covered, two groups are partially covered, and the remaining two groups seem to have been poorly covered by the Physics panel's expertise.

5.4.2.2.3 Distances between barycenters

Figure 22 and Table 23 show the distances between the Physics panel's barycenter and the different Physics groups barycenters. The Physics panel is very near to PHYS-B group, while PHYS-F group is 9.56 times and PHYS-I is 24.3 times further away from the panel than PHYS-B. PHYS-B group (0.025), PHYS-C (0.078), PHYS-E (0.033), and PHYS-H (0.040) are found closest to the panel's coordinates, while PHYS-D (0.123), PHYS-F (0.239) and PHYS-G (0.383) are still moderately close.

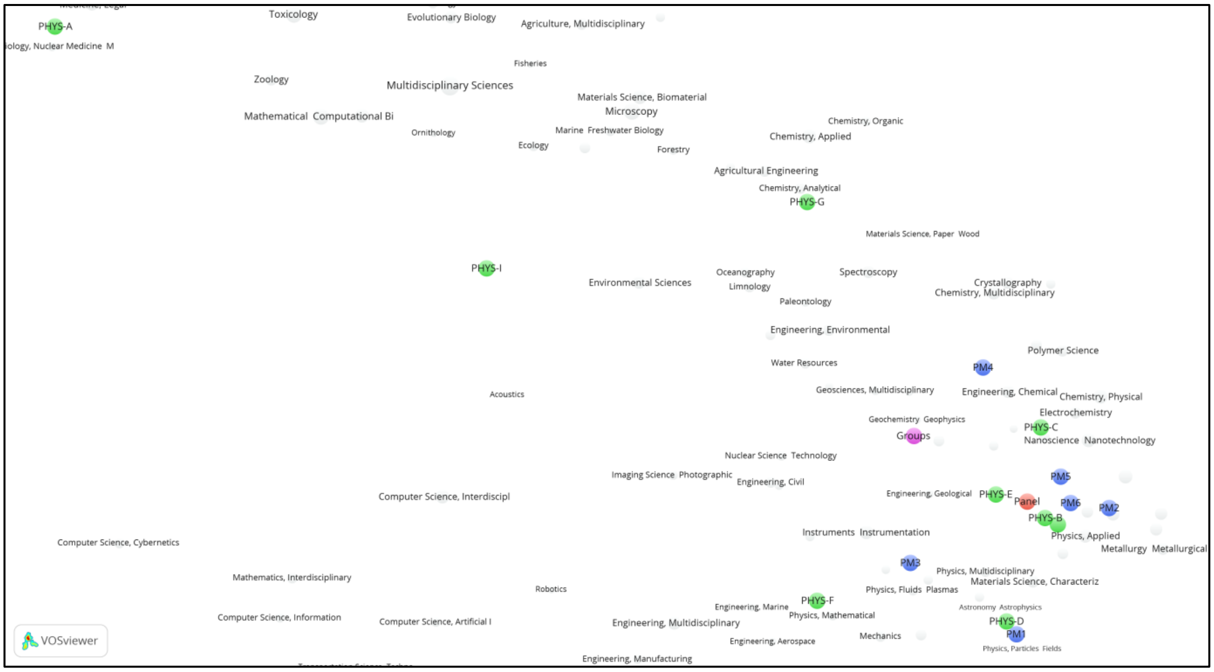


Figure 22: Barycenter overlay map of Physics panel, panel members (PM) and research groups

It should be noted that PHYS-D emerged as well covered and PHYS-F and PHYS-G as partially covered by the panel’s expertise from the comparative individual group vs. panel profile. Furthermore, although PHYS-D is situated moderately far away from the panel’s coordinates, PM1 is located in the immediate (0.017) neighborhood of PHYS-D, with the majority of the publications of PHYS-D and PM1 belonging to the same SCs.

Similar observations can be made for the other moderately close groups, PHYS-G and PHYS-F, which also have individual panel members in their immediate neighborhood, i.e., PM4 (0.249) and PM3 (0.104), respectively, and also have the majority of their publications in the same WoS SCs as these two panel members. Further, PHYS-A (1.115), and PHYS-I (0.607) are located at a considerable distance from the panel’s coordinates, have no individual panel members in their neighborhoods, and are poorly covered by the panel’s expertise.

Table 23 shows that the distances between PHYS-A and PM3 (1.041) and PM4 (1.020), and between PHYS-I and PM3 (0.532) and PM4 (0.522) are smaller compared to other panel members. The average of the shortest distances of the barycenters of the Physics panel members to the barycenters of the Physics research group is 0.232 (SD 0.0337).

5.5. Conclusion

We have explored not only overlap of expertise between research groups and an expert panel but also applied the barycenter method to calculate the distances between groups and panel (members). The barycenter method is well compatible with the WoS SCs-based overlay mapping, since it offers a simple way of representing the location of the panel and groups in a global science map based on WoS SCs. Each map of science "contains a projection from a specific perspective" (Leydesdorff & Rafols, 2009). Therefore, different layout techniques may produce different outputs. An exploration of two different layout techniques from the similarity matrix exposes that the Kamada-Kawai map is fairly strongly correlated with the VOS-map.

Overlay maps constitute an interesting tool to visualize the position of panel and group publications in a fixed map based on WoS SCs. The results reveal a number of discrepancies in WoS SCs between panel and group publications in both the Chemistry and Physics departments. This could be expected, since panel members are selected primarily because of their expertise and not necessarily because of the match thereof with the research in the groups. Overall, group publications are found in a wider range of SCs than panel publications, which might be due to the interdisciplinary orientation of some of the groups.

In chapter IV, we found that correlation coefficients and cosine similarity point to a varied – ranging from high to low – overlap of expertise between the Chemistry panels and research groups, and between the Physics panel and groups. The barycenter analysis showed that six Chemistry groups and five Physics groups are in fact well covered by the respective panels' expertise and are located close to the panel's coordinates while the remaining groups are not, although this gap is sometimes filled by the expertise of individual panel members. Furthermore, in some cases, neither the individual panel members nor the panels (as a whole) are situated close to the groups, in which case the panel seems to possess only partial expertise to evaluate these research groups. These barycenter findings are hence well in line with the results of the comparative analysis of individual group versus panel profile. Overall, the Chemistry panel, with an average barycenter distance of the nearest panel member to the research groups of 0.087 seemed to be better aligned with the research interest of the units under assessment studied in this paper. Note that the conclusion from plain correlations is the opposite (see chapter IV for details). This confirms the necessity of a method that moves beyond correlation coefficients, since they do not capture relatedness between SCs. The application of the barycenter method in the VOS-map and Kamada-Kawai allows to identify the Euclidean distances between the panel, combined research groups, individual panel members and individual research groups. It also allows calculation of average distances, comparison of distance and visual exploration of the barycenters of the map. Thus, the barycenter method provides information about the relevance of the expertise of an individual panel member to the assessment of both individual and combined research groups in a coherent way.

A similar, though less pronounced difference emerges from the comparison of the distance between the combined Chemistry groups and the Chemistry panel (0.105, see Table 21), and that between the combined Physics groups and the panel (0.135, see Table 23). These findings clearly demonstrate that in both cases, the majority of the panel publications appears in the categories in which group publications are found, while the groups have publications in a substantial number of WoS SCs that have no panel publications. There is a visible discrepancy between panel and group publications as far as WoS SCs are concerned. Overall, group publications are found in a wider range of subject categories than panel publications, which might be due to the interdisciplinary orientation of some of the groups.

In this investigation, we used distances between barycenters as a determinant for the correspondence between the publications by the group of panel members and the publications of a research group. Within this framework, a distance of zero would mean a perfect correspondence. As pointed out by a reviewer of the original paper (Rahman et al., 2015) one could envision other frameworks. One such framework would measure the correspondence between these two sets of publications by the similarity-weighted cosine measure as introduced in (Zhou et al., 2012). In this framework, perfect correspondence would be obtained when the similarity is one. We believe that this too is a valid approach in particular because the barycenter and the weighted-similarity approach, as illustrated in (Zhou et al., 2012), use the same input. Further investigations will have to show which of these two leads in practice to the best results (discussed in chapter VII).

A limitation arises from the question whether it is really relevant to have panel and groups publishing in the same subject categories, since one category may comprise a wide array of different subfields and topics. At present, this question cannot be answered with the methods outlined in this chapter, but instead would require a journal level analysis, as journals cover more closely related subfields and topics. A subsequent analysis will hence focus on overlay maps at the journal level (Leydesdorff & Rafols, 2012; Leydesdorff, Rafols, et al., 2013), with special attention to the quantification of similarity between groups and panel at this level for different disciplines (discussed in chapter VI). This comprehensive approach should allow us to define which overlap leads to the best standards for evaluation and hence permit us to propose the most appropriate expert panel composition for a collection of units of assessment. More generally, the matching of research expertise in several contexts might benefit from a comprehensive informetric approach.

Chapter VI: Measuring the match between evaluators and evaluees: Cognitive distances between panel members and research groups at the journal level [4]

6.1 Introduction

In this chapter, we study the problem of composing an expert panel, such that the individual panel members' expertise covers the specific subdomains in the discipline where the units of assessment have publications. In the chapter V, we explored expertise overlap between panel and research groups through publishing in the same or similar WoS SCs (Rahman et al., 2014, 2015). Since one subject category may comprise a wide array of different subfields and topics (Bornmann et al., 2011), it is up for discussion how relevant it is to have panel members and research group members publishing in the same subject categories. As journals cover more closely related subfields and topics (Tseng & Tsay, 2013), we present a journal level analysis to explore the issue.

The analysis relies on the journal similarity matrix and the overlay map derived from it. Science overlay maps (Rafols et al., 2010) have received considerable attention from the field of informetrics (Grauwin & Jensen, 2011; Gorjiara & Baldock, 2014; Boyack & Klavans, 2014a; Fields, 2015; Chen, Arsenault, Gingras, & Lariviere, 2015). We present two informetric methods to assess the cognitive distances between research groups in the Department of Biomedical Sciences, Veterinary Sciences, Pharmaceutical Sciences, Biology, and the respective expert panels based on research evaluations carried out at the University of Antwerp. We have used the data collected in the frame of research evaluation by the University of Antwerp. We explore the cognitive distance between expert panel and research groups. The research questions are:

[4] This chapter is based on Rahman, Guns, Leydesdorff & Engels (2016).

1) How can one quantify the cognitive distances between two entities using the journals in which they have published? How can one estimate the uncertainty inherent to these cognitive distances?

2) To what extent was each individual research group's expertise covered by the panel's expertise?

3) To what extent does each individual panel member's expertise cover the individual research groups?

6.2 Data

In this chapter, we consider data from the research assessments of all the research groups belonging to four departments of the University of Antwerp, Belgium.

Table 25: Publication profile of the panel members

Panel member code	Number of journals	Number of publications	Panel member code	Number of journals	Number of publications
Biomedical Sciences			***Pharmaceutical Sciences***		
BIOM-PM1	78	153	PHAR-PM1	39	122
BIOM-PM2	81	201	PHAR-PM2	93	351
BIOM-PM3	79	261	PHAR-PM3	91	259
BIOM-PM4	86	240	PHAR-PM4	67	124
BIOM-PM5	37	74	PHAR-PM5	86	180
BIOM-PM6	35	109	All Panel members together	**300**	**1,036**
BIOM-PM7	68	194			
BIOM-PM8	32	101			
All Panel members together	**395**	**1,333**			
Veterinary Sciences			***Biology***		
VETE-PM1	50	313	BIOL-PM1	48	146
VETE-PM2	66	121	BIOL-PM2	49	177
VETE-PM3	46	272	BIOL-PM3	35	76
VETE-PM4	53	131	BIOL-PM4	49	185
All Panel members together	**200**	**837**	BIOL-PM5	76	262
			All Panel members together	**217**	**786**

Table 26: Publication profile of the research groups

Group code	Number of journals	Number of publications	Group code	Number of journals	Number of publications
Biomedical Sciences (2006-2013)			***Pharmaceutical Sciences (2001-2008)***		
BIOM-A	55	96	PHAR-A	22	40
BIOM-B	27	43	PHAR-B	32	62
BIOM-C	47	107	PHAR-C	35	61
BIOM-D	95	201	PHAR-D	17	32
BIOM-E	34	70	PHAR-E	42	64
BIOM-F	17	27	PHAR-F	21	34
BIOM-G	115	241	PHAR-G	31	67
BIOM-H	29	50	PHAR-H	27	39
BIOM-I	55	89	PHAR-I	10	29
BIOM-J	27	47	PHAR-J	9	11
BIOM-K	43	74	All groups together	**180**	**376**
BIOM-L	11	12			
BIOM-M	67	164			
BIOM-N	43	114			
BIOM-O	32	60			
All groups together	**476**	**1,234**			
Veterinary Sciences (2006-2013)			***Biology (2004-2010)***		
VETE-A	102	144	BIOL-A	53	168
VETE-B	33	41	BIOL-B	33	58
VETE-C	21	52	BIOL-C	75	212
All groups together	**146**	**231**	BIOL-D	68	176
			BIOL-E	69	169
			BIOL-F	35	58
			BIOL-G	139	280
			BIOL-H	42	67
			BIOL-I	52	86
			All groups together	**372**	**1,158**

These are the 2011 assessment of the nine research groups of the department of Biology, the 2014 assessment of 15 research groups belonging to the department of Biomedical Sciences, the 2009 assessment of the 10 research groups of the department of Pharmaceutical Sciences, and the 2014 assessment of the three research groups of the Veterinary Sciences department. The group names will be standardized using the first four letters of the corresponding department, for example BIOM-A for Biomedical Sciences group A, VETE-C for Veterinary Sciences group C, etc. The reference period encompasses eight years preceding the evaluation. We considered all the articles, letters, notes, proceedings papers, and reviews by the research groups published during the reference period and included in the SCIE and SSCI of the WoS in the evaluation.

Table 25 lists the number of publications of the research groups. The numbers reported for all groups together are smaller than the sum of the individual research groups' publication or journal counts, because of joint publications between groups.

The entire WoS publication record of the individual panel members up to the year of assessment was taken into account. The Veterinary Sciences and Biomedical Sciences panels were composed of four and eight members respectively. Both the Pharmaceutical Sciences and Biology panels were composed of five members including the chair. There are no co-authored publications between panel members and research groups in any of the cases. None of the panel members has co-authored publications with another member of the same panel. Table 26 lists the number of publications of the research groups.

6.3 Methods

6.3.1 Journal similarity matrix and maps

Our method is based on the assumption that the cognitive distance between entities decreases as they have more publications in the same or similar journals, since journals cover closely related subfields and topics. The similarity between journals should be taken into account: if a panel member publishes in different journals than the research groups, they may still have relevant expertise, if their publications are in similar or closely related journals. This requirement rules out a number of approaches, including direct comparison of the top n journals in which two entities have published and correlations between journal portfolios (discussed in chapter IV).

We have harvested data from Thomson Reuters' (currently Clarivate Analytics) WoS JCR of the Science and Social Science Editions 2011. An aggregated journal-journal citation matrix of 10,675 journals[5] was constructed with a grand total of 35,295,459 citations over the entire matrix, which was subsequently normalized in the citing direction. The distances between journals are calculated using the cosine similarity between their citing distributions respectively

[5] The Science and Social Science Editions 2011 contain 8,281 and 2,943 journals respectively. Of these journals, 549 are contained in both databases.

(see Leydesdorff, Rafols, & Chen (2013) for details). The resulting journal similarity matrix can be considered as an adjacency matrix, and thus is equivalent to a weighted network where similar journals are linked and link weights increase with similarity strength. At the moment, it is not yet entirely clear how intense citation traffic around journals such as PLoS ONE (Leydesdorff & de Nooy, 2016) affects the journal similarity matrix.

The journal similarity matrix consists of $10{,}675^2 = 113{,}955{,}625$ cells. The matrix was stored using the HDF5 format (Hierarchical Data Format version 5), which was found to be the most efficient way of storing the data in terms of speed and memory requirements.

We used the full title of the journals for matching journals in the panel's publication list with journals in the research groups' publication lists. However, journals are not static entities and may undergo a name or organizational changes over time. Possible changes include:

- The journal title is changed, shortened or extended;
- Two or more journals merge into a new journal;
- One journal splits into two or more new journals;
- A journal is excluded from the WoS, discontinued, or not listed during the construction of the aggregated journal-journal citation matrix.

While cross-matching, we found 165 journals in our data set that belong to any of the above-mentioned categories. We developed the following guidelines to handle these uniformly:

- If journal A is renamed to B then treat both as equivalent.
- If journals A1 and A2 are merged into journal B, we treat both A1 and A2 as equivalent to B.
- If journal X splits into multiple journals, we look up which research groups or panel members have publications in journal X and determine which of the new journals best corresponds to the specialty of the authors, then change all occurrences of the journals in the WoS exported data with the best fitting latter journals. This was necessary in 15 cases; each time the decision was quite clear.

- If a journal is discontinued or excluded from WoS, or not included in the aggregated journal-journal citation matrix and there is no equivalent for some other reason, then it is removed from the sample.

From the journal similarity matrix, one can construct a global journal map (Leydesdorff & Rafols, 2012), in which similar journals are located more closely together. When used as a portfolio map, the size of the nodes depends on the number of publications in each node, and helps to compare the degree of overlap of multiple entities visually (Leydesdorff, Heimeriks, & Rotolo, 2016). The overlay of research group and panel publications can be visualized on the global journal map based on the retrieved publications data, using the visualization program VOSviewer (van Eck & Waltman, 2010). However, in the process of visualization, the multi-dimensional space is reduced to a projection in two dimensions. Moreover, comparison of overlay maps is difficult, specifically when the journals are located (very) closely to one another or when a panel member or research group has published in many different journals. Therefore, we will explore two methods to create a 'profile' of a panel member or research group: (i) barycenters on the overlay map (Rahman et al., 2015), and (ii) similarity-adapted publication vectors (SAPVs) (Rousseau, Rahman, Guns, & Engels, 2016). Subsequently, we can determine and compare the distances between entities, with overlay maps providing additional qualitative context.

6.3.2 Barycenter and distance calculation

Our barycenter method is based on the journal map. The barycenter is an entity's weighted average location on the map. More specifically, an entity's barycenter is the center of weight (Rousseau, 1989a, 1989b, 2008; Jin & Rousseau, 2001) of the journals in which it has published, where a journal's weight is the entity's number of publications in that journal. The barycenter is defined as the point $C = (C_1, C_2)$, where

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \ ; \ C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} \tag{13}$$

Here, $L_{j,1}$ and $L_{j,2}$ are the horizontal and vertical coordinates of journal j on the map, m_j is the number of publications in journal j, and $T = \sum_{j=1}^{N} m_j$ is the total number of publications of the entity.

The Euclidean distance between points $C = (C_1, C_2)$ and $D = (D_1, D_2)$ is calculated as follows:

$$d = \sqrt{(C_1 - D_1)^2 + (C_2 - D_2)^2}. \tag{14}$$

Many different algorithms and layout techniques have been developed for visualization of matrices. Rahman et al., (2015) found that at least two strongly different techniques – Kamada-Kawai (Kamada & Kawai, 1989) and VOS (van Eck & Waltman, 2007; van Eck et al., 2010) – yielded very similar results in terms of barycenter distances. The journal map used in this chapter was created using the VOS algorithm as implemented in VOSviewer (van Eck & Waltman, 2010). Subsequently, we determine and compare the cognitive distance between entities, with overlay maps providing additional qualitative context through visual comparison. In the Results section, we present several overlay maps (see Figure 23, Figure 24, Figure 25, and Figure 26) including barycenters and corresponding confidence regions (see section 6.3.4 for details). These maps are zoomed in to better highlight places of interest, hence independent of the zoom level of the figures.

6.3.3 Similarity-adapted publication vectors and distance calculation

Similarity-adapted publication vectors (SAPVs), a regular publication vector simply contains publication counts per journal/SC (Rousseau, Rahman, Guns, & Engels, 2016) , in a SAPV these counts are adapted to account for similarity between journals. We will use normalized SAPVs, such that there is scale invariance and publication vectors of entities of varying size can be meaningfully compared.

We calculate SAPVs for each entity, starting from the original journal similarity matrix, where N = 10,675 is the number of rows or columns in the matrix. Based on their respective SAPVs, the distance can be calculated between the expert panel, panel members, groups, and separate groups.

A similarity-adapted publication vector is determined as the vector $\mathrm{C} = (C_1, C_2, \dots, C_N)$, where:

$$C_k = \frac{\sum_{j=1}^{N} s_{kj} m_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} m_j} = \frac{(S * M)_k}{\|S * M\|_1} \tag{15}$$

Here $s_{j,k}$ denotes the k-th coordinate of journal j and m_j is the number of publications in journal j. The numerator of Equation (15) is equal to the k-th element of $S * M$, the multiplication of the similarity matrix S and the column matrix of publications $M = \left(m_j\right)_j$. The denominator is the L_1-norm of the unnormalized vector.

Subsequently, we determine the distance between the expert panel as a whole and individual panel members on the one hand, and the department (the combined groups), and individual groups on the other. The Euclidean distance between vectors a and b in $\mathbf{R}^N$ is:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + \dots + (a_N - b_N)^2} \tag{16}$$

Although the matrix and vectors are large, the calculation of SAPV and distances is relatively fast, due to the use of efficient matrix procedures implemented in NumPy and SciPy.[6]

Both the SAPV method and barycenter method can be used to determine an entity's 'profile'. One can then calculate the distance between profiles as an indicator of cognitive distance. For each research group, we find the shortest distance to one of the panel members. We use the average and standard deviation of the shortest distances as a comparative measure. All the distances are shown up to the third decimal. The distances are arbitrary units on a ratio scale (Egghe & Rousseau, 1990). Hence, one can meaningfully compare them in terms like 'x is twice as large as y'.

[6] http://www.numpy.org and http://scipy.org

6.3.4 Bootstrapping and confidence intervals

The barycenter and SAPV methods determine cognitive distance on the basis of the journals in which the groups and panel members have published. However, such information is not entirely deterministic; it is, for instance, dependent on the database used as well as environmental factors like the speed with which a journal processes a submission. It logically follows that small differences in Euclidean distances bear little meaning. To study this problem in a more systematic way, we employ a bootstrapping approach in order to determine 95% confidence intervals (CIs) to each Euclidean distance (both between barycenters and between SAPVs). If two CIs do not overlap, the difference between the distances is statistically significant at the 0.05 level. Although it is possible for overlapping CIs to have a statistically significant difference between the corresponding distances, the difference between the distances is less likely to have practical meaning.

Bootstrapping (Efron & Tibshirani, 1998) is a simulation-based method for estimating standard error and confidence intervals. Bootstrapping depends on the notion of a ***bootstrap sample***. To determine a bootstrap sample for a panel member or research group with N publications, we randomly sample with replacement N publications from its set of publications. In other words, the same publication can be chosen multiple times. Some publications in the original data set will not occur in the bootstrap data set, whereas others will occur once, twice or even more times. From the bootstrap sample, one can calculate a *bootstrap replication*, in our case a barycenter using formula (13) or SAPV using formula (15).

By generating a large amount of independent bootstrap samples (in our case 1000) and each time calculating the bootstrap replication, we can approximate the variability within the data set. Since we have a two-sample problem (distance between two entities; Efron & Tibshirani, 1998, Ch. 8), we calculate the distances between pairs of bootstrap replications, from which we obtain a CI using a bootstrap percentile approach (Efron & Tibshirani, 1998, Ch. 13). A more detailed explanation and implementation of our method is available on Github (Guns, 2016a, 2016b).

The bootstrap replications of barycenters are also used to add a 95% confidence region for each barycenter to the maps. For each barycenter, we have a cloud of 1000 points (bootstrapped barycenters) surrounding it. The confidence region is an ellipse that covers 95% of the bootstrapped barycenters and is obtained using an implementation by Kington (2014). The larger the confidence region, the less stable the barycenter is. Although the CI of the distance between two barycenters and their confidence regions are related, the two should not be conflated. In particular, we stress that overlapping confidence regions as seen in e.g. Figure 23 does not correspond to overlap between CIs for distances.

6.4 Results

We present the results in four parts. In the first (section 6.4.1) and the second part (section 6.4.2), we will discuss the results of Euclidean distances between barycenters and distances between SAPVs respectively. In the third part (section 6.4.3), we discuss the CIs of both the methods. However, for the intelligibility we show all the relevant tables of the Euclidean distance of barycenter and SAPV in the section 6.4.1 and 6.4.2, where the CIs are included through the typography of the values. In the last part (section 6.4.4), we make a comparison between the two methods.

6.4.1 Barycenter and distances

For each discipline, the barycenters of the panel, panel members, individual research groups and department, as well as Euclidean distances between barycenters are calculated. For each research group, we also calculate the average shortest distance to one of the panel members. The visualizations of barycenters and their confidence regions are added to the overlay maps. In the Table 27 to Table 30, the shortest distances between a group and a panel member are bold and underlined. The distances whose CIs overlap with that of the shortest distance are in bold.

Table 27 provide data on the distances between the Biology panel's barycenter and individual research groups. The average of the shortest distances between the Biology groups and panel members is 0.09. The Biology panel as a whole is closer to BIOL-I (0.087) and BIOL-G (0.065).

Table 27: Euclidean distances between barycenters of Biology individual research groups, panel members, panel and groups together in the journal VOS-map

	Groups	BIOL- A	BIOL- B	BIOL-C	BIOL D	BIOL-E	BIOL -F	BIOL- G	BIOL- H	BIOL- I
Panel	0.136	0.128	0.242	0.271	0.220	0.208	0.136	0.087	0.262	0.087
PM1	0.072	0.154	**0.125**	**0.198**	**0.105**	0.169	0.239	**0.056**	**0.146**	**0.164**
PM2	0.087	**0.016**	0.249	**0.168**	0.190	**0.090**	0.257	**0.091**	**0.227**	0.217
PM3	0.248	0.223	0.336	0.382	0.326	0.316	**0.029**	0.199	0.368	**0.075**
PM4	0.148	0.205	**0.163**	0.279	0.175	0.245	0.187	**0.110**	**0.211**	**0.106**
PM5	0.253	0.195	0.374	0.373	0.348	0.297	**0.104**	0.211	0.390	**0.145**

Average shortest distance is 0.09 (SD 0.05).

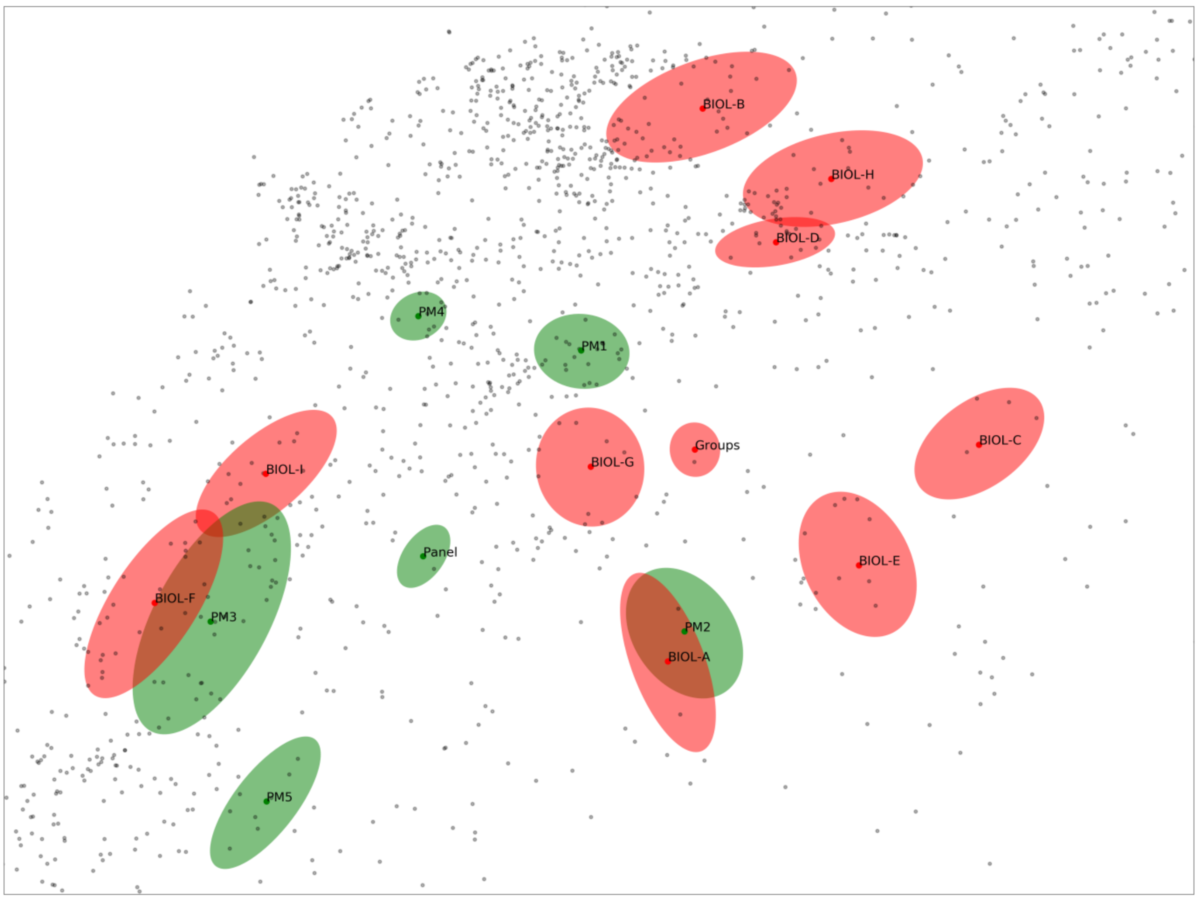


Figure 23: Barycenter overlay map of Biology panel, panel members (PM), research groups and research groups together (groups) with their confidence regions

Table 28: Euclidean distances between barycenters of Biomedical Sciences individual research groups, panel members, panel and groups together in the journal VOS-map

	Groups	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
Panel	0.177	0.225	0.132	0.146	0.109	0.263	0.064	0.396	0.354	0.133	0.303	0.268	0.383	0.312	0.371	0.282
PM1	0.265	0.350	0.180	0.224	**0.110**	**0.242**	**0.081**	0.473	0.319	**0.159**	0.445	0.387	**0.471**	0.397	0.436	0.344
PM2	0.085	**0.176**	**0.038**	**0.046**	0.201	0.177	**0.119**	**0.302**	0.267	**0.234**	**0.297**	**0.208**	**0.294**	**0.221**	**0.272**	**0.181**
PM3	0.413	0.390	0.397	0.397	0.241	0.530	0.303	0.611	0.621	**0.194**	0.356	0.438	0.586	0.527	0.599	0.522
PM4	0.389	0.391	0.355	0.365	**0.168**	0.479	0.243	0.600	0.568	**0.119**	0.390	0.440	0.580	0.515	0.582	0.498
PM5	0.149	**0.250**	**0.058**	**0.107**	**0.183**	**0.144**	**0.095**	0.348	**0.233**	**0.227**	0.371	**0.280**	**0.348**	**0.274**	0.311	0.220
PM6	0.189	**0.295**	0.177	0.184	0.383	**0.072**	0.295	**0.236**	**0.086**	0.426	0.442	**0.291**	**0.258**	**0.207**	**0.187**	**0.135**
PM7	0.251	0.367	0.173	0.217	0.282	**0.103**	0.209	0.395	**0.148**	0.331	0.500	0.385	**0.407**	0.342	0.348	0.271
PM8	0.275	**0.171**	0.363	0.314	0.497	0.445	0.445	**0.238**	0.502	0.504	**0.154**	**0.140**	**0.199**	**0.213**	**0.271**	**0.281**

Average shortest distance is 0.132 (SD 0.06).

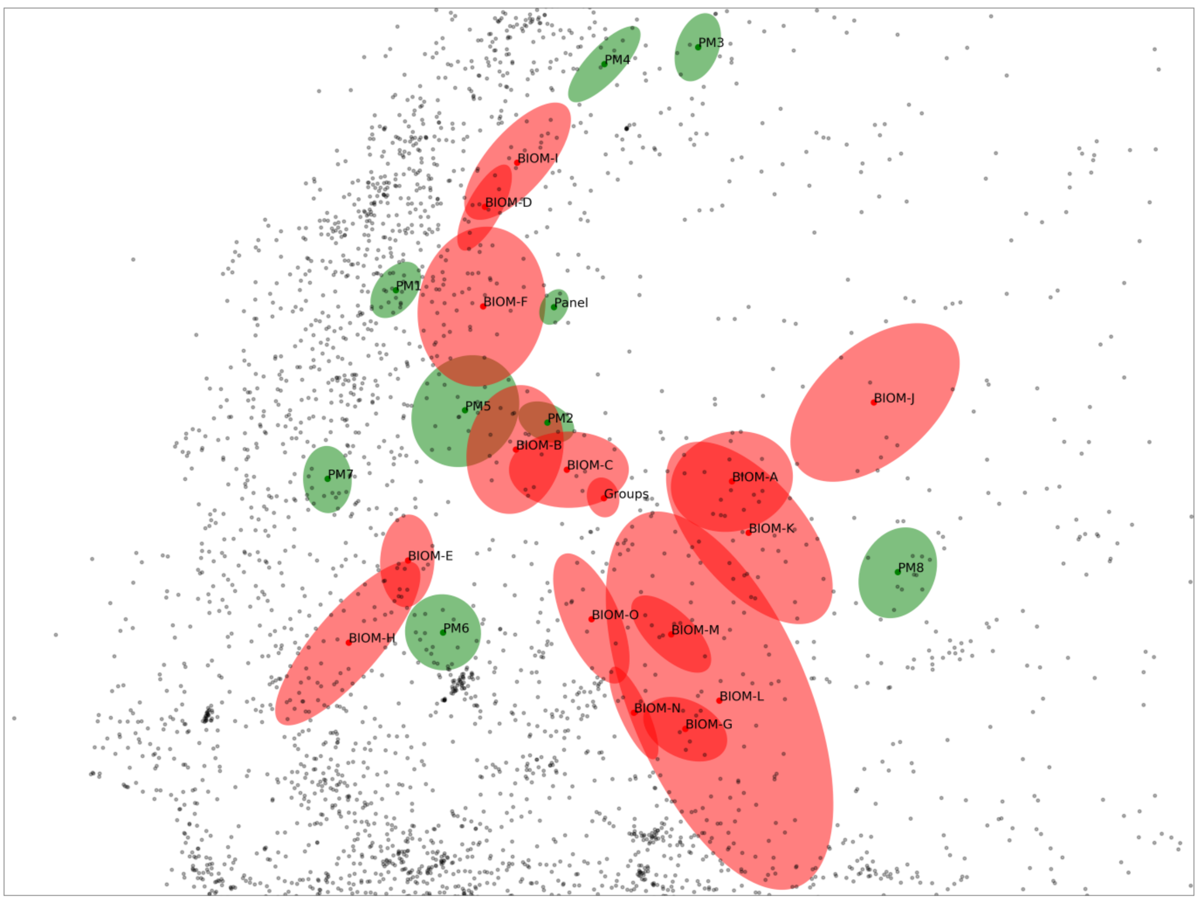


Figure 24: Barycenter overlay map of Biomedical Sciences panel, panel members (PM), research groups and research groups together (groups) with their confidence regions

Table 29: Euclidean distances between barycenters of Pharmaceutical Sciences research groups, panel members, panel and groups together in the journal VOS-map

	Groups	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
Panel	0.078	0.410	0.240	0.536	0.096	0.325	0.239	0.381	0.120	0.769	0.495
PM1	0.559	**0.101**	0.267	1.017	0.413	0.807	**0.271**	**0.262**	0.471	1.251	0.972
PM2	0.268	0.750	0.581	**0.205**	0.428	**0.021**	0.579	0.689	0.398	**0.429**	**0.162**
PM3	0.156	**0.339**	**0.163**	0.610	**0.043**	0.402	**0.162**	**0.332**	**0.110**	0.844	**0.573**
PM4	0.160	**0.332**	**0.161**	0.616	**0.052**	0.408	**0.160**	**0.322**	**0.120**	0.850	**0.577**
PM5	0.318	**0.186**	**0.057**	0.773	**0.170**	0.566	**0.062**	**0.242**	**0.233**	1.008	0.735

Average shortest distance is 0.143 (SD 0.124).

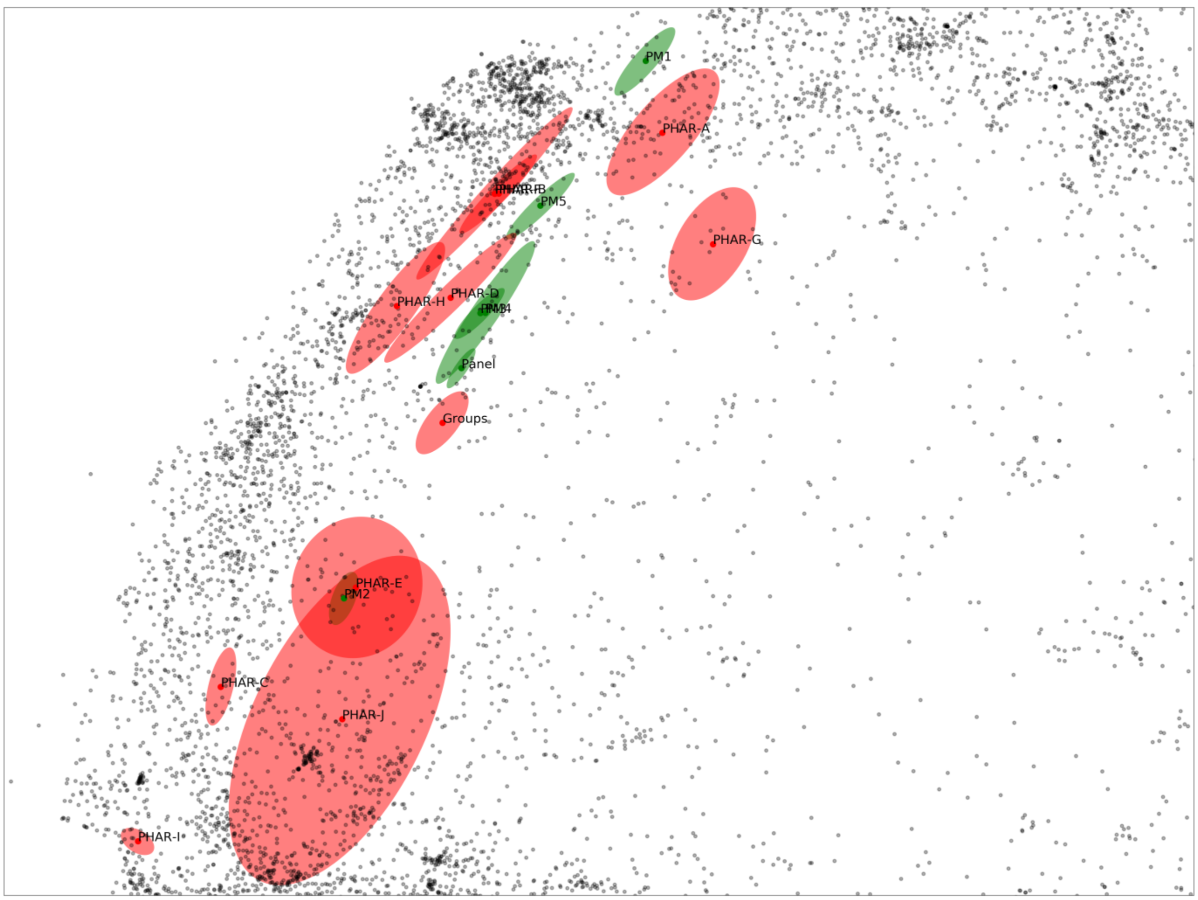


Figure 25: Barycenter overlay map of Pharmaceutical Sciences panel, panel members (PM), research groups and research groups together (groups) with their confidence regions.

Table 30: Euclidean distances between barycenters of Veterinary Sciences individual research groups, panel members, panel and groups together in the journal VOS-map

	Groups	VETE-A	VETE-B	VETE-C
Panel	0.092	0.179	0.076	0.156
PM1	0.178	0.260	**0.160**	**0.124**
PM2	0.088	**0.141**	**0.108**	**0.227**
PM3	0.195	0.273	**0.182**	**0.145**
PM4	0.306	0.272	0.310	0.469

Average shortest distance is 0.124 (SD 0.013).

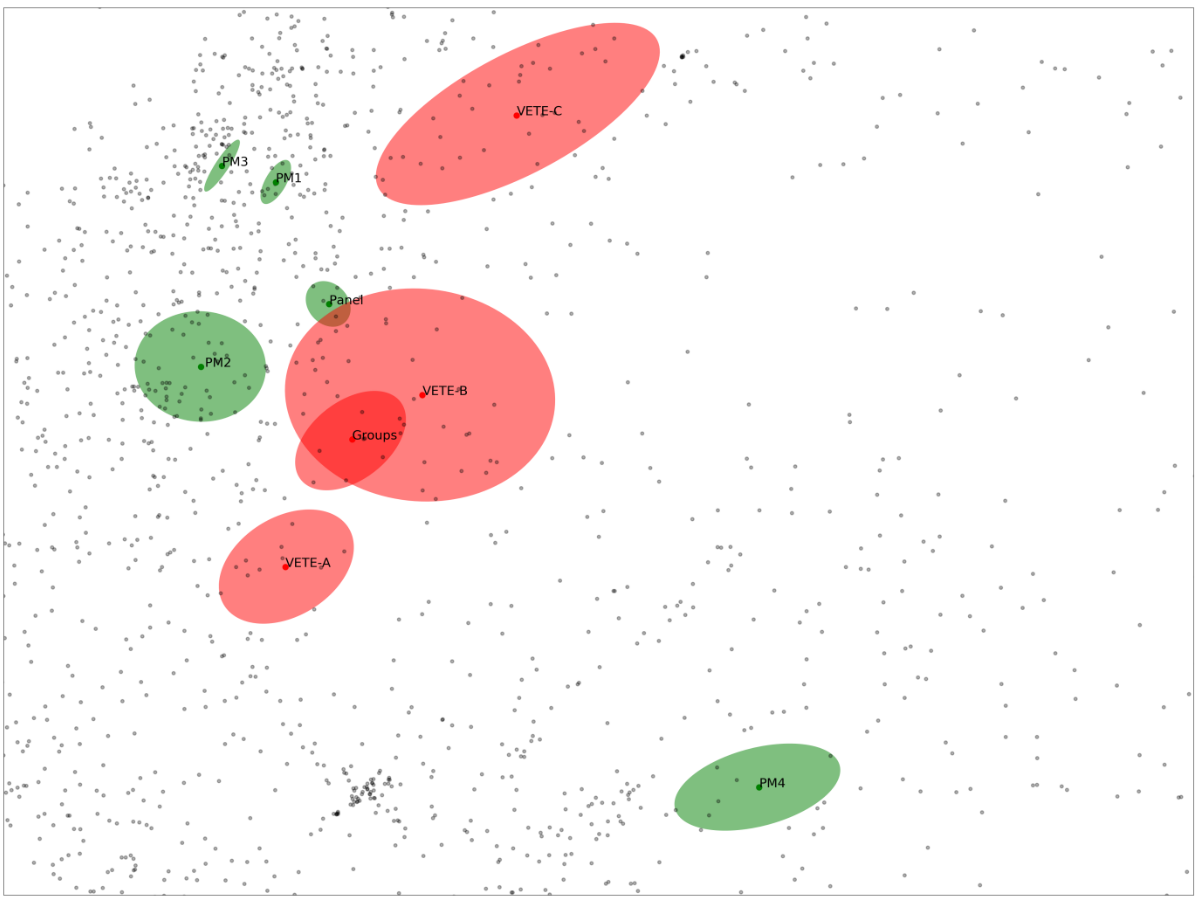


Figure 26: Barycenter overlay map of the Veterinary Sciences panel, panel members (PM), research groups and research groups together (groups) with their confidence regions

BIOL-B (0.242), BIOL-C (0.271), BIOL-D (0.228) and BIOL-H (0.262) are the furthest from the panel. BIOL-A and BIOL-E are found at a considerable distance from the panel's barycenter but PM2 is in their immediate neighborhood. Similar conclusions can be drawn from the visualization in Figure 23. Here, 'PM' stands for 'panel member', 'Panel' represents the barycenter location of the publication profile of the entire panel, and 'Groups' does the same for the research groups taken together (the department). The advantage of the visual representation consists in providing an easily interpretable overview of how the panel and research groups relate, which is much less straightforward from a table of distances.

Table 28 provides data on the distances between the Biomedical Sciences panel's barycenters and its members on the one hand and those of the department and individual research groups on the other. Figure 24 visualizes the situation. The Biomedical panel is very near to BIOM-F (0.064), while BIOM-G (0.396), BIOM-H (0.354), BIOM-L (0.383), and BIOM-N (0.371) are almost 5 to 6 times farther away from the panel than BIOM-F. BIOM-C (0.146), BIOM-D (0.109), BIOM-I (0.133) groups are situated comparatively close to the panel's coordinates, while BIOM-E (0.263) is found at a considerable distance from the panel's barycenter.

In Table 28, the average of the shortest distance between the Biomedical Sciences groups and panel members is 0.132 (SD 0.06) and can be used as a measure of the fit between the expertise of the Biomedical Sciences panel and the research groups. Groups BIOM-G, BIOM-H, BIOM-M, and BIOM-N are situated moderately far away from the panel's coordinates, but PM2 and PM6 are located in their immediate neighborhood.

Table 29 provides data on the distances between the Pharmaceutical sciences panel's barycenter and individual research groups. Figure 25 visualizes the situation. The average of the shortest distance between the Pharmaceutical groups and panel members is 0.143. PHAR-C (0.536) and PHAR-I (0.769) are 5.58, and 8.01 times farther away respectively from the panel than PHAR-D (0.096).

PHAR-B (0.240), PHAR-F (0.239), PHAR-H (0.120) are situated comparatively close to the panel's coordinates, while PHAR-A (0.410) and PHAR-J (0.495) are found at a considerable distance from the panel's barycenter. The case of PHAR-A reinforces our assertion that the mere overlap of journals is not sufficient to quantify the cognitive distance: although 60% of the journals in which this group has published are also covered by the panel, it is located relatively far away from the panel. PHAR-I and the panel do not share any common journals. PHAR-I is located far away from the panel as a whole as well as from any individual panel member. In summary, the Pharmaceutical Sciences panel appears to cover most research groups adequately, with the exception of two.

Table 30 provides data on the distances between the Veterinary science panel's barycenter and those of the individual research groups. The Veterinary panel is the closest to VETE-B, while VETE-A is 1.9 times and VETE-C is 1.7 times farther away from the panel than VETE-B. The overlay map (Figure 26) shows that the panel members are generally quite close to the research groups. Only PM4 is located a bit further away from the groups. Although the fit in this case is fairly good, an even better fit could be obtained if PM4 were replaced with a different person with publications in journals that are more closely related to the groups' publication profile.

6.4.2 Similarity-adapted publication vectors and distances

For each discipline, the SAPV of the panel, panel members, individual research groups and department, as well as Euclidean distances between SAPVs are calculated. For each research group, we also calculate the average shortest distance to one of the panel members.

The Biology panel is closer to BIOL-A (0.005) and BIOL-G (0.006), while BIOL-B (0.010) and BIOL-C (0.012) are at least 2 times farther away from the panel (Table 31). The average of the shortest distances between the Biology groups and panel members is 0.006. Table 32 provides data on the Euclidean distances between (SAPVs of) Biomedical Science research groups, panel and panel members. BIOM-F, and BIOM-I are in the immediate neighborhood of the panel while BIOM-N (0.010) is located farthest away from the panel. PM2 and PM5 are closer to nine and ten research groups respectively, while PM8 is situated moderately far away from all the research groups.

Table 31: Euclidean distances between SAPV of Biology individual research groups, panel members, panel and groups together in the journal similarity matrix

	Groups	BIOL- A	BIOL- B	BIOL-C	BIOL D	BIOL-E	BIOL -F	BIOL- G	BIOL- H	BIOL- I
Panel	0.004	0.005	0.010	0.012	0.007	0.008	0.007	0.006	0.010	0.008
PM1	0.004	0.007	**0.009**	**0.013**	**0.004**	0.008	0.008	**0.006**	**0.009**	0.008
PM2	0.005	**0.003**	**0.010**	**0.015**	0.009	**0.005**	0.012	**0.005**	**0.011**	0.012
PM3	0.009	0.011	0.015	**0.013**	0.011	0.014	**0.003**	0.013	0.015	**0.004**
PM4	0.007	0.006	**0.009**	0.016	0.009	**0.006**	0.014	**0.004**	**0.010**	0.013
PM5	0.009	0.009	0.013	**0.012**	0.012	0.012	0.008	0.011	0.014	0.009

Average shortest distance is 0.006 (SD 0.003).

Table 32: Euclidean distances between SAPV of Biomedical Sciences individual research groups, panel members, panel and groups together in the journal similarity matrix

	Groups	**BIOM-A**	**BIOM-B**	**BIOM-C**	**BIOM-D**	**BIOM-E**	**BIOM-F**	**BIOM-G**	**BIOM-H**	**BIOM-I**	**BIOM-J**	**BIOM-K**	**BIOM-L**	**BIOM-M**	**BIOM-N**	**BIOM-O**
Panel	0.004	0.004	0.004	0.004	0.006	0.005	0.003	0.009	0.008	0.003	0.008	0.005	0.009	0.006	0.010	0.007
PM1	0.006	0.007	0.006	0.006	0.007	0.007	**0.003**	0.011	**0.009**	**0.002**	0.009	**0.007**	**0.011**	0.008	**0.012**	**0.009**
PM2	0.005	**0.004**	0.006	0.007	0.008	0.007	**0.003**	**0.008**	0.010	0.005	**0.005**	**0.006**	**0.009**	**0.006**	**0.010**	**0.007**
PM3	0.007	0.007	0.006	0.007	0.008	0.008	0.006	0.011	0.011	0.006	**0.008**	**0.006**	**0.011**	0.008	**0.011**	**0.009**
PM4	0.007	0.007	0.007	0.007	0.007	0.008	**0.004**	0.011	0.010	**0.002**	0.009	0.007	**0.011**	0.009	**0.012**	**0.009**
PM5	0.004	**0.005**	**0.002**	**0.003**	0.007	0.006	0.005	**0.009**	0.009	0.005	**0.008**	**0.004**	**0.009**	**0.006**	**0.009**	**0.006**
PM6	0.006	0.008	0.008	0.007	0.008	**0.003**	0.009	**0.009**	**0.006**	0.009	0.012	0.008	**0.011**	0.009	**0.011**	**0.009**
PM7	0.007	0.008	0.008	0.007	**0.005**	0.007	0.009	**0.010**	**0.007**	0.008	0.012	0.008	**0.011**	0.009	**0.011**	0.009
PM8	0.011	0.009	0.012	0.013	0.014	0.013	0.013	**0.009**	0.014	0.014	0.011	0.011	**0.010**	0.010	**0.010**	0.010

Average shortest distance is 0.005 (SD 0.002).

Table 33: Euclidean distances between SAPV of Pharmaceutical Sciences individual research groups, panel members, panel and groups together in the journal similarity matrix

	Groups	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
Panel	0.003	0.009	0.008	0.007	0.009	0.004	0.007	0.017	0.005	0.013	0.011
PM1	0.013	**0.011**	**0.011**	0.017	0.015	0.015	**0.008**	**0.020**	0.012	0.021	0.020
PM2	0.005	0.012	**0.010**	**0.005**	0.011	**0.004**	**0.011**	**0.018**	**0.008**	**0.011**	**0.008**
PM3	0.006	**0.010**	**0.009**	0.009	**0.007**	0.007	**0.008**	**0.018**	**0.007**	0.015	**0.013**
PM4	0.006	**0.010**	**0.008**	0.009	0.011	**0.007**	**0.006**	**0.018**	**0.007**	0.014	**0.012**
PM5	0.007	**0.007**	**0.008**	0.010	0.012	0.008	**0.007**	**0.017**	**0.007**	0.017	**0.014**

Average shortest distance is 0.008 (SD 0.004).

Table 34: Euclidean distances between SAPV of Veterinary Sciences individual research groups, panel members, panel and groups together in the journal similarity matrix

	Groups	VETE-A	VETE-B	VETE-C
Panel	0.007	0.008	0.005	0.006
PM1	0.011	0.013	0.010	**<u>0.005</u>**
PM2	0.005	**<u>0.005</u>**	**<u>0.005</u>**	0.011
PM3	0.015	0.016	0.013	0.013
PM4	0.010	0.010	0.010	0.015

Average shortest distance is 0.005 (SD 0.000).

The average of the shortest distances between the Biomedical Sciences groups and panel members is 0.005 (SD 0.002), which can be used as a measure of the fit between the expertise of the panel members and the research groups. In Table 31 to Table 34, the shortest distances between a group and a panel member are bold and underlined. The distances whose CIs overlaps with that of the shortest distance are in bold.

Table 33 provides data on the distances between the Pharmaceutical Sciences panel and individual research groups. The average shortest distances between the panel and individual research groups is 0.008 (SD 0.042). PHAR-E (0.004) and PHAR-H (0.005) are closer to the panel while PHAR-I (0.013) is located moderately far away from all panel members except PM2. PHAR-I (0.011) and the panel do not share any common journals, but PM2 is also closer to this group than other panel members. The Veterinary panel is the closest to VETE-B (0.005). The average shortest distances between the panel and individual research groups is 0.005 (SD 0.002). In the Veterinary department, the panel members are quite close to the research groups except for PM3 and PM4 (Table 34). PM3 and PM4 could be replaced with other potential panel members who have publications in journals that are more closely related to the groups' publication profile to obtain a better panel fit.

6.4.3 Confidence intervals

To get an idea of the reliability of our barycenter and SAPV distances, we apply a bootstrapping approach to obtain 95% CIs. Comparison of the CIs can then inform the analysis. Specifically, if two distances are not equal but their CIs overlap, the difference may not be meaningful.

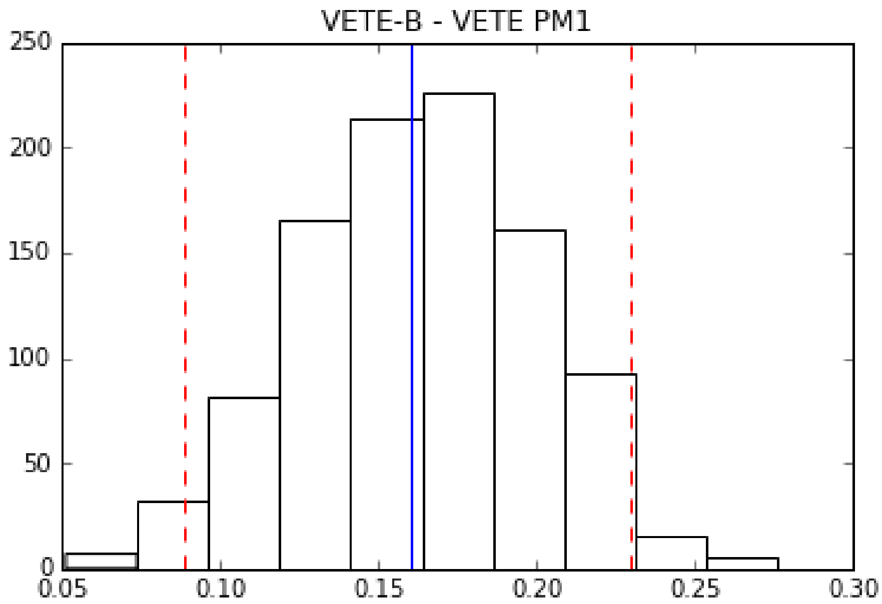


Figure 27: Histogram of 1000 bootstrapped distances between the barycenters of VETE-B and VETE-PM1 (Veterinary Sciences)

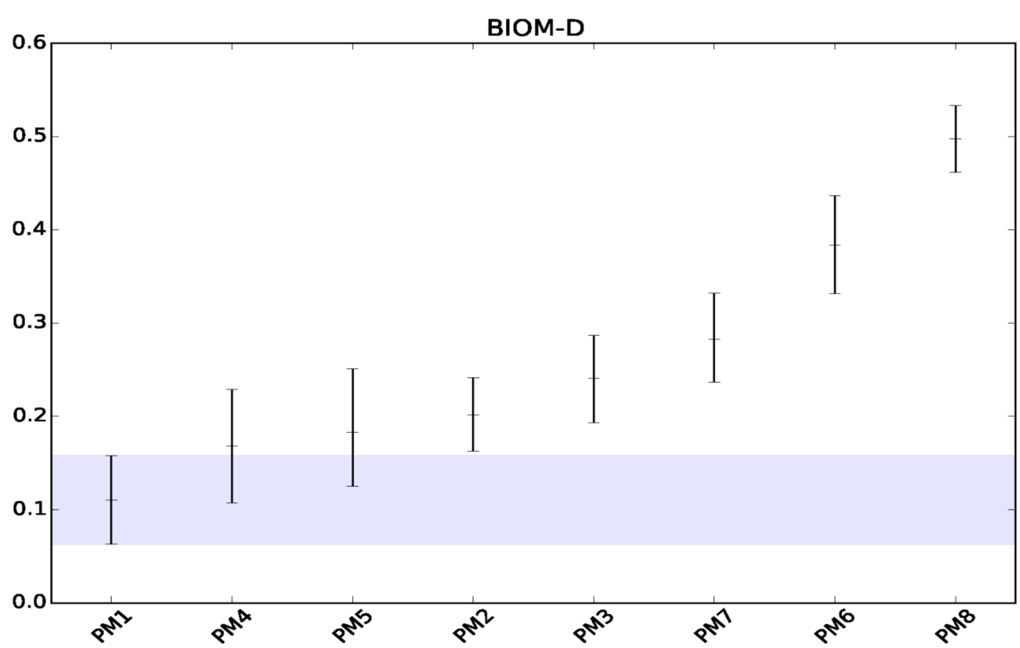


Figure 28: Confidence intervals for barycenter distances for Biomedical Sciences research group D

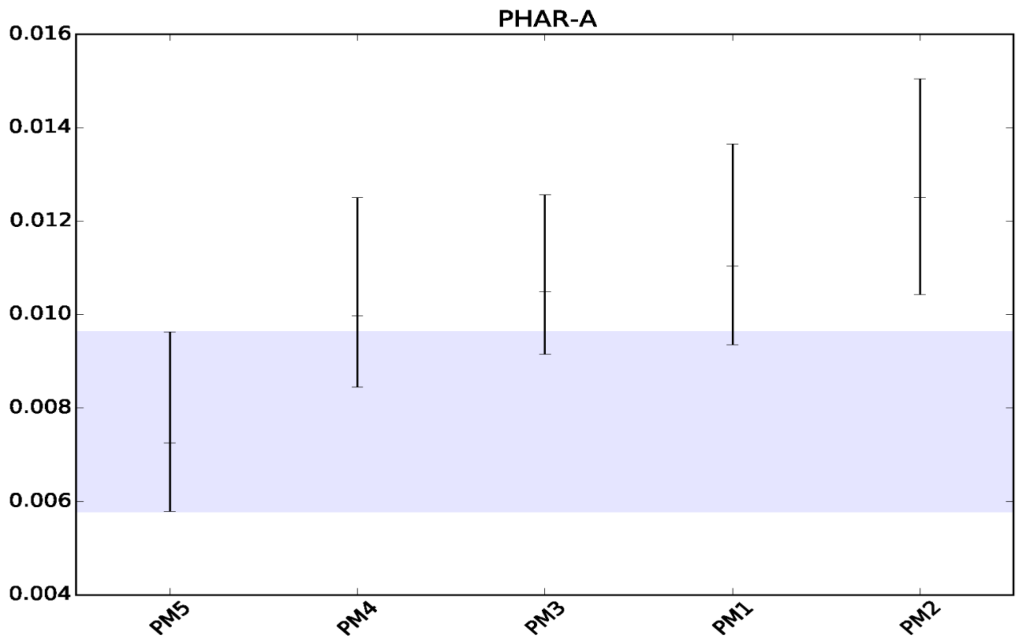


Figure 29: Confidence intervals for SAPV distances for Pharmaceutical Sciences research group A

Table 35: Percentage of overlapping CI's for barycenters and SAPVs in each of the four disciplines

Department	Barycenter method	SAPV method
Biology	28%	28%
Biomedical Sciences	36%	34%
Pharmaceuticals Sciences	43%	55%
Veterinary Sciences	44%	0%

As explained in the Methods section, we calculate distances for 1000 bootstrap samples. The resulting distances tend to be normally distributed, as illustrated in Figure 27. In Figure 27, the blue line indicates the empirically found distance; the dashed red lines indicate the CI. A similar image emerges for all disciplines and for both barycenters and SAPVs. It can be seen that the CI is a reliable approximation of the variability across the bootstrap samples. In Figure 28 and Figure 29, the highlighted part indicates the confidence interval of the shortest distance to the research group.

We illustrate the interpretation of the CIs using a few examples. Our focus will be on the task of finding the panel members that are cognitively closest to a given research group. Figure 28 displays the CIs for the distances between the barycenter of BIOM-D and the barycenters of all panel members in Biomedical Sciences. Ignoring the panel as a whole, the panel member for which we find the closest distance to BIOM-D is PM1 but we cannot simply conclude that this panel member is cognitively closest to the group: both PM4 and PM5 have CIs that partially overlap with PM1. Hence, PM4 and PM5 should be treated as viable alternatives to PM1 if one is seeking a panel member with expertise similar to that of research group BIOM-D.

Likewise, Figure 29 displays CIs for SAPV distances, using the example of Pharmaceutical Sciences research group PHAR-A. In this case, it turns out that the differences between the panel members are relatively small. The result is that, with the exception of PM2, all panel members are eligible candidates. CI plots like Figure 28 and Figure 29 are available in the appendix for all research groups and for both barycenters (appendix A) and SAPVs (Appendix B).

We calculated the rate of overlap of CIs in the case of the barycenter approach and the case of the SAPV approach in all the four departments (see Table 35) in order to get a feel of the extent they might give rise to different conclusions. Overall, the degree of overlap due to the CIs of the barycenter approach seems similar to that of the SAPV approach.

6.4.4 Comparison between two methods

To more directly compare the results, we obtained from both methods, we calculated the Pearson correlation coefficient (r) and the Spearman rank correlation coefficient (ρ) between the distances obtained through the barycenter method and SAPV method. The correlation calculation is based on all distances between research groups and individual panel members. Correlations for the Biomedical department ($r = 0.60$, $\rho = 0.56$), Biology department ($r = 0.73$, $\rho = 0.71$), Pharmaceutical department ($r = 0.63$, $\rho = 0.62$) and Veterinary department ($r = 0.64$, $\rho = 0.66$) are moderately strong (Figure 30).

We now turn to the question how the barycenter method and the SAPV method compare. Both try to quantify the cognitive distance by determining the Euclidean distance between representations or ‘profiles’ of an entity, but the way these profiles are obtained is quite different. The barycenter method has the benefit of visualization, but the reduction of dimensionality that is inherent to creating a two-dimensional map may cause distortions in some cases. In this respect, the SAPV distances are the most reliable measure. We hypothesize that this advantage plays a larger role at the journal level than it did at the level of WoS SCs, since there are many more dimensions in the former case. In general, we recommend using the SAPV method for distance calculation and consider the barycenter approach more appropriate for visual exploration.

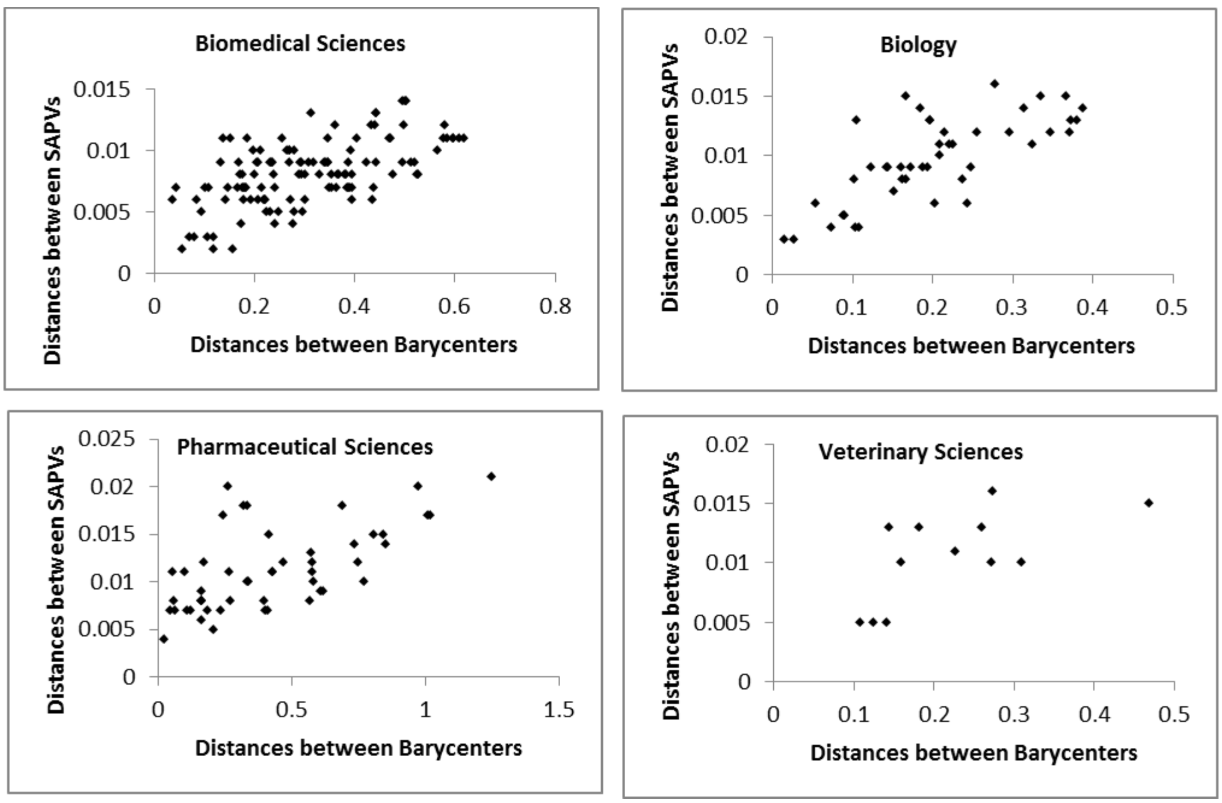


Figure 30: Scatter plot of the barycenter and SAPV distances between groups and individual panel members in the Biology, Biomedical Sciences, Pharmaceutical Sciences, and Veterinary Sciences departments

From the discussion on the composition of the four expert panels, it follows that a group can be far away from the panel as a whole. However, some individual panel members may have sufficient expertise to evaluate a single group, as indicated by publications in closely related or similar journals.

For example, as discussed in section 6.4.1 and shown in Figure 24, the barycenter of PM8 for Biomedical Sciences is in the immediate neighborhood of research groups BIOM-A, BIOM-J, BIOM-K and BIOM-L, while other panel members are farther away from them. On the other hand, according to the SAPV method, BIOM-PM8 is situated moderately far away from all the research groups. In the same way, the barycenter of VETE-PM4 is far away from all the groups, while in the SAPV method this is the case for PM3. These examples illustrate that, while the two methods are clearly correlated, they may yield rather different results at the level of individual groups or panel members.

Even if a research group has no publications in the journals where the panel has publications, the panel might be able to evaluate the research group. For example, as discussed in section 6.4.1 and 6.4.2, there is no overlap between the journal portfolio of group PHAR-I and the Pharmaceutical Sciences panel, but PM2 is still fairly close to this research group (Figure 25) both in the barycenter and the SAPV method (Table 29 and Table 33).

Both methods give the opportunity to assess the composition of the panel in terms of cognitive distance if one or more panel members are replaced and compare the relative contribution of each potential panel member to the panel fit as a whole, by observing the changes to the distance between the panel's and the groups'. In future research, we intend to compare these methods, as well as some others, with external data to gain more insight in their 'practical' merits (discussed in chapter VIII).

6.5 Conclusion

We have considered two potential methods of determining the match between research groups and expert panel members based on the journals in which they have published: distances (including confidence intervals) between barycenters on the map and distances (also including confidence intervals) between SAPVs. Both the barycenter and SAPV methods hold serious advantages over a simple comparison of publication portfolios. Visualizations in the form of overlay maps can provide an intuitive picture of an entity's publication profile and include information on journal similarity, but they are less suited for actually distinguishing between, say, a few different panel members. In these cases, we have argued, distances between profiles that take similarity into account (like barycenters and SAPVs) constitute a method with more 'actionable' results.

6.5.1 Discussion

A research group may deliberately hire other professionals, e.g., a biology research group might hire a physicist or computer scientist who continues to publish in their own discipline. In that case, the group's publication profile may change somewhat. We argue that it is the choice of the research group whether or not to include such publications in their research group profile during the period of research evaluation. As the formation of expert panel considers the focus of the research groups, the application of the barycenter and SAPV methods are not affected.

In our case, the panel members have no prior involvement with the research groups, but the barycenter method and SAPV method can also be applied if the panel members have already collaborated with a research group or unit of assessment. The involvement of the panel member with the research group may result in a much better panel fit, but the research assessment itself might be subject to bias. However, such influence is outside the scope of the thesis, as the formulation of criteria for selection of the panel members depends on the objectives of the concerned authority.

The scope of journals can vary significantly; some journals focus on rather specific topics, whereas others, such as PLoS ONE, are multidisciplinary in nature. One might therefore question whether journals are the adequate level of analysis. We suggest two possible routes for future research in this regard. First, it would be interesting if a comparison could be made between an analysis that considers all journals and one that leaves out multidisciplinary or otherwise broadly scoped journals. Second, one could replace journals with clusters of cognitively related articles. For instance, one could use the CWTS (Centre for Science and Technology Studies) article-level classification (Waltman & van Eck, 2012), which groups related articles together on the basis of direct citations regardless of the journal in which they were published. While we consider this an interesting idea, we also point out that it harbors its own set of theoretical and practical problems.

6.5.2 Normative implications

Our proposed expert panel composition methods based on journal data allow the panel composition authority to see in advance about the panel's fit to the research groups that are going to be evaluated. The distance between units of assessment can be used as an indicator of cognitive distance. Therefore, the concerned authority will have the opportunity to replace outliers among the panel members to make the panel fit well with the research groups to be evaluated. For example, the authority can find a best-fitting expert panel by replacing a more distant panel member with a potential panel member located closer to the groups, in addition to the other panel member to cover the expertise of the PHAR-I research group. Also, the distances between panel members and research groups could be used to facilitate the division of labor among the panel members. In our opinion, adequate coverage can be considered a necessary condition for the quality of an evaluation. Both the barycenter and SAPV methods for measuring cognitive distance can be used to inform the process of expert panel composition for a collection of research groups.

Chapter VII: Measuring cognitive distance between publication portfolios[7]

7.1 Introduction

In this chapter, we address the research question: How can we obtain, using publication data, a meaningful distance or proximity measure which represents the cognitive distance or proximity between two units? This is in fact a rephrased version of a problem we discussed in the chapter V, where we asked, 'How can one quantify the cognitive distances (overlap of expertise) between two entities (e.g., a research group and a panel) using the WoS SCs to which their publications belong?'

In our investigation, entities or units are either experts, panels of experts, or research groups. One can easily think of other informetric contexts in which the calculation of cognitive distances is relevant, e.g. the search of suitable peer reviewers for the evaluation of journal submissions, for grant applications or in hiring/promotion decisions, the exploration of potential collaborations, and distinguishing between different 'modalities' of interdisciplinarity (Molas-Gallart, Rafols & Tang, 2014). Rafols, et al., (2010) suggest several possible uses of overlay maps in research management that depend on cognitive distance, such as benchmarking and comparing the research profiles of organizations, and exploring complementarities and possible collaborations. In this regard, they point out that "successful collaborations tend to occur in a middle range of cognitive distance, whereupon collaborators can succeed at exchanging or sharing complementary knowledge or capabilities, while still being able to understand and coordinate with one another." Our quantitative approaches are complementary to visual approaches like overlay maps (Leydesdorff, & Rafols, 2009; Rafols, et al., 2010; Leydesdorff, Carley, et al., 2013).

[7]This chapter is based on Rousseau, Guns, Rahman & Engels (2017).

In this chapter, we focus on theoretical-logical aspects of the calculation of cognitive distance. As an application and to keep a clear link with chapter V, we re-use the data and framework. In that chapter, publications were assigned to WoS SCs. We admit that the use of WoS SCs was a convenience approach, which has meanwhile been refined by applying a journal level approach in the chapter VI. More precisely, instead of assigning publications to WoS SCs, publications were assigned to the journal in which they were published.

7.2 Measuring cognitive distance

Here, we consider the publication portfolio of the involved researchers to reflect the position of the unit in cognitive space and, hence, to determine cognitive distance. Expressed in general terms we measure cognitive distance between units based on how often they published in the same or similar journals. Similarity between journals can be measured in a direct way or via the WoS SCs to which they belong. Details are provided further on.

In the case study presented in this chapter, similarity is determined by the citation-based similarity of WoS SCs to which journals belong. The research groups are either research groups in physics or in chemistry working at the University of Antwerp, Belgium. For details, we refer to chapter V.

One can think of other informetric ways to determine cognitive distance between scientists. Wang & Sandström (2015) for example use bibliographic coupling and topic modelling to determine cognitive distance between publication portfolios. Besides using publication portfolios, one could also measure cognitive distance between patent portfolios, in terms of conference participation, in terms of diplomas, and so on. Moreover, cognitive distance is relevant in many other social and political contexts as well, e.g. when hiring employees, when comparing the programs of political parties, or to understand cultural differences.

We recall chapter III that in order to obtain meaningful cognitive distances these values must be scale-invariant. This means that the distance between points P and Q must be the same as the distance between the points P and cQ, where c is a strictly positive number. Indeed: the total output of a research group can be several orders of magnitude larger than that of one expert. For the applications, we have in mind this difference must not play a role in determining cognitive distances. Scale-invariance can be obtained through normalization as illustrated (for 3 dimensions) in Figure 31. All points situated on the straight line through the origin are represented by the same point in the plane with equation x + y + z = 1.

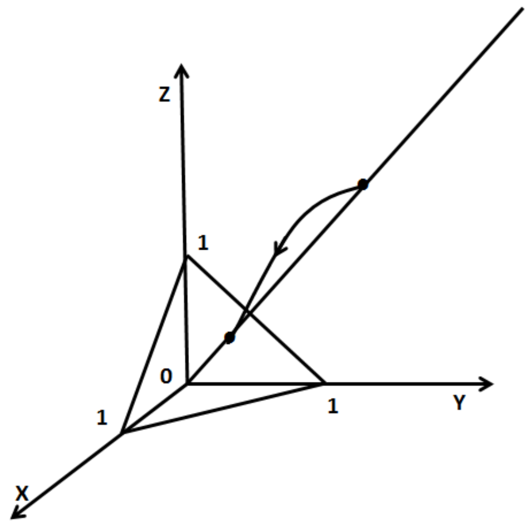


Figure 31: Normalization, leading to a scale invariant approach

This is so-called L_1-normalization: by dividing each coordinate by the sum of all coordinates one obtains a new array for which the sum of all coordinates is one (taking into account that no coordinate is negative). One could equally well divide by an array's Euclidean length (so-called L_2-normalization) but as we do not see an advantage for any of the two approaches, we applied L_1-normalization as is done in diversity studies.

7.3 Representing researchers' publication profiles

Researchers' publication profiles and their (dis)similarities will be represented in five different ways: a benchmark, two methods using barycenters (one in two and one in three dimensions), a fourth method using SAPVs and a fifth one using weighted cosine similarities (WCS). The benchmark and the last two values are applied in N dimensions, where N denotes the total number of SCs. In each case we start from a publication vector $M = (m_j)_j$, with j=1,…,N. The coordinates of this vector are the number of publications belonging to category j. Each panel member and each research group has a corresponding publication vector. In the applications only publications during a specific publication window and included in the WoS are considered, but the approach is independent of the used publication window or data source.

Throughout the remainder of the chapter, we will work with the example of determining cognitive distances between expert panels and their members on the one hand and research groups on the other (in the context of research evaluation). However, we stress the fact that the methods presented are more general and can also be applied in other contexts and for other purposes.

7.3.1 First method: The benchmark

Scientists and research groups are represented as N-dimensional publications vectors. As a start (benchmark) we just calculate the Euclidean distance between the L_1-normalized arrays of each panel member and each research group. Recall that the Euclidean distance between two vectors $a = (a_n)_{n=1,\ldots,k}$ and $b = (b_n)_{n=1,\ldots k}$ in $\mathbf{R}^k$, for any strictly positive integer k, is given as:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_k - b_k)^2} \tag{17}$$

In this chapter, we will use formula (17) for k = 2, k=3 and k = N.

7.3.2 Second and third method: Barycenters

To answer our research question, the second method uses a 2-dimensional base map. We note that this base map can be considered to be universal and hence has nothing to do with the concrete data at hand. Each SC has a place on this map, characterized by corresponding coordinates, denoted as $(L_{j,1}, L_{j,2})$, $j = 1, \ldots, N$. In the application that will follow, the 2-dimensional barycenter approach is based on a VOS (visualization of similarities) (Van Eck & Waltman, 2007) map (taken from Leydesdorff et al., 2013), but other 2-dimensional mappings are feasible. Now for each panel member and for each research group a barycenter derived from their publication profiles is calculated. Coordinates of these barycenters (in 2-dimensions) are given as

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \ ; \ C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} \tag{18}$$

where m_j is the number of publications of the unit under investigation (panel member, research group) belonging to category j; this category j has coordinates ($L_{j,1}$, $L_{j,2}$) in the base map and $T = \sum_{j=1}^{N} m_j$. We note that in the case study performed further on, T is larger than the total number of publications as full counting of WoS SCs has been used, which means that publications belonging to multiple WoS SCs are counted multiple times. Euclidean distances between units, as represented by their barycenters, can be calculated leading to quantitative results answering our research question.

The barycenter method explained above and in particular formulae (18) satisfy the scale-invariance requirement as multiplying all m_js with the same strictly positive factor leads to the same barycenter. Although it is convenient to perform visualization and to determine cognitive distance in the plane, there is no theoretical reason to perform these acts in two dimensions. Likewise, there are no strong reasons to do both in the same dimension. The barycenter method can, at least in theory, be applied in any strictly positive dimension smaller than or equal to N. Not wanting to go too deep into this largely theoretical issue we will just check how results for our case studies compare in 2-dimensions and 3-dimensions, leading to the third method, namely the use of barycenters in three dimensions.

For 3-dimensions, we again use the VOS algorithm, but now resulting in a 3-dimensional base map. This map was based on the network in http://www.leydesdorff.net/overlaytoolkit /map10.paj and obtained using Pajek, which implements the VOS algorithm both in 2 and 3 dimensions.

Again, each SC has a place on this map, characterized by corresponding coordinates, denoted as $(L_{j,1}, L_{j,2}, L_{j,3})$, $j = 1, \dots, N$, and for each panel member and for each research group a barycenter derived from their publication profiles is calculated. Coordinates in 3-dimensions are given as

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \ ; \ C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} ; \ C_3 = \frac{\sum_{j=1}^{N} m_j L_{j,3}}{T} \tag{19}$$

The meaning of the symbols T and m_j in formulae (19) is the same as in formulae (18).

7.3.3 Fourth method: Similarity-adapted publication vectors

In this method, we used a matrix of similarity values between the WoS SCs as made available by Rafols, et al., (2010) at http://www.leydesdorff.net/overlaytoolkit/map10.paj. These authors created a matrix of citing to cited SCs based on the SCIE and SSCI that was cosine-normalized in the citing direction. The result is a symmetric N×N similarity matrix (here, N=224) which we denote by $S = (s_{ij})_{ij}$.

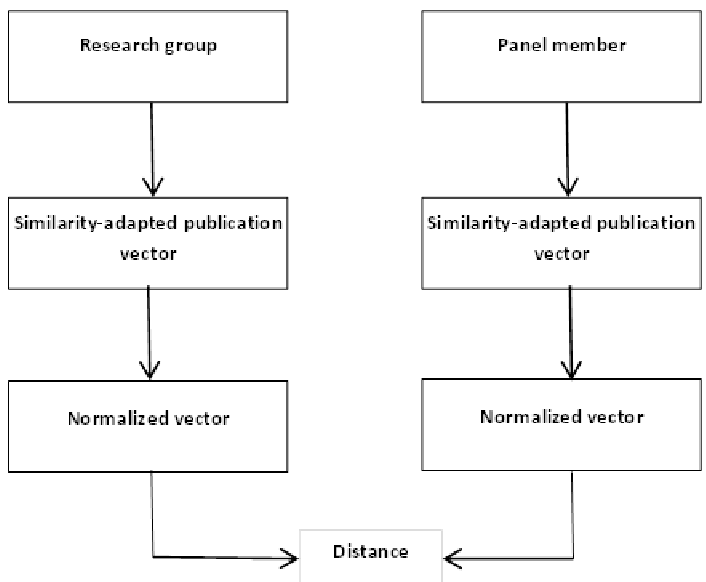


Figure 32: Workflow for determining distances between SAPVs

The multiplication $S * M$, i.e. applying the linear map with matrix representation S to the publication vector M leads to a new vector which we termed a similarity-adapted publication vector, SAPV in short. If we ignore similarity then S is the identity matrix and publication columns stay unchanged. We consider the SAPV method to be quite interesting as it provides a solution to the problem that WoS SCs overlap and are sometimes poorly defined, the SC *Information Science & Library Science* being a well-known example.

We determined the distance for SAPVs through normalization. It suffices though, to follow the workflow shown in Figure 32. Hence, a normalized SAPV of a research group or panel member is determined as the vector $C = (C_1, C_2, \dots, C_N)$, with coordinates C_k determined as:

$$C_k = \frac{\sum_{j=1}^{N} s_{kj} m_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} m_j} = \frac{(S * M)_k}{\|S * M\|_1} \tag{20}$$

where s_{kj} denotes the similarity value between the k-th and the j-th WoS SC, and m_j is the number of publications in WoS SC j of the research group or the panel member. The numerator of Equation (20) is equal to the k-th element of $S * M$, the multiplication of the similarity matrix S and the column matrix of publications $M = \left(m_j\right)_j$. The denominator is the L_1-norm of the unnormalized vector. We observe that the L_1-norm of the normalized vector C is indeed equal to 1.

7.3.4 Fifth method: Weighted cosine similarity

Finally, we mention a weighted cosine similarity (WCS) method. The WCS between panel member (PM) k and research group m, according to Zhou et al. (2012) is:

$$\frac{\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)}{\sqrt{\left(\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} M_j^k s_{ji}\right)\right) \cdot \left(\sum_{i=1}^{N} R_i^m \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)\right)}}$$

$$= \frac{\left(M^k\right)^t * S * R^m}{\sqrt{\left(M^k\right)^t * S * M^k} \cdot \sqrt{\left(R^m\right)^t * S * R^m}} \tag{21}$$

The numerator is nothing but the matrix multiplication: $\left(M^k\right)^t * S * R^m$, where ^{t} denotes matrix transposition, S is the similarity matrix, M^k denotes the column matrix of publications of panel member k and R^m denotes the column matrix of publications of research group m. Similarly, the two products under the square root in the denominator are: $\left(M^k\right)^t * S * M^k$ and $\left(R^m\right)^t * S * R^m$. The result is the WCS value between panel member k and research group m. Formula (21) is clearly scale-invariant: multiplying M^k or R^m with a fixed constant does not change the result. Note that if S is the identity matrix (similarity is not taken into account), formula (21) reduces to regular cosine similarity. A similarity or proximity can be considered as the opposite of a distance: the higher the similarity the better the match – the closer the distance – between a panel member and a research group. This value too is calculated for each panel member and each research group. We note that this fifth method may lead to mathematical problems when applied in general vector spaces, but that these do not occur in the particular framework used in this article (in mathematical terms: we work in the positive cone $(\mathbf{R}^{+})^{N}$, where **R** denotes the real numbers). Details are provided in chapter III (see section 3.2.5).

7.4 Results

As in chapter V, we calculate the cognitive distance between different research groups and panel members. Group names have been standardized using the first four letters of the corresponding department, for example, CHEM-A for chemistry research group A, PHYS-B for physics research group B. The panel member names are standardized as PM1, PM2 etc., but refer to different colleagues depending on the panel in question.

Yet, another problem must be solved before we can really state that one panel member is closer to a research group than another. Small differences in distance or similarity bear little meaning and should not be used to make claims that, for instance, one panel member is a 'better' choice than another. We therefore use a bootstrapping method (Efron & Tibshirani, 1998) leading to 95% CIs for distances and similarities. Details of the bootstrapping method we applied and explained in Chapter III and chapter VI respectively. A more detailed explanation can be found online (Guns, 2016a, 2016b). If the CI of the panel member who is closest to a given research group overlaps with that of the panel member who ranks second (and maybe even with the panel members ranking third or fourth) we say that there is no (statistical) difference in cognitive distance.

In order to facilitate a comparison between the five methods, results for the barycenter method in 2-dimension (Table 38 and Table 39) are re-considered and information about the calculated CIs is added. Hence, we begin the presentation of shortest distances between panel members and research groups with the benchmark case (Table 36 and Table 37), followed by the 3-dimesnional barycenter case (Table 40 and Table 41), the SAPV method (Table 42 and Table 43) and finally the WCS method (Table 44 and Table 45).

Table 36: Euclidean distances in N-dimensions between normalized publication arrays of research groups and panel members of the Chemistry department

	CHEM -A	CHEM -B	CHEM -C	CHEM -D	CHEM -E	CHEM -F	CHEM -G	CHEM - H	CHEM -I	CHEM - J	CHEM - K	CHEM - L
PM1	0.607	0.697	0.646	**0.459**	0.627	0.743	0.656	0.652	0.674	0.646	0.607	0.667
PM2	0.507	0.565	**0.402**	0.588	**0.300**	**0.240**	0.316	**0.377**	**0.269**	**0.356**	**0.445**	0.531
PM3	0.540	0.573	**0.381**	0.598	**0.279**	**0.405**	**0.288**	**0.257**	**0.242**	**0.350**	**0.468**	0.561
PM4	0.542	0.601	**0.441**	0.608	**0.331**	**0.340**	**0.217**	**0.372**	**0.336**	**0.360**	**0.464**	0.556
PM5	**0.180**	**0.157**	0.482	0.604	0.500	0.659	0.547	0.499	0.515	0.520	**0.500**	**0.368**
PM6	0.715	0.762	0.726	**0.255**	0.693	0.809	0.738	0.731	0.749	0.729	0.693	0.745
PM7	0.684	0.770	0.741	0.758	0.732	0.825	0.746	0.744	0.761	0.741	0.713	0.739

Table 37: Euclidean distances in N dimensions between normalized publication arrays of research groups and panel members of the Physics department

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.716	0.793	0.699	**0.114**	0.519	0.786	0.730	0.806	0.662
PM2	0.953	0.466	0.788	1.048	0.801	1.008	0.956	0.457	0.899
PM3	0.639	0.741	0.654	0.819	0.634	0.759	0.701	0.705	0.621
PM4	0.600	0.663	0.476	0.738	0.481	**0.663**	**0.278**	0.662	0.523
PM5	**0.510**	0.376	**0.171**	0.667	**0.296**	**0.559**	0.494	0.410	**0.387**
PM6	0.618	**0.224**	0.388	0.736	**0.379**	**0.576**	0.568	**0.241**	0.531

Table 38: Euclidean distances between barycenters of research groups and panel members of the Chemistry department using the 2-dimensional WoS SCs map

	CHEM -A	CHEM -B	CHEM -C	CHEM -D	CHEM -E	CHEM -F	CHEM -G	CHEM - H	CHEM -I	CHEM - J	CHEM - K	CHEM - L
PM 1	0.167	**0.129**	0.217	**0.165**	0.329	0.337	0.179	0.165	0.111	0.394	0.454	**0.127**
PM 2	0.350	0.342	0.362	**0.129**	**0.079**	**0.090**	0.145	0.215	0.199	**0.259**	**0.228**	0.342
PM 3	0.171	0.161	0.192	**0.129**	0.252	**0.263**	**0.053**	**0.061**	**0.020**	**0.269**	0.330	0.161
PM 4	0.269	0.262	0.280	**0.108**	**0.158**	**0.170**	**0.063**	**0.134**	0.121	**0.232**	**0.250**	0.263
PM 5	**0.056**	**0.055**	**0.091**	0.232	0.367	0.378	0.154	**0.093**	**0.099**	**0.315**	0.411	**0.057**
PM 6	0.302	0.276	0.335	**0.027**	**0.175**	**0.181**	0.161	0.210	0.156	0.366	0.370	0.275
PM 7	**0.116**	**0.072**	0.172	0.235	0.395	0.404	0.216	0.178	0.144	0.410	0.491	**0.070**

Table 39: Euclidean distances between barycenters of research groups and panel members of the Physics department using the 2-dimensional WoS SCs map

	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
PM 1	1.173	0.123	0.215	<u>**0.017**</u>	0.145	**0.208**	0.495	0.120	0.664
PM 2	1.195	0.067	0.109	0.158	0.118	0.316	0.443	0.056	0.688
PM 3	**1.041**	0.146	0.194	0.116	0.113	<u>**0.104**</u>	**0.387**	0.157	**0.532**
PM 4	<u>**1.020**</u>	0.168	**0.085**	0.263	0.132	0.295	<u>**0.249**</u>	0.179	<u>**0.522**</u>
PM 5	**1.136**	**0.046**	<u>**0.055**</u>	0.159	<u>**0.069**</u>	0.281	**0.385**	**0.050**	**0.629**
PM 6	**1.157**	<u>**0.031**</u>	0.084	0.138	**0.078**	0.280	**0.412**	<u>**0.026**</u>	0.649

Table 40: Euclidean distances between barycenters of research groups and panel members of the Chemistry department using the 3-dimensional WoS SCs map

	CHEM -A	CHEM -B	CHEM -C	CHEM -D	CHEM -E	CHEM -F	CHEM -G	CHEM - H	CHEM -I	CHEM - J	CHEM - K	CHEM - L
PM1	**0.037**	**0.032**	**0.043**	**0.033**	0.064	**0.059**	**0.018**	<u>**0.006**</u>	**0.014**	**0.043**	0.103	**0.033**
PM2	0.110	0.108	0.114	**0.045**	<u>**0.017**</u>	<u>**0.022**</u>	0.062	0.075	0.063	**0.060**	<u>**0.035**</u>	0.110
PM3	0.051	0.047	0.056	**0.019**	**0.050**	**0.044**	<u>**0.006**</u>	**0.015**	<u>**0.007**</u>	**0.040**	0.090	0.048
PM4	0.069	0.063	0.074	<u>**0.012**</u>	**0.037**	**0.032**	**0.013**	**0.033**	**0.023**	**0.050**	0.084	0.064
PM5	**0.030**	**0.027**	**0.034**	**0.040**	0.069	**0.064**	0.028	**0.007**	**0.019**	<u>**0.038**</u>	0.103	**0.029**
PM6	0.057	0.052	0.062	**0.013**	**0.044**	**0.038**	**0.007**	**0.021**	**0.010**	**0.039**	0.085	0.054
PM7	<u>**0.023**</u>	<u>**0.016**</u>	<u>**0.028**</u>	**0.049**	0.080	**0.075**	0.034	**0.018**	**0.030**	**0.053**	0.117	<u>**0.017**</u>

Table 41: Euclidean distances between barycenters of research groups and panel members of the Physics department using the 3-dimensional WoS SCs map

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.453	0.054	0.084	<u>**0.011**</u>	0.067	0.064	0.162	0.048	0.257
PM2	**0.408**	**0.007**	0.032	0.043	**0.016**	0.044	**0.112**	**0.008**	**0.211**
PM3	**0.392**	0.024	0.037	0.050	0.026	<u>**0.013**</u>	**0.105**	0.026	**0.196**
PM4	<u>**0.361**</u>	0.049	**0.018**	0.091	0.035	0.061	<u>**0.062**</u>	0.054	<u>**0.163**</u>
PM5	**0.393**	0.014	<u>**0.017**</u>	0.056	<u>**0.003**</u>	0.041	**0.096**	0.019	**0.195**
PM6	**0.409**	<u>**0.006**</u>	0.034	0.040	**0.017**	0.041	**0.113**	<u>**0.004**</u>	**0.211**

Table 42: Euclidean distances between SAPVs of research groups and panel members of the Chemistry department using the similarity matrix of WoS SCs

	CHEM -A	CHEM -B	CHEM -C	CHEM -D	CHEM -E	CHEM -F	CHEM -G	CHEM - H	CHEM -I	CHEM - J	CHEM - K	CHEM - L
PM 1	0.081	0.079	0.108	**0.061**	0.124	0.119	0.116	0.104	0.093	0.129	0.141	0.085
PM 2	0.082	0.074	0.079	**0.054**	**0.036**	**0.032**	0.055	**0.046**	**0.036**	**0.075**	**0.071**	0.070
PM 3	0.082	0.074	0.080	0.066	**0.057**	**0.058**	**0.040**	**0.040**	**0.042**	**0.075**	**0.086**	0.073
PM 4	0.106	0.099	0.104	0.085	0.064	0.070	**0.027**	0.063	0.071	**0.085**	0.094	0.091
PM 5	**0.015**	**0.013**	**0.034**	0.074	0.100	0.102	0.077	**0.053**	**0.050**	**0.082**	0.096	**0.024**
PM 6	0.093	0.087	0.111	**0.025**	0.085	0.080	0.096	0.090	0.080	0.113	0.116	0.088
PM 7	0.068	0.068	0.097	0.072	0.128	0.125	0.113	0.099	0.089	0.125	0.140	0.075

Table 43: Euclidean distances between SAPVs of research groups and panel members of the Physics department using the similarity matrix of WoS SCs

	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
PM 1	0.376	0.358	0.373	**0.098**	0.328	0.301	0.371	0.358	0.367
PM 2	0.172	0.019	0.038	0.272	0.054	0.127	0.115	**0.019**	0.133
PM 3	**0.156**	0.065	0.080	0.256	0.069	**0.100**	0.116	0.063	**0.111**
PM 4	**0.144**	0.060	0.039	0.271	0.051	0.129	**0.066**	0.063	**0.103**
PM 5	**0.157**	0.023	**0.016**	0.271	**0.044**	0.125	**0.095**	0.027	**0.115**
PM 6	**0.165**	**0.012**	0.035	0.258	**0.037**	**0.111**	**0.106**	**0.015**	0.125

Table 44: WCS values of research groups and panel members of the Chemistry department using the similarity matrix of WoS SCs

	CHEM -A	CHEM -B	CHEM -C	CHEM -D	CHEM -E	CHEM -F	CHEM -G	CHEM - H	CHEM -I	CHEM - J	CHEM - K	CHEM - L
PM1	0.709	0.667	0.445	**0.922**	0.469	0.449	0.395	0.440	0.507	0.323	0.273	0.661
PM2	0.670	0.713	0.726	0.675	**0.914**	**0.945**	0.837	**0.847**	**0.947**	**0.703**	**0.527**	0.713
PM3	0.594	0.655	0.673	0.569	**0.839**	0.831	0.866	**0.880**	**0.894**	**0.711**	**0.403**	0.604
PM4	0.459	0.517	0.504	0.484	0.781	0.777	**0.951**	0.758	0.769	**0.626**	0.315	0.549
PM5	**0.983**	**0.990**	**0.842**	0.669	0.581	0.475	0.614	0.747	0.758	**0.573**	**0.512**	**0.933**
PM6	0.613	0.600	0.377	**0.973**	0.545	0.519	0.391	0.410	0.484	0.294	0.280	0.603
PM7	0.758	0.713	0.503	0.850	0.460	0.439	0.440	0.494	0.550	0.373	0.290	0.700

For each research group, we determine the panel member at the shortest distance. The number in the row corresponding to this panel member is indicated in bold and underlined. Distances whose confidence intervals overlap with that of the shortest distance are in bold (same column). We use the same way of showing results for all the tables.

Table 45: WCS values of research groups and panel members of the Physics department using the similarity matrix of WoS SCs

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.030	0.155	0.043	**0.996**	0.561	0.508	0.028	0.154	0.052
PM2	**0.151**	**0.982**	0.920	0.127	0.806	0.513	0.543	**0.977**	**0.497**
PM3	**0.220**	0.714	0.625	0.211	0.668	0.526	0.440	0.762	**0.544**
PM4	**0.182**	0.729	0.829	0.129	0.757	0.436	**0.895**	0.741	**0.479**
PM5	**0.182**	0.965	**0.986**	0.158	**0.852**	0.475	0.656	0.957	**0.567**
PM6	**0.164**	**0.989**	0.930	0.272	**0.903**	**0.643**	0.631	**0.985**	**0.516**

Table 44 and Table 45 contain the WCS results, where we recall that this is a similarity approach (not a distance-based one) and hence largest values refer to entities that are closest.

7.5 Correlations between distances/similarities based on the five methods

We calculated the Pearson correlation coefficient (r) and the Spearman rank correlation coefficient (ρ) between distances/similarities based on the five methods. These calculations are based on all distances between research groups and individual panel members. For calculations involving WCS we show absolute values, as distances and similarities are each other's opposites, and hence correlations are negative.

Table 46: Chemistry: Pearson and Spearman correlations for all cognitive distances between research groups and individual panel members

Pearson / Spearman	Benchmark	Barycenter 2D	Barycenter 3D	SAPV	WCS
Benchmark	1.00	0.38	0.09	0.72	0.72
Barycenter 2D	0.34	1.00	0.81	0.75	0.64
Barycenter 3D	0.06	0.82	1.00	0.42	0.31
SAPV	0.67	0.72	0.42	1.00	0.92
WCS	0.67	0.62	0.30	0.92	1.00

Table 47: Physics: Pearson and Spearman correlation for all cognitive distances between research groups and individual panel members

Pearson / Spearman	Benchmark	Barycenter 2D	Barycenter 3D	SAPV	WCS
Benchmark	1.00	0.12 (0.34)	0.22 (0.27)	0.50 (0.56)	0.63 (0.54)
Barycenter 2D	0.37(0.48)	1.00	0.99 (0.99)	0.29 (0.87)	0.60 (0.89)
Barycenter 3D	0.34(0.38)	0.94 (0.96)	1.00	0.35 (0.81)	0.61 (0.85)
SAPV	0.60(0.56)	0.64 (0.94)	0.71 (0.86)	1.00	0.86 (0.97)
WCS	0.65(0.58)	0.71 (0.91)	0.74 (0.83)	0.94 (0.97)	1.00

In Table 46 and Table 47, the upper triangle refers to Pearson correlations while the lower triangle refers to Spearman correlations. Clearly, SAPV and WCS results in Table 46 and Table 47 are highly correlated. This also applies to correlations between barycenter in two and three dimensions.

Values between brackets in Table 47 are correlations calculated after removal of PHYS-D and PM1; an explanation for doing this is provided further. Correlations for the benchmark case (ignoring all similarities) and the other approaches are moderate at best.

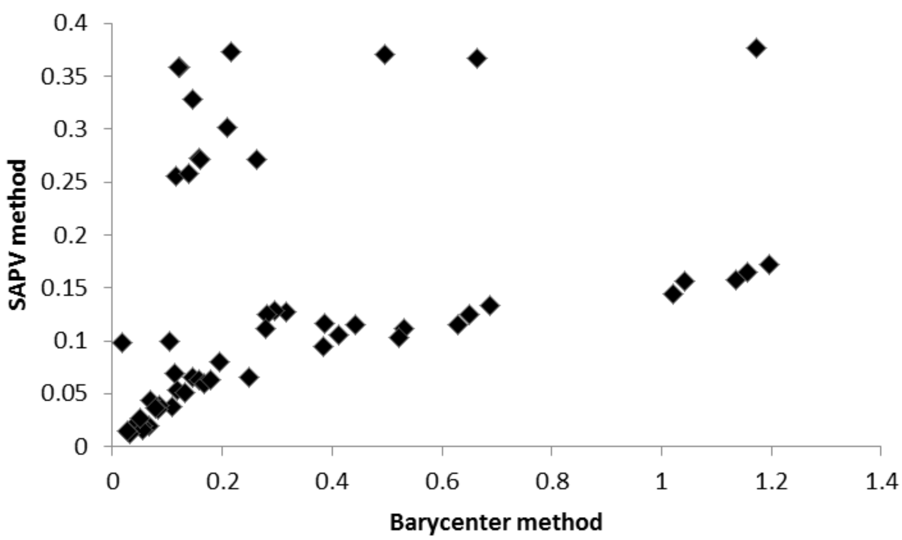


Figure 33: Scatter plot of the cognitive distances between research groups and individual panel members for the 2-dimensional barycenter and SAPV methods in the physics department

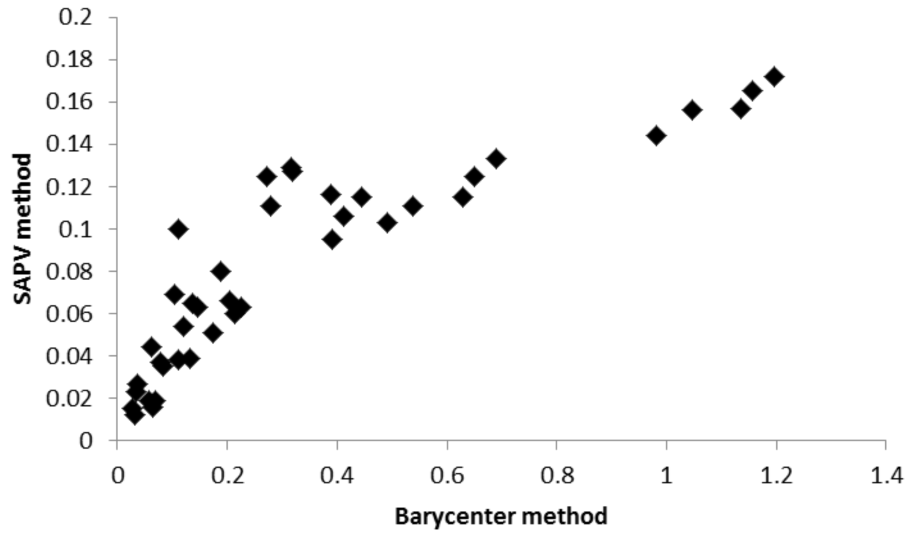


Figure 34: Scatter plot of the cognitive distances between research groups and individual panel members obtained by the 2-dimensional barycenter and SAPV methods in the physics department excluding PHYS-D and PM1

Not surprisingly, the two N-dimensional approaches (SAPV and WCS) are more correlated with the benchmark case than the lower dimensional ones. Correlations between the 2-dimensional and the 3-dimensional approach are high in all cases.

This illustrates that the number of dimensions chosen has only limited influence on the results based on barycenters. Most other correlations can be described as moderate to high. For chemistry we note, however, that the correlations between barycenter 3-dimension on the one hand, and SAPV and WCS on the other, are lower than expected. Moreover, these values are lower than for the 2-dimension case. We were not able to find an explanation for this unexpected difference. We further note a low correlation between SAPV and the barycenter methods in physics. For this case, however, we found a convincing explanation. Figure 33 illustrates what happened.

This low Pearson correlation is due to the 13 points (including two times two points that overlap and cannot be seen) in the upper half of Figure 33. All these points correspond to distances involving research group PHYS-D and PM1 (but not both). This group and this panel member are active in the same field (*Physics, Particles & Fields*) and have different scientific interests than the other groups or panel members: 99.1% of PM1's publications belong to the SC *Physics, Particles & Fields*, while for PHYS-D, this SC covers 83.6% of its publications. Moreover, their publications cover only four (117 publications) and seven (269 publications) WoS SCs respectively while other panel members cover 12 to 26 WoS SCs, and other research groups 26 to 50 SCs. Figure 34 presents the same data as Figure 33, but leaves out distances involving PHYS-D and PM1. In this case, all correlations increase considerably.

A more detailed comparison between the five methods follows in the next section.

7.6 Comparison between the five methods

A comparison would be easy if a gold standard existed. Clearly, it does not, but we used the labour division decided upon by the panel chair as a proxy. Prior to a site visit (see Engels et al., 2013 for details), the panel chair appointed a main assessor for each of the research groups to be evaluated. This main assessor studied the profile and performance of the research group in detail, asked the majority of questions during the site visit and wrote the (first draft of) the final assessment of the research group.

For chemistry, the barycenter methods score slightly better than SAPV and WCS, while for physics there is hardly any difference between the four (even five) methods. Especially in the case of chemistry, we have several cases where most confidence intervals overlap. The barycenter method in 3-dimension clearly has very low discriminatory power leading to cases where all confidence intervals overlap (CHEM-F and CHEM-J). In these cases, the 3-dimension barycenter cannot distinguish between panel members.

We see that for some research groups the five methods and the chosen assessor coincide (taking confidence intervals into account). This perfect result was attained for CHEM-B, CHEM-E, CHEM-J, PHYS-C, PHYS-D, PHYS-G and PHYS-H; while only the benchmark case missed PHYS-A and PHYS-L.

Hence, this is the case for 3 of the 12 chemistry groups and for 4 (or 6) of the 9 physics groups. The smaller number of perfect results in chemistry is largely due to the WCS method. For some other groups, no method leads to the chosen assessor. This is the case for CHEM-A, CHEM-K, and PHYS-E.

Table 48 and Table 49 show the research groups, the corresponding main assessor, and the panel members with the closest distance (for the five methods). The first one in each cell is the panel member closest to the corresponding research group; the others are panel members whose distances are statistically not different from this shortest distance.

Assuming that panel chairs assigned the best suited panel member as main assessor, a perfect method would always rank this main assessor first. However, remember that neither have panel members and research groups ever collaborated nor do they belong to the same university, so this assumption does not necessarily always hold in practice.

Mainly due to the overlapping confidence intervals the barycenter method in 3-dimension is the only one which included the main assessor for CHEM-C and CHEM-I (and the benchmark has PM3 as closest to CHEM-C). In all these negative cases, the results obtained by the five methods largely agree. A possible explanation for this surprising result might simply be that the panel chair included other factors - than pure scientific affinity - in the decision to assign a panel member to a research group.

In the case of chemistry where the suggested labor division was partly contested by PM3, PM5 is identified as the closest to CHEM-C. A possible explanation for this specific case could be that PM5 was already the main assessor for two groups so that, for purely practical reasons, PM3 became the main assessor of CHEM-C.

Table 48: Chemistry: Top ranked panel members according to five methods

Research group	Main assessor	Benchmark	Barycenter 2D	Barycenter 3D	SAPVs	WCS
CHEM-A	PM6	PM5	PM5-PM7	PM7- PM5- PM1	PM5	PM5
CHEM-B	PM5	PM5	PM5-PM7- PM1	PM7- PM5- PM1	PM5	PM5
CHEM-C	PM7/PM3	PM3-PM2- PM4	PM5	PM7- PM5- PM1	PM5	PM5
CHEM-D	PM2	PM6- PM1	PM6-PM4- PM3-PM2- PM1	PM4- PM6- PM3- PM1- PM5- PM2- PM7	PM6-PM2- PM1	PM6-PM1
CHEM-E	PM2	PM3-PM2- PM4	PM2-PM4- PM6	PM2- PM4- PM6- PM3	PM2-PM3	PM2-PM3
CHEM-F	PM3	PM2-PM4- PM3	PM2-PM6- PM4-PM3	PM2- PM4- PM6- PM3- PM1- PM5- PM7	PM2-PM3	PM2
CHEM-G	PM3	PM4-PM3	PM3-PM4	PM3- PM6- PM4- PM1	PM4-PM3	PM4
CHEM-H	PM5	PM3-PM4- PM2	PM4-PM3- PM5	PM1- PM5- PM3- PM7- PM6- PM4	PM3-PM2- PM5	PM3-PM2- PM4
CHEM-I	PM4	PM3-PM2- PM4	PM3-PM5	PM3- PM6- PM1- PM5- PM4- PM7	PM2-PM3- PM5	PM2-PM3
CHEM-J	PM4	PM3-PM2- PM4	PM4-PM2- PM3-PM5	PM5- PM6- PM3- PM1- PM4- PM7- PM2	PM3-PM2- PM5-PM4	PM3-PM2- PM4-PM5
CHEM-K	PM6	PM2-PM4- PM3-PM5	PM2-PM4	PM2	PM2-PM3	PM2- PM5- PM3-
CHEM-L	PM1	PM5	PM5-PM7- PM1	PM7- PM5- PM1	PM5	PM5
score		**7/12 (2/12)**	**8/12 (4/12)**	**10/12 (3/12)**	**7/12 (2/12)**	**3/12 (2/12)**

Table 49: Physics: Top ranked panel members according to five methods

Research group	Main assessor	Benchmark	Barycenter 2D	Barycenter 3D	SAPVs	WCS
PHYS-A	PM3	PM5	PM4-PM3-PM5-PM6	PM4- PM3-PM5- PM2-PM6	PM4-PM3-PM5-PM6	PM3-PM5-PM4-PM6-PM2
PHYS-B	PM2	PM6	PM6-PM5	PM6- PM2	PM6	PM6-PM2
PHYS-C	PM5	PM5	PM5-PM4	PM5- PM4	PM5	PM5
PHYS-D	PM1	PM1	PM1	PM1	PM1	PM1
PHYS-E	PM4	PM5-PM6	PM5-PM6	PM5- PM2-PM6	PM6-PM5	PM6-PM5
PHYS-F	PM1	PM5-PM6-PM4	PM3-PM1	PM3	PM3-PM6	PM6
PHYS-G	PM4	PM4	PM4-PM3-PM5-PM6	PM4- PM5-PM3- PM2-PM6	PM4-PM5-PM6	PM4
PHYS-H	PM6	PM6	PM6-PM5	PM6- PM2	PM6-PM2	PM6-PM2
PHYS-I	PM3	PM5	PM4-PM3-PM5	PM4- PM5-PM3- PM2-PM6	PM4-PM3-PM5	PM5-PM3-PM6-PM2-PM4
Score		**4/9 (4/9)**	**7/9 (4/9)**	**7/9 (4/9)**	**6/9 (4/9)**	**7/9 (4/9)**

Considering now the individual panel members, we see that some are close to several research groups, while others are not close to any. For chemistry, we see that, according to the 2-dimensional barycenter method PM4 and PM5 are close to seven research groups, while PM2, PM3 and PM5 are closest to seven research groups according to the SAPV method. PM5 is closest to six research groups according to the WCS method. Clearly, PM5 was an essential panel member. According to the two barycenter-based methods, all chemistry panel members are closest to at least three groups, but according to the SAPV and the WCS method PM7 is closest to none.

For physics, PM5 and PM6 are closest to at least four research groups, and this for the four similarity-based methods. PM2 is closest to none according to the 2D barycenter method, but closest to four groups according to the WCS method. We observe the special role of PM1 in physics who is the only one closest to PHYS-D and this according to the five methods. This observation confirms the results seen in the correlation analysis. It, moreover, contains a warning that correlation analyses may suggest wrong conclusions. In this case, the poor correlations between the results obtained by the SAPV method and those obtained by the barycenter methods for groups and panel members that have no real importance (they are cognitively unrelated) should not distract from the generally better correlations for pairs that matter.

7.7 Conclusion

In this chapter, we showed that, besides using barycenters in a two- and three-dimensional base map, it is possible to derive cognitive distances in N-dimensions using the SAPVs and WCS methods. Our approach is rather general: it can in principle be applied to all cases where units produce publications, which can be situated on a base map or counted in relation to a similarity matrix. Of course, other approaches are also possible, such as the one proposed by Wang & Sandström (2015) which is based on bibliographic coupling and topic modelling. Operationalizing the notion of cognitive distance is essential to several topics in informetrics, e.g. peer review processes, evaluation procedures, exploration of collaboration, and the study of interdisciplinarity. Indeed, cognitive distance could also be derived from other objects than publications, such as patents. Cognitive distance is also of essence in other contexts such as hiring decisions, political programs, and cultural differences.

As pointed out in this chapter, calculating cognitive distances between units should be scale-invariant. Barycenters in a two- and three dimensional base maps satisfy this requirement. We note though that distances in a 2- or 3-dimensional map are artificial; for instance, Pajek uses coordinates in the interval [0, 1] (this also applies to its VOS implementation), whereas coordinates in VOSviewer may refer to a wider interval. Hence, only comparisons between distances and not their absolute values have meaning. Proper normalization in N dimensions also leads to scale-invariant distances.

We have shown that the barycenter method is relatively insensitive to the number of dimensions in which it is used. Yet, especially in 3D the barycenter method has little discriminatory power. Distances between normalized SAPVs in N dimensions are probably less distorted and hence more meaningful. A similar observation applies to the WCS method. Hence, our preference, based on mathematical logic, goes to the SAPVs and WCS methods. Yet, WCS scores badly in the case of chemistry, so that our final preference goes to the SAPV method. Admitting that in our case studies the barycenter methods score slightly better and that differences between the results obtained by different methods are rather small, it is obvious that the result of this comparison should not be generalized. In future research, we intend to make a similar empirical comparison for more disciplines.

In a chapter V, besides using a VOS map, we also investigated if a map based on the algorithm by Kamada and Kawai (1989) could be used. We found out however that a Kamada-Kawai map (in two and in three dimensions) can yield very different results, depending on the random seed used. For this reason, we turned to a VOS map, which is much more stable. We hope that this warning will prevent colleagues from making wrong inferences.

Finally, our investigations led to two unsolved problems. The first one is the unexplained low correlation between the barycenter method in 3-dimension and the SAPV and WCS methods for chemistry. We checked all calculations related to the barycenter method in 3-dimesnion but did not detect any error. Moreover, consequent investigations related to other departments, in particular the biomedical sciences, gave similar low correlations. The second problem is the use of the main assessor, as appointed by the panel chair, as a “gold standard”. We admit that this is a problematic approach, since it relies on assumptions that are not always met. Yet, for the moment, we have not found a better solution.

Chapter VIII: Cognitive distances between evaluators and evaluees in research evaluation: A comparison between three informetric methods at the journal and subject category aggregation level [8]

8.1 Introduction

As far as we are aware there was, prior to 2013, no method to measure and quantify congruence of expertise or cognitive distance between panels and research groups in discipline-specific research evaluation (Engels et al., 2013). We started to study the problem of quantifying cognitive distance, such that individual panel members' expertise covers the research domains in the discipline where the units of assessment (in our case: research groups) have publications. In the chapter V to VII, we focused on determining the cognitive distances between publication portfolios of an expert panel and research groups (Rahman, Guns, Rousseau, & Engels, 2015; Rahman, Guns, Leydesdorff, & Engels, 2016; Rousseau, Guns, Rahman, & Engels, 2017), while Wang & Sandström (2015) used bibliographic coupling and topic modelling to determine cognitive distance.

More specifically, we explored different ways of quantifying the cognitive distance between panel members' and research groups' publication profile in discipline-specific research evaluation. For this we consider all the publications of the research groups and panel members indexed in the WoS and pursue an investigation at two levels of aggregation: WoS SCs and journals. For this purpose, we used the similarity matrix of WoS SCs and a 2-dimensional base map derived from it (for details see Leydesdorff & Rafols, 2009; Rafols et al., 2010; Leydesdorff, Carley, et al., 2013) and also the similarity matrix of journals and its 2-dimensional base map (for details see Leydesdorff & Rafols, 2012; Leydesdorff, Rafols, & Chen, 2013).

[8] This chapter is based on Rahman, Guns, Rousseau & Engels (2017).

Hence, we proposed five different approaches namely a barycenter approach using WoS SCs and journals (chapter V and VI), SAPV approach using WoS SCs and journals (chapter VI and VII) and a WCS approach using WoS SCs (chapter VII). The SAPV and WCS methods use the similarity matrix of WoS SCs/journals while the barycenter method uses the respective two-dimensional base map derived from the similarity matrix of WoS SCs/journals. So far, we have not yet applied the WCS method at the journal level. In this chapter, we cover that gap. Hence, three methods and two levels of aggregation lead to six informetric approaches to inform cognitive distances between evaluators and evaluees in research evaluation.

Until now, we have not compared the two levels of aggregation. More generally, a systematic comparison and test of all six approaches has not yet been carried out. This chapter fills this gap. Hence, we set the following research questions:

i) What are the correlations between the different approaches? Which aspect (method vs level of aggregation) has the largest influence on the correlation?

ii) To what extent do the approaches agree in matching the panel member at the closest cognitive distance from a research group?

iii) How accurate are the approaches in matching the main assessor for each research group? How accurate are they to *uniquely* match the main assessor?

Firstly, we look at the influence of the level of aggregation and the number of dimensions for determining cognitive distances. Secondly, we explore whether or not all the methods indicate the same panel member as the one at the shortest cognitive distance from a research group. Finally, we investigate if there is any difference between the proposed methods to find the previously assigned main assessor.

8.2 Data

The data in this chapter stem from the research assessment during the period 2009 – 2014 of six departments belonging to the University of Antwerp. The same panel evaluates all research groups in a department.

Table 50: Publication statistics of the research groups and panels

Name of the Department	Assessment year	Research groups				Panel			
		No. of research groups	No. of journals	No. of publications	No. of WoS SCs	No. of panel members	No. of journals	No. of publications	No. of WoS SCs
Biology	2011	9	372	1158	90	5	217	786	54
Biomedical Sciences	2014	15	476	1234	103	8	395	1333	80
Chemistry	2009	12	300	920	94	7	248	2150	66
Pharmaceutical Sciences	2009	10	180	376	67	5	300	1036	68
Physics	2010	9	353	1739	108	6	204	1104	46
Veterinary Sciences	2014	3	146	231	61	4	200	837	55

A research group consists of one professor assisted by junior and/or senior researchers (PhD students and postdocs), or of a group of professors and a number of researchers working with them. These evaluations consider the entire research groups' scientific activity for a specific period, typically eight years preceding the year of evaluation. All articles, letters, notes, proceeding papers, and reviews by the research groups published during the reference period are included in the evaluation. In this article, we consider only the publications that are indexed in the SCIE and the SSCI of the WoS.

Table 50 lists the publication statistics of the research groups during the eight years preceding their evaluation. Altogether, there are 58 research groups in six departments. The number of publications per department ranges from 231 to 1739. In total, these publications appeared in 146 to 476 different journals and are distributed over 61 to 108 WoS SCs. Sometimes different research groups collaborated.

The ADOC of the University of Antwerp organizes research evaluations. Each department can suggest potential panel chairs and panel members, who have the rank of full professor and have a considerable record of accomplishment. Preferably, they have experience with research evaluations, are editors or board members of reputed journals, and have academic management experience. ADOC checks the publication profile and curriculum vitae of the potential panel chair and panel members and ensures that they do not have co-publications or joint projects with the research groups that are evaluated. In addition, they may not have had an appointment as visiting professor at the University of Antwerp, and cannot be a member of an expert panel for the Research Foundation Flanders to avoid any potential bias. ADOC can also make suggestions when the scientists proposed by the departments are not acceptable. Together the panel members have to cover all the sub-domains in the evaluated department. The panel chairs have the last word about the panel composition. The composed panel is presented to the bureau of the university's research council, which has to ratify the composition.

Table 50 also shows that in total, there are 35 panel members involved in the evaluation of the six departments. As publications reflect the expertise of their authors (Rybak et al., 2014), the entire publication profile of the panel members are included, up to the year of assessment. The number of panel members ranges from 4 to 8 for each department. The number of publications per panel ranges from 786 to 2150. In total, these publications appeared in 200 to 395 different journals and are distributed over 46 to 80 WoS SCs. There is no shared authorship between panel members and research groups in any of the cases. None of the panels has any co-authored publications among the respective panel members except for two Chemistry panel members who have two publications in collaboration.

8.3 Methods

Our approaches are based on the assumption that for the evaluation of a research group by a panel member, the shorter the cognitive distances between them the better the fit between the two. Since the analysis is based on Clarivate Analytics' (formerly Thomson Reuters') WoS data, only publications in journals included in the WoS are taken into account. To identify cognitive distances, we consider the journals and WoS SCs in which publications have appeared. An important characteristic of our approaches is that they take into account the similarity between WoS SCs and between journals: if the publications of a panel member and a research group appear in different yet similar or closely related journals, they may still cover the same research areas. Clarivate Analytics has assigned one or more subject categories to WoS indexed journals based on 'subjective, heuristic methods' and has received criticism for being crude for some research areas (Pudovkin & Garfield, 2002). However, WoS SCs cover all disciplines and are generally used by bibliometric practitioners (Rehn et al., 2014; Leydesdorff & Bornmann, 2015).

We use a global map of science based on WoS SCs data made available at http://www.leydesdorff.net/overlaytoolkit/map10.paj (Leydesdorff & Rafols, 2009; Rafols et al., 2010; Leydesdorff, Carley, et al., 2013). These authors created a matrix of citing to cited WoS SCs based on the SCIE and SSCI, which was subsequently normalized in the citing direction. The file 'map10. paj' contains a weighted network of WoS SCs.

We also use a global map of science based on journal similarity available at http://www.leydesdorff.net/journals11. We have received the similarity matrix data from Loet Leydesdorff in the context of a joint paper (Rahman, Guns, Leydesdorff and Engels 2016). The journal similarity matrix can be considered as an adjacency matrix, and thus is equivalent to a weighted network where similar journals are linked and link weights increase with similarity strength (see Leydesdorff, Rafols, & Chen (2013) for details). However, as some of the journals underwent name or other changes over time, we had to find a way to handle these changes in a uniform way. For detailed guidelines, we refer to chapter VI.

We now explain how the three methods – SAPV, barycenter, and WCS – are calculated. Throughout the discussion, N denotes the number of SCs (224) or the number of journals. There are 10,673 journals in the map, and 10,675 journals in the similarity matrix based on JCR 2011.

8.3.1 Similarity-adapted publication vector method

A regular publication vector counts per WoS SC or journal, whereas in a SAPV these counts are adapted to account for similarity between WoS SCs or journals. We use normalized SAPVs, such that there is scale invariance and publication vectors of entities of varying size can be meaningfully compared.

We calculate SAPVs for each entity, starting from the original publication vector and similarity matrices. Based on their respective SAPVs, the distance can be calculated between two entities. A similarity-adapted publication vector is determined as the vector $\mathrm{C} = (C_1, C_2, \ldots, C_N)$, where*:*

$$C_k = \frac{\sum_{j=1}^{N} s_{kj} m_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} m_j} = \frac{(S * M)_k}{\|S * M\|_1} \qquad (22)$$

Here, $s_{j,k}$ denotes the k-th coordinate of SC or journal j and m_j is the number of publications in SC or journal j. The numerator of Equation (22) is equal to the k-th element of $S * M$, the multiplication of the similarity matrix S and the column matrix of publications $M = \left(m_j\right)_j$. The denominator is the L_1-norm of the unnormalized vector.

8.3.2 Barycenter method

A barycenter is an entity's weighted average location on a map. More specifically, an entity's barycenter is the center of weight (Rousseau, 1989a, 1989b, 2008; Jin & Rousseau, 2001) of the WoS SCs or journals in which it has publications. The barycenter is defined as the point $C = (C_1, C_2)$, where

$$C_1 = \frac{\sum_{j=1}^{N} m_j L_{j,1}}{T} \; ; \; C_2 = \frac{\sum_{j=1}^{N} m_j L_{j,2}}{T} \tag{23}$$

Here, $L_{j,1}$ and $L_{j,2}$ are the horizontal and vertical coordinates of SC or journal j on the map, m_j is the number of publications in SC or journal j of the unit under investigation (panel member, research group), and $T = \sum_{j=1}^{N} m_j$ is the total number of publications of the entity. Note that, in case of WoS SCs, T is larger than the total number of publications as we use full counting: if a publication appears in a journal belonging to two categories, it will be counted twice.

Subsequently, we determine the Euclidean distance between the barycenters or the SAPVs of the panel members and individual research groups. The Euclidean distance between two vectors a = $(a_n)_{n=1,\ldots,k}$ and b = $(b_n)_{n=1,\ldots k}$ in $\mathbf{R}^k$, for any strictly positive integer k, is given as:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_k - b_k)^2} \tag{24}$$

In this chapter, we use formula (24) for k = 2 for the barycenter method and k = N for the SAPV method.

8.3.3 Weighted cosine similarity method

Finally, we consider a weighted similarity method (generalized cosine similarity). The WCS between panel member (PM) k and research group m is (Zhou et al., 2012):

$$\frac{\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)}{\sqrt{\left(\sum_{i=1}^{N} M_i^k \left(\sum_{j=1}^{N} M_j^k s_{ji}\right)\right) \cdot \left(\sum_{i=1}^{N} R_i^m \left(\sum_{j=1}^{N} R_j^m s_{ji}\right)\right)}}$$

$$= \frac{\left(M^k\right)^t * S * R^m}{\sqrt{\left(M^k\right)^t * S * M^k} \cdot \sqrt{\left(R^m\right)^t * S * R^m}} \tag{25}$$

The numerator is the matrix multiplication: $\left(M^k\right)^t * S * R^m$, where ^{t} denotes matrix transposition, S is the similarity matrix, M^k denotes the column matrix of publications of panel member k and R^m denotes the column matrix of publications of research group m. Similarly, the two products under the square root in the denominator are: $\left(M^k\right)^t * S * M^k$ and $\left(R^m\right)^t * S * R^m$. The result is the similarity between panel member k and research group m.

The Euclidean distances and similarity values are calculated for each panel member and each research group. The shorter the distance or the larger the similarity the closer the cognitive distance. In the ‘Results’ section, we present the cognitive distances in table form. All values are shown up to the third decimal. Cognitive distances are expressed as arbitrary units on a ratio scale (Egghe & Rousseau, 1990). Hence, we can compare them in terms like ‘x is twice as large as y’.

8.3.4 Bootstrapping and confidence intervals

We further calculated 95% confidence intervals (CIs) for each Euclidean distance (both between barycenters and SAPVs) and similarity (for WCS) by applying a bootstrapping approach (Efron & Tibshirani, 1998). If two CIs do not overlap, the difference between the distances is statistically significant at the 0.05 level. Although it is possible for overlapping CIs to have a statistically significant difference between the corresponding distances, the difference between the distances is less likely to have practical meaning. If the CI of two or more panel members overlaps, we treat them as interchangeable unless explicitly stated otherwise.

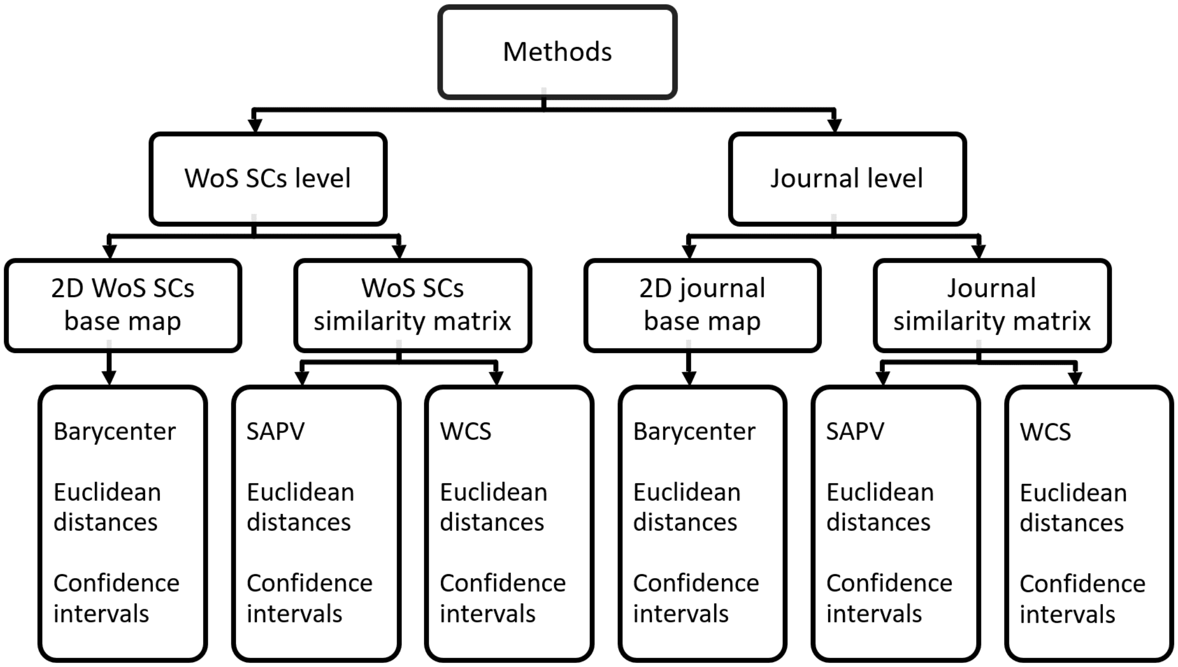


Figure 35: Main components of the six approaches at a glance

In applying the bootstrap for barycenters and SAPV distances, we generate 1000 independent bootstrap samples and for each sample calculate a bootstrap replication (barycenter or SAPV). Since we have a two-sample problem (distance between two entities), we calculate the distances between pairs of bootstrap replications, from which we obtain a CI using a bootstrap percentile approach (Efron & Tibshirani, 1998, Ch. 13). To apply the bootstrap to WCS, we again generated 1000 independent bootstrap samples. For each pair of samples, we calculated the similarity, from which we again obtain a CI using bootstrap percentiles. A more detailed explanation and implementation of our method is available on Github (Guns, 2016a, 2016b)

Figure 35 illustrates the main components of the six approaches at a glance. We have used two levels of aggregation – WoS SCs and journals. For each level of aggregation, there is a similarity matrix (N dimensions, with N the number of WoS SCs or journals) and a 2-dimensional base map derived from the similarity matrix. The SAPV and WCS methods operate at the level of N dimensions, whereas the barycenter method uses the 2-dimensional base map. We calculate

Euclidean distances between SAPVs (in N dimensions) or barycenters (in 2 dimensions) of entities, i.e. panel members and research groups. For the WCS case, we do not calculate a distance but a similarity between entities. Furthermore, a bootstrapping method is applied to determine confidence intervals for the distance or similarity between two entities.

8.3.5 Comparison of the approaches

To answer the first research question of this chapter, we calculate Spearman's rank-order correlation between the results/values of each pair of the six approaches. The distances/similarity values between the individual panel members and individual research groups have been included in the correlation calculation. Since the barycenter and SAPV approaches are distance-based rather than similarity-based, we determine the correlation using the distances between barycenters and between SAPVs, and the dissimilarity of individual research groups and panel members using a normalized weighted cosine dissimilarity = 1 – WCS which can more easily be compared with the other two. For the sake of simplicity, the results are shown under the heading "WCS method".

We created a heat map with hierarchical clustering based on the correlation results. For the clustering we used average linkage clustering with the UPGMA (unweighted pair group method with arithmetic mean) algorithm (Sokal & Michener, 1958).

The heat map is a two-dimensional representation of data where the values are represented by colors. It provides a visual summary of the results. The hierarchical clustering directly shows which approaches are more closely related.

To answer the second research question of this chapter – to what extent do the approaches agree in finding the panel member at the closest cognitive distance – we first explore whether the methods agree regarding the first ranked panel member ignoring the CIs overlap.

Table 51: Illustrations of procedures A and B

Research groups	Main assessor	Ranking (left to right) according to SAPV method in journals approach				
Procedure A						
BIOL-A	PM2	<u>**PM2**</u> 1	PM4	PM1	PM5	PM3
BIOL-B	PM5	PM1 0	PM4	PM2	<u>**PM5**</u>	PM3
BIOL-C	PM3	**PM5** 1	**PM1**	<u>**PM3**</u>	**PM2**	PM4
Procedure B						
BIOL-A	PM2	<u>**PM2**</u> 1	PM4	PM1	PM5	PM3
BIOL-B	PM5	PM1 0	PM4	PM2	<u>**PM5**</u>	PM3
BIOL-C	PM3	**PM5** 0.25	**PM1**	<u>**PM3**</u>	**PM2**	PM4

Concerning the third research question of this chapter, we recall that during the research evaluation exercises at the University of Antwerp, the panel chair of each panel decides which panel member should evaluate which research group (see Engels et al., 2013 for details). This panel member is referred to as the research group's main assessor. Lacking other information and practical considerations, it seems logical that in each case the closest PM is assigned to each group. Hence, we simply compare the closest PM with the main assessor.

For each approach, we ranked all the panel members in decreasing order of distance or in increasing order of similarity to the research group. We use two procedures (procedure A and procedure B, see Table 51) to compare the actual main assessor, assigned by the panel chair, to the panel member(s) recommended by our approaches. Procedure A focuses on how accurate the approaches are to identify the main assessor for each research group, whereas Procedure B focuses on how accurate the approaches are to uniquely identify the main assessor. For the sake of clarity, we underline and show in bold the main assessor in our approaches. We also show in bold the panel members whose confidence intervals overlap with the main assessor's.

In procedure A, we assign a score of 1 if the main assessor ranks first; a score of one is also assigned if the CI of the panel member who ranks first overlaps with the CI of the main assessor. If neither of these cases applies a zero score is assigned. For example, PM2 is the main assessor of BIOL-A and ranks first for the SAPV method applied to journals. There is no other panel member whose CI overlaps with PM2 in that case. Consequently, we assign a score of 1. Further, PM3 is the main assessor of BIOL-C; PM5 ranks first, but its CI overlaps with PM3's, hence also here a score of one is assigned. Considering BIOL-B we see that PM5 is the main assessor, but the CI of PM1, ranking first, does not overlap with PM5's CI. Hence, a zero score is assigned. In this procedure, even if the main assessor ranks last but the CI of the first ranked panel member overlaps with the CIs of the others, including the last ranked PM, a score of 1 is assigned.

For procedure B, we assign a score of 1 if the main assessor ranks first and has no overlapping CI with other PMs and zero otherwise. For example, PM2 is the main assessor of BIOL-A and ranks first in that case. There are no panel members whose CIs overlap with PM1. Therefore, a score of 1 is assigned to this case. On the other hand, PM5 is the main assessor of BIOL-B but PM1 ranks first in that case. Therefore, this case does not warrant any score. In case of overlapping CIs among the closest *n* PMs, one of which is the main assessor, we assign a score of 1/n. For example, PM3 is the main assessor of BIOL-C and ranks first in that case. The CIs of PM5, PM1 and PM2 overlap with PM3. Therefore, we assign a score of 1/4 = 0.25 in this case. The rationale here is that in this case, we randomly pick one of these *n* PMs, and hence we have a chance of 1/n of picking the main assessor.

The final score is the sum of all individual scores and ranges between zero and the total number of research groups in the department.

8.4 Results

For all six departments, the SAPVs of the panel members and individual research groups are calculated using the journal and WoS SCs similarity matrices by applying formula (22). We also calculate barycenters using the journal and WoS SC 2-dimensional base maps by applying formula (23). We determine the Euclidean distance between two SAPVs and two barycenters by applying formula (24).

Table 52: Euclidean distances between SAPVs of Biology panel members and individual research groups using the similarity matrix of WoS SCs

	BIOL- A	BIOL- B	BIOL-C	BIOL D	BIOL-E	BIOL -F	BIOL- G	BIOL- H	BIOL- I
PM1	0.057	<u>**0.052**</u>	<u>**0.059**</u>	<u>**0.036**</u>	0.046	0.076	**0.033**	<u>**0.066**</u>	0.071
PM2	<u>**0.017**</u>	0.091	0.073	0.089	<u>**0.023**</u>	0.071	**0.038**	0.104	0.073
PM3	0.062	0.129	0.082	0.114	0.083	<u>**0.015**</u>	0.080	0.139	<u>**0.023**</u>
PM4	0.048	**0.070**	0.079	0.067	0.041	0.082	<u>**0.028**</u>	**0.085**	0.081
PM5	0.039	0.111	**0.065**	0.105	0.058	0.046	0.061	0.120	0.052

Table 53: Euclidean distances between barycenters of Biology panel members and individual research groups using 2-dimensional base map of WoS SCs

	BIOL-A	BIOL-B	BIOL-C	BIOL-D	BIOL-E	BIOL-F	BIOL-G	BIOL-H	BIOL-I
PM1	0.344	<u>**0.075**</u>	<u>**0.075**</u>	<u>**0.093**</u>	0.282	0.201	0.200	<u>**0.123**</u>	**0.132**
PM2	<u>**0.042**</u>	0.409	0.317	0.444	<u>**0.088**</u>	0.317	0.165	0.454	0.353
PM3	**0.288**	0.263	0.223	0.275	0.274	<u>**0.016**</u>	0.195	0.310	<u>**0.065**</u>
PM4	0.217	0.191	**0.113**	0.220	**0.166**	0.143	<u>**0.078**</u>	0.242	0.130
PM5	0.109	0.324	0.241	0.353	**0.120**	0.17	**0.093**	0.374	0.215

Table 54: WCS values of the Biology panel members and individual research groups using the similarity matrix of WoS SCs

	BIOL- A	BIOL- B	BIOL-C	BIOL-D	BIOL-E	BIOL -F	BIOL- G	BIOL- H	BIOL- I
PM1	0.780	<u>**0.889**</u>	**0.674**	<u>**0.948**</u>	0.804	0.723	**0.886**	<u>**0.817**</u>	0.741
PM2	<u>**0.969**</u>	0.686	0.540	0.607	<u>**0.972**</u>	0.545	**0.910**	0.597	0.514
PM3	0.639	0.350	0.472	0.467	0.489	<u>**0.977**</u>	0.538	0.282	<u>**0.944**</u>
PM4	0.864	**0.773**	0.552	0.730	0.866	0.562	<u>**0.928**</u>	0.689	0.548
PM5	0.814	0.538	<u>**0.683**</u>	0.458	0.746	0.739	0.723	0.533	0.670

Table 55: WCS values of the Biology panel members and individual research groups using the similarity matrix of journals

	BIOL-A	BIOL-B	BIOL-C	BIOL-D	BIOL-E	BIOL-F	BIOL-G	BIOL-H	BIOL-I
PM1	0.609	<u>**0.567**</u>	0.267	<u>**0.890**</u>	0.558	0.729	0.726	<u>**0.613**</u>	0.728
PM2	<u>**0.816**</u>	**0.430**	0.199	0.453	<u>**0.896**</u>	0.33	<u>**0.824**</u>	0.429	0.314
PM3	0.325	0.270	0.243	0.566	0.248	<u>**0.940**</u>	0.330	0.267	<u>**0.900**</u>
PM4	0.643	**0.450**	0.174	0.516	0.629	0.308	**0.770**	**0.469**	0.303
PM5	0.610	0.341	<u>**0.461**</u>	0.321	0.463	0.427	0.463	0.374	0.366

Finally, WCS values are calculated using the journal and WoS SCs similarity matrices by applying formula (25). In this chapter, we show the results of the Biology department as an example. The result of the rest of the departments is shown in the appendix C.

Table 52 shows the results for the SAPV method using the WoS SC similarity matrix, Table 53 shows the result of the barycenter method using the 2-dimensional WoS SCs, Table 54 shows the result of the WCS method using the WoS SC similarity matrix and Table 55 shows the result of the WCS method using the journal similarity matrix. For the comparison between the approaches, we reuse the results of the SAPV and barycenter methods at the level of journals, that were previously obtained in the chapter VI.

Table 52 and Table 53 show, for each research group the panel member at the shortest distance. Similarly, Table 54 and Table 55 show, for each research group the panel member with the highest similarity. In both cases, the number in the row corresponding to this panel member is indicated in bold and underlined. For the former, distances whose confidence intervals overlap with that of the shortest distance are in bold (same column). For the latter, similarities whose confidence intervals overlap with that of the highest similarities are in bold (same column).

8.4.1 Correlation coefficients between six approaches

We explore how the six approaches are correlated. The heat map (Figure 36) represents the hierarchical clustering based on correlation coefficient between six approaches in the Biology department. Similar heat maps for other departments are available in the appendix D. The heat maps show that there are two clusters, except in the biology department, the 'barycenter' (2-dimensional) cluster and the 'similarity matrix cluster' (N-dimensional). We find that, in general the same methods at different levels of aggregation (journals and WoS SCs) are highly correlated.

At the WoS SC level of aggregation, the heat maps suggest that the correlation between the barycenter and the SAPV method is moderate to strong (range 0.61 to 0.73). A similar correlation (range r = 0.56 to 0.71) was found between the barycenter method and the WCS method, while the correlation between the SAPV and WCS methods is strong to very strong (r = 0.75 to 0.95).

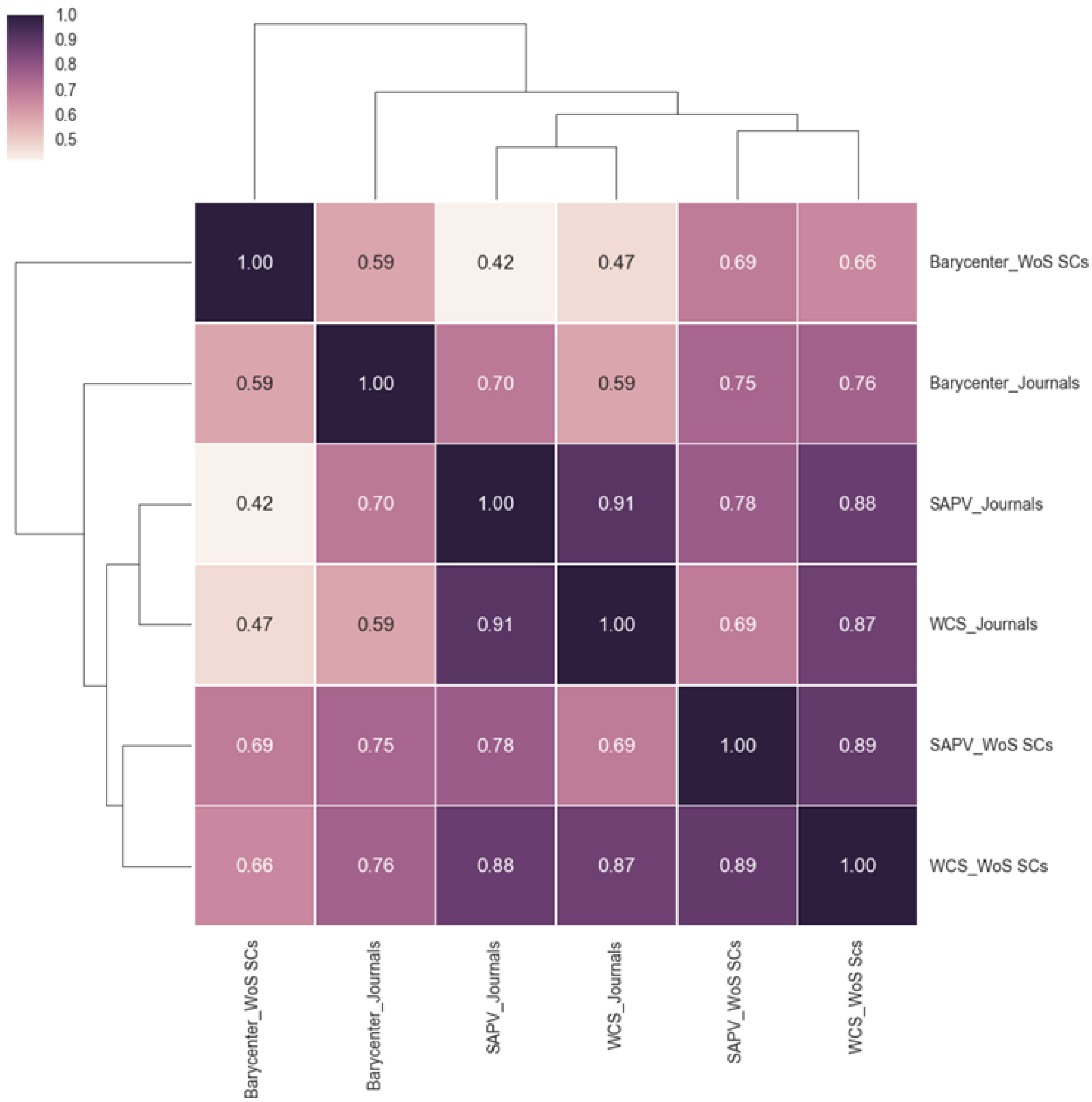


Figure 36: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Biology department

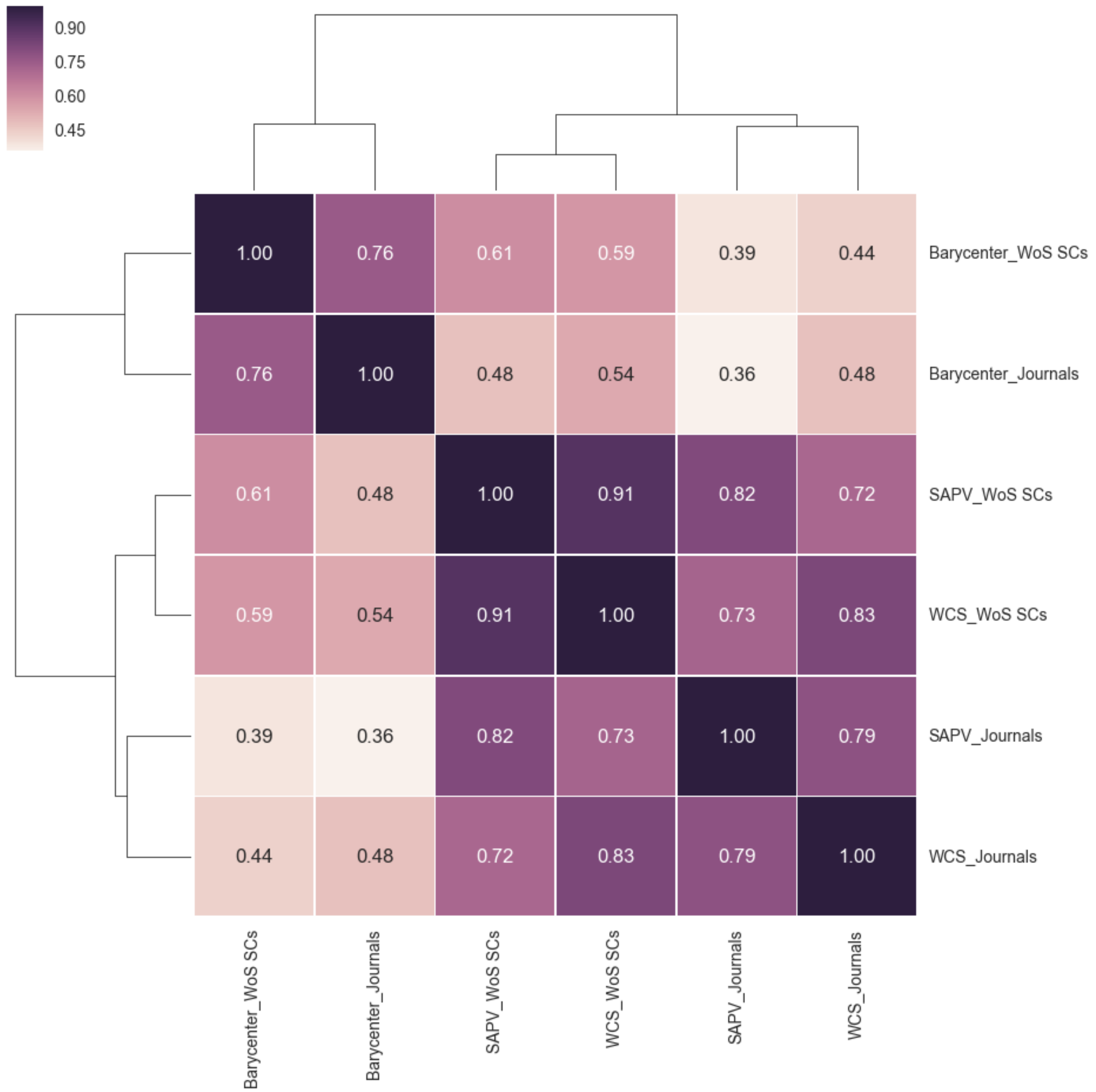


Figure 37: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the six departments

The correlation between the barycenter approaches at both levels of aggregation is strong (between 0.80 and 0.92) except for a moderate correlation (r = 0.59) for Biology. In addition, the correlation for SAPV is strong (range r = 0.78 to 0.93) as well, except for a moderate correlation (r = 0.68) in Pharmaceutical Sciences. Finally, the correlation for WSC is strong (r = 0.71 to 0.90) in all disciplines. In total, we find a strong correlation for 16 out of 18 cases.

Further, at the journal level of aggregation, the correlation between the barycenter and the SAPV method is moderate to strong (0.56 to 0.71), and between the barycenter and the WCS method is low to moderate (0.36 to 0.68) except for Veterinary Science where the correlation is strong (r = 0.80). Again, the correlation between the SAPV and the WCS methods is strong to very strong (r = 0.85 to 0.91).

We combined all the cognitive distances of the six approaches of the six departments and calculated the correlation between them. Figure 37 shows the heat map and the hierarchical clustering based on correlation coefficients between six approaches in the six departments. It also shows that there are two clusters: the 'barycenter' cluster and the 'similarity matrix' cluster. When the same method is used correlations between WoS SCs and journal level of aggregation are strong. However, the correlation between N-dimensional and 2-dimensional cases is low to moderate.

This finding suggests that different levels of aggregation tend to yield rather similar results. The influence of dimensionality (2-dimensions for barycenter versus N-dimensions for SAPV and WCS) is substantial, however. From here, we can conclude that the level of aggregation has a minor influence for determining cognitive distances in all the proposed six approaches, but the dimension matters.

8.4.2 Agreement between the approaches

To answer the second research question of this chapter, we explore whether the approaches agree regarding the panel member at the closest cognitive distance to each group. Note that, in this case, we ignore CIs. Without taking CIs into account, the analysis is stricter than if we take CIs into account. Table 56 shows the panel members with the closest cognitive distance (first ranked) to the research groups in the six approaches.

Table 56 shows that there is a clear difference between 2-dimensional and N-dimensional approaches. At the journal level of aggregation, the SAPV and WSC methods agree in all but five cases (91% match, research groups BIOL-G, PHYS-A, BIOM-I, CHEM-C, and PHAR-F) being exceptions.

Table 56: First ranked panel members for six approaches

	Journal level of aggregation			WoS SC level of aggregation		
Research groups	**SAPV**	**WCS**	**Barycenter**	**SAPV**	**WCS**	**Barycenter**
Biology department						
BIOL-A	PM2	PM2	PM2	PM2	PM2	PM2
BIOL-B	PM1	PM1	PM1	PM1	PM1	PM1
BIOL-C	PM5	PM5	PM2	PM1	PM5	PM1
BIOL-D	PM1	PM1	PM1	PM1	PM1	PM1
BIOL-E	PM2	PM2	PM2	PM2	PM2	PM2
BIOL-F	PM3	PM3	PM3	PM3	PM3	PM3
BIOL-G	PM4	PM2	PM1	PM4	PM4	PM4
BIOL-H	PM1	PM1	PM1	PM1	PM1	PM1
BIOL-I	PM3	PM3	PM3	PM3	PM3	PM3
Biomedical Sciences department						
BIOM-A	PM2	PM2	PM8	PM2	PM8	PM2
BIOM-B	PM5	PM5	PM2	PM5	PM5	PM5
BIOM-C	PM5	PM5	PM2	PM5	PM5	PM5
BIOM-D	PM7	PM7	PM1	PM7	PM7	PM5
BIOM-E	PM6	PM6	PM6	PM1	PM6	PM5
BIOM-F	PM1	PM1	PM1	PM1	PM1	PM5
BIOM-G	PM2	PM2	PM6	PM2	PM2	PM8
BIOM-H	PM6	PM6	PM6	PM6	PM6	PM6
BIOM-I	PM4	PM1	PM4	PM4	PM4	PM4
BIOM-J	PM2	PM2	PM8	PM2	PM2	PM3
BIOM-K	PM5	PM5	PM8	PM2	PM5	PM2
BIOM-L	PM5	PM5	PM8	PM2	PM2	PM8
BIOM-M	PM5	PM5	PM6	PM2	PM2	PM8
BIOM-N	PM5	PM5	PM6	PM8	PM2	PM8
BIOM-O	PM5	PM5	PM6	PM2	PM2	PM8
Chemistry department						
CHEM-A	PM 5	PM5	PM 7	PM 5	PM5	PM 5
CHEM-B	PM 5	PM5	PM 7	PM 5	PM5	PM 5
CHEM-C	PM 5	PM3	PM 5	PM 5	PM5	PM 5
CHEM-D	PM6	PM6	PM 6	PM 6	PM6	PM 6
CHEM-E	PM2	PM2	PM 2	PM 2	PM2	PM 2
CHEM- F	PM4	PM4	PM 2	PM 2	PM2	PM 2
CHEM-G	PM4	PM4	PM 3	PM 4	PM4	PM 3
CHEM-H	PM3	PM3	PM 5	PM 3	PM3	PM 3
CHEM-I	PM4	PM4	PM 3	PM 2	PM2	PM 3
CHEM-J	PM4	PM4	PM 4	PM 3	PM3	PM 4
CHEM-K	PM2	PM2	PM 4	PM 2	PM2	PM 2
CHEM-L	PM5	PM5	PM 5	PM 5	PM5	PM 5

	Journal level of aggregation			WoS SC level of aggregation		
Research groups	**SAPV**	**WCS**	**Barycenter**	**SAPV**	**WCS**	**Barycenter**
Pharmaceuticals Sciences department						
PHAR-A	PM5	PM5	PM1	PM5	PM5	PM1
PHAR-B	PM4	PM4	PM5	PM2	PM2	PM4
PHAR-C	PM2	PM2	PM2	PM2	PM2	PM2
PHAR-D	PM3	PM3	PM3	PM2	PM2	PM4
PHAR-E	PM2	PM2	PM2	PM2	PM2	PM2
PHAR-F	PM4	PM1	PM5	PM4	PM5	PM4
PHAR-G	PM5	PM5	PM5	PM5	PM5	PM3
PHAR-H	PM4	PM4	PM3	PM2	PM2	PM2
PHAR-I	PM2	PM2	PM2	PM2	PM2	PM2
PHAR-J	PM2	PM2	PM2	PM2	PM2	PM2
Physics department						
PHYS-A	PM4	PM3	PM4	PM 4	PM3	PM 4
PHYS-B	PM6	PM6	PM6	PM 6	PM6	PM 6
PHYS-C	PM5	PM5	PM5	PM 5	PM5	PM 5
PHYS-D	PM1	PM1	PM1	PM 1	PM1	PM 1
PHYS-E	PM6	PM6	PM5	PM 6	PM6	PM 5
PHYS-F	PM6	PM6	PM1	PM 3	PM6	PM 3
PHYS-G	PM4	PM4	PM4	PM 4	PM4	PM 4
PHYS-H	PM6	PM6	PM2	PM 6	PM6	PM 6
PHYS-I	PM5	PM5	PM4	PM 4	PM5	PM 4
Veterinary department						
VETE-A	PM2	PM2	PM2	PM2	PM2	PM2
VETE-B	PM2	PM2	PM2	PM2	PM2	PM2
VETE-C	PM1	PM1	PM1	PM1	PM1	PM3

Furthermore, the barycenter method agrees in 30 cases (52%) with the SAPV method and in 27 cases (47%) with the WCS method. Barycenter, SAPV and WCS methods agree in 27 cases (47%). Similarly, at the WoS SCs level of aggregation, the SAPV and the WCS methods agree in 49 cases (84%). The barycenter method agrees in 41 cases (71%) with SAPV and 34 cases (58%) with WCS. Barycenter, the SAPV and the WCS methods agree in 33 cases (57%).

We also explore whether the same method agrees at both levels of aggregation. Table 56 shows that the SAPV method agrees in 52 cases (90% matches), the WCS method in 53 cases (91% matches), and the barycenter method in 34 cases (59% matches). As the SAPV and WCS methods are in N dimensions, we find that they agree in 47 out of 58 cases (81%) across both levels of aggregation. Hence, we conclude that for finding the first ranked panel members the

SAPV and WCS approaches agree in most of the cases at both levels of aggregation, while the barycenter approaches yield considerably different results.

8.4.3 Finding previously assigned panel members

We have panel members' assignment data for all departments, with the exception of Pharmaceutical Sciences. Therefore, the analysis is based on the other five departments. As we stated in the method section, we calculated the total score according to procedure A. Table 89 and Table 90 (see appendix E) show the calculations of the Biology department panel ranked positions and highlighted the panel members whose CIs overlap with shortest distance panel members in all the six approaches. Similar tables for other departments are shown in the appendix E. Table 57 summarizes the outcomes of all the six departments.

In Table 57, the higher the score, the better the approach replicates the original panel member assignment to research groups. With the exception of Biomedical Sciences, the barycenter method (at both levels of aggregation) scores the same as or higher than the other two methods.

This is also reflected in the total score. Contrary to what one might expect, the SAPV method scores higher at the level of WoS SCs than at the level of journals. For the WCS method, the level of aggregation does not make a difference. It is evident that the barycenter method performs better than the other two in terms of finding the main assessor for each research group.

Table 57: The distribution of total scores for six approaches according to procedure A

		Journal level of aggregation			WoS SCs level of aggregation		
Department	**No. of groups**	**SAPV**	**WCS**	**Barycenter**	**SAPV**	**WCS**	**Barycenter**
Biology	9	6	5	6	5	5	6
Biomedical Sciences	15	9	8	7	9	8	7
Chemistry	12	5	5	8	7	3	8
Physics	9	4	6	8	6	7	7
Veterinary Sciences	3	2	2	2	3	3	3
Total:	**48**	**26**	**26**	**31**	**30**	**26**	**31**

Table 58: The distribution of total scores for six approaches according to procedure B

Department	No. of groups	Journal level of aggregation			WoS SCs level of aggregation		
		SAPV	WCS	Barycenter	SAPV	WCS	Barycenter
Biology	9	4.08	4.50	4.08	4.33	4.33	3.83
Biomedical Sciences	15	3.84	3.93	2.11	2.97	3.04	1.70
Chemistry	12	2.99	2.49	2.52	3.41	1.75	2.52
Physics	9	3.50	4.83	4.95	3.41	4.40	3.33
Veterinary Sciences	3	2.00	1.33	0.66	2.00	1.88	0.83
Total:	**48**	**16.41**	**17.08**	**14.32**	**16.12**	**15.40**	**12.21**

We also calculated the total score according to procedure B. Table 99 and Table 100 (see appendix E) show the calculations of the Biology department panel ranked positions and highlighted the panel members whose confidence intervals overlap with shortest distance panel members in all the six approaches. Similar tables for other departments are shown in the appendix E. Table 58 summarizes the outcomes of all the six departments. In Table 58, the higher the score, the better the approach replicates the original panel member assignment to research groups. Table 58 shows that the journal level analysis scores higher than the WoS SCs level. Moreover, the N-dimensional approaches score higher than 2-dimension approaches at both levels of analysis. However, at both levels, the barycenter method always scores lower than the other methods. This result is what one expects theoretically: using journals is a more refined method than using WoS SCs, and performing calculations in N dimensions yields a more precise outcome than performing calculations in two dimensions.

From the two procedures, we can conclude that in our case studies, the barycenter methods are, generally speaking, better able to find the main assessor. However, the methods based on barycenters are also less discriminatory, in that they tend to have more overlapping CIs. Simultaneously, all the methods score higher at the journal level than at the WoS SC level in uniquely identifying the main assessor. In addition, the SAPV and WCS methods score higher than the barycenter methods at both levels of aggregation.

8.5 Discussion

Our proposed approaches quantify the shortest cognitive distance between a research group and PMs. Simultaneously, they can be used to rank the PMs based on cognitive distances. If the CIs of some PMs overlap, the differences between them are relatively small and work as an indicator to assign the next potential PM to evaluate a research group. The methods can be used *ex ante* to inform the process during which potential PMs are identified and invited, as well as while the review process takes place (in view of division of labor within a panel) or ex post (to assess the appropriateness of a panel). The quantitative methods can support and inform experts during panel composition, similar to how scientometric indicators can support and inform peer review-based evaluations themselves.

If any of the proposed approaches totally agrees with the previous assignment of a main assessor, we may state that the panel chair or the research affairs department has rightly identified the expertise match between a panel member and research group. However, that is not the case in any of the six approaches. The major reason is that the panel assignment was based on a qualitative judgement, whereas our methods use a quantitative approach based on the publication portfolio of panel and research groups. Panel members and panel chairs are chosen following the suggestions of research groups and the research affairs department. Panel chairs have the list of panel members and their curricula vitae, and the research activity profile of the research groups as a means to come to a decision. The chair needs to reach a decision to assign a panel member to one or two research groups based on the match of the expertise with the research group. As there is no formal method to match expertise, the panel chairs distributed the workload based on their own tacit knowledge. In all the cases except for Veterinary Sciences, there are more research groups than panel members. Hence, one panel member can be close to multiple research groups, but due to practical considerations of workload distribution, the panel chair may not assign the panel member to more than two research groups. Therefore, another panel member who is intellectually further from a particular research group may be assigned to that research group for purely practical reasons.

We observed that some panel members are never the closest to any research group. This is for example the case for PM1 in the Chemistry panel and PM4 in the Veterinary Sciences panel. This raises the question why these members were included in a panel. We have observed that generally panel members are not assigned to more than two research groups, with two exceptions: PM8 of the Biomedical Sciences panel was assigned to three research groups and PM3 of the veterinary panel was not assigned to any research group. Our approaches can help to inform the assignment of panel members by quantifying the cognitive distance between individual panel members and research groups. The proposed approaches rank the panel members based on cognitive distances and indicate the panel members who are at a comparable distance from the research group through the overlap of CIs. The overlap of CIs of the shortest cognitive distance panel members with other potential panel members helps to assign next potential panel members to a research group. Even if a research group has no publications in the WoS SCs or journals where the panel has publications, the panel might be able to evaluate the research group (discussed in chapter V and VI).

Asking research group members and/or panel members for their personal opinions might be an alternative method to determine panel members that are cognitively closest to a given research group. As the research evaluations mentioned in this article were done three to eight to years ago, this was not practically possible for our case studies.

Knowing cognitive distances between entities is an important aspect in panel composition, but in itself it is not sufficient. For instance, our approaches do not consider the aspect of cohesion (Casey-Campbell & Martens, 2009). Cohesion is the common bond that drives colleagues to remain together and to cooperate (Salas, Grossman, Hughes, & Coultas, 2015). In some cases, a panel member could be included in a panel for other reasons than their specific research expertise in relation to the research groups. For example, there might be a selection of a panel chair based on his/her expertise in the discipline in general (e.g. PM1 in the Chemistry panel). S/he may not be the closest panel member based on publication profiles to any of the research groups that will be evaluated. A reason could be that the panel member plays an important role for the cohesion of the panel. Hence, cohesion may be an indicator in expert panel composition, to be applied in a step-wise manner, once the chair has been selected.

We note that the 2-dimensional base maps at the level of WoS SCs as well as journals are publicly available. In addition, a similarity matrix of WoS SCs is also readily available. A journal similarity matrix, on the other hand, is not available openly. The JCR of 2014 contain 11,149 journals in the SCIE and SSCI (Leydesdorff, Bornmann, & Zhou, 2016). This constitutes an increase by 474 journals compared to the journal similarity matrix we used in this article. Since evaluations are retrospective, it is not necessary to always have the most recent journal similarity matrix. Moreover, journals are not static entities and may undergo name changes over time, split in different new journals, or two or more journals can be merged together (see chapter VI section 6.3.1 for details). However, any changes to the journal similarity matrix will have – at least in theory - a direct impact on the cognitive distances obtained. It is a topic for further investigation to find out to what extent the cognitive distances differ and the CIs overlap if a different base map or similarity matrix (based on different years) are used for the same panel and research groups.

We have used similarity matrices and base maps derived from them based on data available during the construction of the matrices. If the similarity matrix changed over the years, and we keep the same panel and research groups publication data, this might result in different cognitive distances. Moreover, if we use a different similarity matrix (for example, based on Scopus data) and retrieve the same panel and research groups' data, we can expect different results as well, because the similarity matrix and the data will not be the same. An interesting follow-up investigation could therefore be based on Scopus data (e.g., Leydesdorff, de Moya-Anegón, & Guerrero-Bote, 2010; Leydesdorff, Moya-Anegón, & Guerrero-Bote, 2015). Hence, although there is a practical stability problem, the methods we introduced have general applicability.

8.6 Conclusion

The SAPV and WCS methods use the N-dimensional similarity matrix of journals or WoS SCs while the barycenter method uses a 2-dimensional base map derived from the respective similarity matrices. The approaches proposed in this chapter allow the concerned authority to assess how well the expertise of panel members corresponds with the research interests of the groups to be evaluated (Rahman et al., 2015, 2016; Rousseau et al., 2017). In this chapter, we focused on the question which of the approaches best reflect cognitive distance, how much influence the level of aggregation (journals and WoS SCs) plays, and how much the dimensionality matters. The results show that the level of aggregation (journals and WoS SCs) has only minor influence for determining cognitive distances in all the proposed six approaches, whereas the influence of the number of dimensions is substantial. The results also show that the number of dimensions plays a role in the case of identifying shortest cognitive distance. While the SAPV and WCS methods agree at both levels of aggregation, the barycenter method yields different results to identify the panel members at the shortest cognitive distance.

We find that the barycenter method scores highest at both levels of aggregation to identify the previously assigned main assessor. This finding is aligned with our earlier finding that the barycenter method has less discriminatory power than the other methods at WoS SCs level of aggregation (discussed in chapter VII). When it comes to uniquely identifying the main assessor, all methods score better at the journal level than at the WoS SC level.

The proposed approaches can be tested in any future scenario where X panel members need to be chosen out of N candidates. Panel composition based on different approaches can then be matched with the opinion of the panel chair. Concrete differences can then be discussed, leading to a better panel composition. In addition, the opinion of the respective panel members can be taken into account beforehand, so that the main assessor is indeed the most qualified person for the job.

Chapter IX: Conclusion

In this conclusion, we first give a summary of the findings. Subsequently, we will discuss the salient features of the proposed methods and their possible implications for practice. We conclude with an overview of the limitations of the research and suggestions for further study.

9.1 Summary of the main findings

We formulated two main research questions (with sub-questions) in the first chapter of this thesis. Chapter I, section 1.2.2 provides an overview of the questions as well as the relevant chapters where each (sub-)question is answered in full detail. Here, we summarize the main findings and answers to the research questions.

The thesis focuses on the problem of composing an expert panel in such a way that the panel members' expertise is congruent with the research groups' expertise. More specifically, our work has focused on the important problem of finding appropriate informetric methods to gauge the cognitive distance between panels and research groups. We recall that at the start of our research project no such methods were available.

Our first main research question was, **how can we measure cognitive distance between two entities using publication data, especially between an expert panel and the research groups under evaluation?**

As a preliminary exploration, in Chapter IV we focused on correlation coefficients and cosine similarity to measure the strength and direction (positive/negative) of the association between the publication profiles of research groups and panels. We determined the correlation between the publication outputs of two entities using Pearson's correlation coefficient, Spearman's rank correlation coefficient and top-down correlation, and use cosine similarity to determine the similarity.

The correlation coefficients and cosine similarity are strong to moderate at the level of WoS SCs, yet low to negative at the level of journals. This difference can be explained, at least partially, by the higher number of journals compared to WoS SCs: other things being equal, the probability that two publications belong to the same set (SC or journal) is much higher for SCs than it is for journals. We argue that such correlations and similarity measures are insufficient, as they do not take into account the *relatedness* of WoS SCs or journals.

If a panel member has many publications in WoS SCs or journals that are closely related to those a research group has published in, his/her expertise may still be relevant to evaluating the group, even if s/he has no publications in exactly the same SCs or journals. Consequently, a comparison of publication profiles that does not take WoS SC similarity or journal similarity into account might yield distorted results. Therefore, a method that does take similarity into account is necessary to identify the match between the panel and research groups. In the following text, we summarized the main finding of the (sub-) questions:

The first sub-question was:

i) **How can we visualize the expertise of two entities (e.g., a research group and a panel) using publication data?**

To answer this sub-question, we explore the usefulness of overlay mapping to gauge cognitive distance between the expertise of two entities. We use a base map of science based on WoS SCs and created overlays that visually represent the publication profiles of panels and research groups. Based on the overlay maps, one can visually compare and estimate the cognitive distance between the research groups and panel (discussed in chapter V). We also use a base map of science at the aggregation level of journals to visualize the expertise of entities. The corresponding maps are shown in the technical reports that are available online (see section 3.2.8 of chapter III). These overlay mapping techniques provide an answer to the first sub-question. Based on the overlay maps, we can visually compare individual research groups, and entire department publication profiles with the respective panel profiles. While this comparison provides worthwhile information to get an overall view, it is less suitable for, e.g., comparing which of the potential panel members would be the best fit to evaluate the research group.

The second sub-question was:

ii) How can one quantify the cognitive distances (overlap of expertise) between two entities (e.g., a research group and a panel) using the WoS SCs to which their publications belong?

To answer the second sub-question, we develop methods to gauge cognitive distance or proximity between two entities. First, we calculate the Euclidean distance between L_1-normalized arrays (publication vectors) of each panel member and each research group. In this way, a perfect match between the entities would mean a distance of zero. We refer to this method as a benchmark in N-dimensions. At the initial stage, we found that correlation and cosine similarity are insufficient since they do not consider the relatedness of WoS SCs or journals (discussed in chapter IV). Similarly, the benchmark method does not consider the similarity or relatedness of WoS SCs or journals. Comparing the benchmark results with those of our proposed SAPV and WCS methods (which do consider the similarity of, in this case, WoS SCs), we find that the correlations between the benchmark results with other methods are at best moderate. As was logically expected, we observed that the methods that consider the relatedness perform better than the benchmark (see details in chapter VII).

Second, we apply the barycenter method. We use a two-dimensional global map of science based on WoS SCs. This map is the same base map that we used to create the overlay maps to answer the first sub-question. In our case, the barycenter method calculates the center of weight of the SCs in which the research groups and/or the panel members have publications. Here, a SC's weight is the research groups and/or the panel members' number of publications. We calculate the barycenter for each individual panel member and research group, for the expert panel as a whole and for the combined research groups of a discipline. Subsequently, we calculate the Euclidean distances between the barycenters. These distances can be compared in terms of percentage or fraction. We use Euclidean distances between barycenters as an indicator of cognitive distance between two entities. The shorter the cognitive distance between the entities the better the match.

In addition, we projected the barycenters on the global map of science to visualize the cognitive distances between the entities. The two-dimensional barycenter method is well compatible with the overlay mapping techniques (discussed in chapter V).

The barycenter method can be applied in any strictly positive number of dimensions smaller than or equal to N. The two-dimensional barycenter method has the advantage that it can be straightforwardly visualized, but there is no inherent reason to prefer two dimensions over any other number $\leq N$. More generally, one may ask what the influence of the number of dimensions on the results of the barycenter method is. We explored this factor by calculating three-dimensional barycenters (and corresponding distances) and comparing the results with those of two-dimensional barycenters. We found that there is a strong correlation between the two and three-dimensional barycenter method and both methods yield very similar results. For this reason, we considered only the two-dimensional barycenter method in subsequent chapters.

Furthermore, we introduced the SAPV method. We used a matrix of similarity values between the WoS SCs. While a regular publication vector contains publication counts per SCs, in an SAPV these counts are adapted to account for the similarity between the SCs. An SAPV is the result of the multiplication of the similarity matrix with the publication vector of the research groups or panel members. The Euclidean distance between SAPVs is again an indicator of cognitive distance between two entities (discussed in chapter VII).

Moreover, we proposed the use of the WCS method. Like the SAPV method, its input consists of publication vectors of the entities and a similarity matrix. A similarity-weighted generalization of regular cosine similarity, this method yields a similarity value between publication vectors. We calculated this value for each combination of panel member and research group. We consider the similarity as the opposite of distance. The higher the similarity, the better the match – the closer the cognitive distance (discussed in chapter VII).

In summary, in this thesis five methods are developed: the benchmark, two methods using barycenters (in two and three dimensions), a fourth method based on SAPVs and a fifth one using WCS. The benchmark and the last two methods are applied in N dimensions, where N denotes the total number of SCs. As an answer to the second sub-research question, we consider the benchmark method insufficient since it does not consider similarity between SCs. We also exclude the three-dimensional barycenter method for two reasons. First, the two-dimensional and three-dimensional barycenter yield almost the same result. Therefore, it makes sense to just look at the simpler of the two, the two-dimensional barycenter. Second, the two-dimensional barycenter method is well compatible with visualizing the barycenter's location (discussed in chapter VII). We conclude that the two-dimensional barycenter method applied to WoS SC base map, and the SAPV and the WCS methods applied to the similarity matrix of WoS SCs provide an answer to the second sub-question (discussed in chapter V and VII).

The third sub-question was:

iii) How can one quantify the cognitive distances between two entities using the journals in which they have published?

It is open for discussion to what extent publishing in the same WoS SC can be regarded as a sign of cognitive proximity, as one SC may comprise a wide array of different subfields and topics. At this point, we propose a journal level of aggregation, as most journals tend to cover more closely related subfields and topics.

To answer the third sub-question, we reapply the barycenter, SAPV, and WCS methods. This time instead of assigning publications to WoS SCs, publications were assigned to the journal in which they were published. Instead of a WoS SC map a journal map is used to calculate barycenters. Likewise, instead of a WoS SC similarity matrix, a journal similarity matrix is used. The Euclidean distances between barycenters or SAPVs are again treated as indicators of cognitive distance (discussed in chapter VI). The results of the WCS method at the journal level are discussed in the technical reports (see section 3.2.8 of chapter III). Thus, the application of the barycenter, SAPV, and WCS methods at the journal level answer the third sub-question.

It is mentionable that the journals are dynamic entities in respect of title changes, shortened or extended, merging two or more journals together, or one split a journal into two or more journals. Moreover, a journal can be excluded from or added to the WoS or may be discontinued (see section 6.3.1 of chapter VI). On the other hand, WoS SCs are mostly static entities and are more stable than journals.

The fourth sub-question was:

iv) How can one estimate the uncertainty inherent to these cognitive distances?

If the differences in distance or similarity between a research group and two or more panel members are small, it bears little meaning to claim that one panel member is a better choice than another. Moreover, the distance or similarity between the panel member and research groups is determined based on the journals and WoS SCs in which the research groups and panel members have published. However, there are certain factors like publication processing time, number of volumes of journal, required time for indexing, that may influence the results. Therefore, small differences in distances or similarity are likely unstable – e.g., a similar exercise carried out using another database or other publication years might reverse the ranking – and do not reflect any real difference between the units involved. To identify meaningful differences in the distance or similarity we used a bootstrapping method, leading to 95% confidence intervals (CIs) for distances (benchmark, barycenter two-dimensional and three-dimensional, and SAPV method) and similarities (WCS). With bootstrapping, we use our sample data to generate multiple random samples, based on which we estimate the CIs of the distances. If two CIs do not overlap, the difference between the distances is statistically significant at the 0.05 level. If the CIs of two or more panel members overlap, we treat the panel members as interchangeable. The confidence intervals are thus an operationalization of the inherent uncertainty related to our distances, and provide an answer to the fourth sub-question (discussed in chapter VI and VII).

Our second main research question was, **how do the proposed approaches relate?**

The first sub-question was:

i) **What are the correlations between the different approaches? Which aspect (method vs level of aggregation) has the largest influence on the correlation?**

Over the course of this PhD project, we have proposed the barycenter, SAPV and WCS methods. The barycenter method is two-dimensional (based on maps of science), whereas the SAPV and WCS methods are N-dimensional (based on a full similarity matrix). We used these three methods at two levels of aggregation – WoS SCs and journals. This leads to six different approaches, all of which are based on the publication profile of research groups and panel. We systematically compare how these six approaches relate.

In order to answer the first sub-question, we compared the six approaches to identify the relative influence of the level of aggregation and the number of dimensions. The number of dimensions refers to the two-dimensional barycenter method and the N-dimensional SAPV and WCS methods. We calculated Spearman's rank-order correlation between the results of each pair of approaches, using the distances/similarity values between the individual panel members and individual research groups. The results showed that the same methods at different levels of aggregation (journals and WoS SCs) are highly correlated in most of the cases. Moreover, the level of aggregation has minor influence on determining cognitive distances, but dimensionality (two-dimensions versus N-dimensions) has a greater influence.

The second sub-question was:

ii) **To what extent do the approaches agree in matching the panel member at the closest cognitive distance from a research group?**

In order to answer the second sub-question, we explored whether or not all the methods indicate the same panel member as the one at the shortest cognitive distance from a research group. We ranked the panel members according to the closest cognitive distance to each group and compare only the first ranked panel member in each approach. In this case we did not consider CIs. The results showed that the SAPV and WCS methods agree in most cases at both levels of aggregation on which panel member has the closest cognitive distance to the group to be evaluated, whereas the barycenter approaches often differ.

The third sub-question was:

iii) How accurate are the approaches in matching the main assessor for each research group? How accurate are they to *uniquely* match the main assessor?

In order to answer the third sub-question, we consider the panel chair of each panel has assigned the closest panel member to each research group during the research evaluation exercises at the University of Antwerp. This panel member is referred to as the research group's main assessor.

Here, we compared the closest panel member in our approaches with the main assessor. We use two procedures to compare the actual main assessor, assigned by the panel chair, to the panel member(s) recommended by our approaches. The first procedure identifies how accurate the approaches are to identify the main assessor for each research group while the second procedure identifies how accurate the approaches are to uniquely identify the main assessor. In both the procedures, we consider the CIs overlap with the main assessor's. We find that the barycenter method performs better than the two other methods in terms of finding the main assessor for each research group and suggests more potential evaluators, whereas SAPV and WCS are more precise. At the same time, to uniquely identifying the main assessor all the methods perform better at the journal level of aggregation than the WoS SCs level of aggregation. In addition, the SAPV and WCS methods perform better than the barycenter method at both levels of aggregation to uniquely identifying the main assessor.

The above findings of the sub-questions provide an answer to the second research question (discussed in chapter VIII).

9.2 Policy recommendations

This thesis and the literature review indicate that in research evaluation it is of paramount importance that the panel members are considered trustworthy experts who are able to provide valuable, relevant recommendations and suggestions that can lead to improved research quality. Therefore, it is vital to assign the most appropriate expert to the relevant research group in a discipline specific research evaluation. The common process of assignment of panel members to research groups without any systematic method can lead to evaluations where the matching of expertise between the evaluators and the units under assessment is sub-optimal. The policy makers or the concerned authority involved in panel member assignment to research groups can consider cognitive distance between the research groups and panel members based on their publication profile to solve the panel member assignment problem. As our proposed approaches quantify the degree to which the expertise of the panel members matches with the expertise of the research groups these approaches can help to improve the practice of expert panel composition, i.e. ex-ante, in preparation of a research evaluation. The approaches can also help to assign the appropriate panel member to a research group, i.e. once a panel has been composed.

For example, in a given situation X panel members need to be chosen out of N candidate panel members to evaluate Y research groups. The proposed approaches allow to calculate the match between a research group and a potential panel member in terms of cognitive distance (i.e. the shorter the distance or the higher the similarity the better the match). The approaches rank the panel members based on cognitive distances and indicate the panel members who are at a comparable distance from the research group through the overlap of CIs. The overlap of CIs of the shortest cognitive distance panel members with other potential panel members solves the problem of assigning next potential panel members to a research group. A research group can be far away from the panel as a whole. However, one or more individual panel members may have sufficient expertise to evaluate a research group as indicated by publications in closely related WoS SCs or similar journals (discussed in Chapter V and VI). The approaches allow the concerned authority to assess and improve the composition of the panel in terms of cognitive distance, e.g. by virtually replacing one or more potential panel members and comparing the relative contribution of each potential panel member to the panel fit as a whole, by observing the changes to the distance between the panel's and the groups' (discussed in chapter V – VII). In addition, the proposed approaches can also identify the cognitive distance and similarity between the combined groups and the panel.

Which approach should be used? The findings showed that all the proposed methods can be used to quantify the cognitive distances and CIs can be calculated. However, the level of aggregation and method should be the choice of the concerned authority based on their respective policy. The classification in WoS SCs has received criticism for grouping a wide array of different subfields and topics while journals usually cover more closely related subfields and topics (discussed in chapter V and VI). We note that the two-dimensional base maps at the level of WoS SCs as well as journals are publicly available. In addition, a similarity matrix of WoS SCs is also readily available. A journal similarity matrix, on the other hand, is not publicly available. In addition, the journals are not static entities and change over time. Therefore, manual work to track these changes may be needed (discussed in chapter VI). Hence, the WoS SC level of aggregation can be seen as a convenience approach but the journal level approach is more refined.

The SAPV and WCS methods are more precise as both used the similarity matrix while the barycenter method used the two-dimensional map derived from the similarity matrix. In deriving a two-dimensional map from an N-dimensional similarity matrix, some loss of information is inevitable. Hence, the cognitive distances obtained with the barycenter method are distorted to the number of dimensions while distances stemming from the SAPV and WCS methods are more meaningful based on mathematical logic (discussed in chapter VII). However, the barycenter method has advantages as it can be visualized. In the practice of research evaluation, projecting the location of panel members and research group on a map and at the same time looking at the corresponding distances can help the people responsible for the evaluation to have an immediate impression and grasp of the situation.

Our proposed approaches are not dependent on any specific map or matrices. The methods can be applied to any local map (or subject or topic specific map) or similarity matrix for their own purposes. We are aware that the proposed methods are best suited to evaluations that cover a longer period with a larger set of publications and may not be suitable for assessing individuals, grants, individual research projects, or the process of selecting potential reviewers for submitted manuscripts.

9.3 Limitations of this study

A first limitation is that we use the data, similarity matrix, and base map derived from the WoS only. WoS is interdisciplinary and covers all scientific areas, but it only covers what it considers to be "best" journals and concentrates on English language ones. We have used the similarity matrix and base map based on a certain year of data (JCR 2011) that was available at the start of our study. If we used a different similarity matrix (for example, based on Scopus data) and retrieved the same panel and research groups' data, we can expect the results to be different as the similarity matrix and the data will not be the same. However, the methods will remain the same in any situation.

A second limitation we want to mention is that our proposed methods start from journal article profiles of panel members and research groups, assuming that these publication profiles adequately represent their expertise or research interest. Therefore, our proposed approaches might be less acceptable in some fields, such as engineering or computer sciences, where core conferences are important publication outlets for original research (Rahm, 2008), or the social sciences and humanities where a large part of the total output occurs as book publications (Engels, Ossenblok, & Spruyt, 2012). In addition, patents, designs, software, databases and other types of non-journal research outputs are all important markers of expertise that are thus far not considered in the proposed approaches.

A third limitation is that there is no *a priori* answer to the question "what distance between panel and research groups is acceptable for evaluation purposes". This is because the context, objectives and practical setting of an expert panel evaluation may all play a role. Hence, this cannot be decided on beforehand. However, 'the shorter the distances the better the fit of the expert panel' can be suggested as a rule of thumb. At this point, we cannot make any claim regarding acceptable or preferable distances, and hence certainly not about the link between distances and the 'quality' of evaluations.

A fourth limitation is that our study includes only six departments (six evaluation exercises) from one university. Moreover, all departments are in exact and biomedical fields. It is as of yet an open question to what extent similar results would be obtained in other organizational contexts and in other disciplines.

9.4 Suggestions for further study

In this thesis, the departments are in exact and biomedical fields. A subsequent analysis can be done with the departments that belong to social science and humanities faculties. One should be aware that both Scopus and WoS focus mostly on journal contributions rather than monographs, edited volumes or reports that are considered as important channels of communication (Lancho-Barrantes, Guerrero-Bote, & Moya-Anegón, 2010b). In addition, the publication patterns (Engels et al., 2012; Hicks, 2004; Sivertsen, 2016) and their coverage in Scopus and WoS (Rahman et al., 2017) of the social science and humanities should be taken into account.

Since the similarity matrix changes each year, results obtained using maps or matrices based on data from other years will lead to different results. The JCR of 2014 contain 11,149 journals in the SCIE and SSCI (Leydesdorff, Bornmann, et al., 2016). This constitutes an increase by 474 journals compared to the 2011 journal similarity matrix we used in this thesis. Since evaluations are retrospective, it is not necessary to always have the most recent journal similarity matrix. Moreover, journals are not static entities and may undergo name changes over time, split in different new journals, or two or more journals can be merged together (see Chapter VI for details). However, any changes to the journal similarity matrix will have – at least in theory - a direct impact on the cognitive distances obtained. At this moment, we do not know how much such year-to-year changes and differences in projection methods affect the results of our methods.

A subsequent analysis can be done with the similarity matrix based on a different year. It is a topic for further investigation to find out to what extent the cognitive distances differ and the CIs overlap if a different base map or similarity matrix (based on different years) are used for the same panel and research groups. The changes in similarity matrix may be occurring due to two reasons: first, the changes in similarity values between entities (e.g. WoS SCs or journals), and, second, due to changes of entities in the matrix. The first one may take place due to citation traffic change while both the first and the latter may be caused by events such as the reorganization of WoS SCs or adding or deleting journals. How much does a different similarity matrix/map affect the results? One possible approach would be to run simulation experiment in which a large number of fictitious publication profiles for research groups and panel members is created and cognitive distances, using different similarity matrices, are calculated and systematically compared.

Moreover, the use of different mapping techniques to derive two-dimensional base maps gives different results (discussed in chapter VII). An analysis can be done based on two-dimensional base map derived by different techniques. We observed that the cognitive distances obtained in a Kamada-Kawai map differ from the distances derived by VOS map techniques. There are other algorithms or layout techniques available. We can apply the proposed approaches for a panel and research groups on the base maps obtained from different algorithms or layout techniques. Here too a simulation based approach might be beneficial. These further studies (different similarity matrix and different mapping techniques) will lead to new insights into how much differences in the similarity matrix and mapping techniques affect the observed cognitive distances.

We have used similarity matrices and base maps derived from them based on data available during the construction of the matrices. If the similarity matrix changed over the years, and we keep the same panel and research groups publication data, this might result in different cognitive distances. Moreover, if we use a different similarity matrix (for example, based on Scopus data) and retrieve the same panel and research groups' data, we can expect different results as well, because the similarity matrix and the data will not be the same. An interesting follow-up investigation could therefore be based on Scopus data (e.g., Leydesdorff, de Moya-Anegón, & Guerrero-Bote, 2010; Leydesdorff, Moya-Anegón, & Guerrero-Bote, 2015). Hence, although there is a practical stability problem, the methods we introduced have general applicability.

In our case, for the barycenter method using the global map of sciences of WoS SCs, if a publication appears in a journal belonging to two SCs, it has been counted twice. A subsequent analysis using fractional counting in WoS SCs (Bornmann, 2014) can be done.

The scope of journals can vary significantly; some journals focus on rather specific topics, whereas others, such as PLoS ONE, are multidisciplinary in nature. One might therefore question whether journals are the adequate level of analysis. We suggest two possible routes for future research in this regard. First, it would be interesting if a comparison could be made between an analysis that considers all journals and one that leaves out multidisciplinary or otherwise broadly scoped journals. Second, one could replace journals with clusters of cognitively related articles. For instance, one could use the CWTS (Centre for Science and Technology Studies) article-level classification (Waltman & van Eck, 2012), which groups related articles together on the basis of direct citations regardless of the journal in which they were published. While we consider this an interesting idea, we also point out that it harbors its own set of theoretical and practical problems. This article-level classification includes publications from the period 2001-2010. Hence, publications before and after that period cannot be taken into account. In addition, it seems harder to interpret the results of a publication-level analysis, e.g. clusters in this map are not really labeled but only characterized by frequently occurring terms.

Our proposed approaches have not yet been tested against a 'real' gold standard due to absence of it (discussed in chapter VII and VIII). The proposed approaches can be tested in any future case scenario where a panel needs to be chosen. For example, in a situation where X panel members need to be chosen out of N candidate panel members to evaluate Y research groups. The panel composition based on the approaches can then be matched with the opinion of the panel chair and see how they differ in practice and why. In addition, the opinion of the respective panel members can be taken into account on beforehand, whether they consider themselves as the right person to evaluate the assigned research groups or not.

An integrated method for composing an expert panel is yet to be established. Our proposed approaches can become part of such a method. The proposed approaches help the concerned authority to assign panel members to research groups taking into account the cognitive distance and confidence intervals. When an expert panel needs to be composed with K panel members from a set of N potential panel members, a subset can be chosen from a number of subsets (Aggarwal, Imai, Katoh, & Suri, 1991). For example, there are three subsets: subset A (PM1, PM2, PM3, PM5 and PM7), subset B (PM1, PM2, PM4, PM5, and PM6), and subset C (PM2, PM5, PM6, PM7, PM8). These subsets are chosen based on either standalone or a combination of one or more parameters: overlap of confidence intervals, two or more panel members having similar expertise, the panel chair must be there, the cognitive distances between the panel members and the research groups should not be too close or too far, the distances between the panel members should not be too small or spread over so that they can understand each other, etc. From these subsets, one can choose the best subset depend on the goals and needs of the evaluation at hand. Developing and testing such a software-facilitated approach would be a major achievement.

Appendix

Appendix A: Confidence interval plot of barycenter distances

The highlighted part indicates the confidence interval of the shortest distance to the research group.

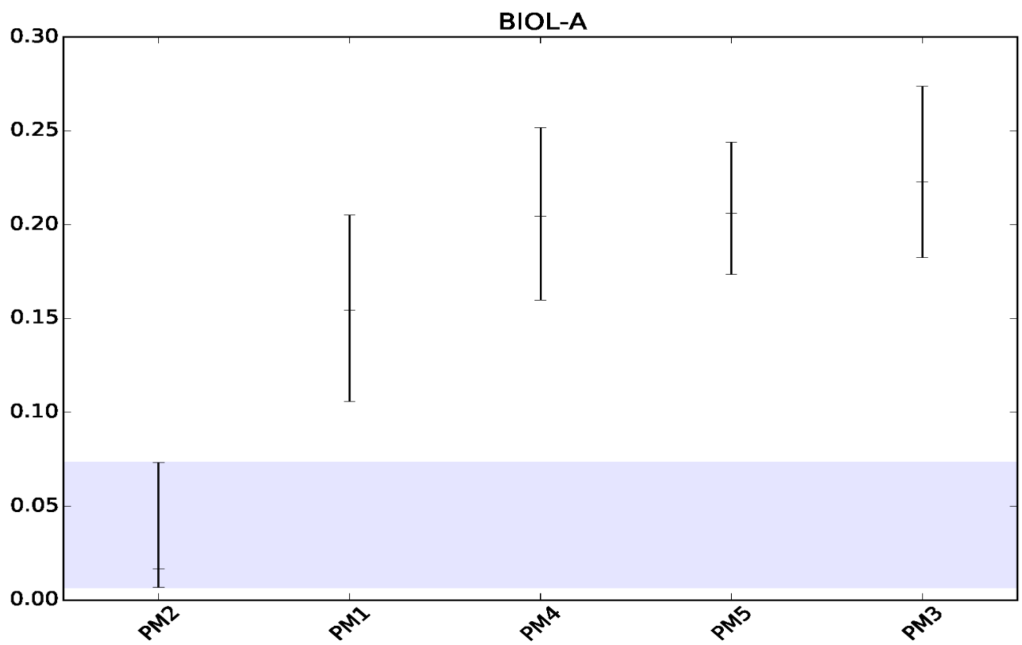


Figure 38: Confidence interval plot of barycenter distances of BIOM-A research group

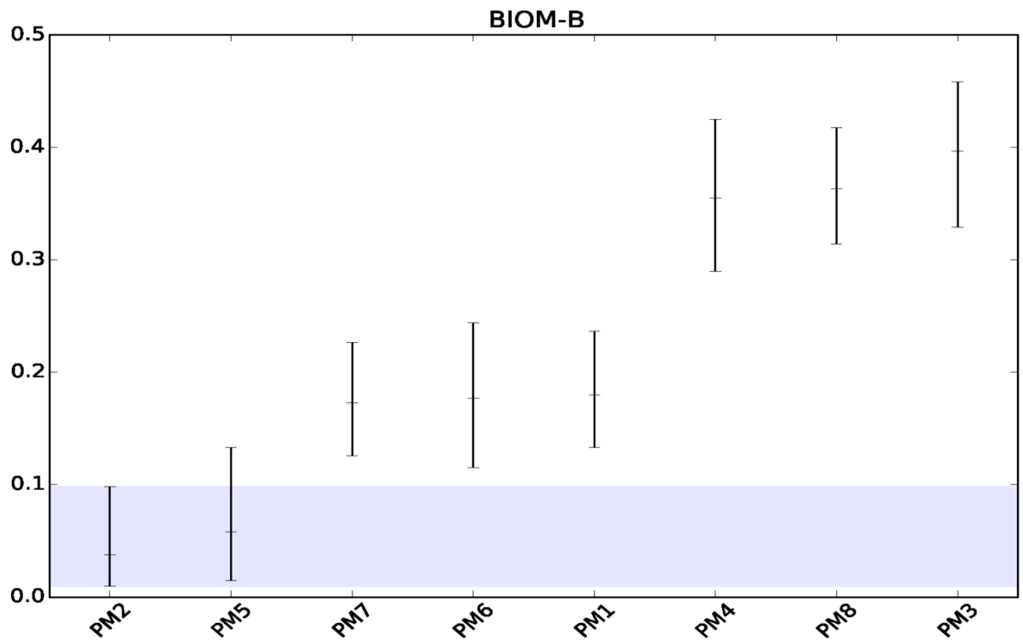


Figure 39: Confidence interval plot of barycenter distances of BIOM-B research group

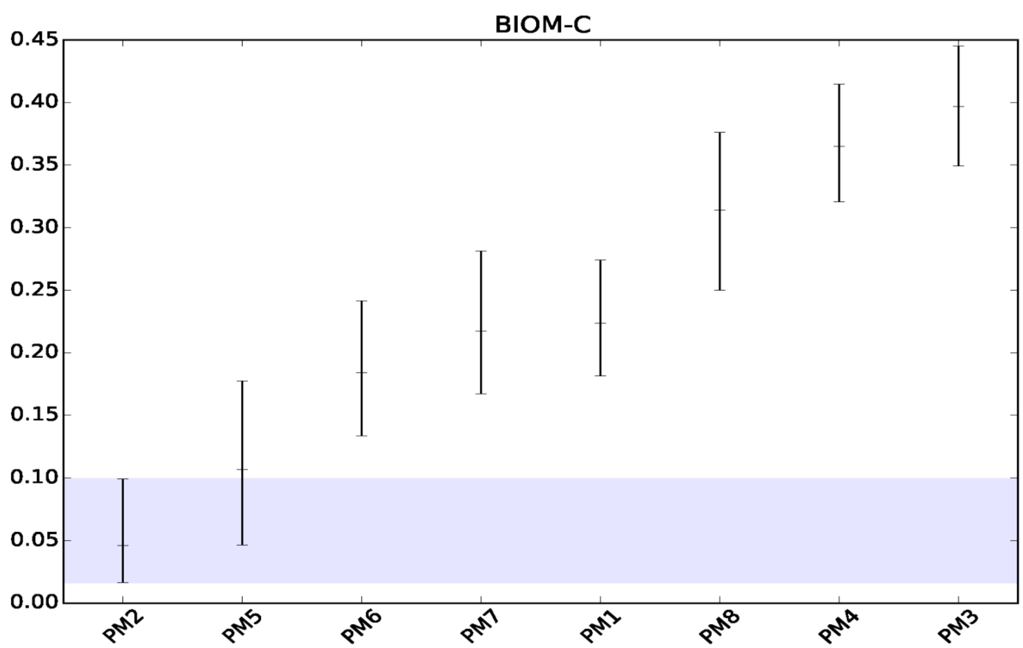


Figure 40: Confidence interval plot of barycenter distances of BIOM-C research group

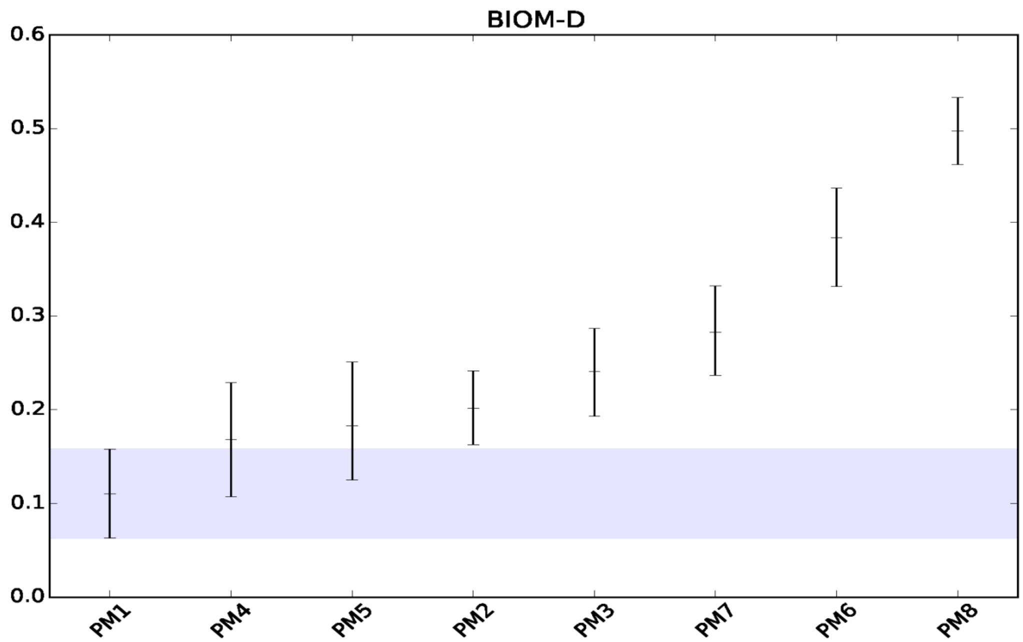


Figure 41: Confidence interval plot of barycenter distances of BIOM-D research group

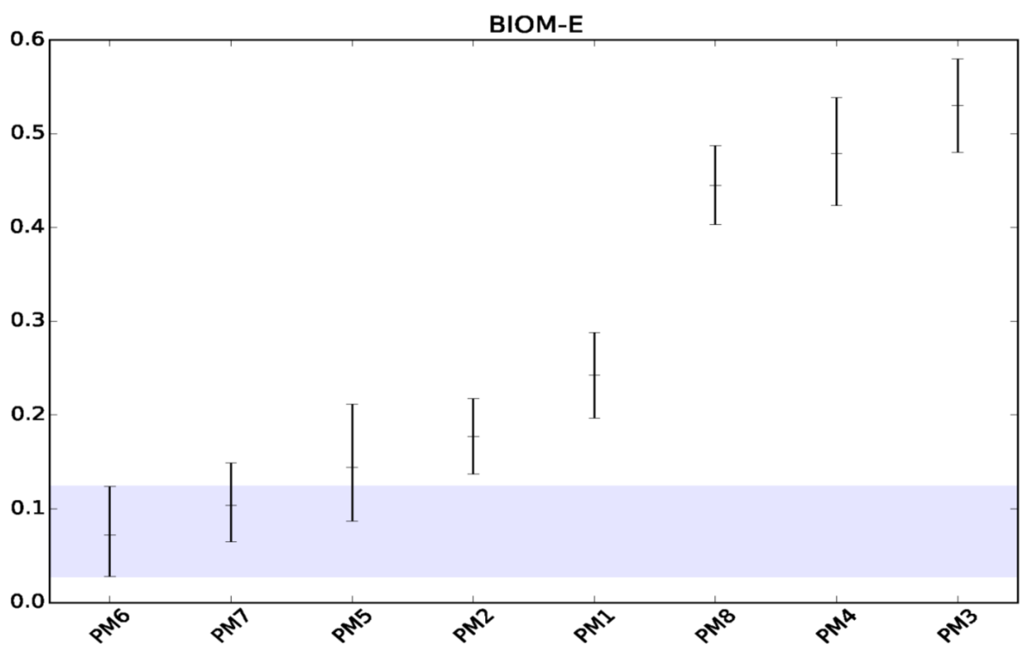


Figure 42: Confidence interval plot of barycenter distances of BIOM-E research group

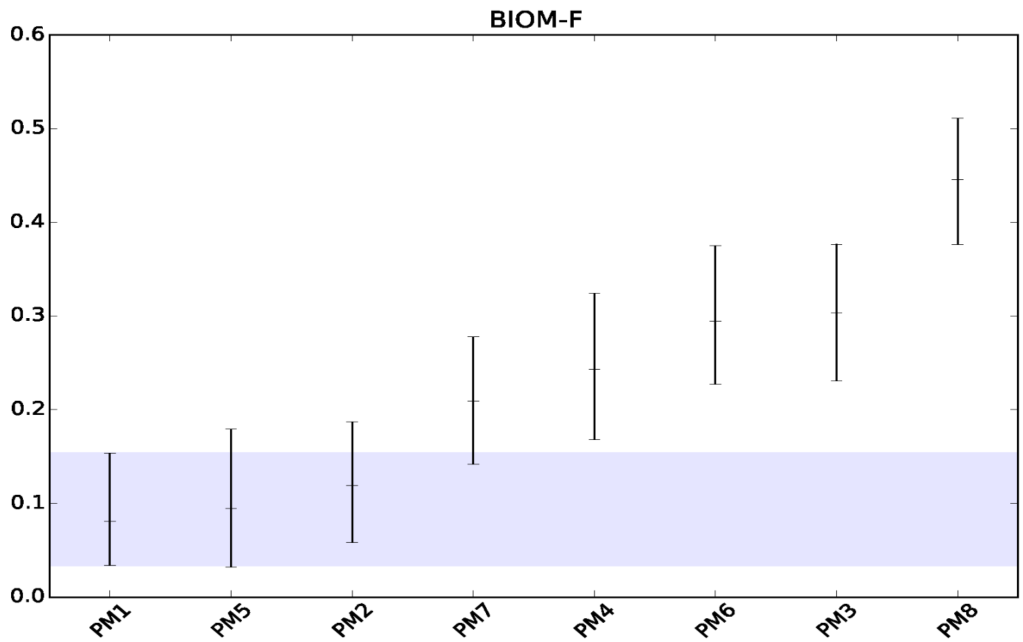


Figure 43: Confidence interval plot of barycenter distances of BIOM-F research group

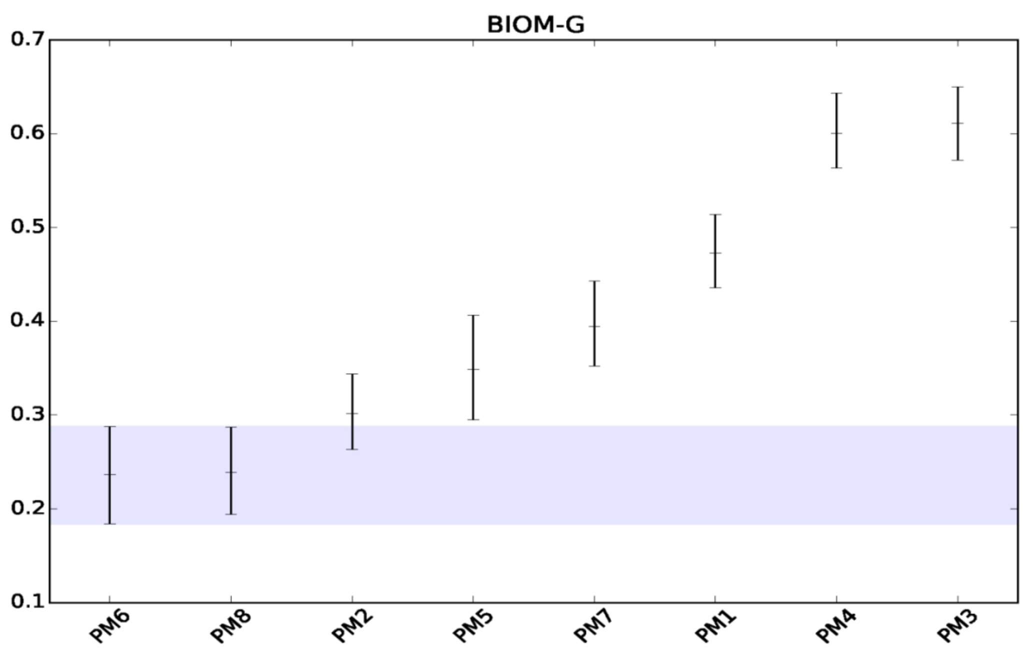


Figure 44: Confidence interval plot of barycenter distances of BIOM-G research group

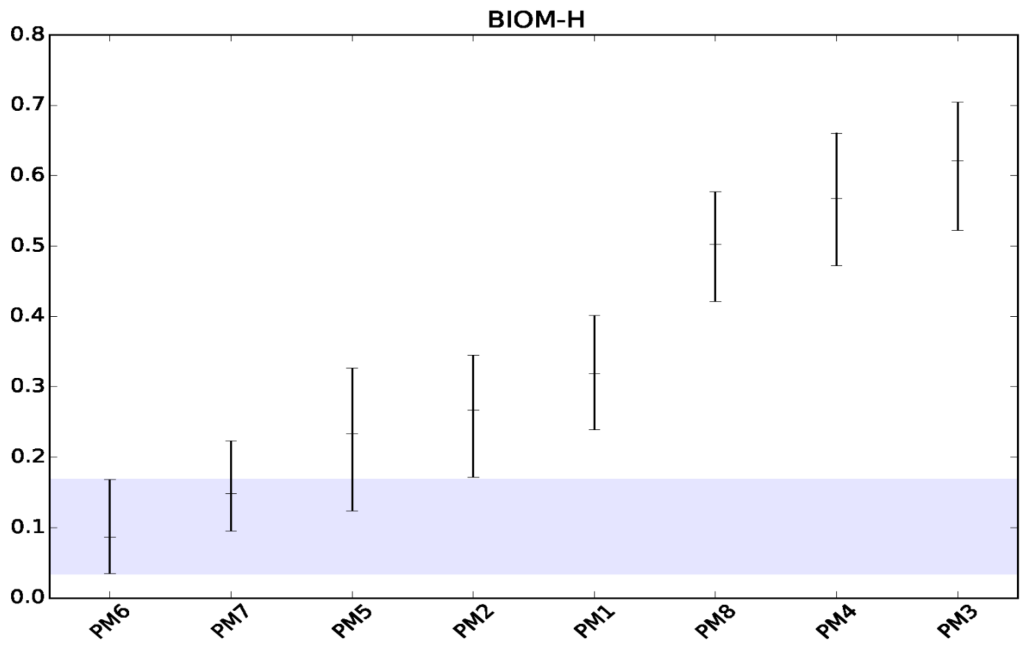


Figure 45: Confidence interval plot of barycenter distances of BIOM-H research group

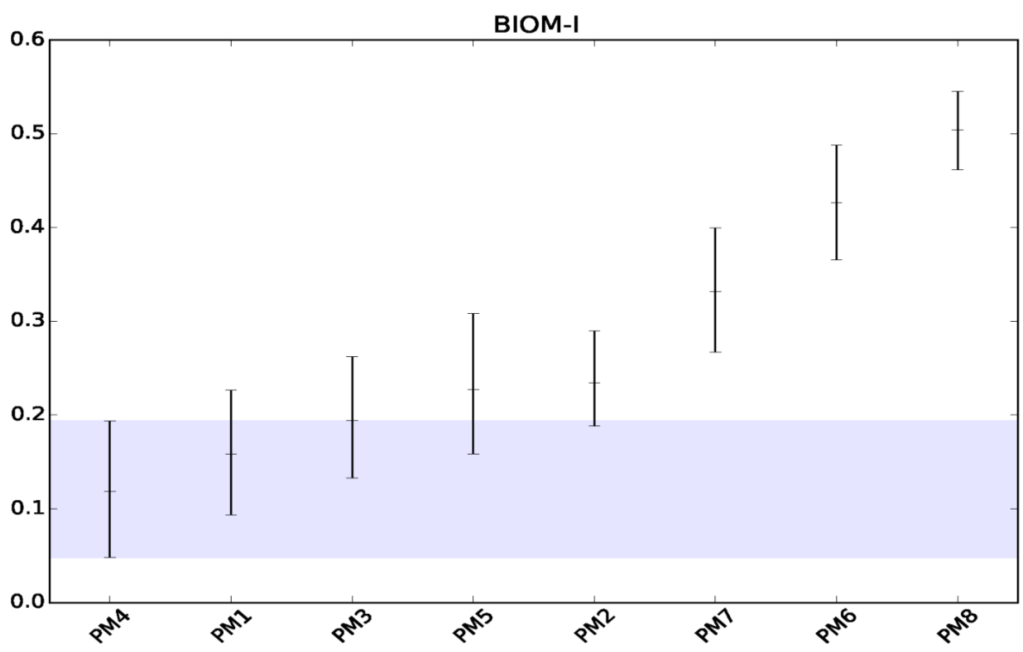


Figure 46: Confidence interval plot of barycenter distances of BIOM-I research group

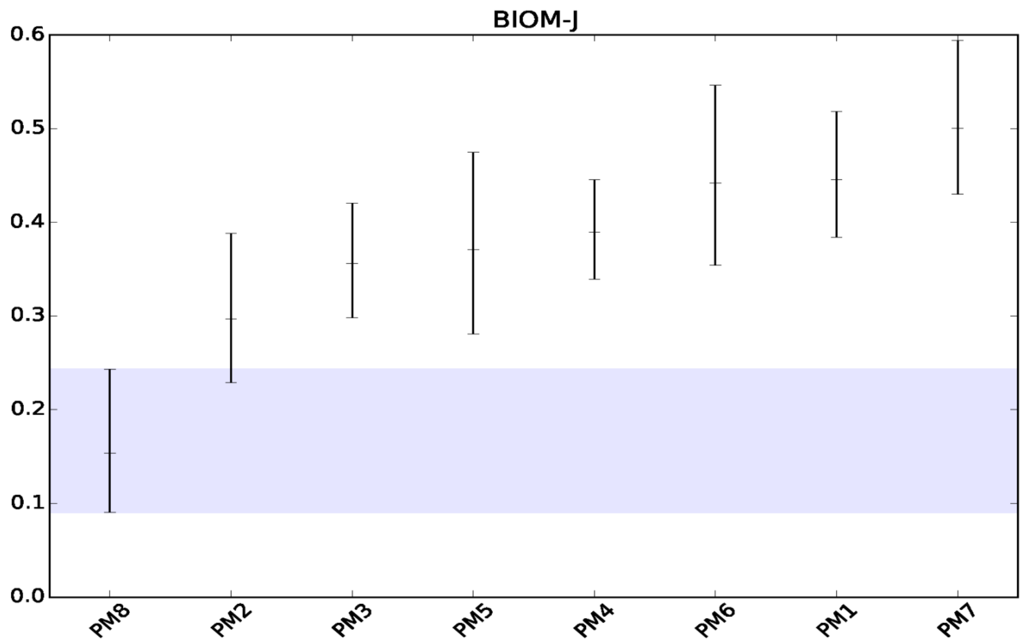


Figure 47: Confidence interval plot of barycenter distances of BIOM-J research group

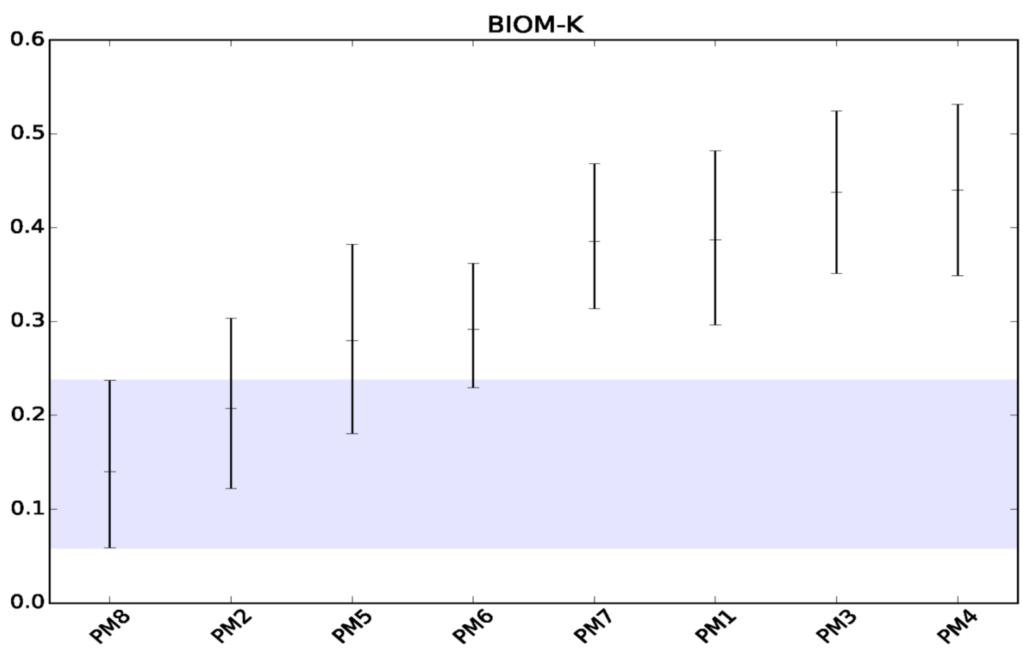


Figure 48: Confidence interval plot of barycenter distances of BIOM-K research group

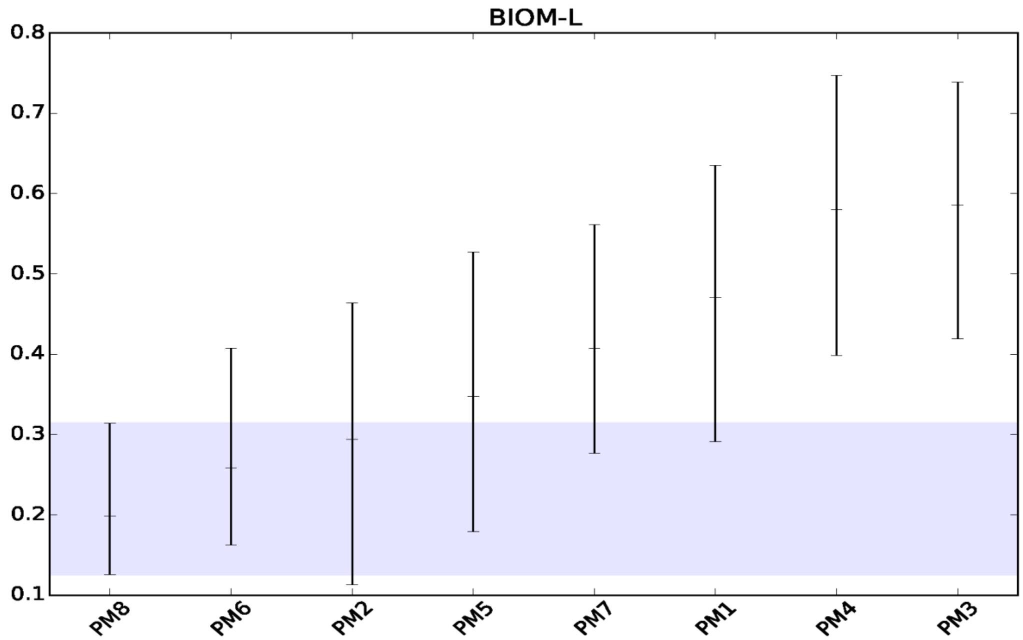


Figure 49: Confidence interval plot of barycenter distances of BIOM-L research group

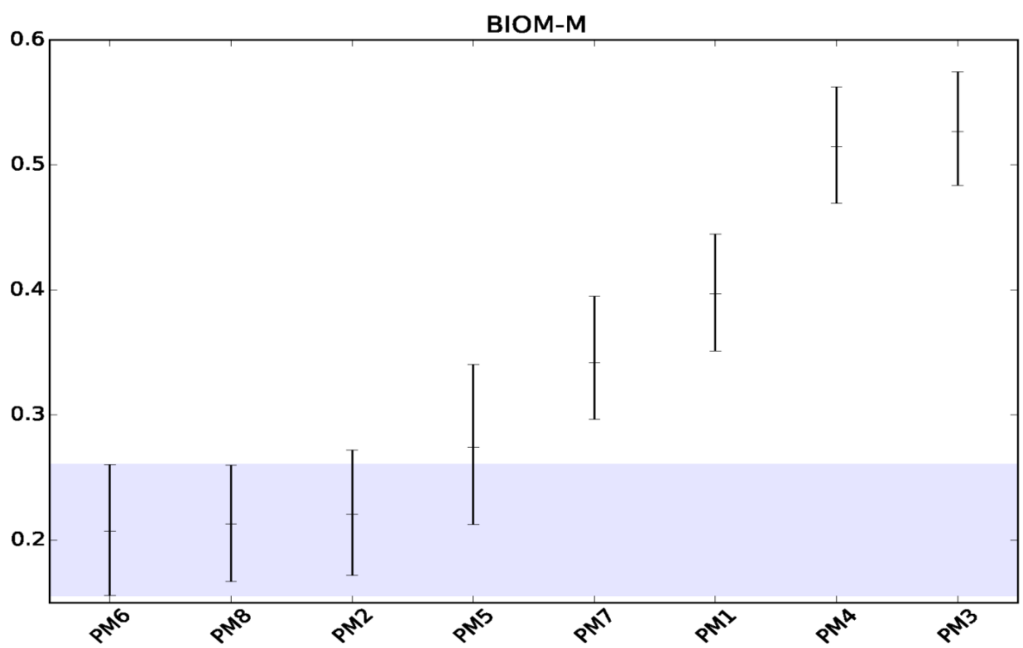


Figure 50: Confidence interval plot of barycenter distances of BIOM-M research group

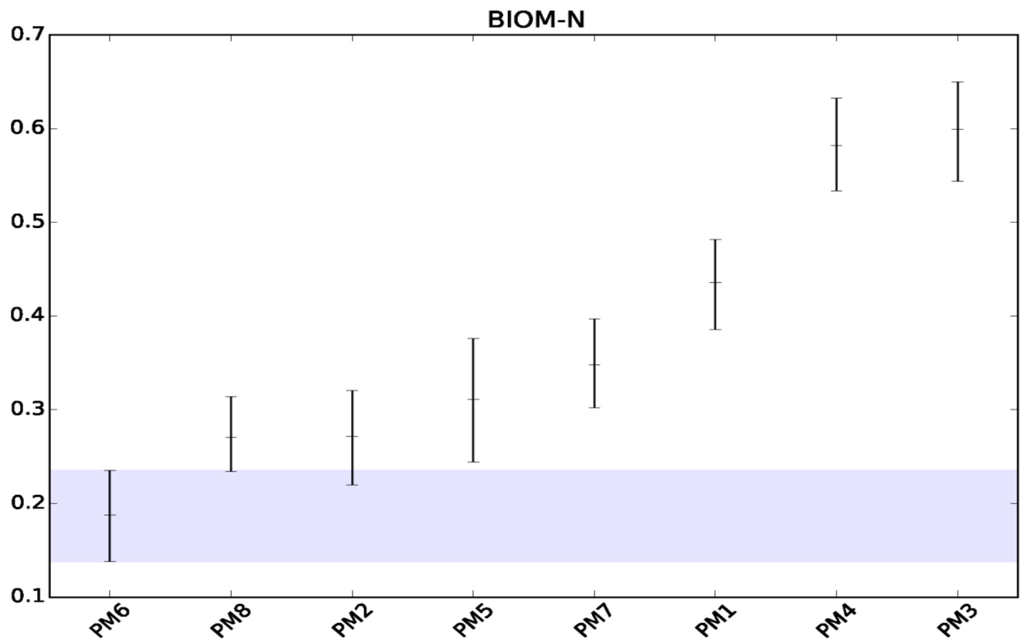


Figure 51: Confidence interval plot of barycenter distances of BIOM-N research group

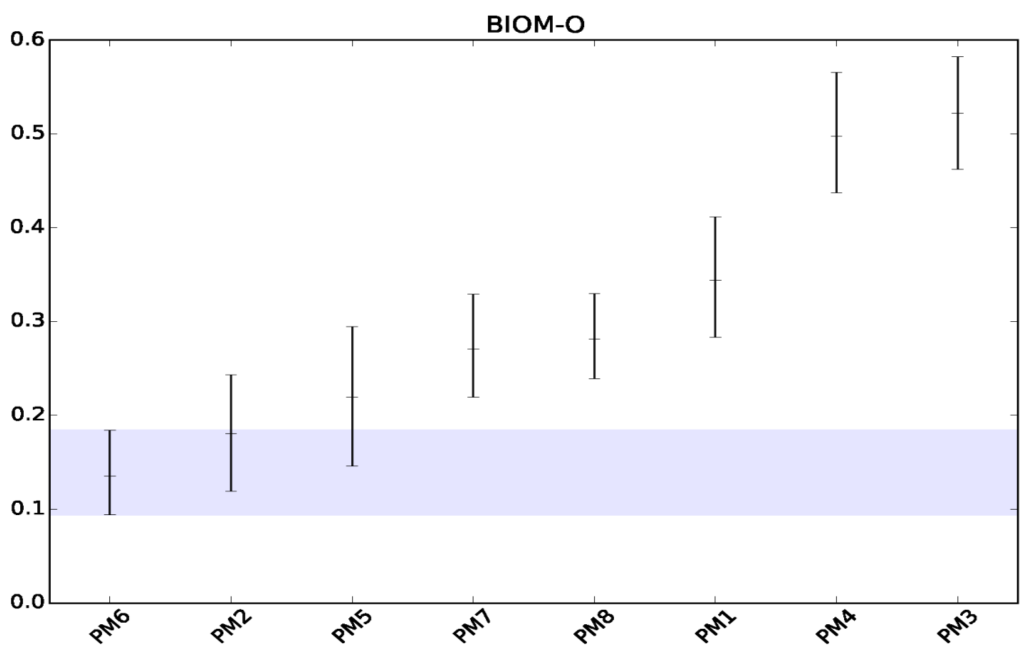


Figure 52: Confidence interval plot of barycenter distances of BIOM-O research group

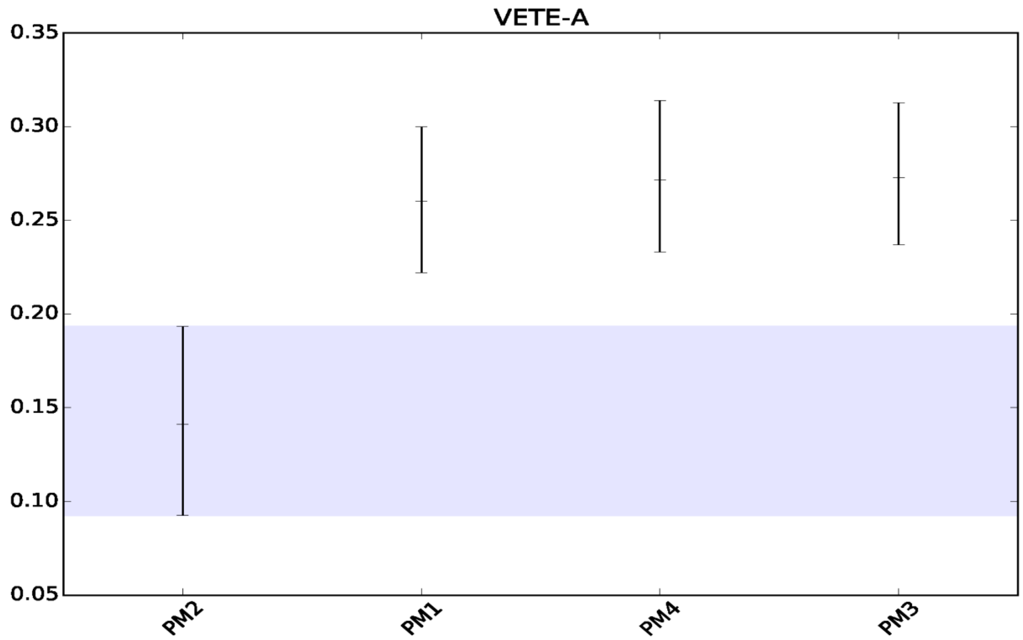


Figure 53: Confidence interval plot of barycenter distances of VETE-A research group

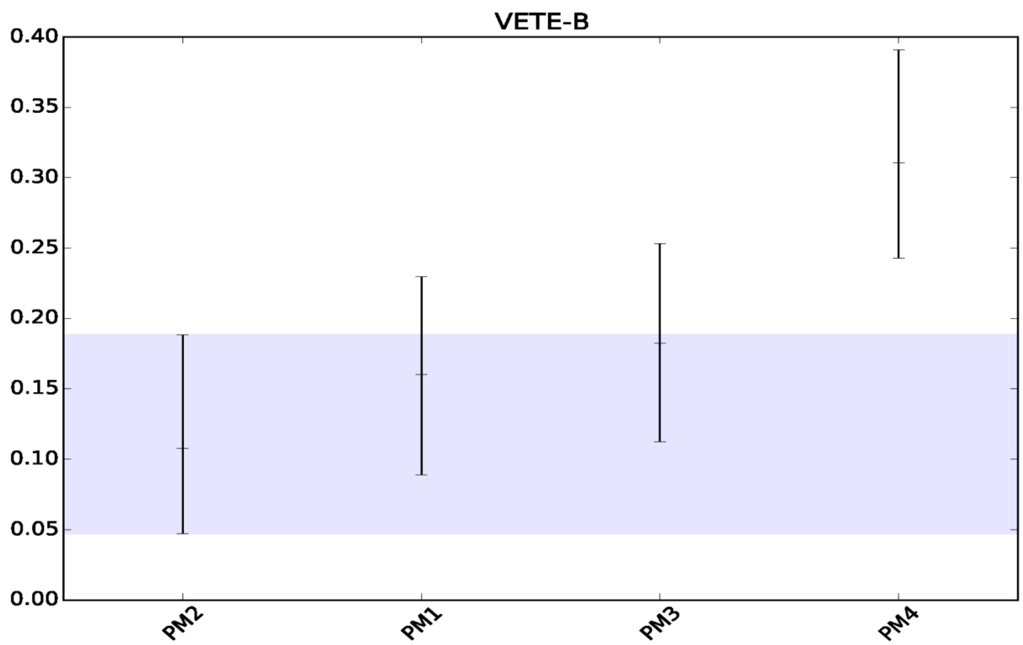


Figure 54: Confidence interval plot of barycenter distances of VETE-B research group

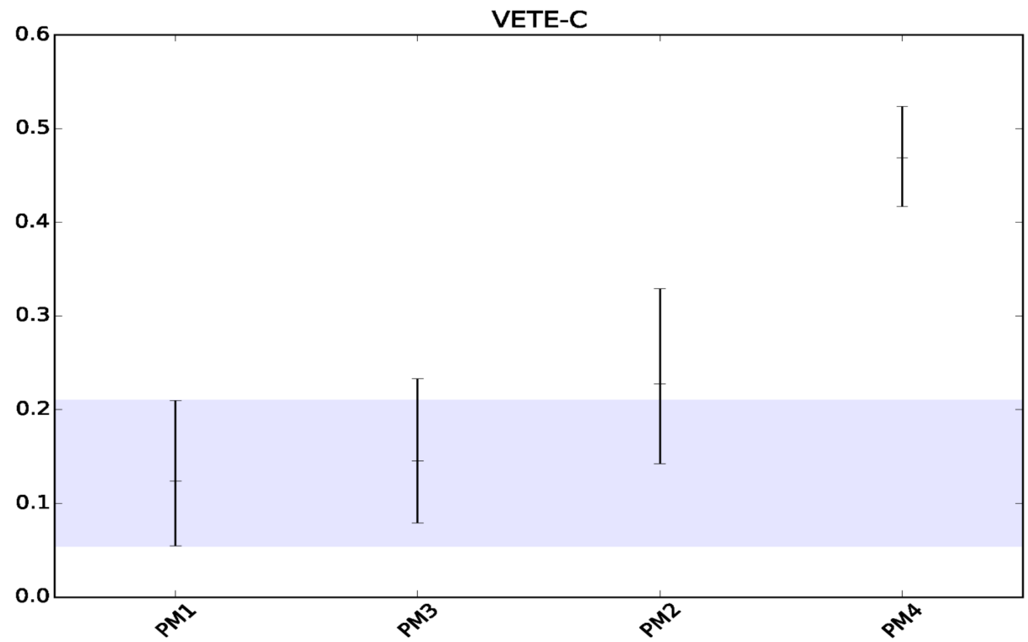


Figure 55: Confidence interval plot of barycenter distances of VETE-C research group

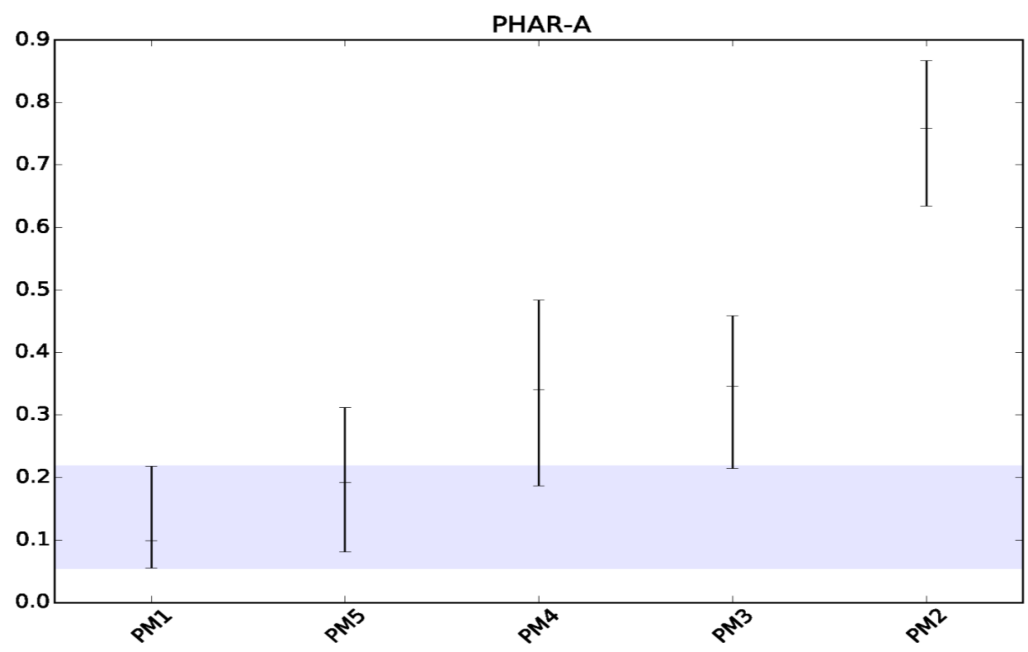


Figure 56: Confidence interval plot of barycenter distances of PHAR-A research group

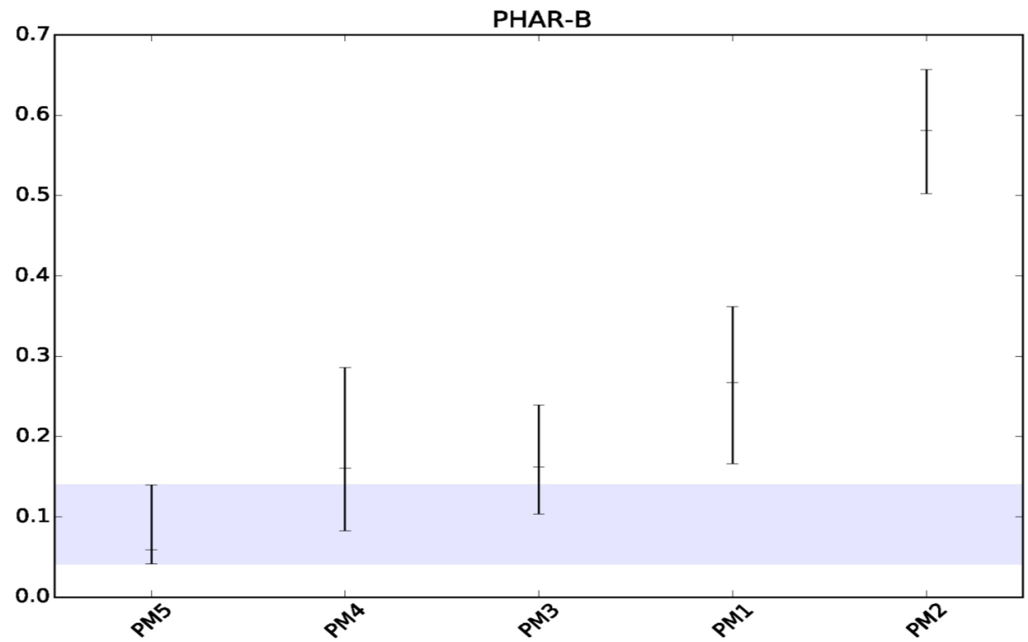


Figure 57: Confidence interval plot of barycenter distances of PHAR-B research group

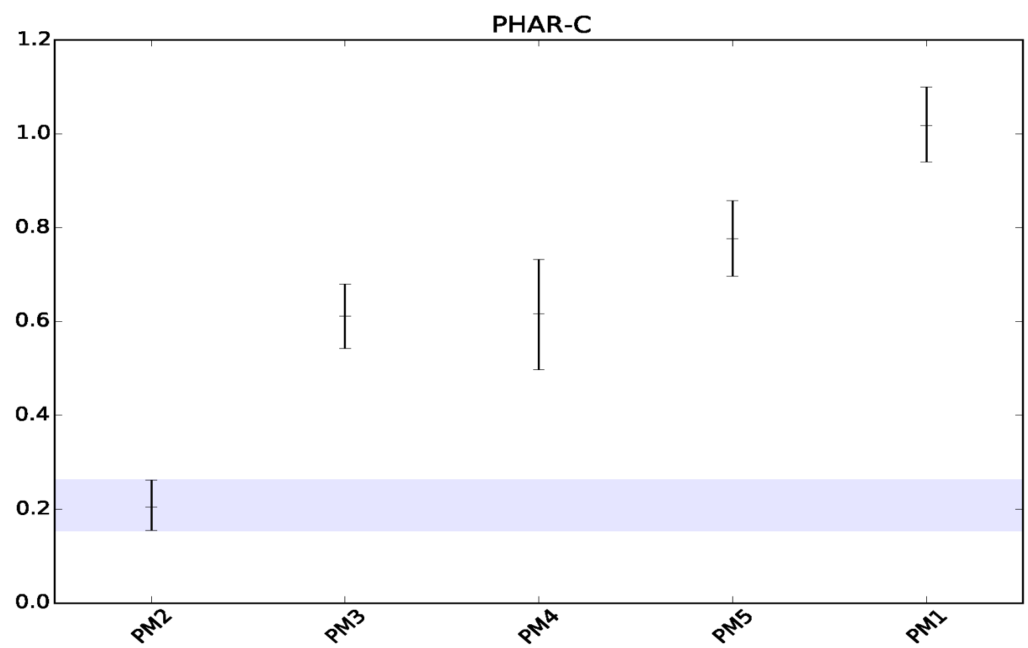


Figure 58: Confidence interval plot of barycenter distances of PHAR-C research group

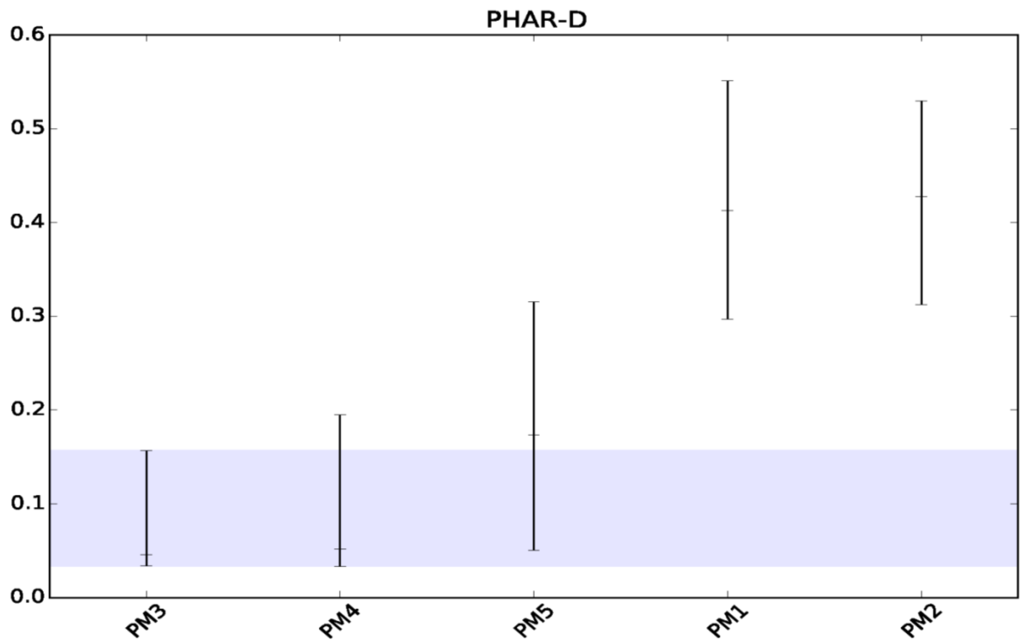


Figure 59: Confidence interval plot of barycenter distances of PHAR-D research group

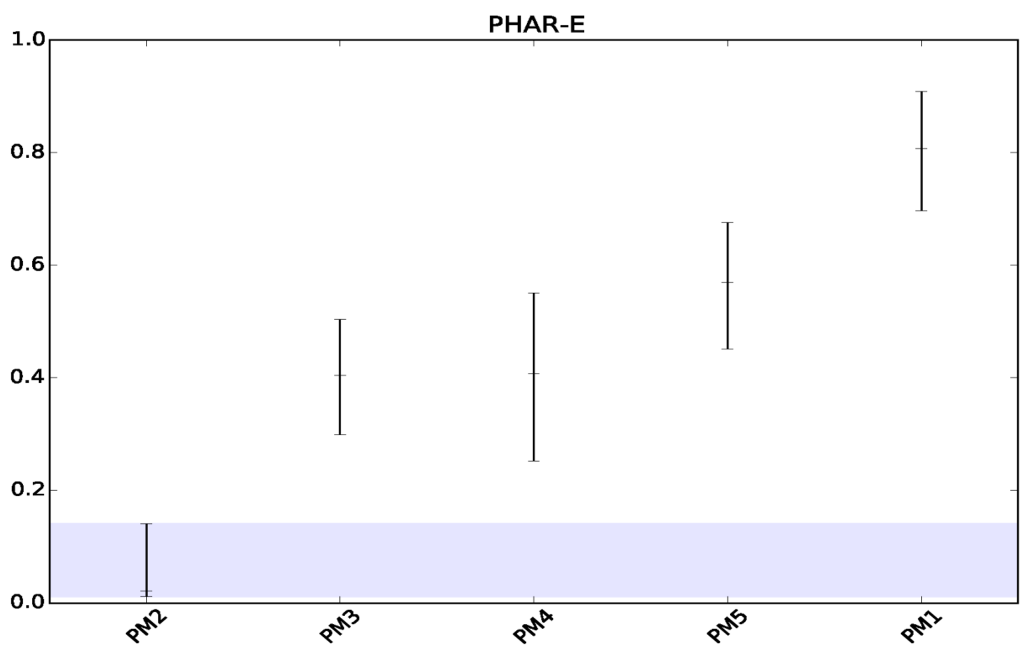


Figure 60: Confidence interval plot of barycenter distances of PHAR-E research group

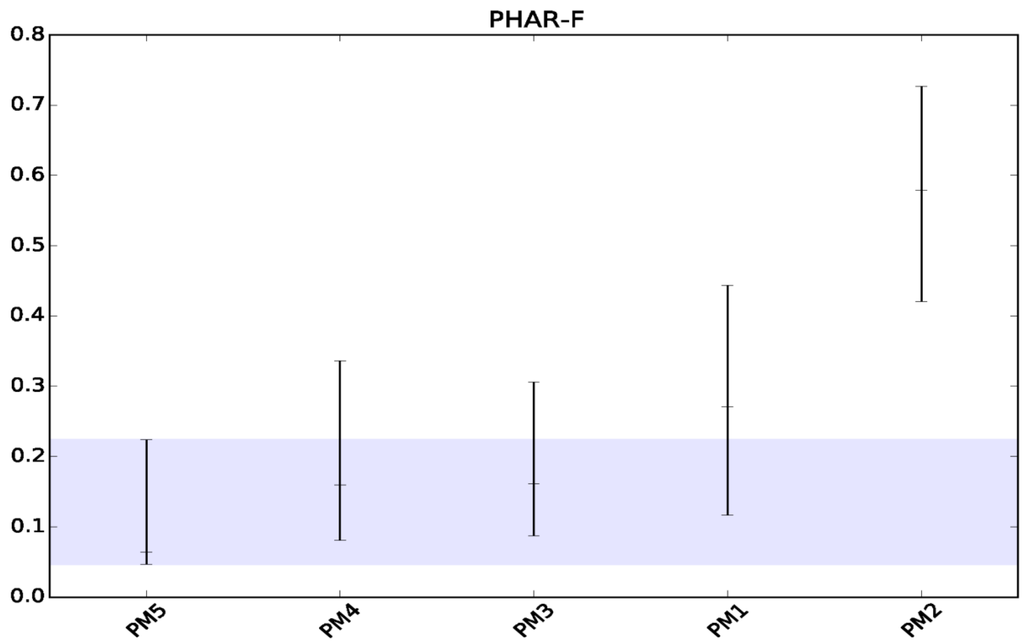


Figure 61: Confidence interval plot of barycenter distances of PHAR-F research group

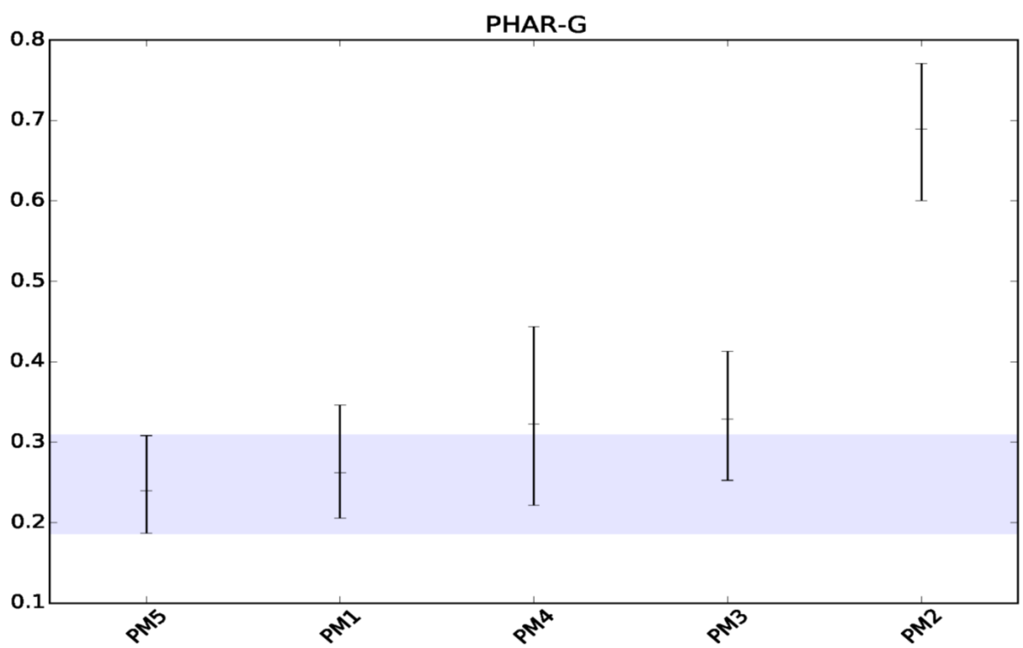


Figure 62: Confidence interval plot of barycenter distances of PHAR-G research group

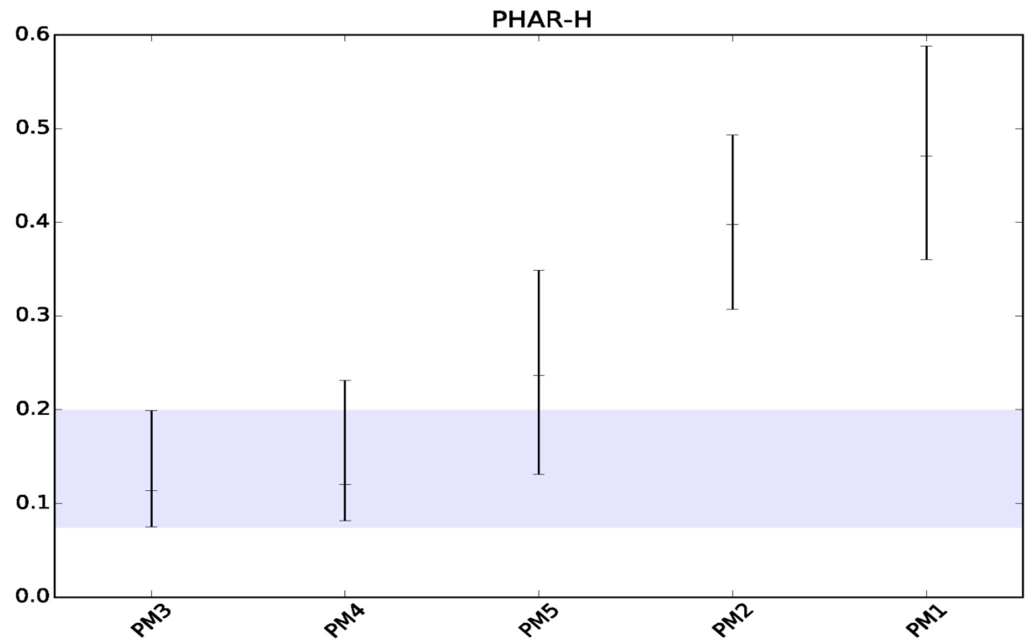


Figure 63: Confidence interval plot of barycenter distances of PHAR-H research group

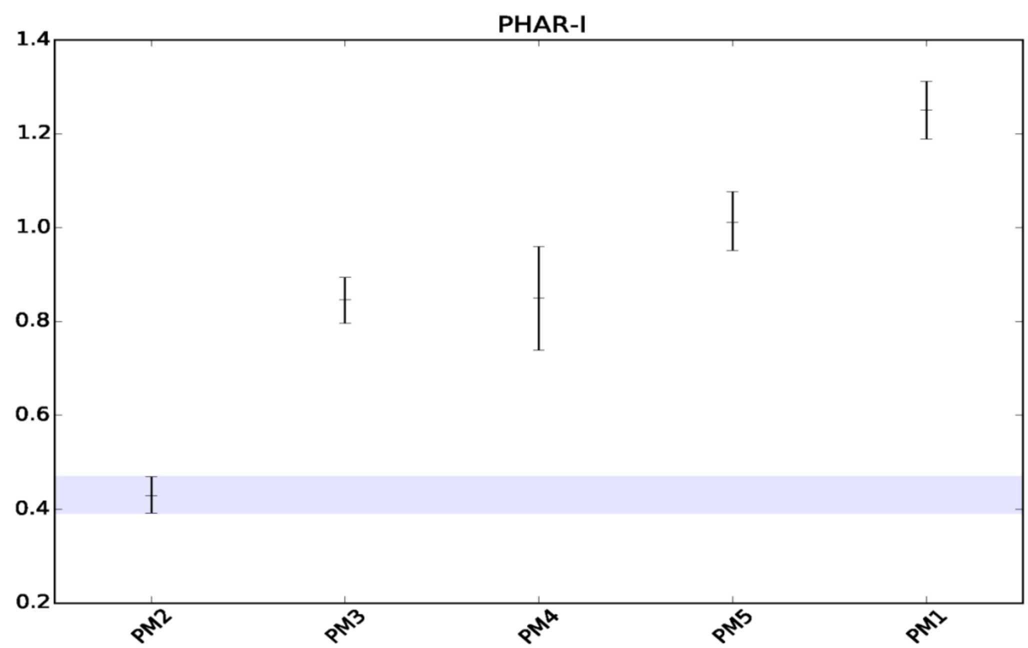


Figure 64: Confidence interval plot of barycenter distances of PHAR-I research group

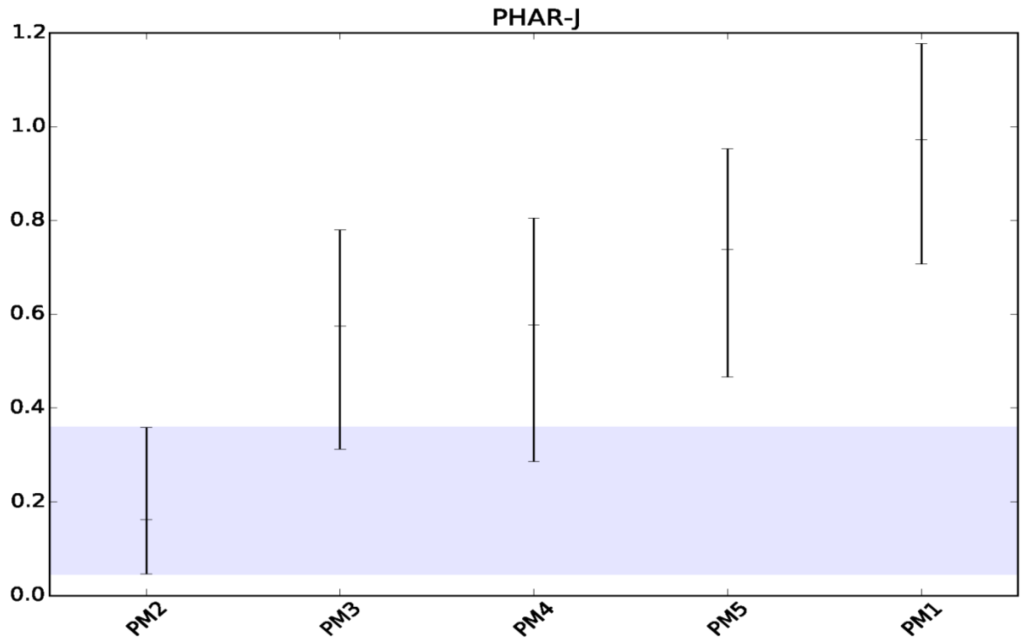


Figure 65: Confidence interval plot of barycenter distances of PHAR-J research group

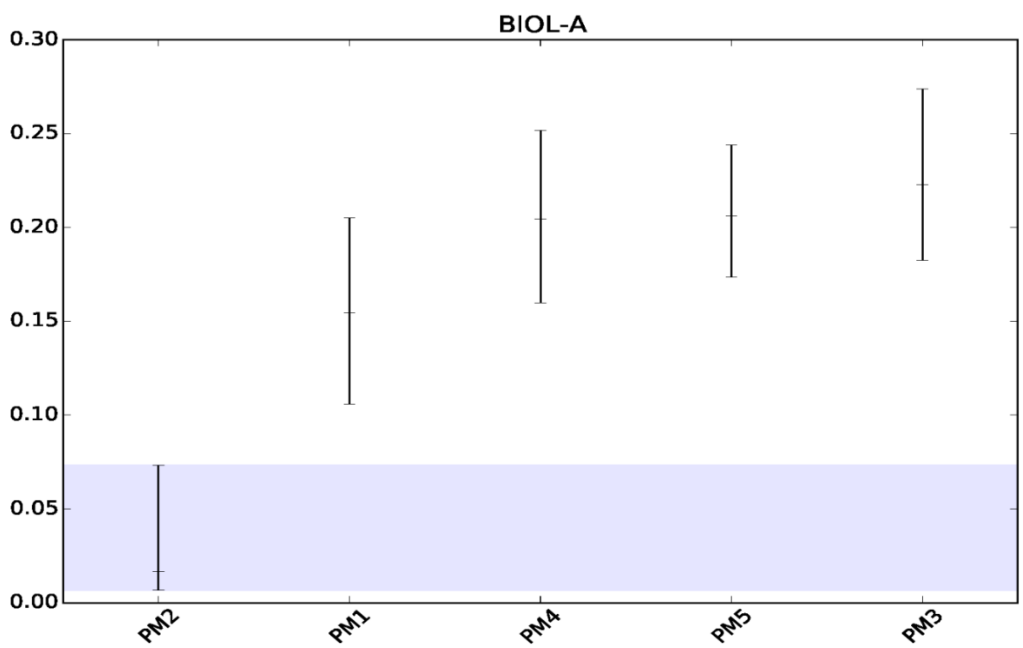


Figure 66: Confidence interval plot of barycenter distances of BIOL-A research group

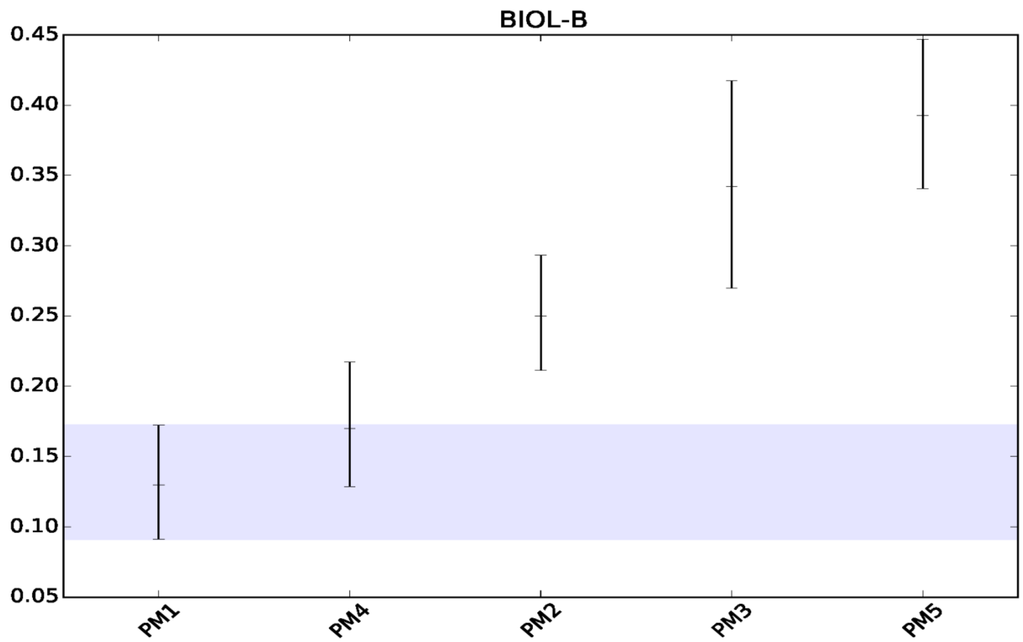


Figure 67: Confidence interval plot of barycenter distances of BIOL-B research group

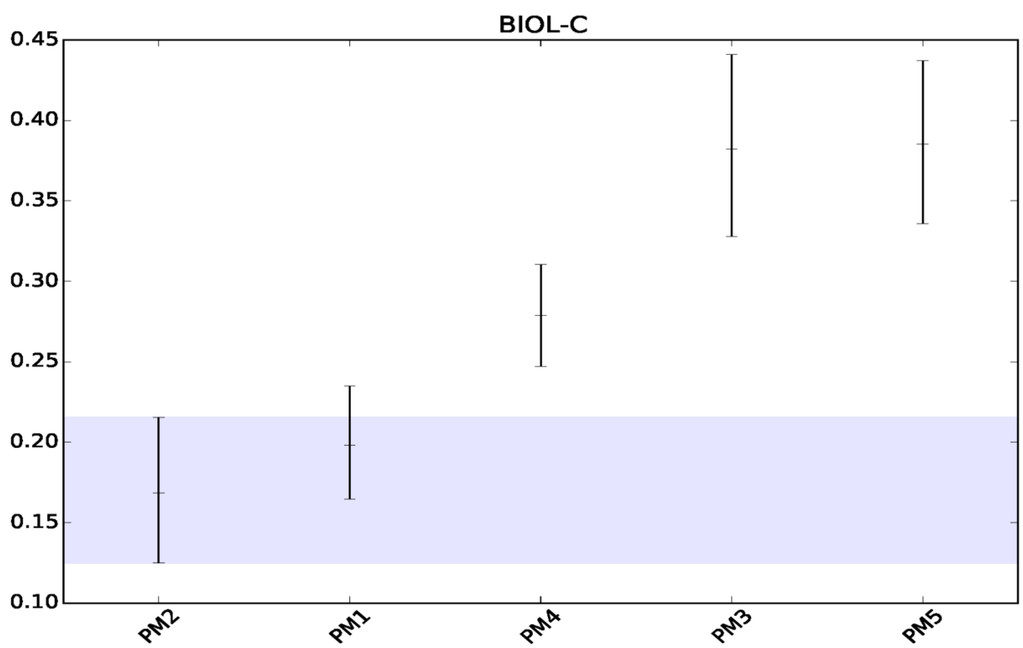


Figure 68: Confidence interval plot of barycenter distances of BIOL-C research group

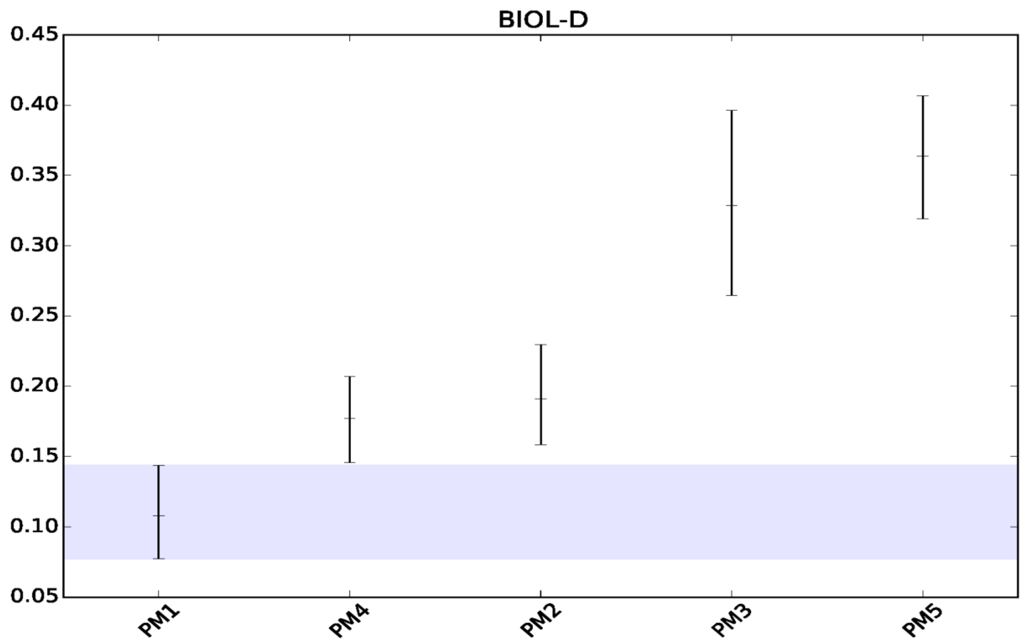


Figure 69: Confidence interval plot of barycenter distances of BIOL-D research group

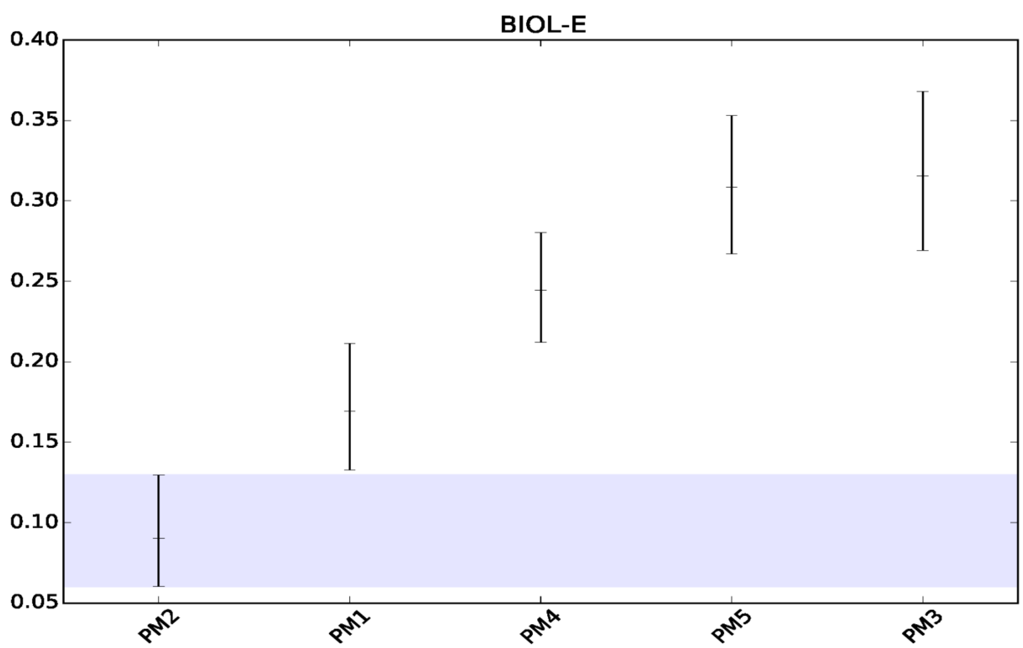


Figure 70: Confidence interval plot of barycenter distances of BIOL-E research group

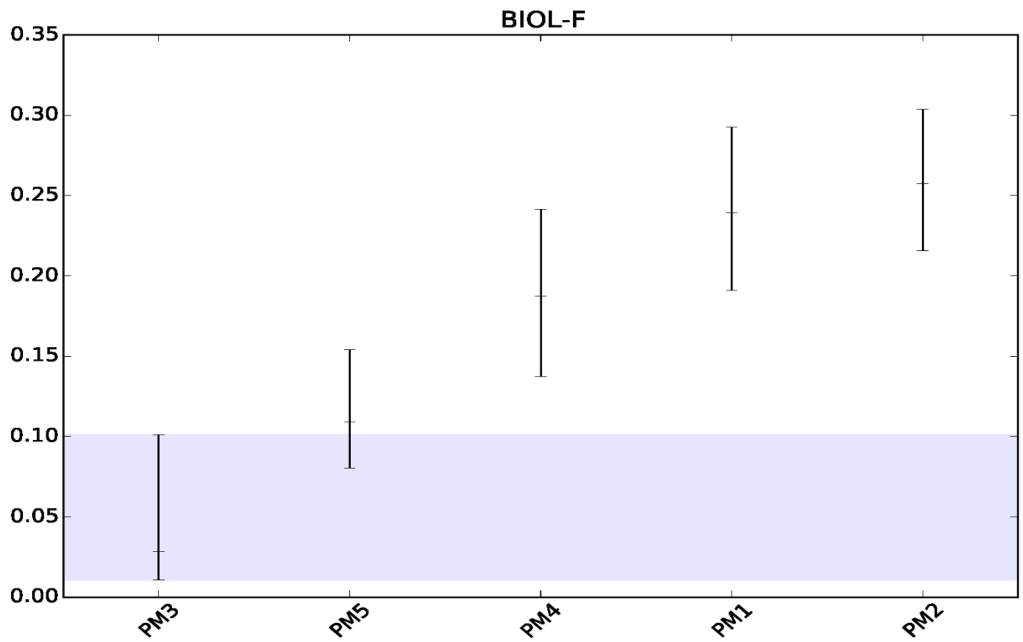


Figure 71: Confidence interval plot of barycenter distances of BIOL-F research group

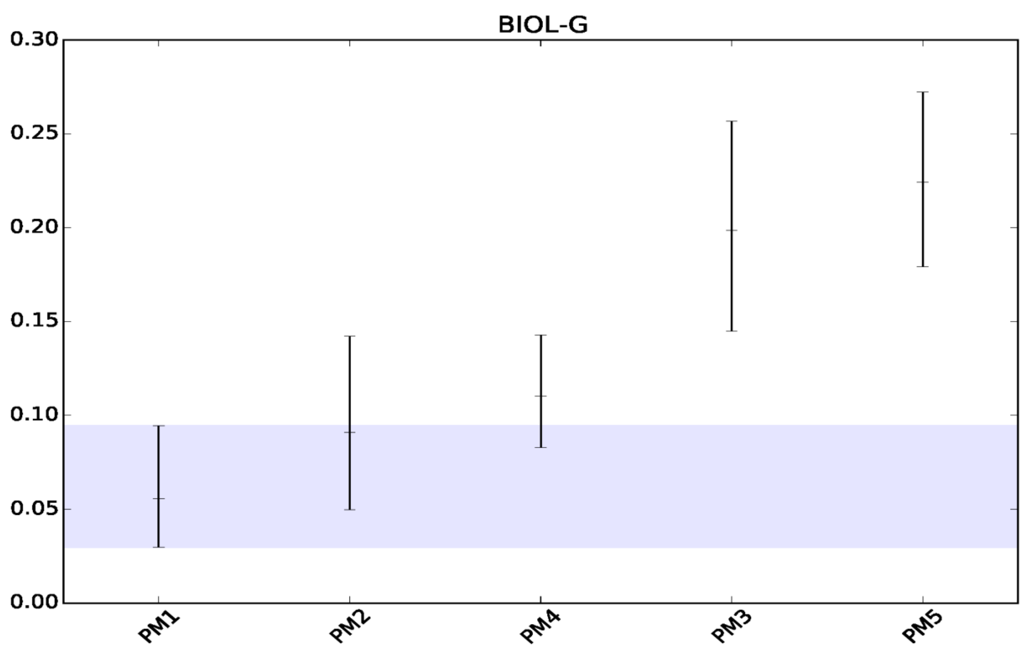


Figure 72: Confidence interval plot of barycenter distances of BIOL-G research group

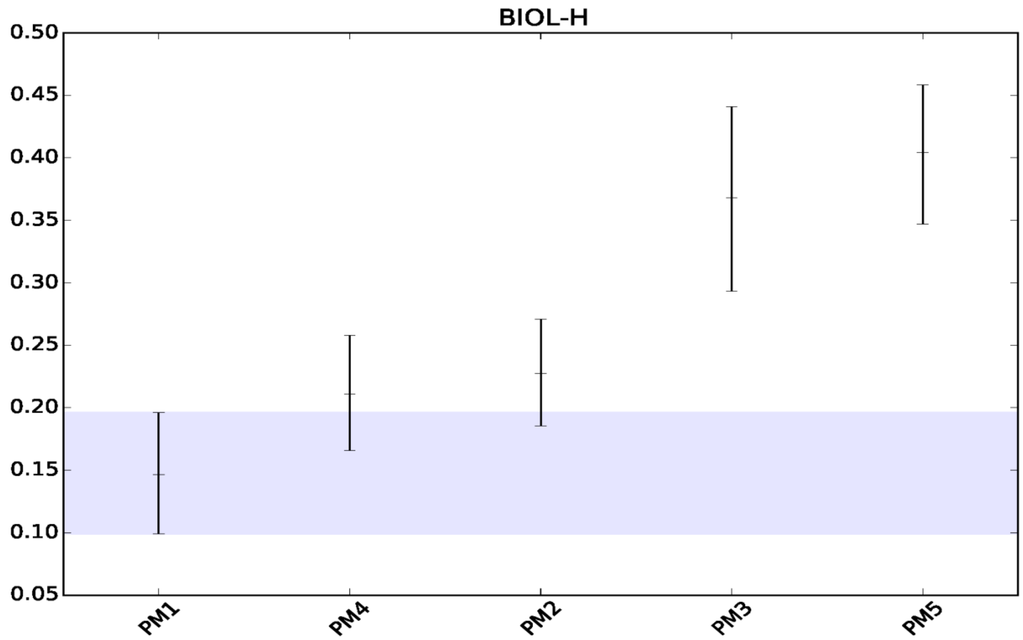


Figure 73: Confidence interval plot of barycenter distances of BIOL-H research group

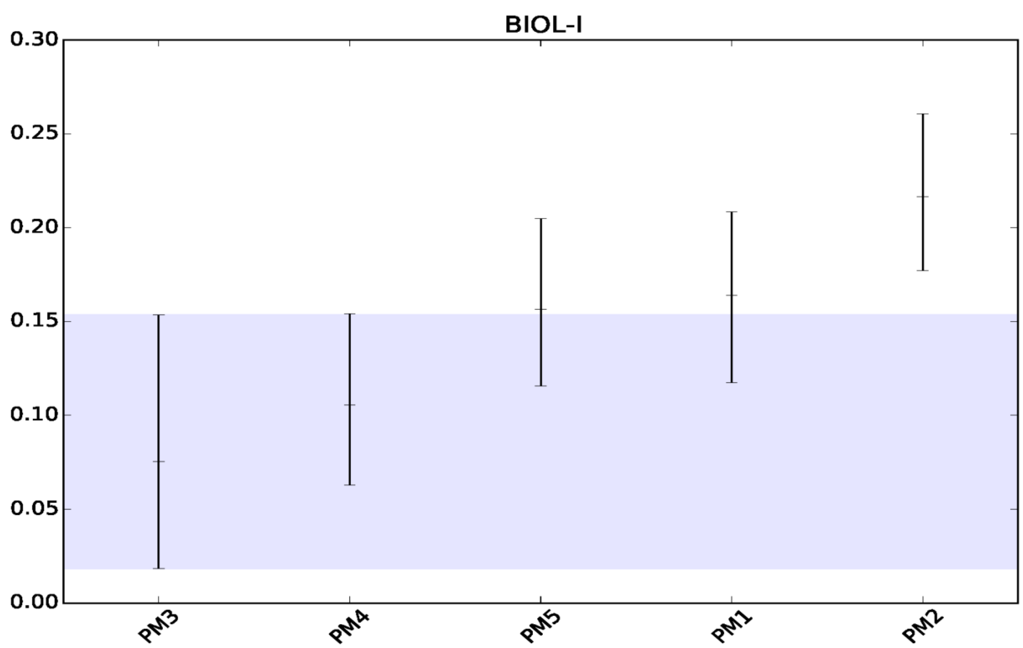


Figure 74: Confidence interval plot of barycenter distances of BIOL-I research group

Appendix B: Confidence interval plot of SAPV distances

The highlighted part indicates the confidence interval of the shortest distance to the research group.

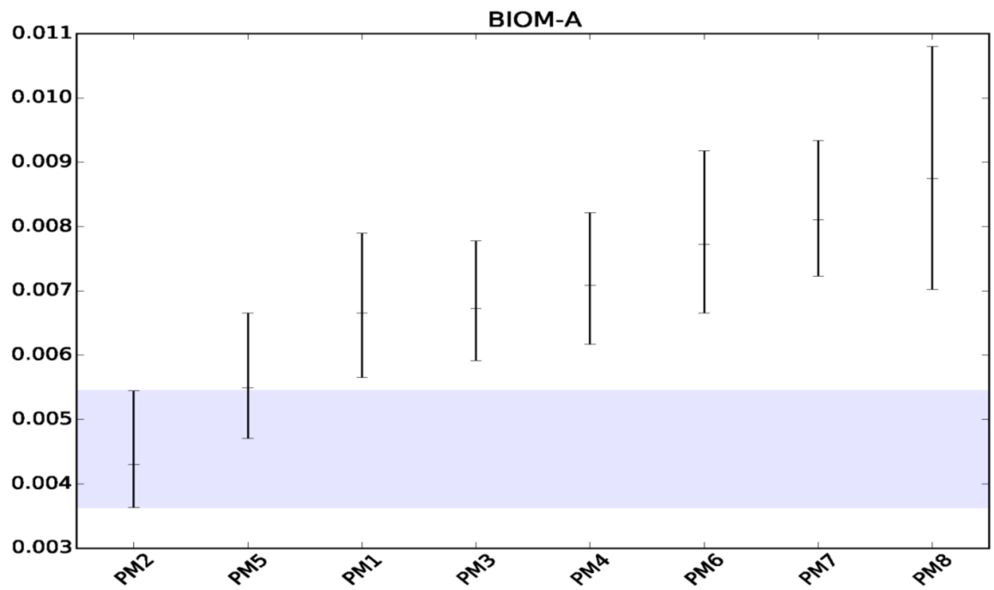


Figure 75: Confidence interval plot of SAPV distances of BIOM-A research group

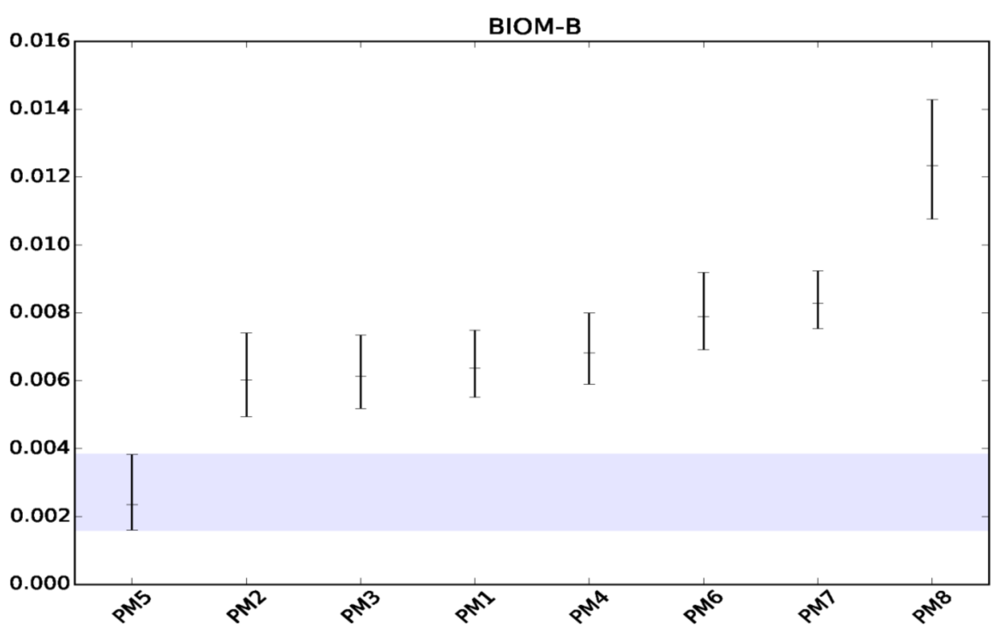


Figure 76: Confidence interval plot of SAPV distances of BIOM-B research group

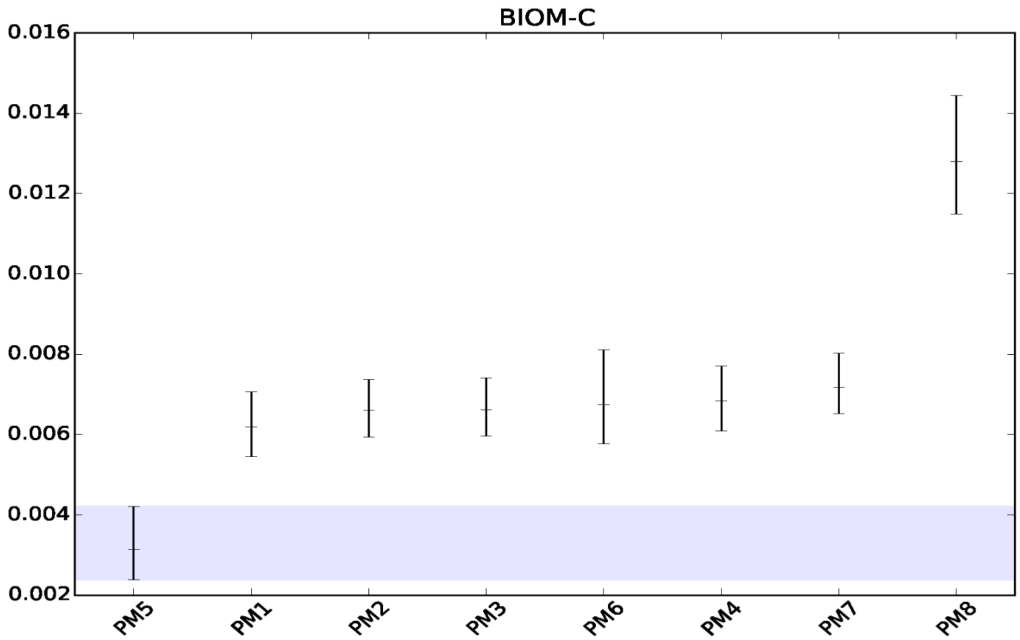


Figure 77: Confidence interval plot of SAPV distances of BIOM-C research group

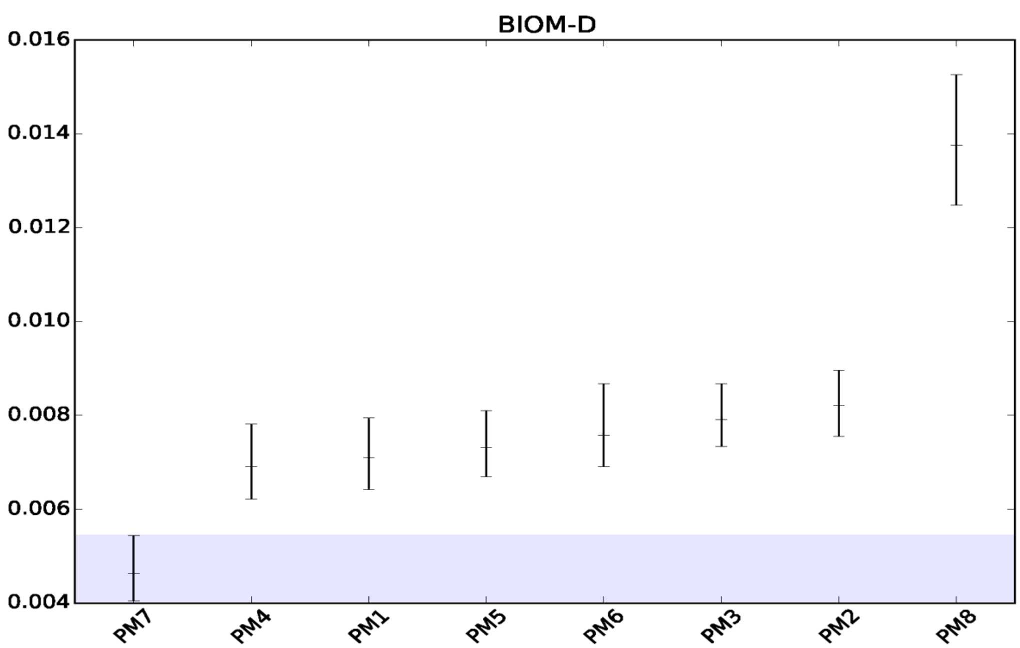


Figure 78: Confidence interval plot of SAPV distances of BIOM-D research group

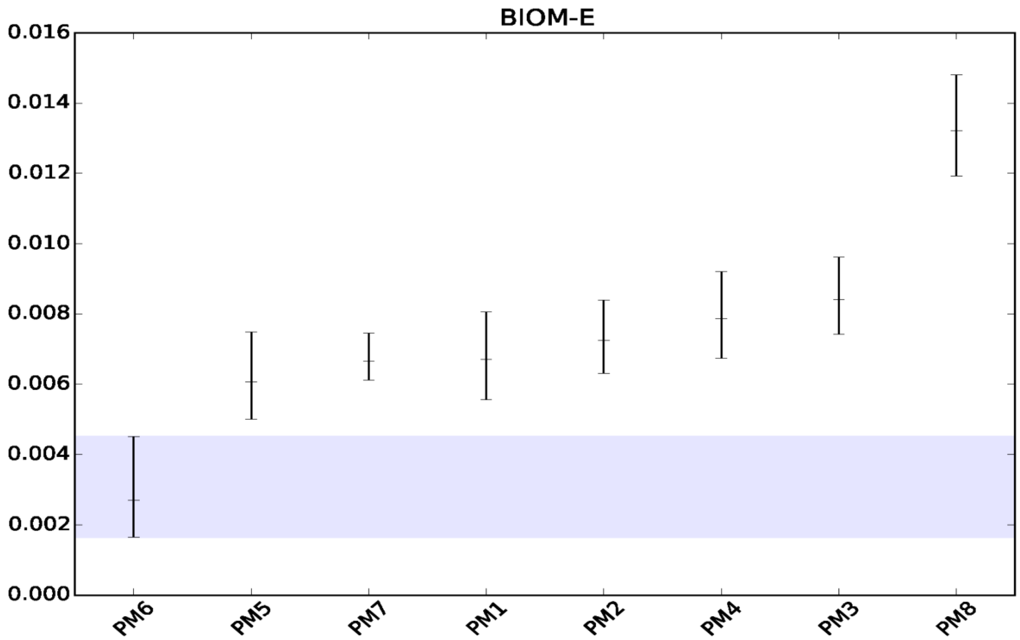


Figure 79: Confidence interval plot of SAPV distances of BIOM-E research group

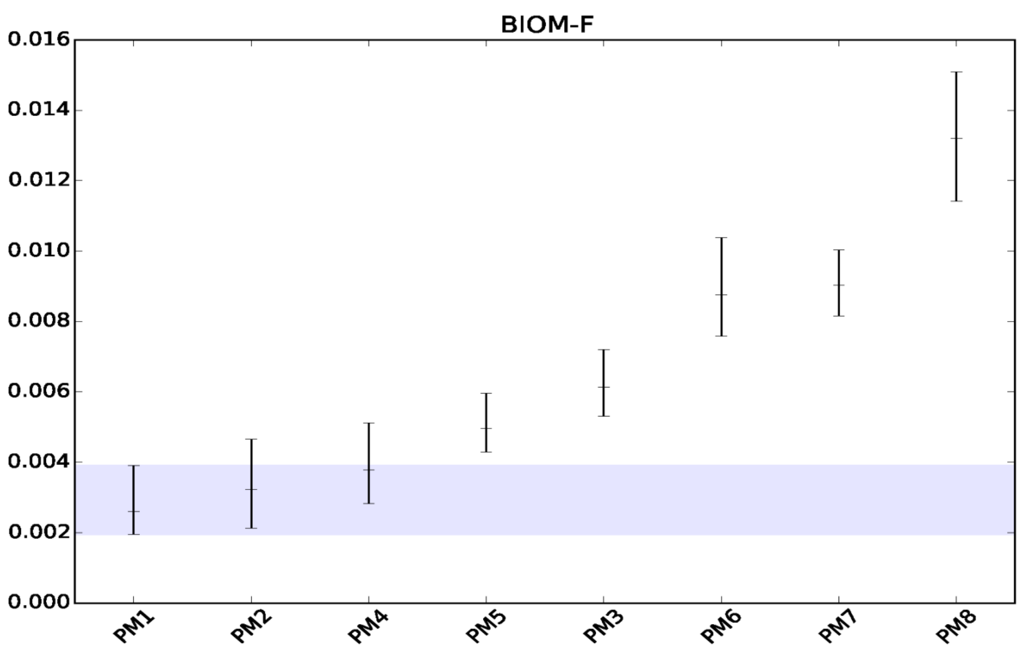


Figure 80: Confidence interval plot of SAPV distances of BIOM-F research group

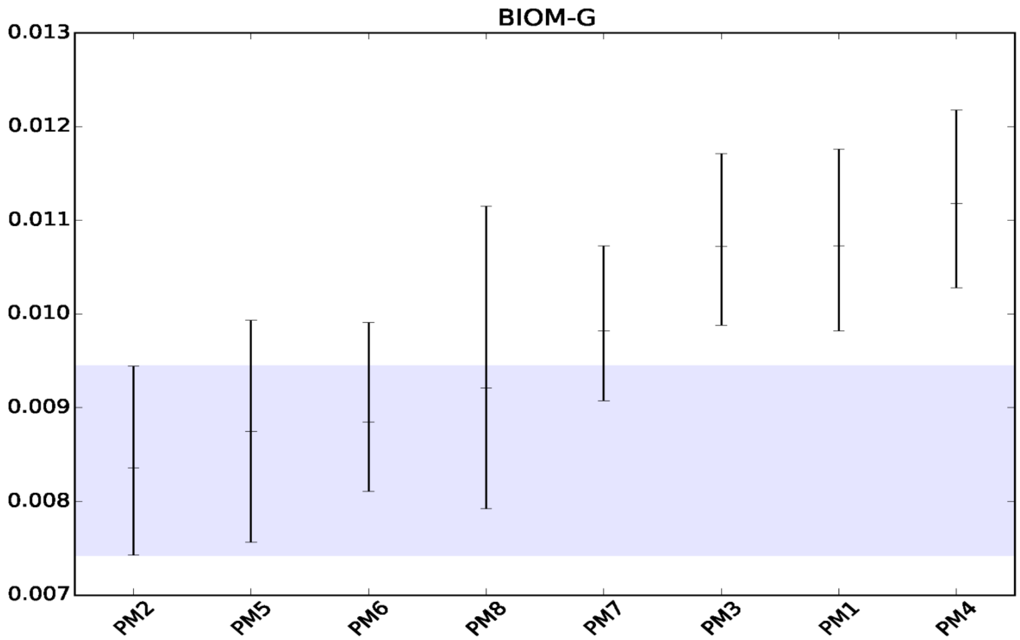


Figure 81: Confidence interval plot of SAPV distances of BIOM-G research group

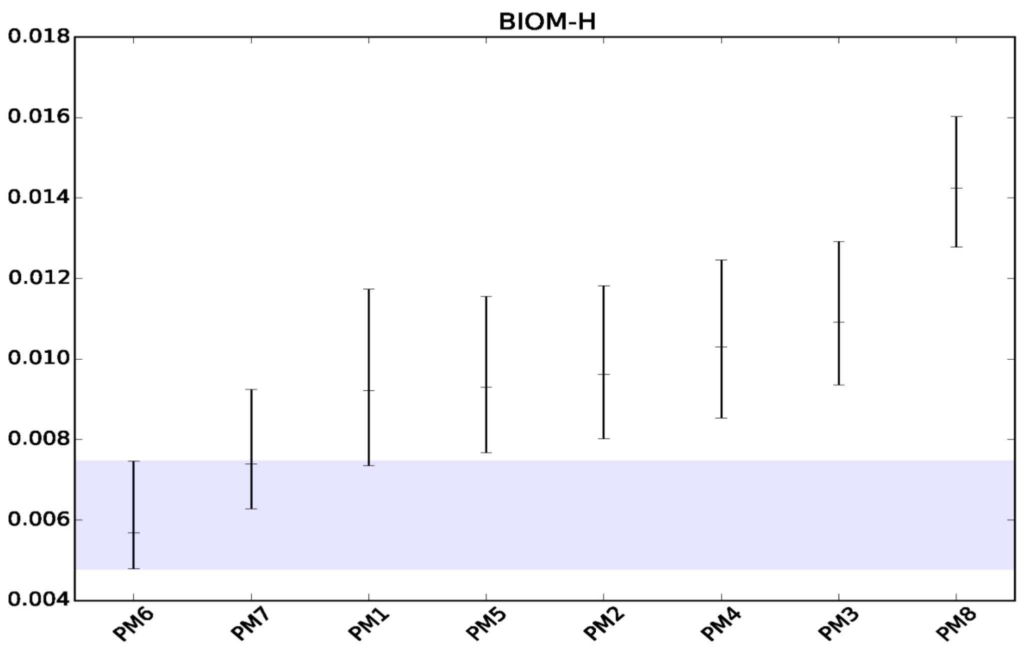


Figure 82: Confidence interval plot of SAPV distances of BIOM-H research group

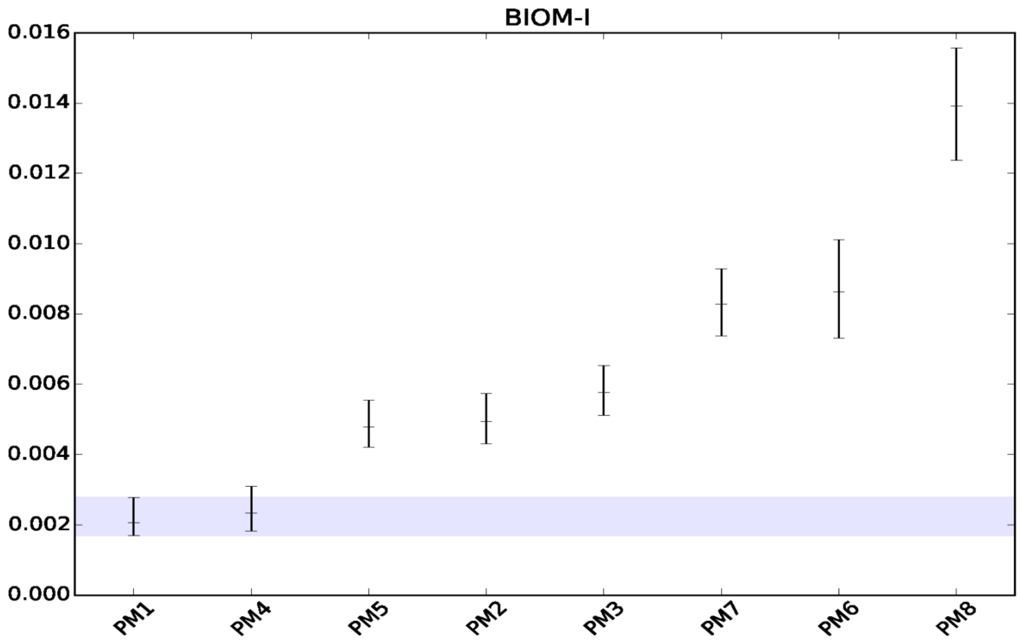


Figure 83: Confidence interval plot of SAPV distances of BIOM-I research group

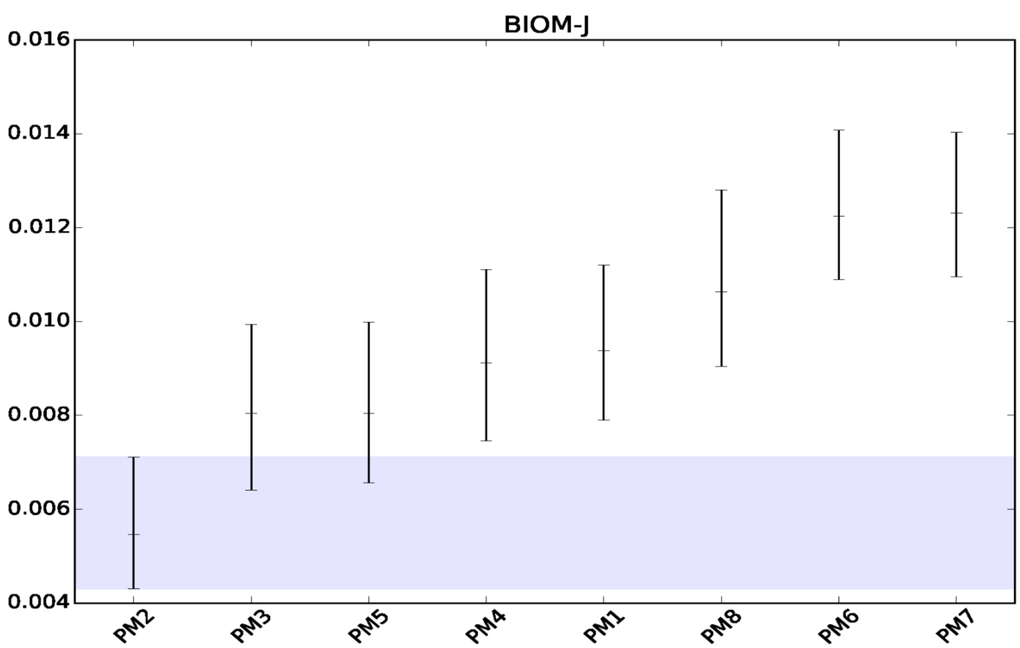


Figure 84: Confidence interval plot of SAPV distances of BIOM-J research group

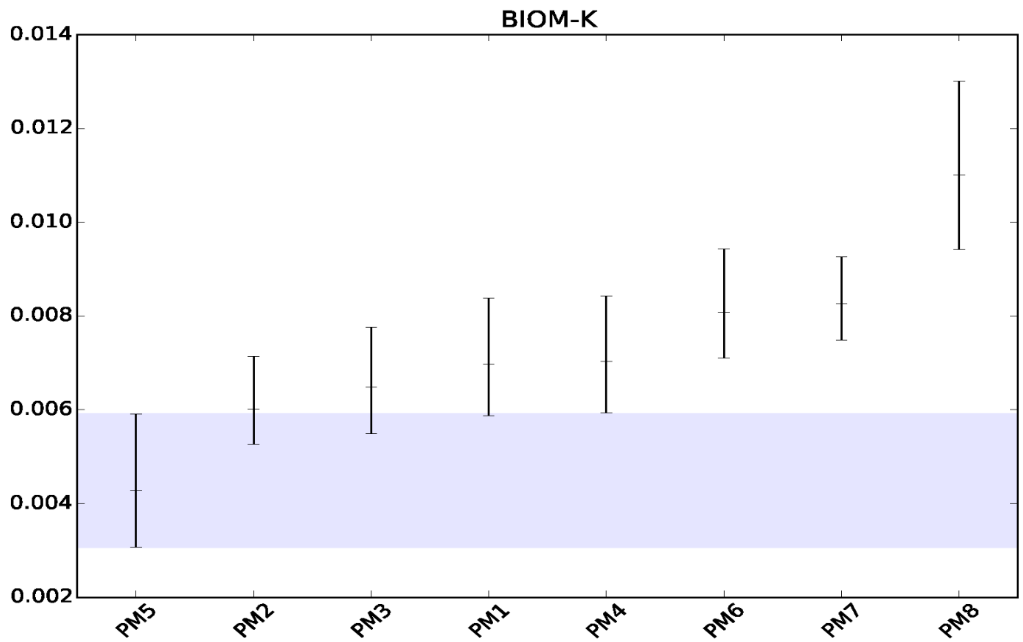


Figure 85: Confidence interval plot of SAPV distances of BIOM-K research group

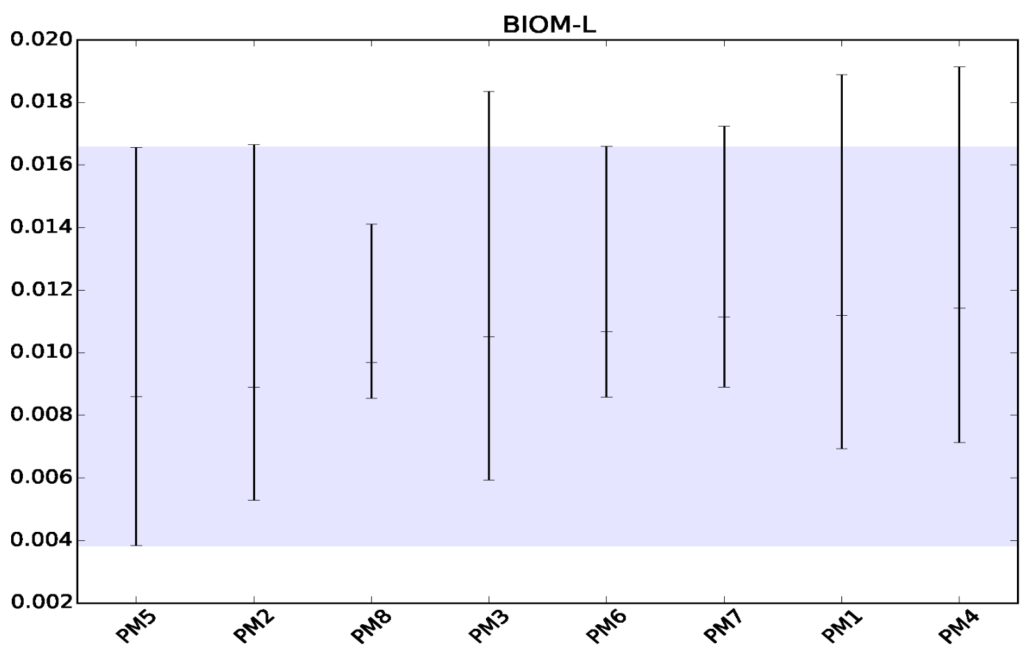


Figure 86: Confidence interval plot of SAPV distances of BIOM-L research group

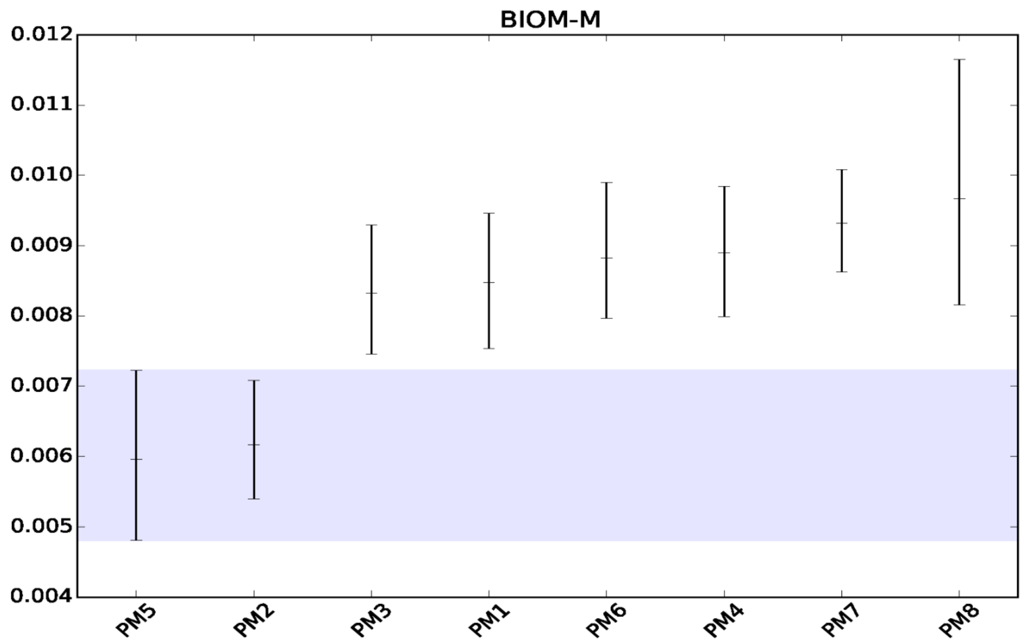


Figure 87: Confidence interval plot of SAPV distances of BIOM-M research group

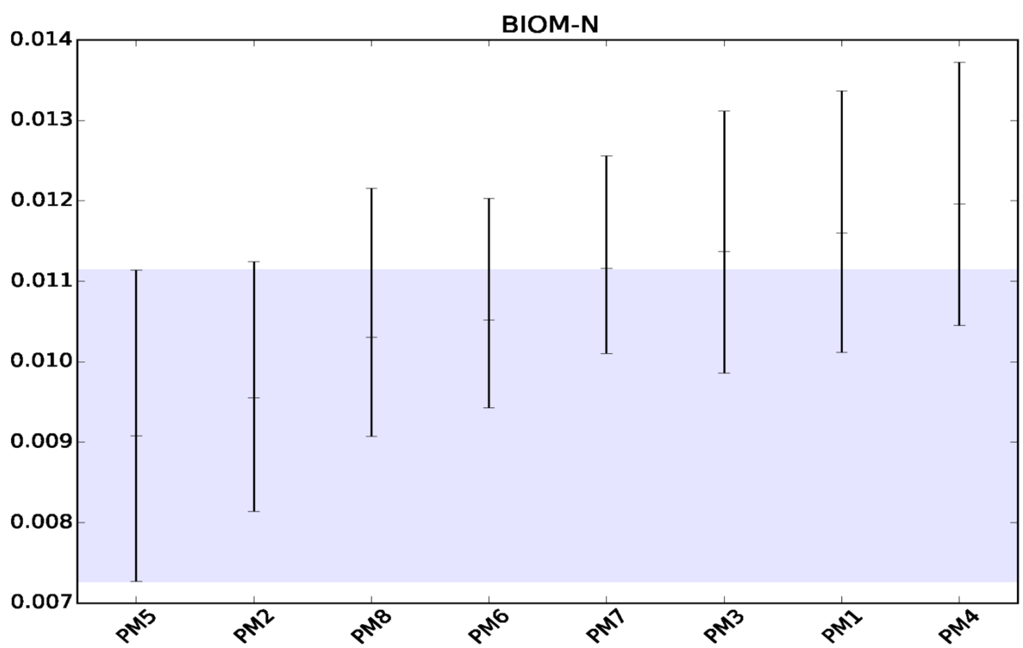


Figure 88: Confidence interval plot of SAPV distances of BIOM-N research group

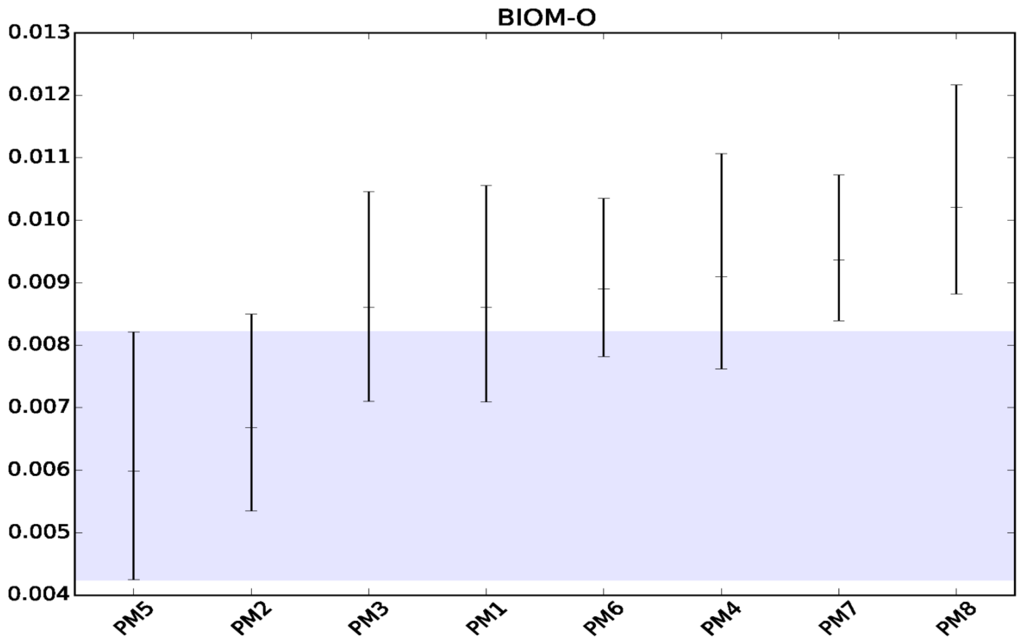


Figure 89: Confidence interval plot of SAPV distances of BIOM-O research group

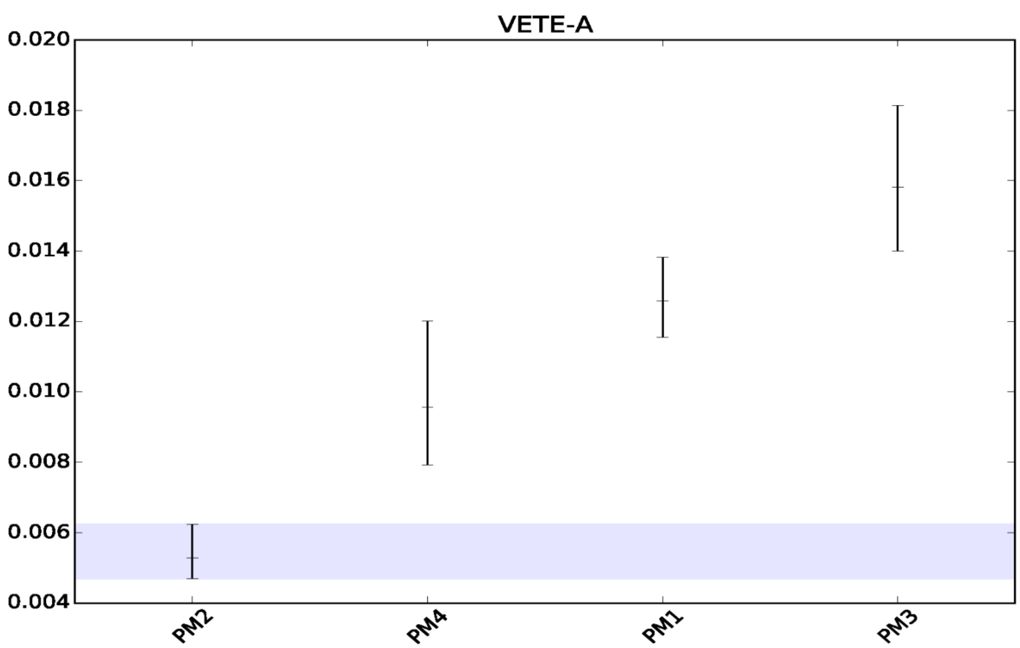


Figure 90: Confidence interval plot of SAPV distances of VETE-A research group

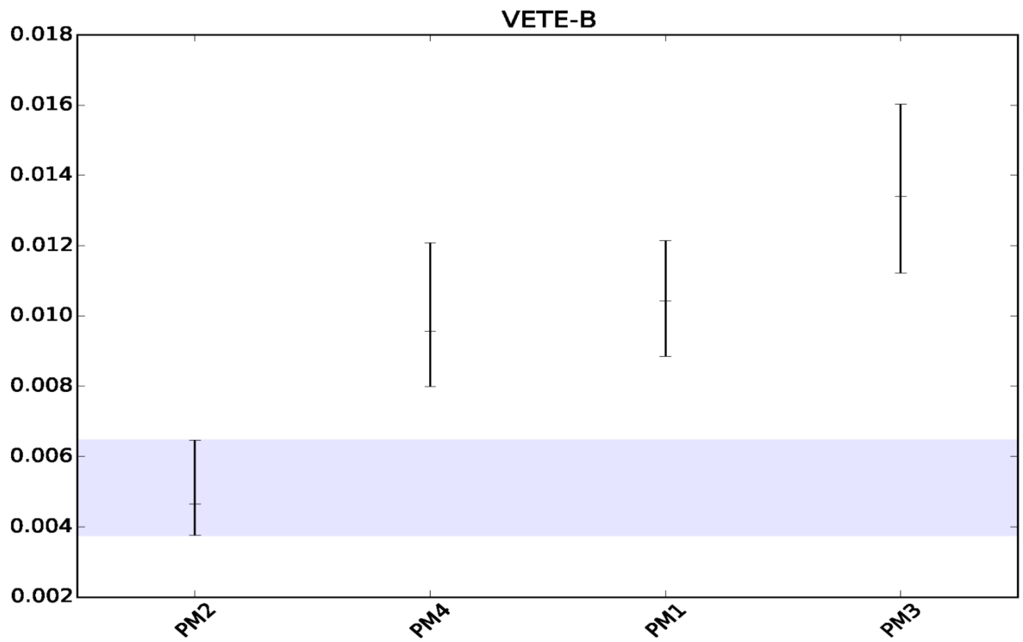


Figure 91: Confidence interval plot of SAPV distances of VETE-B research group

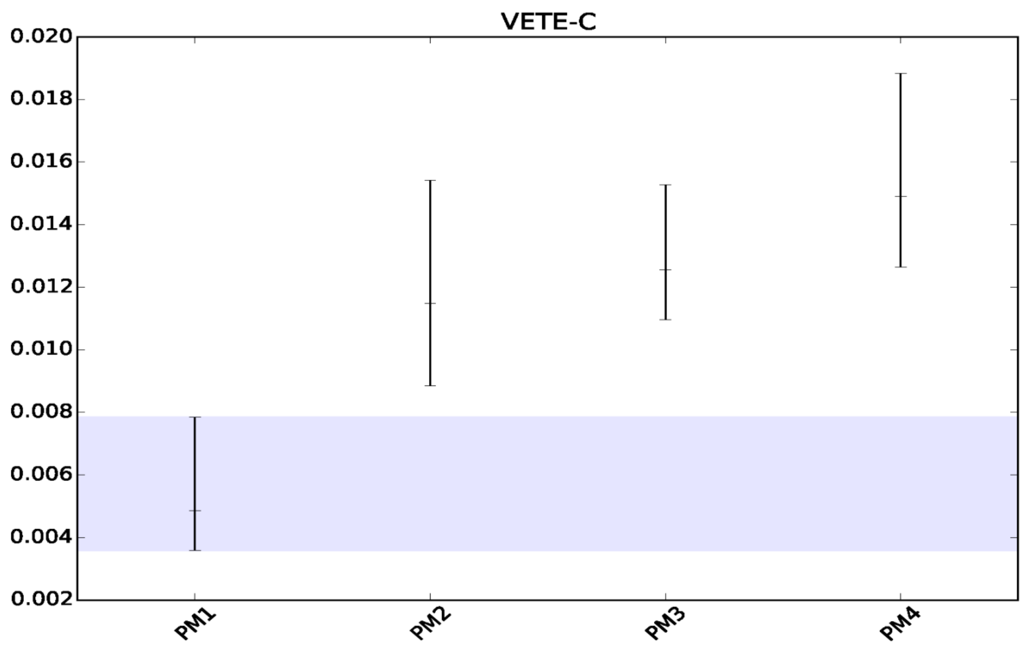


Figure 92: Confidence interval plot of SAPV distances of VETE-C research group

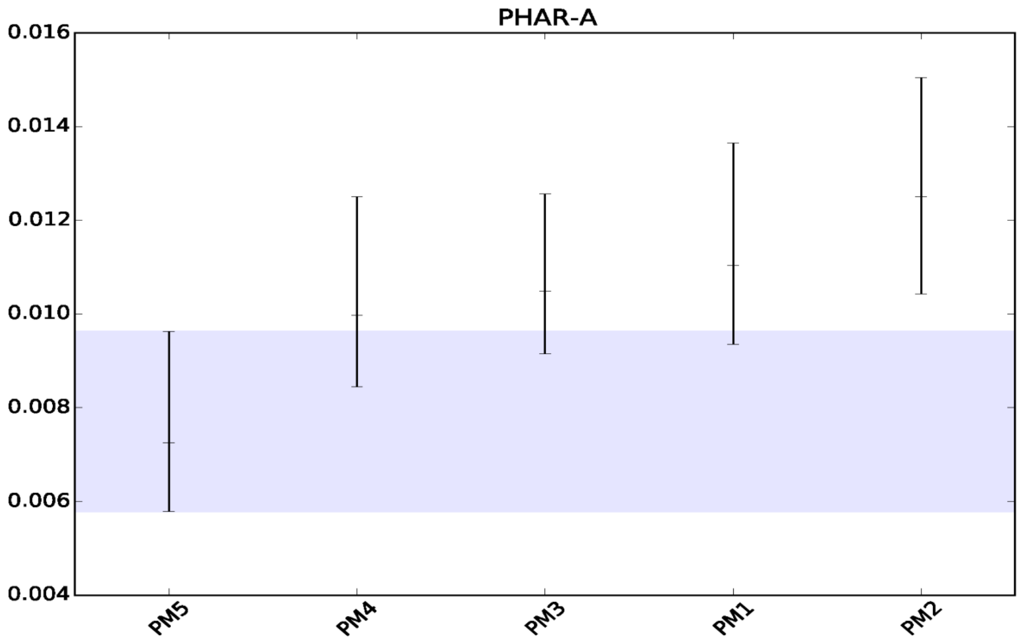


Figure 93: Confidence interval plot of SAPV distances of PHAR-A research group

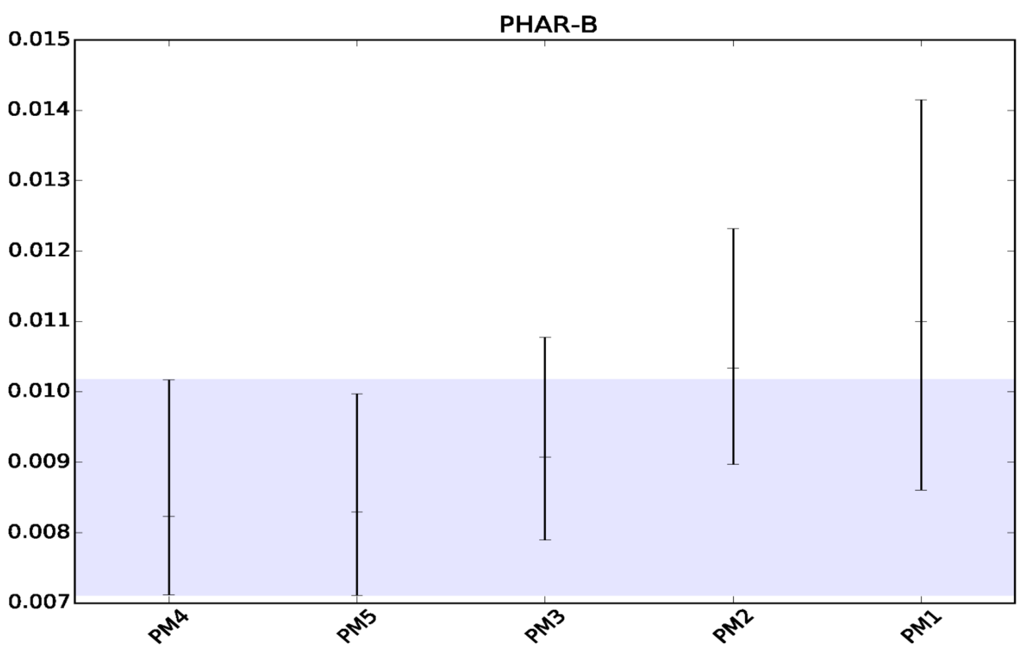


Figure 94: Confidence interval plot of SAPV distances of PHAR-B research group

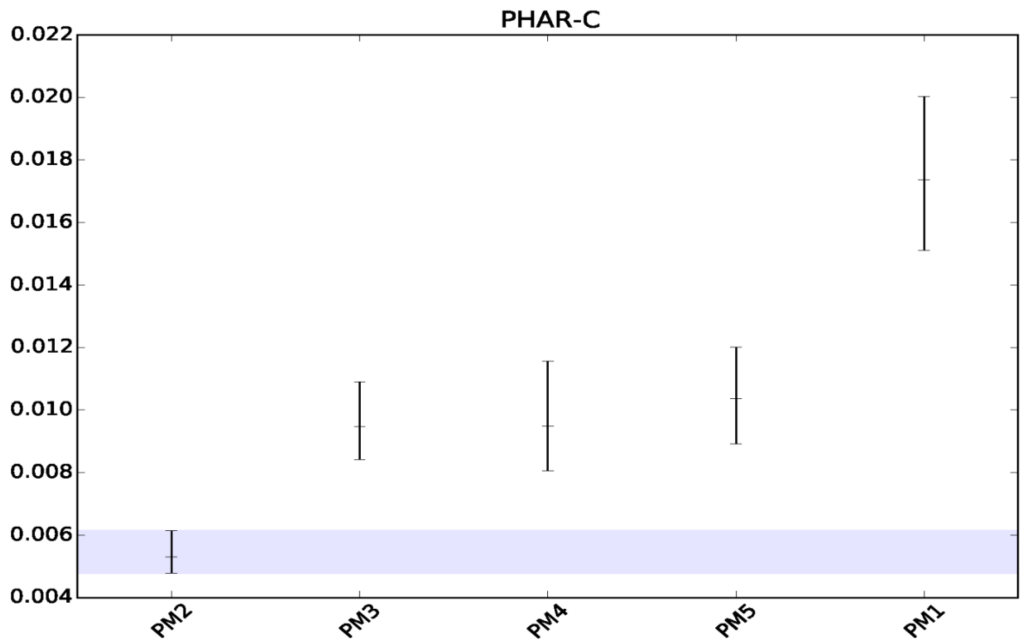


Figure 95: Confidence interval plot of SAPV distances of PHAR-C research group

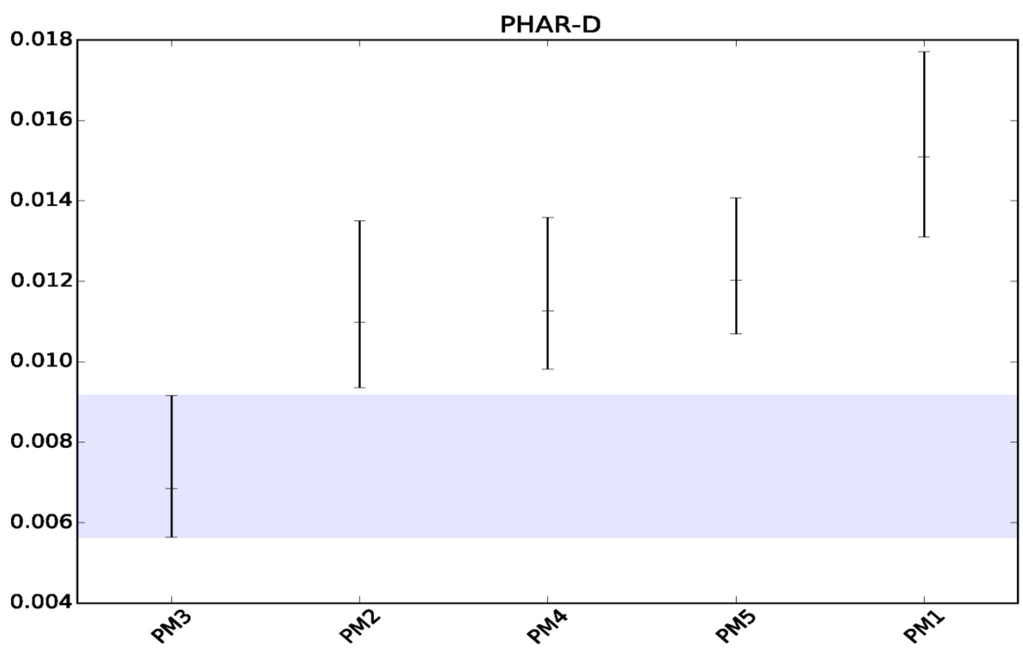


Figure 96: Confidence interval plot of SAPV distances of PHAR-D research group

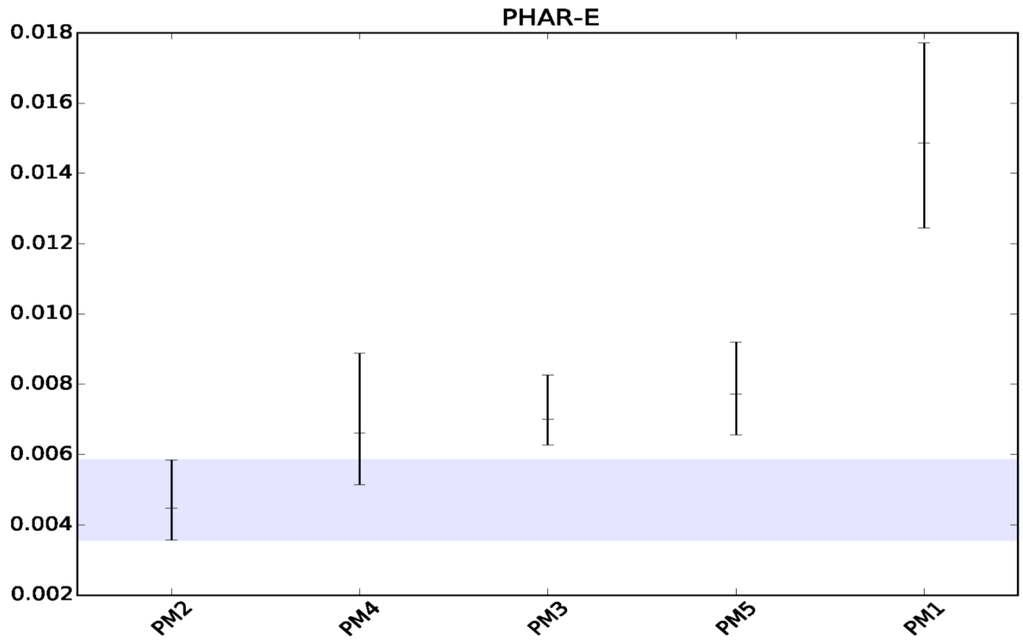


Figure 97: Confidence interval plot of SAPV distances of PHAR-E research group

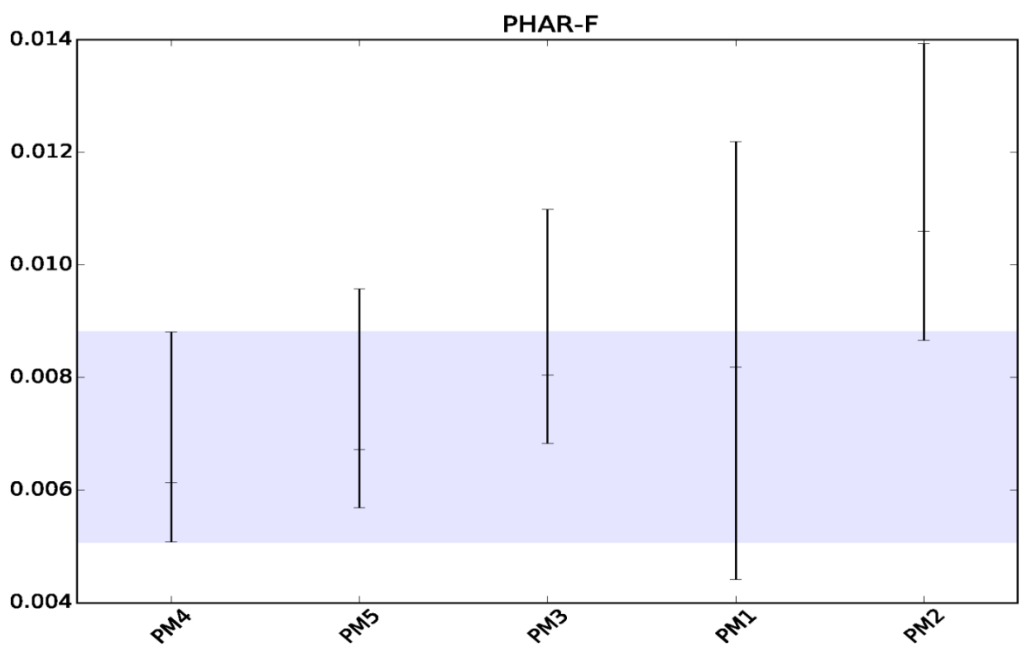


Figure 98: Confidence interval plot of SAPV distances of PHAR-F research group

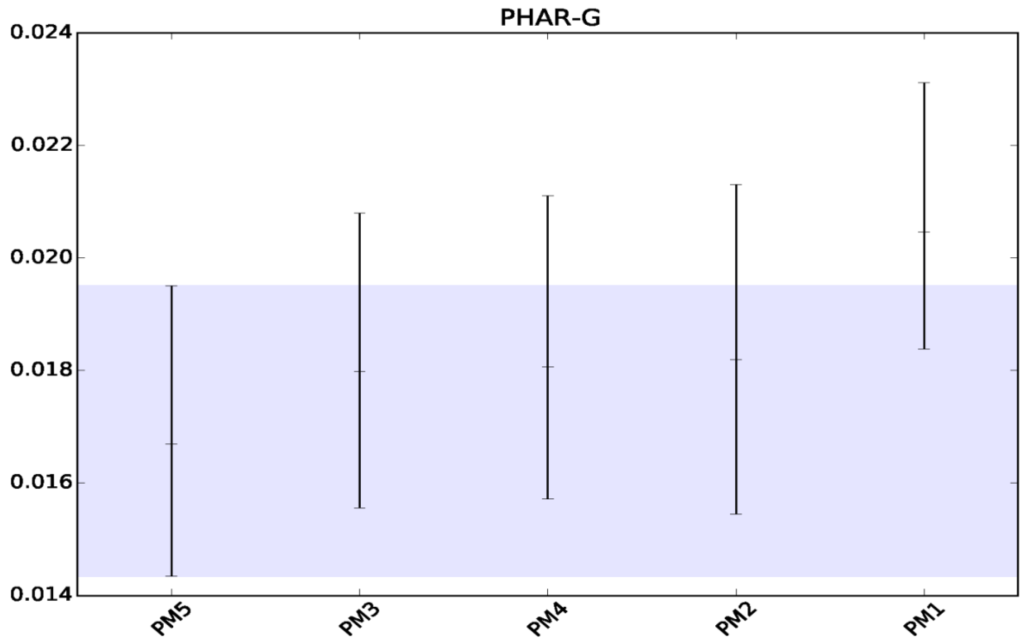


Figure 99: Confidence interval plot of SAPV distances of PHAR-G research group

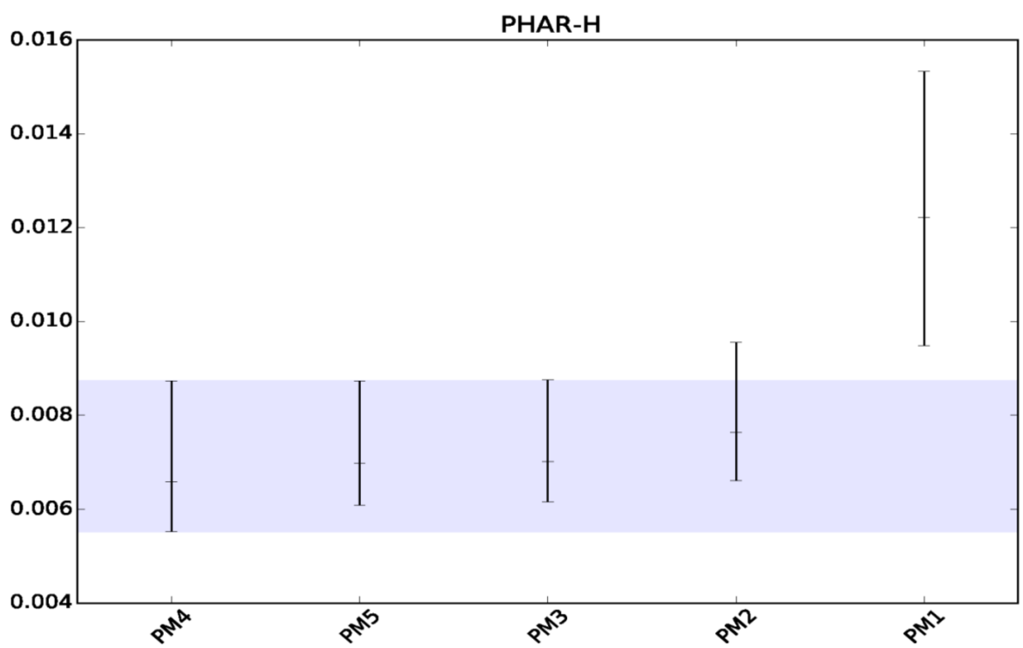


Figure 100: Confidence interval plot of SAPV distances of PHAR-H research group

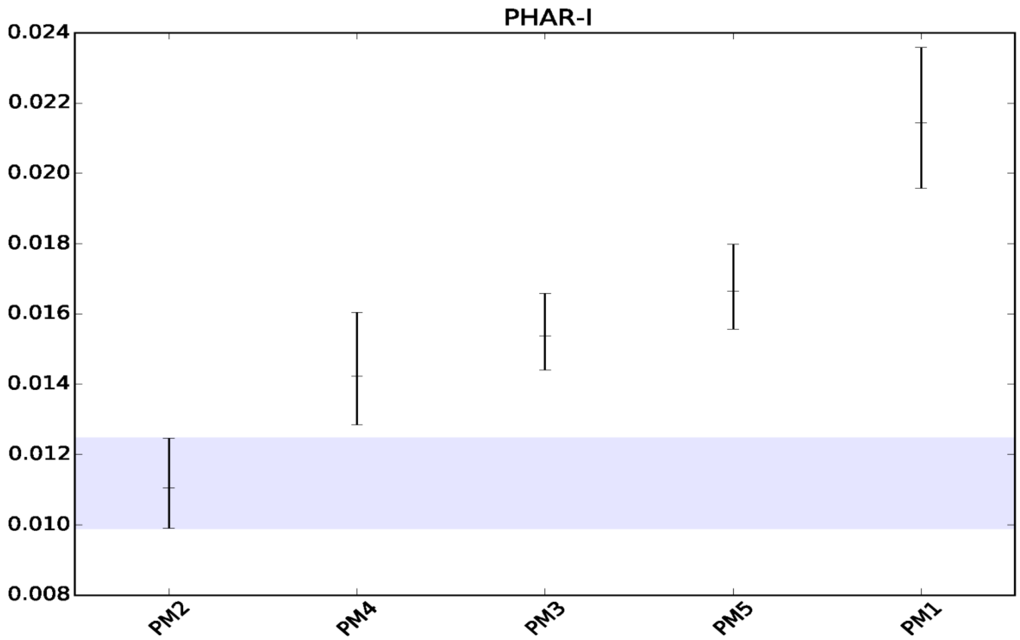


Figure 101: Confidence interval plot of SAPV distances of PHAR-I research group

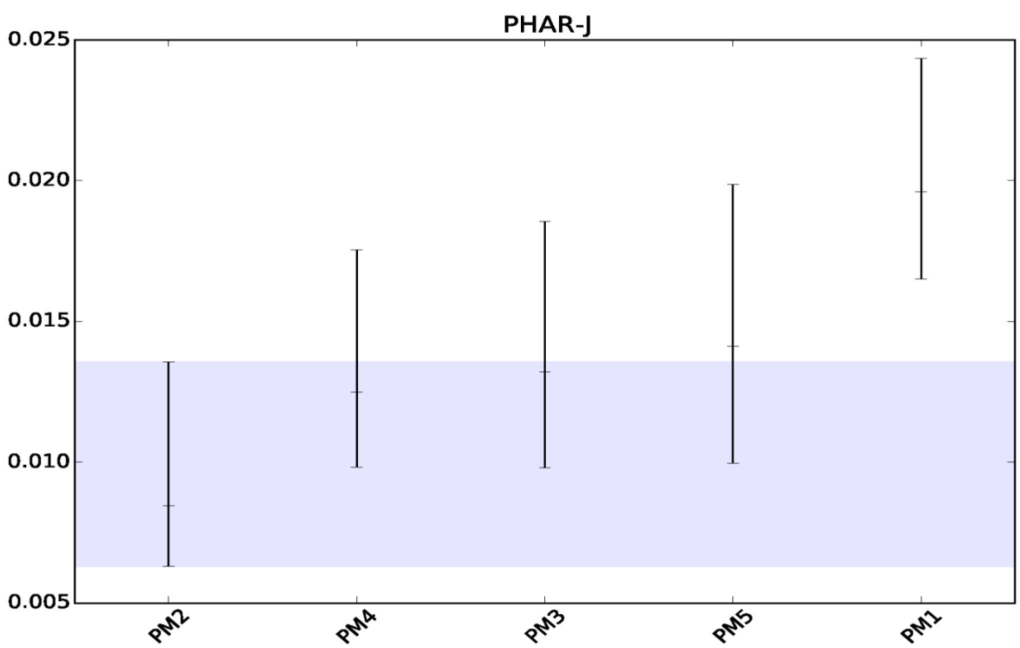


Figure 102: Confidence interval plot of SAPV distances of PHAR-J research group

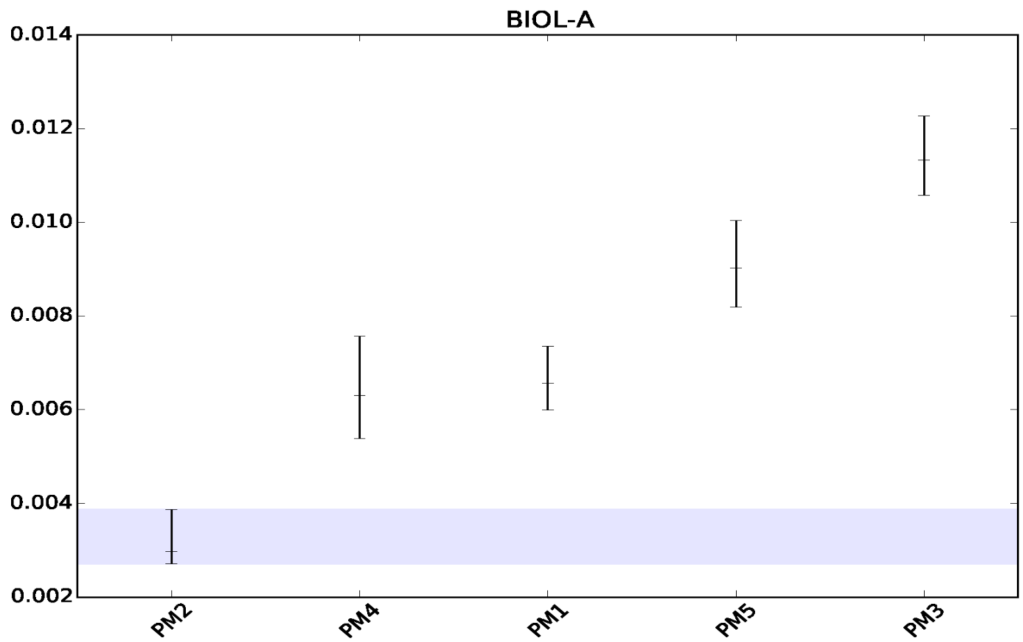


Figure 103: Confidence interval plot of SAPV distances of BIOL-A research group

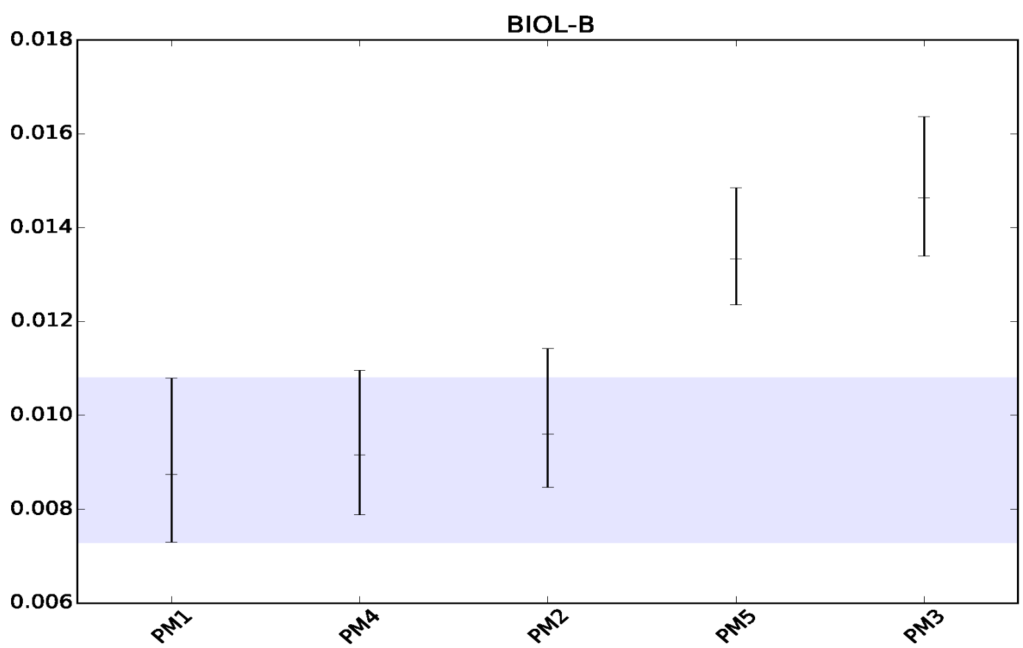


Figure 104: Confidence interval plot of SAPV distances of BIOL-B research group

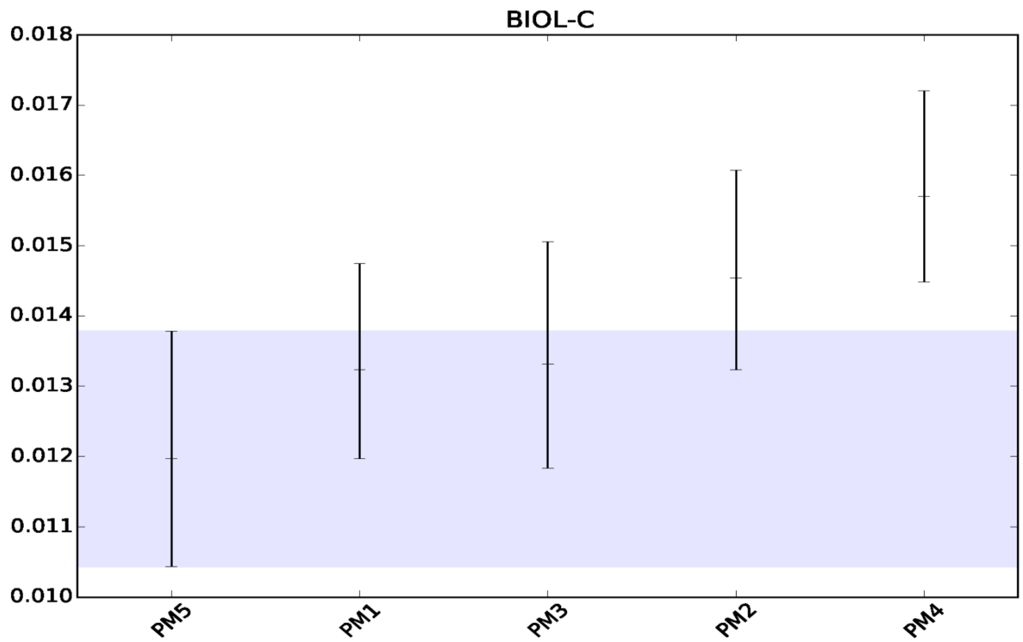


Figure 105: Confidence interval plot of SAPV distances of BIOL-C research group

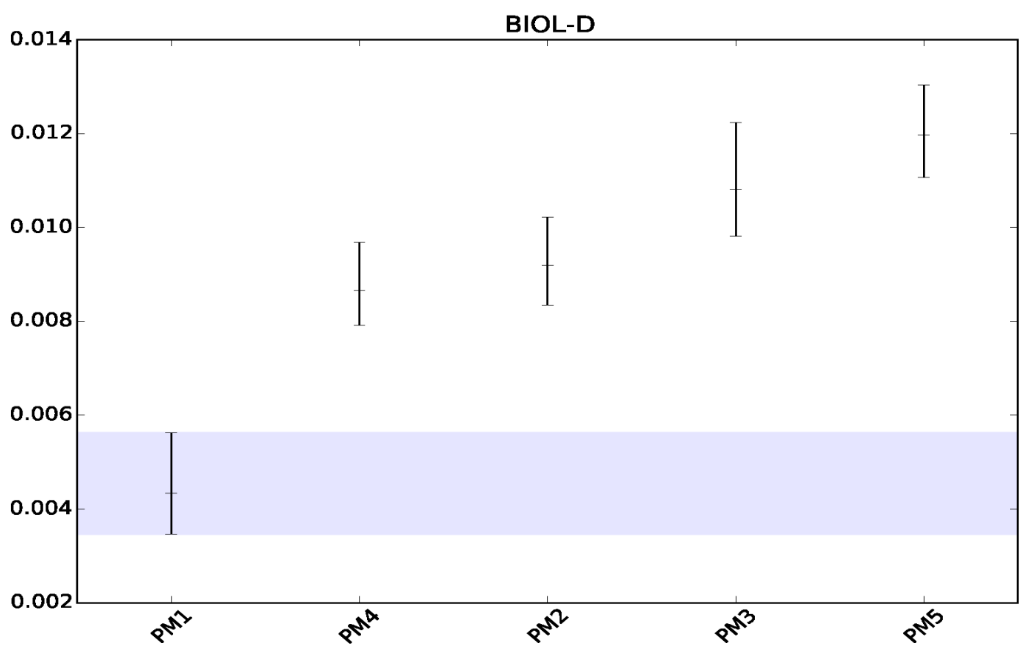


Figure 106: Confidence interval plot of SAPV distances of BIOL-D research group

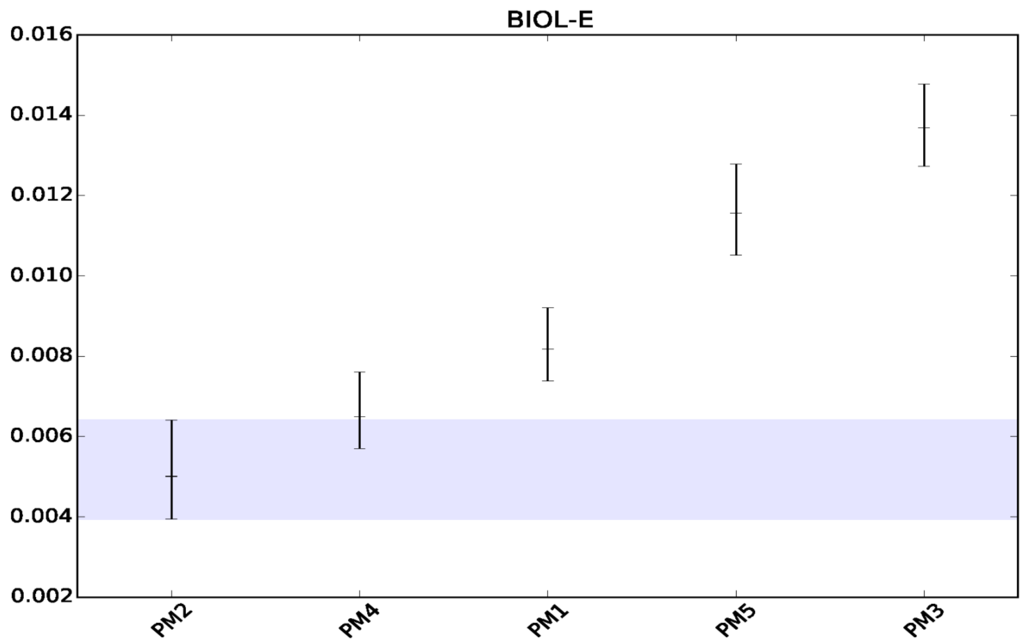


Figure 107: Confidence interval plot of SAPV distances of BIOL-E research group

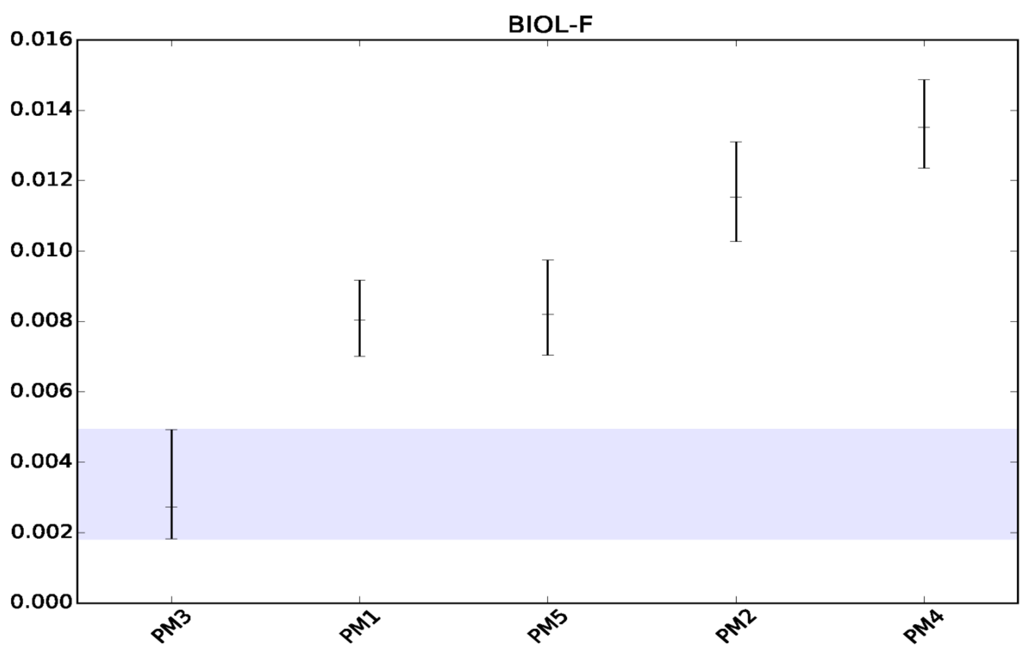


Figure 108: Confidence interval plot of SAPV distances of BIOL-F research group

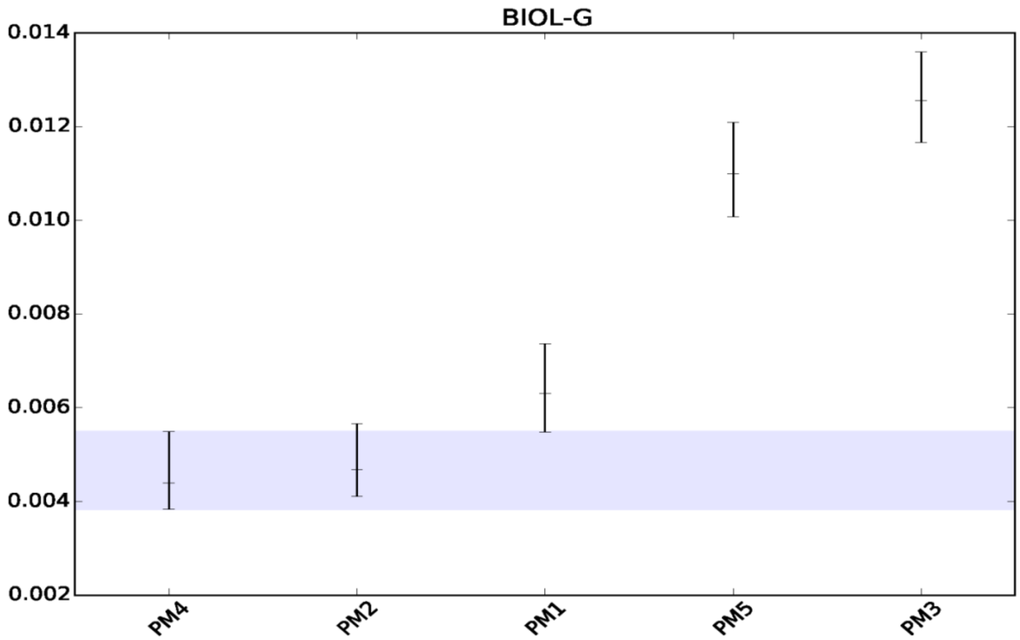


Figure 109: Confidence interval plot of SAPV distances of BIOL-G research group

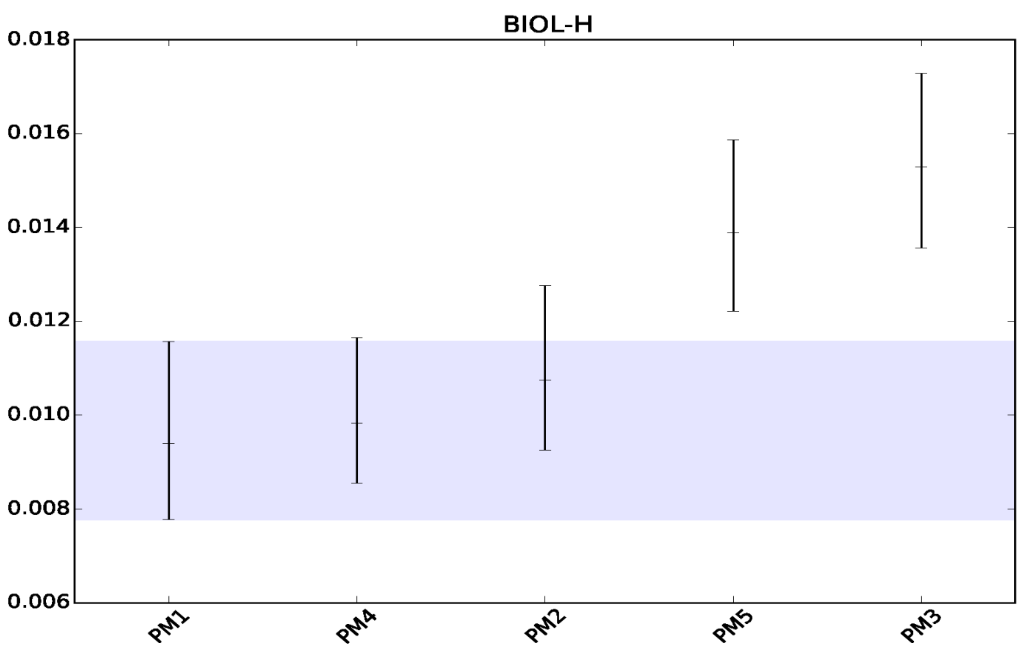


Figure 110: Confidence interval plot of SAPV distances of BIOL-H research group

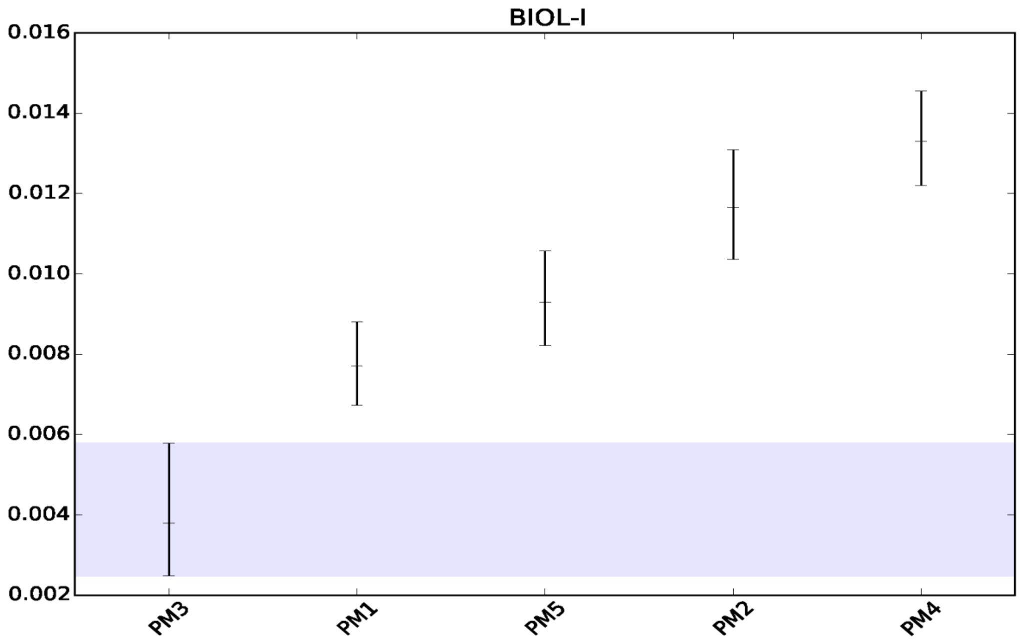


Figure 111: Confidence interval plot of SAPV distances of BIOL-I research group

Appendix C: Result of six approaches

Biomedical Sciences department

The main assessor is bold and underlined in our methods. Distances/similarities whose CIs overlap with that of the shortest distance/ highest similarities are in bold (same column). This is applicable for Table 59 to Table 88.

Table 59: Euclidean distances between barycenters of Biomedical Sciences panel members and individual research groups using the 2-dimensional base map of WoS SCs

	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
PM1	0.251	**0.120**	**0.121**	**0.075**	**0.098**	**0.125**	0.513	0.208	**0.119**	0.309	0.255	0.400	0.446	0.463	0.376
PM2	**<u>0.078</u>**	**0.164**	**0.149**	0.198	0.172	**0.147**	**0.245**	**0.120**	0.345	**0.227**	**<u>0.046</u>**	**0.133**	**0.177**	**0.196**	**0.108**
PM3	0.182	**0.086**	**0.141**	0.096	0.133	**0.113**	0.472	0.240	**0.125**	**<u>0.185</u>**	0.202	0.365	0.404	0.426	0.340
PM4	0.383	0.269	0.306	0.247	0.286	0.291	0.672	0.405	**<u>0.091</u>**	0.355	0.402	0.563	0.604	0.625	0.539
PM5	0.193	**<u>0.063</u>**	**<u>0.069</u>**	**<u>0.017</u>**	**<u>0.047</u>**	**<u>0.067</u>**	0.457	**0.166**	**0.148**	**0.261**	0.197	**0.344**	0.389	0.408	0.320
PM6	**0.111**	**0.153**	**0.122**	0.178	0.146	**0.131**	0.272	**<u>0.073</u>**	0.332	**0.260**	**0.081**	**0.159**	**0.206**	**0.222**	**0.136**
PM7	0.210	**0.132**	**0.075**	0.111	0.074	**0.108**	0.414	**0.086**	0.249	0.327	0.197	**0.303**	0.350	0.364	0.281
PM8	**0.142**	0.272	0.273	0.314	0.296	0.263	**<u>0.161</u>**	0.248	0.445	**0.228**	**0.133**	**<u>0.093</u>**	**<u>0.102</u>**	**<u>0.127</u>**	**<u>0.085</u>**

Table 60: Euclidean distances between SAPVs of Biomedical Sciences panel members and individual research groups using the similarity matrix of WoS SCs

	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
PM1	0.055	**0.027**	0.024	**0.040**	**<u>0.021</u>**	**<u>0.023</u>**	0.078	**0.049**	**0.016**	0.068	**0.038**	**0.067**	0.068	0.085	0.066
PM2	**<u>0.028</u>**	0.042	0.043	0.060	0.042	**0.026**	**<u>0.039</u>**	0.054	0.048	**<u>0.039</u>**	**<u>0.030</u>**	**<u>0.028</u>**	**<u>0.028</u>**	**0.044**	**<u>0.027</u>**
PM3	0.061	**0.027**	0.034	0.059	0.043	**0.044**	0.077	0.069	0.041	0.067	**0.037**	**0.068**	0.066	0.082	0.065
PM4	0.058	0.036	0.037	**0.041**	0.036	**0.034**	0.083	0.06	**<u>0.013</u>**	**0.062**	**0.042**	**0.073**	0.074	0.090	0.072
PM5	0.056	**<u>0.011</u>**	**<u>0.012</u>**	0.047	**0.022**	**0.028**	0.077	0.057	0.023	0.068	**0.032**	**0.065**	0.065	0.082	0.062
PM6	0.042	0.046	0.04	0.050	**0.026**	**0.033**	0.057	**<u>0.033</u>**	0.047	0.063	**0.040**	**0.052**	0.052	0.066	0.051
PM7	0.073	0.064	0.057	**<u>0.033</u>**	0.052	0.061	0.090	0.057	0.059	0.092	0.065	0.086	0.086	0.100	0.085
PM8	**0.046**	0.092	0.092	0.101	0.092	0.079	**0.040**	0.087	0.097	**0.057**	0.075	**0.047**	0.047	**<u>0.030</u>**	**0.047**

Table 61: Euclidean distances between barycenters of Biomedical Sciences panel members and individual research groups using 2-dimensional base map of journals

	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
PM1	0.350	0.180	0.224	**0.110**	0.242	**0.081**	0.473	0.319	**0.159**	0.445	0.387	**0.471**	0.397	0.436	0.344
PM2	**0.176**	**0.038**	**0.046**	0.201	0.177	**0.119**	**0.302**	0.267	**0.234**	**0.297**	**0.208**	**0.294**	**0.221**	**0.272**	**0.181**
PM3	0.390	0.397	0.397	0.241	0.530	0.303	0.611	0.621	**0.194**	0.356	0.438	0.586	0.527	0.599	0.522
PM4	0.391	0.355	0.365	**0.168**	0.479	0.243	0.600	0.568	**0.119**	0.390	0.440	0.580	0.515	0.582	0.498
PM5	**0.250**	**0.058**	**0.107**	**0.183**	**0.144**	**0.095**	0.348	**0.233**	**0.227**	0.371	**0.280**	**0.348**	**0.274**	0.311	**0.220**
PM6	**0.295**	0.177	0.184	0.383	**0.072**	0.295	**0.236**	**0.086**	0.426	0.442	**0.291**	**0.258**	**0.207**	**0.187**	**0.135**
PM7	0.367	0.173	0.217	0.282	**0.103**	**0.209**	0.395	**0.148**	0.331	0.500	0.385	**0.407**	0.342	0.348	0.271
PM8	**0.171**	0.363	0.314	0.497	0.445	0.445	**0.238**	0.502	0.504	**0.154**	**0.140**	**0.199**	**0.213**	**0.271**	0.281

Table 62: Euclidean distances between SAPVs of Biomedical Sciences panel members and individual research groups using the similarity matrix of journals

	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
PM1	0.007	0.006	0.006	0.007	0.007	**0.003**	0.011	**0.009**	**0.002**	0.009	**0.007**	**0.011**	0.008	**0.012**	**0.009**
PM2	**0.004**	0.006	0.007	0.008	0.007	**0.003**	**0.008**	0.010	0.005	**0.005**	**0.006**	**0.009**	**0.006**	**0.010**	**0.007**
PM3	0.007	0.006	0.007	0.008	0.008	0.006	0.011	0.011	0.006	**0.008**	**0.006**	**0.011**	0.008	**0.011**	**0.009**
PM4	0.007	0.007	0.007	0.007	0.008	**0.004**	0.011	0.010	**0.002**	0.009	0.007	**0.011**	0.009	**0.012**	**0.009**
PM5	**0.005**	**0.002**	**0.003**	0.007	0.006	0.005	**0.009**	0.009	0.005	**0.008**	**0.004**	**0.009**	**0.006**	**0.009**	**0.006**
PM6	0.008	0.008	0.007	0.008	**0.003**	0.009	**0.009**	**0.006**	0.009	0.012	0.008	**0.011**	0.009	**0.011**	**0.009**
PM7	0.008	0.008	0.007	**0.005**	0.007	0.009	**0.010**	**0.007**	0.008	0.012	0.008	**0.011**	0.009	**0.011**	0.009
PM8	0.009	0.012	0.013	0.014	0.013	0.013	**0.009**	0.014	0.014	0.011	0.011	**0.010**	0.010	**0.010**	0.010

Table 63: WCS values of the Biomedical Sciences panel members and individual research groups using the similarity matrix of WoS SCs

	BIOM-A	BIOM-B	BIOM-C	BIOM-D	BIOM-E	BIOM-F	BIOM-G	BIOM-H	BIOM-I	BIOM-J	BIOM-K	BIOM-L	BIOM-M	BIOM-N	BIOM-O
PM1	0.692	0.887	0.887	0.765	**0.909**	**0.942**	0.572	**0.705**	**0.976**	0.628	**0.838**	**0.644**	0.642	0.512	0.643
PM2	**0.869**	0.82	0.787	0.614	0.792	**0.934**	**0.881**	**0.629**	0.809	**0.909**	**0.897**	**0.897**	**0.914**	**0.832**	**0.899**
PM3	0.626	0.923	0.883	0.613	0.822	0.805	0.565	0.533	0.845	0.591	**0.85**	**0.616**	0.656	0.535	0.650
PM4	0.650	0.841	0.831	0.772	**0.850**	**0.890**	0.513	**0.619**	**0.984**	0.637	**0.805**	**0.596**	0.582	0.456	0.583
PM5	0.677	**0.987**	**0.973**	0.691	**0.914**	**0.874**	0.609	**0.592**	0.935	0.634	**0.909**	**0.691**	0.714	0.596	0.724
PM6	0.686	0.689	0.703	0.602	**0.915**	0.770	0.629	**0.735**	0.718	0.573	0.712	0.618	0.635	0.550	0.623
PM7	0.484	0.575	0.631	**0.880**	0.664	0.582	0.370	**0.577**	0.655	0.326	0.575	0.372	0.385	0.289	0.385
PM8	**0.907**	0.418	0.397	0.292	0.410	0.546	0.793	0.374	0.374	0.745	0.579	**0.747**	0.761	**0.775**	0.753

Table 64: WCS values of the Biomedical Sciences panel members and individual research groups using the similarity matrix of journals

	BIOM -A	BIOM -B	BIOM -C	BIOM -D	BIOM -E	BIOM -F	BIOM -G	BIOM -H	BIOM -I	BIOM -J	BIOM -K	BIOM -L	BIOM -M	BIOM -N	BIOM -O
PM1	**0.649**	0.569	0.601	0.653	0.671	**0.901**	0.501	**0.541**	**0.955**	0.510	0.637	**0.437**	**0.587**	**0.414**	**0.557**
PM2	**0.791**	0.567	0.541	0.544	0.585	**0.893**	**0.670**	0.477	0.801	**0.828**	0.646	**0.561**	0.712	0.521	0.647
PM3	0.609	0.589	0.583	0.556	0.546	0.694	0.466	**0.391**	0.736	0.564	0.636	**0.470**	0.574	0.411	0.539
PM4	0.576	0.520	0.540	0.636	0.565	**0.846**	0.438	**0.436**	**0.926**	0.488	0.602	**0.390**	0.53	0.369	0.497
PM5	0.599	**0.926**	**0.910**	0.508	0.693	0.669	**0.596**	**0.426**	0.705	0.508	**0.861**	**0.633**	**0.773**	**0.643**	**0.781**
PM6	0.445	0.407	0.474	0.415	**0.906**	0.462	0.382	**0.570**	0.501	0.259	0.423	**0.301**	0.386	0.307	0.376
PM7	0.451	0.376	0.436	**0.779**	0.468	0.484	0.381	**0.452**	0.568	0.292	0.433	**0.313**	0.397	0.306	0.382
PM8	**0.732**	0.190	0.160	0.156	0.158	0.244	0.387	0.130	0.205	0.365	0.281	**0.337**	0.353	0.307	0.314

Chemistry department

Table 65: Euclidean distances between barycenters of Chemistry panel members and individual research groups using the 2-dimensional base map of WoS SCs

	CHEM-A	CHEM- B	CHEM-C	CHEM- D	CHEM-E	CHEM- F	CHEM- G	CHEM- H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM 1	0.167	**0.129**	0.217	**0.165**	0.329	0.337	0.179	0.165	0.111	0.394	0.454	**0.127**
PM 2	0.350	0.342	0.362	**0.129**	**0.079**	**0.090**	0.145	0.215	0.199	**0.259**	**0.228**	0.342
PM 3	0.171	0.161	0.192	**0.129**	0.252	**0.263**	**0.053**	**0.061**	**0.020**	**0.269**	0.330	0.161
PM 4	0.269	0.262	0.280	**0.108**	**0.158**	**0.170**	**0.063**	**0.134**	0.121	**0.232**	**0.250**	0.263
PM 5	**0.056**	**0.055**	**0.091**	0.232	0.367	0.378	0.154	**0.093**	**0.099**	**0.315**	0.411	**0.057**
PM 6	0.302	0.276	0.335	**0.027**	**0.175**	**0.181**	0.161	0.210	0.156	0.366	0.370	0.275
PM 7	**0.116**	**0.072**	0.172	0.235	0.395	0.404	0.216	0.178	0.144	0.410	0.491	**0.070**

Table 66: Euclidean distances between SAPVs of Chemistry panel members and individual research groups using the similarity matrix of WoS SCs

	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM1	0.081	0.079	0.108	**0.061**	0.124	0.119	0.116	0.104	0.093	0.129	0.141	0.085
PM2	0.082	0.074	0.079	**0.054**	**0.036**	**0.032**	0.055	**0.046**	**0.036**	**0.075**	**0.071**	0.070
PM3	0.082	0.074	0.08	0.066	**0.057**	**0.058**	**0.040**	**0.040**	**0.042**	**0.075**	**0.086**	0.073
PM4	0.106	0.099	0.104	0.085	0.064	0.070	**0.027**	0.063	0.071	**0.085**	0.094	0.091
PM5	**0.015**	**0.013**	**0.034**	0.074	0.100	0.102	0.077	**0.053**	**0.050**	**0.082**	0.096	**0.024**
PM6	0.093	0.087	0.111	**0.025**	0.085	0.080	0.096	0.090	0.080	0.113	0.116	0.088
PM7	0.068	0.068	0.097	0.072	0.128	0.125	0.113	0.099	0.089	0.125	0.140	0.075

Table 67: Euclidean distances between barycenters of Chemistry panel members and individual research groups using the 2-dimensional base map of journals

	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM1	**0.115**	**0.103**	0.276	**0.149**	0.422	0.312	0.281	0.272	0.187	0.291	0.483	**0.122**
PM2	0.414	0.404	0.452	0.266	**0.090**	**0.091**	0.217	0.274	0.226	**0.216**	**0.282**	0.332
PM3	0.211	0.202	0.276	**0.138**	0.261	**0.153**	**0.111**	**0.141**	**0.027**	**0.120**	**0.305**	**0.131**
PM4	0.286	0.277	0.332	**0.178**	**0.196**	**0.100**	**0.115**	**0.167**	**0.098**	**0.120**	**0.270**	0.204
PM5	**0.109**	**0.110**	**0.124**	0.241	0.412	0.307	**0.149**	**0.109**	**0.128**	**0.156**	**0.327**	**0.067**
PM6	0.173	0.159	0.316	**0.071**	0.349	**0.242**	0.256	0.263	0.152	**0.266**	0.457	0.139
PM7	**0.096**	**0.083**	0.255	**0.157**	0.423	0.312	0.267	0.255	0.176	0.277	0.468	**0.104**

Table 68: Euclidean distances between SAPVs of Chemistry panel members and individual research groups using the similarity matrix of journals

	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM1	0.015	0.015	0.029	**0.011**	0.027	0.025	0.027	0.028	0.025	0.028	0.030	0.018
PM2	0.017	0.017	0.017	0.022	**0.005**	**0.008**	0.011	**0.010**	**0.010**	**0.012**	**0.010**	0.012
PM3	0.019	0.018	**0.017**	0.024	0.009	**0.011**	0.009	**0.008**	**0.010**	**0.012**	**0.012**	0.014
PM4	0.019	0.019	0.019	0.023	0.011	**0.006**	**0.005**	**0.010**	**0.008**	**0.009**	0.016	0.015
PM5	**0.006**	**0.006**	**0.013**	0.019	0.016	0.017	0.017	0.013	0.014	0.017	0.015	**0.008**
PM6	0.017	0.017	0.030	**0.005**	0.025	0.024	0.026	0.027	0.025	0.027	0.029	0.019
PM7	0.012	0.011	0.026	**0.012**	0.024	0.022	0.024	0.024	0.021	0.024	0.026	0.015

Table 69: WCS values of the Chemistry panel members and individual research groups using the similarity matrix of WoS SCs

	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM1	0.709	0.667	0.445	**0.922**	0.469	0.449	0.395	0.440	0.507	0.323	0.273	0.661
PM2	0.670	0.713	0.726	0.675	**0.914**	**0.945**	0.837	**0.847**	**0.947**	**0.703**	**0.527**	0.713
PM3	0.594	0.655	0.673	0.569	**0.839**	0.831	0.866	**0.880**	**0.894**	**0.711**	**0.403**	0.604
PM4	0.459	0.517	0.504	0.484	0.781	0.777	**0.951**	**0.758**	0.769	**0.626**	0.315	0.549
PM5	**0.983**	**0.990**	**0.842**	0.669	0.581	0.475	0.614	0.747	0.758	**0.573**	**0.512**	**0.933**
PM6	0.613	0.600	0.377	**0.973**	0.545	0.519	0.391	0.410	0.484	0.294	0.280	0.603
PM7	0.758	0.713	0.503	0.850	0.460	0.439	0.440	0.494	0.550	0.373	0.290	0.700

Table 70: WCS values of the Chemistry panel members and individual research groups using the similarity matrix of journals

	CHEM-A	CHEM-B	CHEM-C	CHEM-D	CHEM-E	CHEM-F	CHEM-G	CHEM-H	CHEM-I	CHEM-J	CHEM-K	CHEM-L
PM1	0.670	0.633	0.144	0.730	0.336	0.244	0.179	0.127	0.220	0.099	0.091	**0.580**
PM2	0.373	0.397	**0.432**	0.315	**<u>0.830</u>**	**0.805**	0.649	**0.546**	**0.795**	**0.513**	**<u>0.308</u>**	0.470
PM3	0.235	0.247	**<u>0.549</u>**	0.176	0.558	0.585	0.726	**<u>0.774</u>**	**0.793**	**0.481**	**0.243**	0.284
PM4	0.249	0.281	0.370	0.231	**0.724**	**<u>0.877</u>**	**<u>0.850</u>**	**0.613**	**<u>0.805</u>**	**<u>0.594</u>**	**0.184**	0.319
PM5	**<u>0.941</u>**	**<u>0.944</u>**	**0.452**	0.393	0.371	0.229	0.231	0.285	0.313	0.166	**0.295**	**<u>0.702</u>**
PM6	0.534	0.538	0.110	**<u>0.957</u>**	0.358	0.257	0.165	0.111	0.198	0.099	0.076	0.495
PM7	0.727	0.689	0.186	0.599	0.345	0.261	0.212	0.162	0.253	0.12	0.117	**0.614**

Pharmaceuticals Sciences department

Table 71: Euclidean distances between barycenters of Pharmaceutical Sciences panel members and individual research groups using the 2-dimensional base map of WoS SCs

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	**<u>0.007</u>**	0.266	0.670	**0.318**	0.538	**0.189**	**0.144**	0.353	0.757	0.624
PM2	0.451	**0.194**	**<u>0.232</u>**	**0.130**	**<u>0.095</u>**	0.283	0.358	**<u>0.141</u>**	**<u>0.322</u>**	**<u>0.188</u>**
PM3	**0.162**	**0.118**	0.513	**0.159**	0.379	**0.082**	**<u>0.107</u>**	**0.207**	0.602	0.467
PM4	0.197	**<u>0.076</u>**	0.476	**<u>0.127</u>**	0.344	**<u>0.057</u>**	**0.148**	**0.165**	0.564	0.430
PM5	**0.017**	0.274	0.678	**0.328**	0.547	**0.194**	**0.156**	0.359	0.765	0.632

Table 72: Euclidean distances between SAPVs of Pharmaceutical Sciences panel members and individual research groups using the similarity matrix of WoS SCs

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	**0.033**	0.055	0.097	**0.046**	0.072	**0.046**	**0.091**	0.07	**0.157**	0.094
PM2	0.062	**<u>0.031</u>**	**<u>0.045</u>**	**<u>0.023</u>**	**<u>0.020</u>**	**0.041**	**0.094**	**<u>0.034</u>**	**<u>0.124</u>**	**<u>0.049</u>**
PM3	0.037	**0.036**	0.069	**0.023**	0.044	**0.036**	**0.086**	**0.048**	**0.138**	**0.067**
PM4	**0.036**	**0.035**	0.072	**0.023**	0.046	**<u>0.031</u>**	**0.087**	**0.047**	**0.138**	**0.070**
PM5	**<u>0.018</u>**	0.048	0.091	0.047	0.065	0.047	**<u>0.084</u>**	0.061	**0.156**	**0.085**

Table 73: Euclidean distances between barycenters of Pharmaceutical Sciences panel members and individual research groups using the 2-dimensional base map of journals

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	**<u>0.101</u>**	0.267	1.017	0.413	0.807	**0.271**	**0.262**	0.471	1.251	0.972
PM2	0.750	0.581	**<u>0.205</u>**	0.428	**<u>0.021</u>**	0.579	0.689	0.398	**<u>0.429</u>**	**<u>0.162</u>**
PM3	**0.339**	**0.163**	0.610	**<u>0.043</u>**	0.402	**0.162**	**0.332**	**<u>0.110</u>**	0.844	**0.573**
PM4	**0.332**	**0.161**	0.616	**0.052**	0.408	**0.160**	**0.322**	**0.120**	0.850	**0.577**
PM5	**0.186**	**<u>0.057</u>**	0.773	**0.170**	0.566	**<u>0.062</u>**	**<u>0.242</u>**	**0.233**	1.008	0.735

Table 74: Euclidean distances between SAPVs of Pharmaceutical Sciences panel members and individual research groups using the similarity matrix of journals

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	**0.011**	**0.011**	0.017	0.015	0.015	**0.008**	**0.020**	0.012	0.021	0.020
PM2	0.012	**0.010**	**0.005**	0.011	**0.004**	**0.011**	**0.018**	**0.008**	**0.011**	**0.008**
PM3	**0.010**	**0.009**	0.009	**0.007**	0.007	**0.008**	**0.018**	**0.007**	0.015	**0.013**
PM4	**0.010**	**0.008**	0.009	0.011	**0.007**	**0.006**	**0.018**	**0.007**	0.014	**0.012**
PM5	**0.007**	**0.008**	0.010	0.012	0.008	**0.007**	**0.017**	**0.007**	0.017	**0.014**

Table 75: WCS values of the Pharmaceuticals panel members and individual research groups using the similarity matrix of WoS SCs

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	0.823	**0.819**	0.426	**0.881**	0.694	**0.909**	**0.387**	0.740	0.092	**0.375**
PM2	0.702	**0.902**	**0.747**	**0.956**	0.914	**0.840**	**0.391**	**0.897**	**0.274**	**0.623**
PM3	0.820	**0.884**	**0.632**	**0.941**	0.856	**0.886**	**0.407**	**0.850**	**0.203**	**0.558**
PM4	0.812	**0.877**	0.570	**0.929**	0.836	**0.923**	**0.404**	**0.835**	**0.192**	**0.488**
PM5	**0.962**	0.742	0.440	0.685	0.679	0.719	**0.531**	0.675	0.095	**0.488**

Table 76: WCS values of the Pharmaceuticals Sciences panel members and individual research groups using the similarity matrix of journals

	PHAR-A	PHAR-B	PHAR-C	PHAR-D	PHAR-E	PHAR-F	PHAR-G	PHAR-H	PHAR-I	PHAR-J
PM1	0.391	**0.502**	0.206	0.298	0.445	**0.850**	0.103	**0.521**	0.075	0.125
PM2	0.286	**0.408**	**0.645**	0.469	**0.717**	0.456	**0.161**	**0.518**	**0.299**	**0.478**
PM3	0.266	**0.365**	0.421	**0.871**	0.527	0.442	0.121	**0.462**	0.160	0.263
PM4	0.374	**0.520**	0.458	0.354	0.711	**0.847**	0.128	**0.621**	0.207	0.300
PM5	**0.812**	**0.469**	0.444	0.241	0.578	0.527	**0.271**	0.471	0.153	0.325

Physics department

Table 77: Euclidean distances between barycenters of Physics panel members and individual research groups using the 2-dimensional base map of WoS SCs

	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
PM 1	1.173	0.123	0.215	**0.017**	0.145	**0.208**	0.495	0.120	0.664
PM 2	1.195	0.067	0.109	0.158	0.118	0.316	0.443	0.056	0.688
PM 3	**1.041**	0.146	0.194	0.116	0.113	**0.104**	**0.387**	0.157	**0.532**
PM 4	**1.020**	0.168	**0.085**	0.263	0.132	0.295	**0.249**	0.179	**0.522**
PM 5	**1.136**	**0.046**	**0.055**	0.159	**0.069**	0.281	**0.385**	**0.050**	**0.629**
PM 6	**1.157**	**0.031**	0.084	0.138	**0.078**	0.280	**0.412**	**0.026**	0.649

Table 78: Euclidean distances between SAPVs of Physics panel members and individual research groups using the similarity matrix of WoS SCs

	PHYS-A	PHYS- B	PHYS-C	PHYS- D	PHYS-E	PHYS- F	PHYS- G	PHYS- H	PHYS-I
PM 1	0.376	0.358	0.373	**0.098**	0.328	0.301	0.371	0.358	0.367
PM 2	0.172	0.019	0.038	0.272	0.054	0.127	0.115	**0.019**	0.133
PM 3	**0.156**	0.065	0.080	0.256	0.069	**0.100**	0.116	0.063	**0.111**
PM 4	**0.144**	0.060	0.039	0.271	0.051	0.129	**0.066**	0.063	**0.103**
PM 5	**0.157**	0.023	**0.016**	0.271	**0.044**	0.125	**0.095**	0.027	**0.115**
PM 6	**0.165**	**0.012**	0.035	0.258	**0.037**	**0.111**	**0.106**	**0.015**	0.125

Table 79: Euclidean distances between barycenters of Physics panel members and individual research groups using the 2-dimensional base map of journals

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	1.134	0.154	0.310	**0.030**	0.204	**0.087**	0.813	0.146	0.707
PM2	1.045	**0.025**	0.159	0.127	**0.063**	0.185	0.668	**0.015**	**0.600**
PM3	**0.960**	0.086	0.185	0.151	**0.090**	**0.155**	**0.647**	0.098	**0.527**
PM4	**0.857**	0.301	0.146	0.427	0.252	0.461	**0.369**	0.309	**0.404**
PM5	**0.970**	0.085	**0.074**	0.211	**0.036**	0.251	**0.577**	0.094	**0.519**
PM6	**1.029**	**0.023**	0.142	0.142	**0.045**	0.195	**0.650**	**0.023**	**0.582**

Table 80: Euclidean distances between SAPVs of Physics panel members and individual research groups using the similarity matrix of journals

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.061	0.052	0.057	**0.018**	0.048	0.048	0.060	0.052	0.058
PM2	0.030	0.004	0.016	0.035	0.013	0.017	0.028	**0.004**	0.024
PM3	0.027	0.019	0.020	0.041	0.019	0.022	0.027	0.020	0.022
PM4	**0.021**	0.020	0.011	0.042	0.015	0.020	**0.012**	0.021	**0.015**
PM5	**0.022**	0.009	**0.005**	0.038	**0.008**	**0.015**	0.018	0.010	**0.014**
PM6	0.026	**0.002**	0.011	0.034	**0.008**	**0.012**	0.022	**0.003**	0.019

Table 81: WCS values of the Physics panel members and individual research groups using the similarity matrix of WoS SCs

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.030	0.155	0.043	<u>**0.996**</u>	0.561	0.508	0.028	0.154	0.052
PM2	**0.151**	**0.982**	0.920	0.127	0.806	0.513	0.543	**0.977**	**0.497**
PM3	<u>**0.220**</u>	0.714	0.625	0.211	0.668	0.526	0.440	0.762	**0.544**
PM4	**0.182**	0.729	0.829	0.129	0.757	0.436	<u>**0.895**</u>	0.741	**0.479**
PM5	**0.182**	0.965	<u>**0.986**</u>	0.158	**0.852**	0.475	0.656	0.957	<u>**0.567**</u>
PM6	**0.164**	<u>**0.989**</u>	0.930	0.272	<u>**0.903**</u>	<u>**0.643**</u>	0.631	<u>**0.985**</u>	**0.516**

Table 82: WCS values of the Physics panel members and individual research groups using the similarity matrix of journals

	PHYS-A	PHYS-B	PHYS-C	PHYS-D	PHYS-E	PHYS-F	PHYS-G	PHYS-H	PHYS-I
PM1	0.019	0.186	0.095	<u>**0.990**</u>	**0.688**	0.372	0.035	0.191	0.063
PM2	0.086	**0.982**	0.745	0.244	**0.708**	**0.542**	0.234	**0.975**	0.368
PM3	<u>**0.252**</u>	0.306	0.320	0.102	0.301	0.246	0.131	0.310	0.323
PM4	0.083	0.397	0.582	0.102	0.482	0.310	<u>**0.822**</u>	0.377	0.298
PM5	**0.152**	0.880	<u>**0.927**</u>	0.247	**0.764**	**0.517**	0.389	0.865	<u>**0.531**</u>
PM6	**0.106**	<u>**0.983**</u>	0.779	0.329	<u>**0.795**</u>	<u>**0.665**</u>	0.337	<u>**0.980**</u>	**0.420**

Veterinary Sciences department

Table 83: Euclidean distances between barycenters of Veterinary Sciences panel members and individual research groups using the 2-dimensional base map of WoS SCs

	VETE-A	VETE-B	VETE-C
PM1	**0.138**	**0.111**	**0.059**
PM2	<u>**0.092**</u>	<u>**0.068**</u>	**0.103**
PM3	**0.131**	**0.104**	<u>**0.052**</u>
PM4	**0.114**	**0.140**	0.272

Table 84: Euclidean distances between SAPVs of Veterinary Sciences panel members and individual research groups using the similarity matrix of WoS SCs

	VETE-A	VETE-B	VETE-C
PM1	0.083	0.066	<u>**0.028**</u>
PM2	<u>**0.048**</u>	<u>**0.041**</u>	0.072
PM3	0.106	0.085	0.062
PM4	**0.053**	**0.058**	0.106

Table 85: Euclidean distances between barycenters of Veterinary Sciences panel members and individual research groups using the 2-dimensional base map of journals

	VETE-A	VETE-B	VETE-C
PM1	0.260	**0.160**	**0.124**
PM2	**0.141**	**0.108**	**0.227**
PM3	0.273	**0.182**	**0.145**
PM4	0.272	0.310	0.469

Table 86: Euclidean distances between SAPVs of Veterinary Sciences panel members and individual research groups using the similarity matrix of journals

	VETE-A	VETE-B	VETE-C
PM1	0.013	0.010	**0.005**
PM2	**0.005**	**0.005**	0.011
PM3	0.016	0.013	0.013
PM4	0.010	0.010	0.015

Table 87: WCS values of the Veterinary Sciences panel members and individual research groups using the similarity matrix of WoS SCs

	VETE-A	VETE-B	VETE-C
PM1	0.460	**0.783**	**0.982**
PM2	**0.722**	**0.818**	0.602
PM3	0.321	**0.732**	0.815
PM4	**0.674**	0.616	0.263

Table 88: WCS values of the Veterinary Sciences panel members and individual research groups using the similarity matrix of journals

	VETE-A	VETE-B	VETE-C
PM1	0.268	**0.534**	**0.941**
PM2	**0.690**	**0.729**	0.286
PM3	0.232	**0.496**	0.399
PM4	0.410	0.414	0.115

Appendix D: Heat map with hierarchical clustering

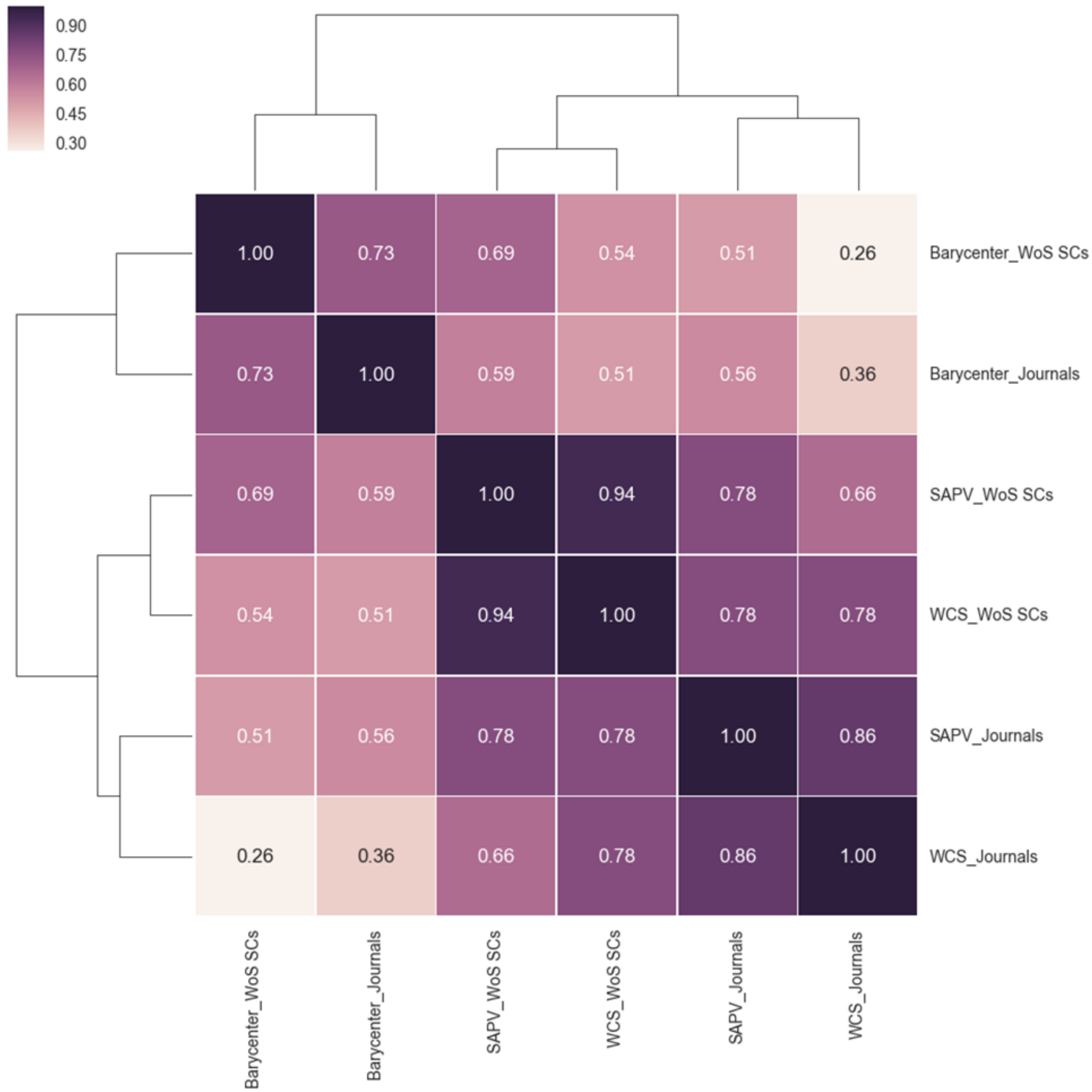


Figure 112: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Biomedical sciences department

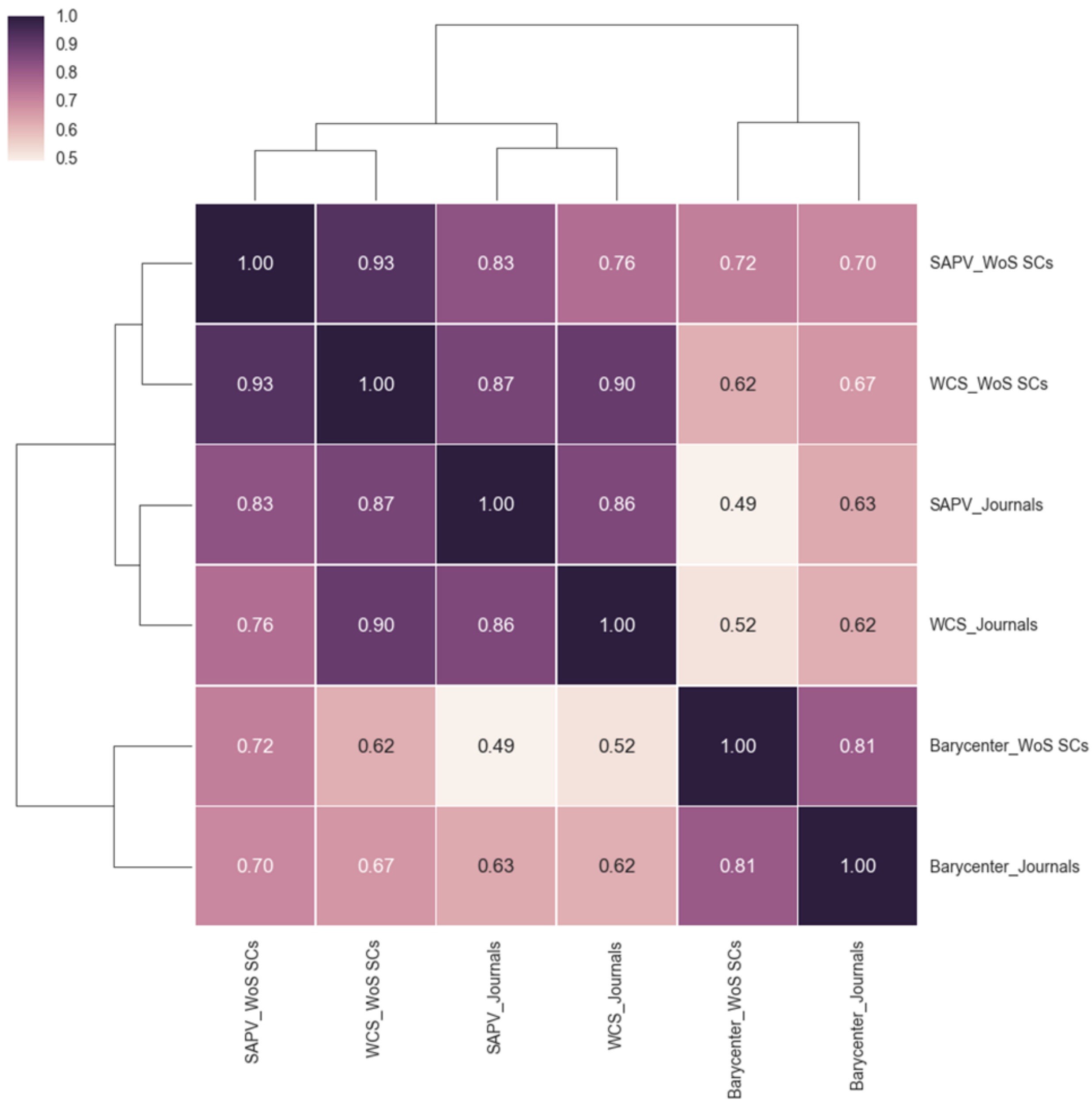


Figure 113: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Chemistry department

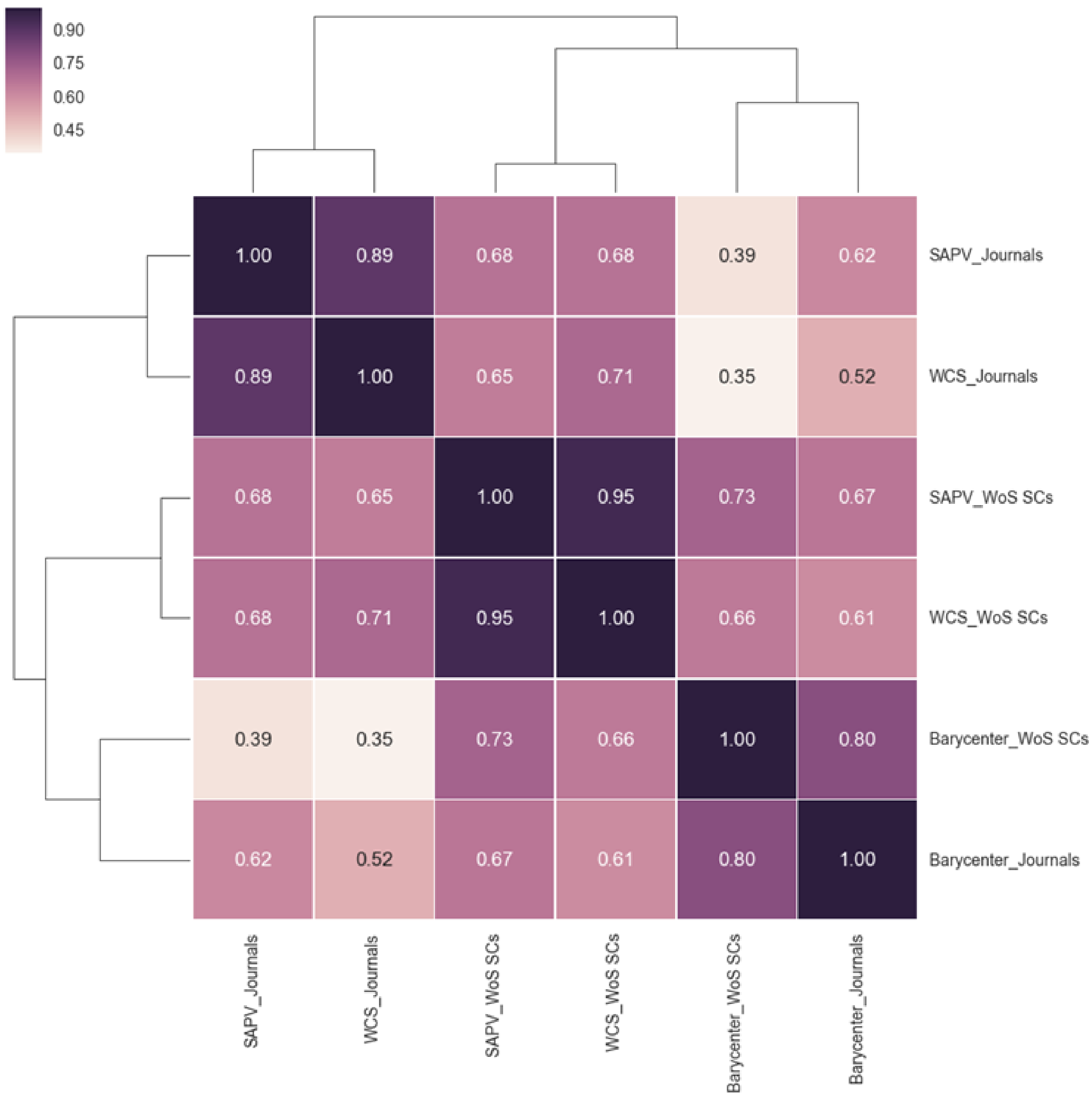


Figure 114: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Pharmaceuticals sciences department

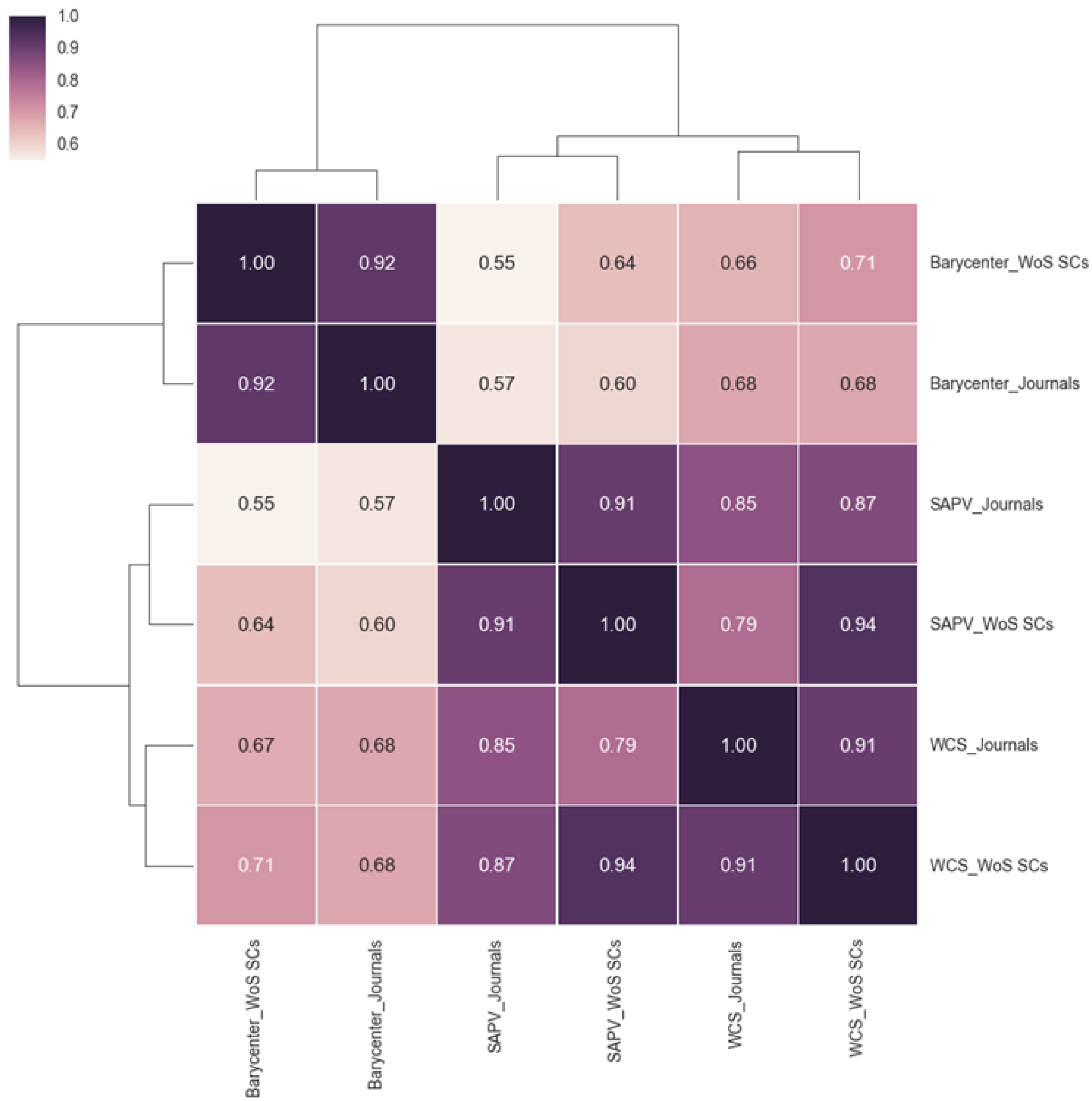


Figure 115: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Physics department

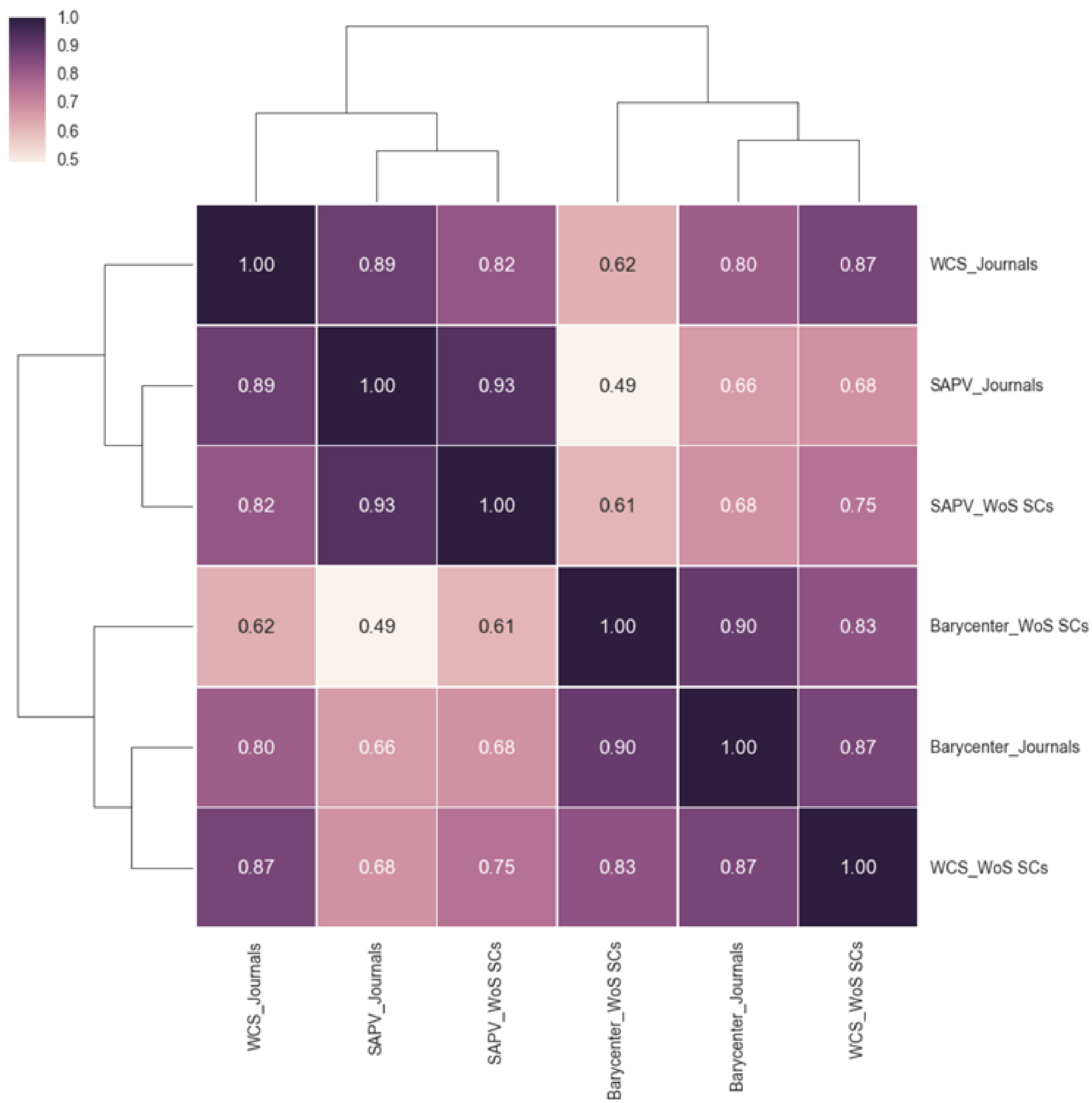


Figure 116: Heat map with hierarchical clustering based on correlation coefficient between six approaches in the Veterinary sciences department

Appendix E: Ranking of panel members according to six approaches

Procedure A

For each research group, we determine the rank of panel member according to their shortest distance (barycenter and SAPV method) or highest similarity (for WCS method). The main assessor is bold and underlined in our methods. Distances/similarities whose confidence intervals overlap with that of the shortest distance/ highest similarities are in bold. This is applicable for Table 89 to Table 108.

Table 89: Biology department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence intervals) at journal level

Research groups	Main assessor	SAPV method in Journals					WCS method in Journals					Barycenter method in Journals				
BIOL-A	PM2	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM4	PM5	PM1	PM3	**PM2**	PM1	PM5	PM4	PM3
		1					1					1				
BIOL-B	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	PM3	**PM5**
		0					0					0				
BIOL-C	PM3	**PM5**	**PM1**	**PM3**	**PM2**	PM4	PM5	PM1	**PM3**	PM2	PM4	PM2	PM1	PM4	PM5	**PM3**
		1					0					0				
BIOL-D	PM1	**PM1**	PM2	PM4	PM3	PM5	**PM1**	PM3	PM4	PM2	PM5	**PM1**	PM4	PM2	PM3	PM5
		1					1					1				
BIOL-E	PM2	**PM2**	**PM4**	PM1	PM5	PM3	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM1	PM4	PM5	PM3
		1					1					1				
BIOL-F	PM3	**PM3**	PM1	PM5	PM2	PM4	**PM3**	PM1	PM5	PM2	PM4	**PM3**	**PM5**	PM4	PM1	PM2
		1					1					1				
BIOL-G	PM4	**PM4**	**PM2**	**PM1**	PM5	PM3	**PM2**	**PM4**	PM1	PM5	PM3	**PM1**	**PM2**	**PM4**	PM3	PM5
		1					1					1				
BIOL-H	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	PM3	**PM5**
		0					0					0				
BIOL-I	PM1	PM3	**PM1**	PM5	PM2	PM4	PM3	**PM1**	PM5	PM2	PM4	**PM3**	**PM4**	**PM5**	**PM1**	PM2
		0					0					1				
	Score	**6**					**5**					**6**				

Table 90: Biology department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs					WCS method in WoS SCs					Barycenter method in WoS SCs				
BIOL-A	PM2	**PM2**	PM5	PM4	PM1	PM3	**PM2**	PM4	PM5	PM1	PM3	**PM2**	**PM5**	PM4	PM3	PM1
		1					1					1				
BIOL-B	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM3	**PM5**	PM2
		0					0					0				
BIOL-C	PM3	PM1	PM5	PM2	PM4	**PM3**	PM5	PM1	PM4	PM2	**PM3**	PM1	PM4	**PM3**	PM5	PM2
		0					0					0				
BIOL-D	PM1	**PM1**	PM4	PM2	PM5	PM3	**PM1**	PM4	PM2	PM3	PM5	**PM1**	PM4	PM3	PM5	PM2
		1					1					1				
BIOL-E	PM2	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM4	PM1	PM5	PM3	**PM2**	**PM5**	**PM4**	PM3	PM1
		1					1					1				
BIOL-F	PM3	**PM3**	PM5	PM2	PM1	PM4	**PM3**	PM5	PM1	PM4	PM2	**PM3**	PM4	PM5	PM1	PM2
		1					1					1				
BIOL-G	PM4	**PM4**	**PM1**	**PM2**	PM5	PM3	**PM4**	**PM2**	**PM1**	PM5	PM3	**PM4**	**PM5**	PM2	PM3	PM1
		1					1					1				
BIOL-H	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM3	**PM5**	PM2
		0					0					0				
BIOL-I	PM1	PM3	PM5	**PM1**	PM2	PM4	PM3	**PM1**	PM5	PM4	PM2	**PM3**	**PM1**	PM4	PM5	PM2
		0					0					1				
	Score	**5**					**5**					**6**				

Table 91: Biomedical Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals								WCS method in Journals								Barycenter method in Journals							
BIOM-A	PM8	PM2	PM5	PM1	PM3	PM4	PM6	PM7	**PM8**	**PM2**	**PM8**	**PM1**	PM3	PM5	PM4	PM7	PM6	**PM8**	**PM2**	**PM5**	**PM6**	PM1	PM7	PM3	PM4
		0								1								1							
BIOM-B	PM8	PM5	PM1	PM2	PM3	PM4	PM6	PM7	**PM8**	PM5	PM3	PM1	PM2	PM4	PM6	PM7	**PM8**	PM2	PM5	PM7	PM6	PM1	PM4	**PM8**	PM3
		0								0								0							
BIOM-C	PM7	PM5	PM1	PM2	PM3	PM4	PM6	**PM7**	PM8	PM5	PM1	PM3	PM2	PM4	PM6	**PM7**	PM8	PM2	PM5	PM6	**PM7**	PM1	PM8	PM4	PM3
		0								0								0							
BIOM-D	PM7	**PM7**	PM1	PM4	PM5	PM2	PM3	PM6	PM8	**PM7**	PM1	PM4	PM3	PM2	PM5	PM6	PM8	PM1	PM4	PM5	PM2	PM3	**PM7**	PM6	PM8
		1								1								0							
BIOM-E	PM6	**PM6**	PM5	PM1	PM2	PM7	PM3	PM4	PM8	**PM6**	PM5	PM1	PM2	PM4	PM3	PM7	PM8	**PM6**	**PM7**	**PM5**	PM2	PM1	PM8	PM4	PM3
		1								1								1							
BIOM-F	PM2	**PM1**	**PM2**	**PM4**	PM5	PM3	PM6	PM7	PM8	**PM1**	**PM2**	**PM4**	PM3	PM5	PM7	PM6	PM8	**PM1**	**PM5**	**PM2**	**PM7**	PM4	PM6	PM3	PM8
		1								1								1							
BIOM-G	PM3	PM2	PM5	PM6	PM8	PM7	PM1	**PM3**	PM4	PM2	PM5	PM1	**PM3**	PM4	PM8	PM6	PM7	PM6	PM8	PM2	PM5	PM7	PM1	PM4	**PM3**
		0								0								0							
BIOM-H	PM6	**PM6**	**PM7**	**PM1**	PM5	PM2	PM4	PM3	PM8	**PM6**	**PM1**	**PM2**	**PM7**	**PM4**	**PM5**	**PM3**	PM8	**PM6**	**PM7**	**PM5**	PM2	PM1	PM8	PM4	PM3
		1								1								1							
BIOM-I	PM1	**PM4**	**PM1**	PM2	PM5	PM3	PM7	PM6	PM8	**PM1**	**PM4**	PM2	PM3	PM5	PM7	PM6	PM8	**PM4**	**PM1**	**PM3**	**PM5**	**PM2**	PM7	PM6	PM8
		1								1								1							
BIOM-J	PM8	PM2	PM3	PM5	PM1	PM4	**PM8**	PM6	PM7	PM2	PM3	PM1	PM5	PM4	**PM8**	PM7	PM6	**PM8**	**PM2**	PM3	PM5	PM4	PM6	PM1	PM7
		0								0								1							
BIOM-K	PM2	**PM5**	**PM2**	**PM3**	**PM1**	PM4	PM6	PM7	PM8	PM5	**PM2**	PM1	PM3	PM4	PM7	PM6	PM8	**PM8**	**PM2**	**PM5**	**PM6**	PM7	PM1	PM3	PM4
		1								0								1							
BIOM-L	PM4	**PM5**	**PM2**	**PM8**	**PM3**	**PM6**	**PM7**	**PM1**	**PM4**	**PM5**	**PM2**	**PM3**	**PM1**	**PM4**	**PM8**	**PM7**	**PM6**	PM8	PM6	PM2	PM5	PM7	PM1	**PM4**	PM3
		1								1								0							
BIOM-M	PM3	PM5	PM2	**PM3**	PM1	PM6	PM4	PM7	PM8	PM5	PM2	PM1	**PM3**	PM4	PM7	PM6	PM8	PM6	PM8	PM2	PM5	PM7	PM1	PM4	**PM3**
		0								0								0							
BIOM-N	PM5	**PM5**	**PM2**	**PM8**	**PM6**	**PM7**	**PM3**	**PM1**	**PM4**	**PM5**	**PM2**	PM1	PM3	PM4	PM6	PM8	PM7	PM6	PM8	PM2	**PM5**	PM7	PM1	PM4	PM3
		1								1								0							
BIOM-O	PM4	**PM5**	**PM2**	**PM3**	**PM1**	**PM6**	**PM4**	PM7	PM8	PM5	PM2	PM1	PM3	**PM4**	PM7	PM6	PM8	PM6	PM2	PM5	PM7	PM8	PM1	**PM4**	PM3
		1								0								0							

Research groups	Main assessor	SAPV method in Journals	WCS method in Journals	Barycenter method in Journals
	Total Score	**9**	**8**	**7**

Table 92: Biomedical Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs								WCS method in WoS SCs								Barycenter method in WoS SCs							
BIOM-A	PM8	**PM2** 1	PM6	**PM8**	PM1	PM5	PM4	PM3	PM7	**PM8** 1	**PM2**	PM1	PM6	PM5	PM4	PM3	PM7	**PM2** 1	**PM6**	**PM8**	PM3	PM5	PM7	PM1	PM4
BIOM-B	PM8	PM5 0	PM1	PM3	PM4	PM2	PM6	PM7	**PM8**	PM5 0	PM3	PM1	PM4	PM2	PM6	PM7	**PM8**	PM5 0	PM3	PM1	PM7	PM6	PM2	PM4	**PM8**
BIOM-C	PM7	PM5 0	PM1	PM3	PM4	PM6	PM2	**PM7**	PM8	PM5 0	PM1	PM3	PM4	PM2	PM6	**PM7**	PM8	**PM5** 1	**PM7**	**PM1**	**PM6**	**PM3**	**PM2**	PM8	PM4
BIOM-D	PM7	**PM7** 1	**PM1**	**PM4**	PM5	PM6	PM3	PM2	PM8	**PM7** 1	PM4	PM1	PM5	PM2	PM3	PM6	PM8	PM5 0	PM1	PM3	**PM7**	PM6	PM2	PM4	PM8
BIOM-E	PM6	**PM1** 1	**PM5**	**PM6**	PM4	PM2	PM3	PM7	PM8	**PM6** 1	**PM5**	**PM1**	**PM4**	PM3	PM2	PM7	PM8	PM5 0	PM7	PM1	PM3	**PM6**	PM2	PM4	PM8
BIOM-F	PM2	**PM1** 1	**PM2**	**PM5**	**PM6**	**PM4**	**PM3**	PM7	PM8	**PM1** 1	**PM2**	**PM4**	**PM5**	PM3	PM6	PM7	PM8	**PM5** 1	**PM7**	**PM3**	**PM1**	**PM6**	**PM2**	PM8	PM4
BIOM-G	PM3	PM2 0	PM8	PM6	**PM3**	PM5	PM1	PM4	PM7	PM2 0	PM8	PM6	PM5	PM1	**PM3**	PM4	PM7	PM8 0	PM2	PM6	PM7	PM5	**PM3**	PM1	PM4
BIOM-H	PM6	**PM6** 1	**PM1**	PM2	PM5	PM7	PM4	PM3	PM8	**PM6** 1	**PM1**	**PM2**	**PM4**	**PM5**	**PM7**	PM3	PM8	**PM6** 1	**PM7**	**PM2**	**PM5**	PM1	PM3	PM8	PM4
BIOM-I	PM1	**PM4** 1	**PM1**	PM5	PM3	PM6	PM2	PM7	PM8	**PM4** 1	**PM1**	PM5	PM3	PM2	PM6	PM7	PM8	**PM4** 1	**PM1**	**PM3**	**PM5**	PM7	PM6	PM2	PM8
BIOM-J	PM8	**PM2** 1	**PM8**	**PM4**	PM6	PM3	PM1	PM5	PM7	PM2 0	**PM8**	PM4	PM5	PM1	PM3	PM6	PM7	**PM3** 1	**PM2**	**PM8**	**PM6**	**PM5**	PM1	PM7	PM4
BIOM-K	PM2	**PM2** 1	**PM5**	**PM3**	**PM1**	**PM6**	**PM4**	PM7	PM8	**PM5** 1	**PM2**	**PM3**	**PM1**	**PM4**	PM6	PM8	PM7	**PM2** 1	**PM6**	**PM8**	PM5	PM7	PM3	PM1	PM4
BIOM-L	PM4	**PM2** 1	**PM8**	**PM6**	**PM5**	**PM1**	**PM3**	**PM4**	PM7	**PM2** 1	**PM8**	**PM5**	**PM1**	PM6	**PM3**	**PM4**	PM7	PM8 0	PM2	PM6	PM7	PM5	PM3	PM1	**PM4**
BIOM-M	PM3	PM2 0	PM8	PM6	PM5	**PM3**	PM1	PM4	PM7	PM2 0	PM8	PM5	**PM3**	PM1	PM6	PM4	PM7	PM8 0	PM2	PM6	PM7	PM5	**PM3**	PM1	PM4
BIOM-N	PM5	PM8 0	PM2	PM6	PM3	**PM5**	PM1	PM4	PM7	PM2 0	PM8	**PM5**	PM6	PM3	PM1	PM4	PM7	PM8 0	PM2	PM6	PM7	**PM5**	PM3	PM1	PM4
BIOM-O	PM4	PM2 0	PM8	PM6	PM5	PM3	PM1	**PM4**	PM7	PM2 0	PM8	PM5	PM3	PM1	PM6	**PM4**	PM7	PM8 0	PM2	PM6	PM7	PM5	PM3	PM1	**PM4**

Research groups	Main assessor	SAPV method in WoS SCs	WCS method in WoS SCs	Barycenter method in WoS SCs
	Total Score	**9**	**8**	**7**

Table 93: Chemistry department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals							WCS method in Journals							Barycenter method in Journals						
CHEM-A	PM6	PM 5	PM 7	PM 1	PM 2	**PM 6**	PM 3	PM 4	PM5	PM7	PM1	**PM6**	PM2	PM4	PM3	PM 7	PM 5	PM 1	**PM 6**	PM 3	PM 4	PM 2
		0							0							0						
CHEM- B	PM5	**PM 5**	PM 7	PM 1	PM 2	PM 6	PM 3	PM 4	**PM5**	PM7	PM1	PM6	PM2	PM4	PM3	**PM 7**	**PM 1**	**PM 5**	PM 6	PM 3	PM 4	PM 2
		1							1							1						
CHEM-C	PM7	PM 5	PM 2	PM 3	PM 4	**PM 7**	PM 1	PM 6	PM3	PM5	PM2	PM4	**PM7**	PM1	PM6	PM 5	**PM 7**	PM 1	PM 3	PM 6	PM 4	PM 2
		0							0							0						
CHEM- D	PM2	PM 6	PM 1	PM 7	PM 5	**PM 2**	PM 4	PM 3	PM6	PM1	PM7	PM5	**PM2**	PM4	PM3	PM 6	PM 3	PM 1	PM 7	PM 4	PM 5	**PM 2**
		0							0							0						
CHEM-E	PM2	**PM 2**	PM 3	PM 4	PM 5	PM 7	PM 6	PM 1	**PM2**	**PM4**	PM3	PM5	PM6	PM7	PM1	**PM 2**	**PM 4**	PM 3	PM 6	PM 5	PM 1	PM 7
		1							1							1						
CHEM- F	PM3	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 6	PM 1	PM4	PM2	**PM3**	PM7	PM6	PM1	PM5	**PM 2**	**PM 4**	**PM 3**	**PM 6**	PM 5	PM 1	PM 7
		1							0							1						
CHEM- G	PM3	PM 4	**PM 3**	PM 2	PM 5	PM 7	PM 6	PM 1	PM4	**PM3**	PM2	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	**PM 5**	PM 2	PM 6	PM 7	PM 1
		0							0							1						
CHEM- H	PM5	PM 3	PM 2	PM 4	**PM 5**	PM 7	PM 6	PM 1	PM3	PM4	PM2	**PM5**	PM7	PM1	PM6	**PM 5**	**PM 3**	**PM 4**	PM 7	PM 6	PM 1	PM 2
		0							0							1						
CHEM-I	PM4	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 1	PM 6	**PM4**	**PM2**	**PM3**	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	**PM 5**	PM 6	PM 7	PM 1	PM 2
		1							1							1						
CHEM-J	PM4	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 6	PM 1	**PM4**	**PM2**	**PM3**	PM5	PM7	PM1	PM6	**PM 4**	**PM 3**	**PM 5**	**PM 2**	**PM 6**	PM 7	PM 1
		1							1							1						
CHEM-K	PM6	PM 2	PM 3	PM 5	PM 4	PM 7	**PM 6**	PM 1	PM2	PM5	PM3	PM4	PM7	PM1	**PM6**	PM 4	PM 2	PM 3	PM 5	**PM 6**	PM 7	PM 1
		0							0							0						
CHEM-L	PM1	PM 5	PM 2	PM 3	PM 4	PM 7	**PM 1**	PM 6	**PM5**	**PM7**	**PM1**	PM6	PM2	PM4	PM3	**PM 5**	**PM 7**	**PM 1**	**PM 3**	PM 6	PM 4	PM 2
		0							1							1						
	Score	**5**							**5**							**8**						

Table 94: Chemistry department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs							WCS method in WoS SCs							Barycenter method in WoS SCs						
CHEM-A	PM6	PM 5	PM 7	PM 1	PM 2	PM 3	**PM 6**	PM 4	PM5	PM7	PM1	PM2	**PM6**	PM3	PM4	PM 5	PM 7	PM 1	PM 3	PM 4	**PM 6**	PM 2
		0							0							0						
CHEM- B	PM5	**PM 5**	PM 7	PM 2	PM 3	PM 1	PM 6	PM 4	**PM5**	PM2	PM7	PM1	PM3	PM6	PM4	**PM 5**	**PM 7**	**PM 1**	PM 3	PM 4	PM 6	PM 2
		1							1							1						
CHEM-C	PM7	PM 5	PM 2	PM 3	**PM 7**	PM 4	PM 1	PM 6	PM5	PM2	PM3	PM4	**PM7**	PM1	PM6	PM 5	**PM 7**	PM 3	PM 1	PM 4	PM 6	PM 2
		0							0							0						
CHEM- D	PM2	**PM 6**	**PM 2**	**PM 1**	PM 3	PM 7	PM 5	PM 4	PM6	PM1	PM7	**PM2**	PM5	PM3	PM4	**PM 6**	**PM 4**	**PM 2**	**PM 3**	**PM 1**	PM 5	PM 7
		1							0							1						
CHEM-E	PM2	**PM 2**	**PM 3**	PM 4	PM 6	PM 5	PM 1	PM 7	**PM2**	**PM3**	PM4	PM5	PM6	PM1	PM7	**PM 2**	**PM 4**	**PM 6**	PM 3	PM 1	PM 5	PM 7
		1							1							1						
CHEM- F	PM3	**PM 2**	**PM 3**	PM 4	PM 6	PM 5	PM 1	PM 7	PM2	**PM3**	PM4	PM6	PM5	PM1	PM7	**PM 2**	**PM 4**	**PM 6**	**PM 3**	PM 1	PM 5	PM 7
		1							0							**1**						
CHEM- G	PM3	**PM 4**	**PM 3**	PM 2	PM 5	PM 6	PM 7	PM 1	PM4	**PM3**	PM2	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	PM 2	PM 5	PM 6	PM 1	PM 7
		1							0							1						
CHEM- H	PM5	**PM 3**	**PM 2**	**PM 5**	PM 4	PM 6	PM 7	PM 1	PM3	PM2	PM4	**PM5**	PM7	PM1	PM6	**PM 3**	**PM 5**	**PM 4**	PM 1	PM 7	PM 6	PM 2
		1							0							1						
CHEM-I	PM4	PM 2	PM 3	PM 5	**PM 4**	PM 6	PM 7	PM 1	PM2	PM3	**PM4**	PM5	PM7	PM1	PM6	PM 3	PM 5	PM 1	**PM 4**	PM 7	PM 6	PM 2
		0							0							0						
CHEM-J	PM4	**PM 3**	**PM 2**	**PM 5**	**PM 4**	PM 6	PM 7	PM 1	**PM3**	**PM2**	**PM4**	**PM5**	PM7	PM1	PM6	**PM 4**	**PM 2**	**PM 3**	**PM 5**	PM 6	PM 1	PM 7
		1							1							1						
CHEM-K	PM6	PM 2	PM 3	PM 4	PM 5	**PM 6**	PM 7	PM 1	**PM2**	**PM5**	**PM3**	PM4	PM7	**PM6**	PM1	PM 2	PM 4	PM 3	**PM 6**	PM 5	PM 1	PM 7
		0							0							0						
CHEM-L	PM1	PM 5	PM 2	PM 3	PM 7	**PM 1**	PM 6	PM 4	**PM5**	PM2	PM7	**PM1**	PM3	PM6	PM4	**PM 5**	**PM 7**	**PM 1**	PM 3	PM 4	PM 6	PM 2
		0							0							1						
	Score	**7**							**3**							**8**						

Table 95: Physics department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals						WCS method in Journals						Barycenter method in Journals					
PHYS-A	PM3	PM4	PM5	PM6	**PM3**	PM2	PM1	**PM3**	**PM5**	**PM6**	PM2	PM4	PM1	**PM4**	**PM3**	**PM5**	**PM6**	PM2	PM1
		0						1						1					
PHYS- B	PM2	PM6	**PM2**	PM5	PM3	PM4	PM1	**PM6**	**PM2**	PM5	PM4	PM3	PM1	**PM6**	**PM2**	PM5	PM3	PM1	PM4
		0						1						1					
PHYS-C	PM5	**PM5**	PM4	PM6	PM2	PM3	PM1	**PM5**	PM6	PM2	PM4	PM3	PM1	**PM5**	PM6	PM4	PM2	PM3	PM1
		1						1						1					
PHYS- D	PM1	**PM1**	PM6	PM2	PM5	PM3	PM4	**PM1**	PM6	PM5	PM2	PM3	PM4	**PM1**	PM2	PM6	PM3	PM5	PM4
		1						1						1					
PHYS-E	PM4	PM6	PM5	PM2	**PM4**	PM3	PM1	PM6	PM5	PM2	PM1	**PM4**	PM3	PM5	PM6	PM2	PM3	PM1	**PM4**
		0						0						0					
PHYS- F	PM1	PM6	PM5	PM2	PM4	PM3	**PM1**	PM6	PM2	PM5	**PM1**	PM4	PM3	**PM1**	**PM3**	PM2	PM6	PM5	PM4
		0						0						1					
PHYS- G	PM4	**PM4**	PM5	PM6	PM3	PM2	PM1	**PM4**	PM5	PM6	PM2	PM3	PM1	**PM4**	PM5	PM3	PM6	PM2	PM1
		1						1						1					
PHYS- H	PM6	**PM6**	**PM2**	PM5	PM3	PM4	PM1	**PM6**	**PM2**	PM5	PM4	PM3	PM1	**PM2**	**PM6**	PM5	PM3	PM1	PM4
		1						1						1					
PHYS-I	PM3	PM5	PM4	PM6	**PM3**	PM2	PM1	PM5	PM6	PM2	**PM3**	PM4	PM1	**PM4**	**PM5**	**PM3**	**PM6**	**PM2**	PM1
		0						0						1					
	Total Score	**4**						**6**						**8**					

Table 96: Physics department’s top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs						WCS method in WoS SCs						Barycenter method in WoS SCs					
PHYS-A	PM3	**PM 4** 1	**PM 3**	**PM 5**	**PM 6**	PM 2	PM 1	**PM3** 1	**PM5**	**PM4**	**PM6**	**PM2**	PM1	**PM 4** 1	**PM 3**	**PM 5**	**PM 6**	PM 1	PM 2
PHYS- B	PM2	PM 6 0	**PM 2**	PM 5	PM 4	PM 3	PM 1	**PM6** 1	**PM2**	PM5	PM4	PM3	PM1	PM 6 0	PM 5	**PM 2**	PM 1	PM 3	PM 4
PHYS-C	PM5	**PM 5** 1	PM 6	PM 2	PM 4	PM 3	PM 1	**PM5** 1	PM6	PM2	PM4	PM3	PM1	**PM 5** 1	PM 6	PM 2	PM 4	PM 3	PM 1
PHYS- D	PM1	**PM 1** 1	PM 3	PM 6	PM 4	PM 5	PM 2	**PM1** 1	PM6	PM3	PM5	PM4	PM2	**PM 1** 1	PM 3	PM 6	PM 5	PM 2	PM 4
PHYS-E	PM4	PM 6 0	PM 5	**PM 4**	PM 2	PM 3	PM 1	PM6 0	PM5	PM2	**PM4**	PM3	PM1	PM 5 0	PM 6	PM 3	PM 2	PM 1	**PM 4**
PHYS- F	PM1	PM 3 0	PM 6	PM 5	PM 2	PM 4	**PM 1**	PM6 0	PM3	PM2	**PM1**	PM5	PM4	**PM3** 1	**PM 1**	PM 5	PM 6	PM 4	PM 2
PHYS- G	PM4	**PM 4** 1	**PM 5**	**PM 6**	PM 2	PM 3	PM 1	**PM4** 1	PM5	PM6	PM2	PM3	PM1	**PM 4** 1	**PM 3**	**PM 5**	**PM 6**	PM 2	PM 1
PHYS- H	PM6	**PM 6** 1	**PM 2**	PM 5	PM 3	PM 4	PM 1	**PM6** 1	**PM2**	PM5	PM3	PM4	PM1	**PM 6** 1	**PM 5**	PM 2	PM 1	PM 3	PM 4
PHYS-I	PM3	**PM 4** 1	**PM 3**	**PM 5**	PM 6	PM 2	PM 1	**PM5** 1	**PM3**	**PM6**	**PM2**	**PM4**	PM1	**PM 4** 1	**PM 3**	**PM 5**	PM 1	PM 6	PM 2
	Total Score	**6**						**7**						**7**					

Table 97: Veterinary Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals				WCS method in Journals				Barycenter method in Journals			
VETE-A	PM4	PM2	PM1	**PM4**	PM3	PM2	**PM4**	PM1	PM3	PM2	PM1	**PM4**	PM3
		0				0				0			
VETE-B	PM2	**PM2**	PM1	PM4	PM3	**PM2**	**PM1**	**PM3**	PM4	**PM2**	**PM1**	**PM3**	PM4
		1				1				1			
VETE-C	PM1	**PM1**	PM3	PM2	PM4	**PM1**	PM3	PM2	PM4	**PM1**	**PM3**	**PM2**	PM4
		1				1				1			
	Total Score	**2**				**2**				**2**			

Table 98: Veterinary Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs				WCS method in WoS SCs				Barycenter method in WoS SCs			
VETE-A	PM4	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM4**	**PM3**	**PM1**
		1				1				1			
VETE-B	PM2	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM1**	**PM3**	PM4	**PM2**	**PM3**	**PM1**	**PM4**
		1				1				1			
VETE-C	PM1	**PM1**	PM3	PM2	PM4	**PM1**	PM3	PM2	PM4	**PM3**	**PM1**	**PM2**	PM4
		1				1				1			
	Score	**3**				**3**				**3**			

Procedure B

Table 99: Biology department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence intervals) at journal level

Research groups	Main assessor	SAPV method in Journals					WCS method in Journals					Barycenter method in Journals				
BIOL-A	PM2	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM4	PM5	PM1	PM3	**PM2**	PM1	PM5	PM4	PM3
		1					1					1				
BIOL-B	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	PM3	**PM5**
		0					0					0				
BIOL-C	PM3	**PM5**	**PM1**	**PM3**	**PM2**	PM4	PM5	PM1	**PM3**	PM2	PM4	PM2	PM1	PM4	PM5	**PM3**
		0.25					0					0				
BIOL-D	PM1	**PM1**	PM2	PM4	PM3	PM5	**PM1**	PM3	PM4	PM2	PM5	**PM1**	PM4	PM2	PM3	PM5
		1					1					1				
BIOL-E	PM2	**PM2**	**PM4**	PM1	PM5	PM3	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM1	PM4	PM5	PM3
		0.50					1					1				
BIOL-F	PM3	**PM3**	PM1	PM5	PM2	PM4	**PM3**	PM1	PM5	PM2	PM4	**PM3**	**PM5**	PM4	PM1	PM2
		1					1					0.50				
BIOL-G	PM4	**PM4**	**PM2**	**PM1**	PM5	PM3	**PM2**	**PM4**	PM1	PM5	PM3	**PM1**	**PM2**	**PM4**	PM3	PM5
		0.33					0.50					0.33				
BIOL-H	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	PM3	**PM5**
		0					0					0				
BIOL-I	PM1	PM3	**PM1**	PM5	PM2	PM4	PM3	**PM1**	PM5	PM2	PM4	**PM3**	**PM4**	**PM5**	**PM1**	PM2
		0					0					0.25				
	Score	**4.08**					**4.50**					**4.08**				

Table 100: Biology department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs					WCS method in WoS SCs					Barycenter method in WoS SCs				
BIOL-A	PM2	**PM2**	PM5	PM4	PM1	PM3	**PM2**	PM4	PM5	PM1	PM3	**PM2**	**PM5**	PM4	PM3	PM1
		1					1					0.50				
BIOL-B	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM3	**PM5**	PM2
		0					0					0				
BIOL-C	PM3	PM1	PM5	PM2	PM4	**PM3**	PM5	PM1	PM4	PM2	**PM3**	PM1	PM4	**PM3**	PM5	PM2
		0					0					0				
BIOL-D	PM1	**PM1**	PM4	PM2	PM5	PM3	**PM1**	PM4	PM2	PM3	PM5	**PM1**	PM4	PM3	PM5	PM2
		1					1					1				
BIOL-E	PM2	**PM2**	PM4	PM1	PM5	PM3	**PM2**	PM4	PM1	PM5	PM3	**PM2**	**PM5**	**PM4**	PM3	PM1
		1					1					0.33				
BIOL-F	PM3	**PM3**	PM5	PM2	PM1	PM4	**PM3**	PM5	PM1	PM4	PM2	**PM3**	PM4	PM5	PM1	PM2
		1					1					1				
BIOL-G	PM4	**PM4**	**PM1**	**PM2**	PM5	PM3	**PM4**	**PM2**	**PM1**	PM5	PM3	**PM4**	**PM5**	PM2	PM3	PM1
		0.33					0.33					0.50				
BIOL-H	PM5	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM2	**PM5**	PM3	PM1	PM4	PM3	**PM5**	PM2
		0					0					0				
BIOL-I	PM1	PM3	PM5	**PM1**	PM2	PM4	PM3	**PM1**	PM5	PM4	PM2	**PM3**	**PM1**	PM4	PM5	PM2
		0					0					0.50				
	Score	**4.33**					**4.33**					**3.83**				

Table 101: Biomedical Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals								WCS method in Journals								Barycenter method in Journals							
BIOM-A	PM8	PM2	PM5	PM1	PM3	PM4	PM6	PM7	**PM8**	**PM2**	**PM8**	**PM1**	PM3	PM5	PM4	PM7	PM6	**PM8**	**PM2**	**PM5**	**PM6**	PM1	PM7	PM3	PM4
		0								0.33								0.25							
BIOM-B	PM8	PM5	PM1	PM2	PM3	PM4	PM6	PM7	**PM8**	PM5	PM3	PM1	PM2	PM4	PM6	PM7	**PM8**	PM2	PM5	PM7	PM6	PM1	PM4	**PM8**	PM3
		0								0								0							
BIOM-C	PM7	PM5	PM1	PM2	PM3	PM4	PM6	**PM7**	PM8	PM5	PM1	PM3	PM2	PM4	PM6	**PM7**	PM8	PM2	PM5	PM6	**PM7**	PM1	PM8	PM4	PM3
		0								0								0							
BIOM-D	PM7	**PM7**	PM1	PM4	PM5	PM2	PM3	PM6	PM8	**PM7**	PM1	PM4	PM3	PM2	PM5	PM6	PM8	PM1	PM4	PM5	PM2	PM3	**PM7**	PM6	PM8
		1								1								0							
BIOM-E	PM6	**PM6**	PM5	PM1	PM2	PM7	PM3	PM4	PM8	**PM6**	PM5	PM1	PM2	PM4	PM3	PM7	PM8	**PM6**	**PM7**	**PM5**	PM2	PM1	PM8	PM4	PM3
		1								1								0.33							
BIOM-F	PM2	**PM1**	**PM2**	**PM4**	PM5	PM3	PM6	PM7	PM8	**PM1**	**PM2**	**PM4**	PM3	PM5	PM7	PM6	PM8	**PM1**	**PM5**	**PM2**	**PM7**	PM4	PM6	PM3	PM8
		0.33								0.33								0.25							
BIOM-G	PM3	PM2	PM5	PM6	PM8	PM7	PM1	**PM3**	PM4	PM2	PM5	PM1	**PM3**	PM4	PM8	PM6	PM7	PM6	PM8	PM2	PM5	PM7	PM1	PM4	**PM3**
		0								0								0							
BIOM-H	PM6	**PM6**	**PM7**	**PM1**	PM5	PM2	PM4	PM3	PM8	**PM6**	**PM1**	**PM2**	**PM7**	**PM4**	**PM5**	**PM3**	PM8	**PM6**	**PM7**	**PM5**	PM2	PM1	PM8	PM4	PM3
		0.33								0.14								0.33							
BIOM-I	PM1	**PM4**	**PM1**	PM2	PM5	PM3	PM7	PM6	PM8	**PM1**	**PM4**	PM2	PM3	PM5	PM7	PM6	PM8	**PM4**	**PM1**	**PM3**	**PM5**	**PM2**	PM7	PM6	PM8
		0.50								0.50								0.20							
BIOM-J	PM8	PM2	PM3	PM5	PM1	PM4	**PM8**	PM6	PM7	PM2	PM3	PM1	PM5	PM4	**PM8**	PM7	PM6	**PM8**	**PM2**	PM3	PM5	PM4	PM6	PM1	PM7
		0								0								0.50							
BIOM-K	PM2	**PM5**	**PM2**	**PM3**	**PM1**	PM4	PM6	PM7	PM8	PM5	**PM2**	PM1	PM3	PM4	PM7	PM6	PM8	**PM8**	**PM2**	**PM5**	**PM6**	PM7	PM1	PM3	PM4
		0.25								0								0.25							
BIOM-L	PM4	**PM5**	**PM2**	**PM8**	**PM3**	**PM6**	**PM7**	**PM1**	**PM4**	**PM5**	**PM2**	**PM3**	**PM1**	**PM4**	**PM8**	**PM7**	**PM6**	PM8	PM6	PM2	PM5	PM7	PM1	**PM4**	PM3
		0.13								0.13								0							
BIOM-M	PM3	PM5	PM2	**PM3**	PM1	PM6	PM4	PM7	PM8	PM5	PM2	PM1	**PM3**	PM4	PM7	PM6	PM8	PM6	PM8	PM2	PM5	PM7	PM1	PM4	**PM3**
		0								0								0							
BIOM-N	PM5	**PM5**	**PM2**	**PM8**	**PM6**	**PM7**	**PM3**	**PM1**	**PM4**	**PM5**	**PM2**	PM1	PM3	PM4	PM6	PM8	PM7	PM6	PM8	PM2	**PM5**	PM7	PM1	PM4	PM3
		0.13								0.50								0							
BIOM-O	PM4	**PM5**	**PM2**	**PM3**	**PM1**	**PM6**	**PM4**	PM7	PM8	PM5	PM2	PM1	PM3	**PM4**	PM7	PM6	PM8	PM6	PM2	PM5	PM7	PM8	PM1	**PM4**	PM3
		0.17								0								0							

Research groups	Main assessor	SAPV method in Journals	WCS method in Journals	Barycenter method in Journals
	Score	3.84	3.93	2.11

Table 102: Biomedical Sciences department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs								WCS method in WoS SCs								Barycenter method in WoS SCs							
BIOM-A	PM8	**PM2**	PM6	**PM8**	PM1	PM5	PM4	PM3	PM7	**PM8**	**PM2**	PM1	PM6	PM5	PM4	PM3	PM7	**PM2**	**PM6**	**PM8**	PM3	PM5	PM7	PM1	PM4
		0.50								0.50								0.33							
BIOM-B	PM8	PM5	PM1	PM3	PM4	PM2	PM6	PM7	**PM8**	PM5	PM3	PM1	PM4	PM2	PM6	PM7	**PM8**	PM5	PM3	PM1	PM7	PM6	PM2	PM4	**PM8**
		0								0								0							
BIOM-C	PM7	PM5	PM1	PM3	PM4	PM6	PM2	**PM7**	PM8	PM5	PM1	PM3	PM4	PM2	PM6	**PM7**	PM8	**PM5**	**PM7**	**PM1**	**PM6**	**PM3**	**PM2**	PM8	PM4
		0								0								0.17							
BIOM-D	PM7	**PM7**	**PM1**	**PM4**	PM5	PM6	PM3	PM2	PM8	**PM7**	PM4	PM1	PM5	PM2	PM3	PM6	PM8	PM5	PM1	PM3	**PM7**	PM6	PM2	PM4	PM8
		0.33								1								0							
BIOM-E	PM6	**PM1**	**PM5**	**PM6**	PM4	PM2	PM3	PM7	PM8	**PM6**	**PM5**	**PM1**	**PM4**	PM3	PM2	PM7	PM8	PM5	PM7	PM1	PM3	**PM6**	PM2	PM4	PM8
		0.33								0.25								0							
BIOM-F	PM2	**PM1**	**PM2**	**PM5**	**PM6**	**PM4**	**PM3**	PM7	PM8	**PM1**	**PM2**	**PM4**	**PM5**	PM3	PM6	PM7	PM8	**PM5**	**PM7**	**PM3**	**PM1**	**PM6**	**PM2**	PM8	PM4
		0.17								0.25								0.17							
BIOM-G	PM3	PM2	PM8	PM6	**PM3**	PM5	PM1	PM4	PM7	PM2	PM8	PM6	PM5	PM1	**PM3**	PM4	PM7	PM8	PM2	PM6	PM7	PM5	**PM3**	PM1	PM4
		0								0								0							
BIOM-H	PM6	**PM6**	**PM1**	PM2	PM5	PM7	PM4	PM3	PM8	**PM6**	**PM1**	**PM2**	**PM4**	**PM5**	**PM7**	PM3	PM8	**PM6**	**PM7**	**PM2**	**PM5**	PM1	PM3	PM8	PM4
		0.50								0.17								0.25							
BIOM-I	PM1	**PM4**	**PM1**	PM5	PM3	PM6	PM2	PM7	PM8	**PM4**	**PM1**	PM5	PM3	PM2	PM6	PM7	PM8	**PM4**	**PM1**	**PM3**	**PM5**	PM7	PM6	PM2	PM8
		0.50								0.50								0.25							
BIOM-J	PM8	**PM2**	**PM8**	**PM4**	PM6	PM3	PM1	PM5	PM7	PM2	**PM8**	PM4	PM5	PM1	PM3	PM6	PM7	**PM3**	**PM2**	**PM8**	**PM6**	**PM5**	PM1	PM7	PM4
		0.33								0								0.20							
BIOM-K	PM2	**PM2**	**PM5**	**PM3**	**PM1**	**PM6**	**PM4**	PM7	PM8	**PM5**	**PM2**	**PM3**	**PM1**	**PM4**	PM6	PM8	PM7	**PM2**	**PM6**	**PM8**	PM5	PM7	PM3	PM1	PM4
		0.17								0.20								0.33							
BIOM-L	PM4	**PM2**	**PM8**	**PM6**	**PM5**	**PM1**	**PM3**	**PM4**	PM7	**PM2**	**PM8**	**PM5**	**PM1**	PM6	**PM3**	**PM4**	PM7	PM8	PM2	PM6	PM7	PM5	PM3	PM1	**PM4**
		0.14								0.17								0							
BIOM-M	PM3	PM2	PM8	PM6	PM5	**PM3**	PM1	PM4	PM7	PM2	PM8	PM5	**PM3**	PM1	PM6	PM4	PM7	PM8	PM2	PM6	PM7	PM5	**PM3**	PM1	PM4
		0								0								0							
BIOM-N	PM5	PM8	PM2	PM6	PM3	**PM5**	PM1	PM4	PM7	PM2	PM8	**PM5**	PM6	PM3	PM1	PM4	PM7	PM8	PM2	PM6	PM7	**PM5**	PM3	PM1	PM4
		0								0								0							
BIOM-O	PM4	PM2	PM8	PM6	PM5	PM3	PM1	**PM4**	PM7	PM2	PM8	PM5	PM3	PM1	PM6	**PM4**	PM7	PM8	PM2	PM6	PM7	PM5	PM3	PM1	**PM4**
		0								0								0							
	Score	**2.97**								**3.04**								**1.70**							

Table 103: Chemistry department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals							WCS method in Journals							Barycenter method in Journals						
CHEM-A	PM6	PM 5	PM 7	PM 1	PM 2	**PM 6**	PM 3	PM 4	PM5	PM7	PM1	**PM6**	PM2	PM4	PM3	PM 7	PM 5	PM 1	**PM 6**	PM 3	PM 4	PM 2
		0							0							0						
CHEM- B	PM5	**PM 5**	PM 7	PM 1	PM 2	PM 6	PM 3	PM 4	**PM5**	PM7	PM1	PM6	PM2	PM4	PM3	**PM 7**	**PM 1**	**PM 5**	PM 6	PM 3	PM 4	PM 2
		1							1							0.33						
CHEM-C	PM7	PM 5	PM 2	PM 3	PM 4	**PM 7**	PM 1	PM 6	PM3	PM5	PM2	PM4	**PM7**	PM1	PM6	PM 5	**PM 7**	PM 1	PM 3	PM 6	PM 4	PM 2
		0							0							0						
CHEM- D	PM2	PM 6	PM 1	PM 7	PM 5	**PM 2**	PM 4	PM 3	PM6	PM1	PM7	PM5	**PM2**	PM4	PM3	PM 6	PM 3	PM 1	PM 7	PM 4	PM 5	**PM 2**
		0							0							0						
CHEM-E	PM2	**PM 2**	PM 3	PM 4	PM 5	PM 7	PM 6	PM 1	**PM2**	**PM4**	PM3	PM5	PM6	PM7	PM1	**PM 2**	**PM 4**	PM 3	PM 6	PM 5	PM 1	PM 7
		1							0.50							0.50						
CHEM- F	PM3	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 6	PM 1	PM4	PM2	**PM3**	PM7	PM6	PM1	PM5	**PM 2**	**PM 4**	**PM 3**	**PM 6**	PM 5	PM 1	PM 7
		0.33							0							0.25						
CHEM- G	PM3	PM 4	**PM 3**	PM 2	PM 5	PM 7	PM 6	PM 1	PM4	**PM3**	PM2	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	**PM 5**	PM 2	PM 6	PM 7	PM 1
		0							0							0.33						
CHEM- H	PM5	PM 3	PM 2	PM 4	**PM 5**	PM 7	PM 6	PM 1	PM3	PM4	PM2	**PM5**	PM7	PM1	PM6	**PM 5**	**PM 3**	**PM 4**	PM 7	PM 6	PM 1	PM 2
		0							0							0.33						
CHEM-I	PM4	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 1	PM 6	**PM4**	**PM2**	**PM3**	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	**PM 5**	PM 6	PM 7	PM 1	PM 2
		0.33							0.33							0.33						
CHEM-J	PM4	**PM 4**	**PM 2**	**PM 3**	PM 5	PM 7	PM 6	PM 1	**PM4**	**PM2**	**PM3**	PM5	PM7	PM1	PM6	**PM 4**	**PM 3**	**PM 5**	**PM 2**	**PM 6**	PM 7	PM 1
		0.33							0.33							0.20						
CHEM-K	PM6	PM 2	PM 3	PM 5	PM 4	PM 7	**PM 6**	PM 1	PM2	PM5	PM3	PM4	PM7	PM1	**PM6**	PM 4	PM 2	PM 3	PM 5	**PM 6**	PM 7	PM 1
		0							0							0						
CHEM-L	PM1	PM 5	PM 2	PM 3	PM 4	PM 7	**PM 1**	PM 6	**PM5**	**PM7**	**PM1**	PM6	PM2	PM4	PM3	**PM 5**	**PM 7**	**PM 1**	**PM 3**	PM 6	PM 4	PM 2
		0							0.33							0.25						
	Score	**2.99**							**2.49**							**2.52**						

Table 104: Chemistry department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs							WCS method in WoS SCs							Barycenter method in WoS SCs						
CHEM-A	PM6	PM 5	PM 7	PM 1	PM 2	PM 3	**PM 6**	PM 4	PM5	PM7	PM1	PM2	**PM6**	PM3	PM4	PM 5	PM 7	PM 1	PM 3	PM 4	**PM 6**	PM 2
		0							0							0						
CHEM- B	PM5	**PM 5**	PM 7	PM 2	PM 3	PM 1	PM 6	PM 4	**PM5**	PM2	PM7	PM1	PM3	PM6	PM4	**PM 5**	**PM 7**	**PM 1**	PM 3	PM 4	PM 6	PM 2
		1							1							0.33						
CHEM-C	PM7	PM 5	PM 2	PM 3	**PM 7**	PM 4	PM 1	PM 6	PM5	PM2	PM3	PM4	**PM7**	PM1	PM6	PM 5	**PM 7**	PM 3	PM 1	PM 4	PM 6	PM 2
		0							0							0						
CHEM- D	PM2	**PM 6**	**PM 2**	**PM 1**	PM 3	PM 7	PM 5	PM 4	PM6	PM1	PM7	**PM2**	PM5	PM3	PM4	**PM 6**	**PM 4**	**PM 2**	**PM 3**	**PM 1**	PM 5	PM 7
		0.33							0							0.20						
CHEM-E	PM2	**PM 2**	**PM 3**	PM 4	PM 6	PM 5	PM 1	PM 7	**PM2**	**PM3**	PM4	PM5	PM6	PM1	PM7	**PM 2**	**PM 4**	**PM 6**	PM 3	PM 1	PM 5	PM 7
		0.50							0.50							0.33						
CHEM- F	PM3	**PM 2**	**PM 3**	PM 4	PM 6	PM 5	PM 1	PM 7	PM2	**PM3**	PM4	PM6	PM5	PM1	PM7	**PM 2**	**PM 4**	**PM 6**	**PM 3**	PM 1	PM 5	PM 7
		0.50							0							**0.25**						
CHEM- G	PM3	**PM 4**	**PM 3**	PM 2	PM 5	PM 6	PM 7	PM 1	PM4	**PM3**	PM2	PM5	PM7	PM1	PM6	**PM 3**	**PM 4**	PM 2	PM 5	PM 6	PM 1	PM 7
		0.50							0							0.50						
CHEM- H	PM5	**PM 3**	**PM 2**	**PM 5**	PM 4	PM 6	PM 7	PM 1	PM3	PM2	PM4	**PM5**	PM7	PM1	PM6	**PM 3**	**PM 5**	**PM 4**	PM 1	PM 7	PM 6	PM 2
		0.33							0							0.33						
CHEM-I	PM4	PM 2	PM 3	PM 5	**PM 4**	PM 6	PM 7	PM 1	PM2	PM3	**PM4**	PM5	PM7	PM1	PM6	PM 3	PM 5	PM 1	**PM 4**	PM 7	PM 6	PM 2
		0							0							0						
CHEM-J	PM4	**PM 3**	**PM 2**	**PM 5**	**PM 4**	PM 6	PM 7	PM 1	**PM3**	**PM2**	**PM4**	**PM5**	PM7	PM1	PM6	**PM 4**	**PM 2**	**PM 3**	**PM 5**	PM 6	PM 1	PM 7
		0.25							0.25							0.25						
CHEM-K	PM6	PM 2	PM 3	PM 4	PM 5	**PM 6**	PM 7	PM 1	**PM2**	**PM5**	**PM3**	PM4	PM7	**PM6**	PM1	PM 2	PM 4	PM 3	**PM 6**	PM 5	PM 1	PM 7
		0							0							0						
CHEM-L	PM1	PM 5	PM 2	PM 3	PM 7	**PM 1**	PM 6	PM 4	**PM5**	PM2	PM7	**PM1**	PM3	PM6	PM4	**PM 5**	**PM 7**	**PM 1**	PM 3	PM 4	PM 6	PM 2
		0							0							0.33						
	Score	**3.41**							**1.75**							**2.52**						

Table 105: Physics department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals						WCS method in Journals						Barycenter method in Journals					
PHYS-A	PM3	PM4	PM5	PM6	**<u>PM3</u>**	PM2	PM1	**<u>PM3</u>**	**PM5**	**PM6**	PM2	PM4	PM1	**PM4**	**<u>PM3</u>**	**PM5**	**PM6**	PM2	PM1
		0						0.33						0.25					
PHYS- B	PM2	PM6	**<u>PM2</u>**	PM5	PM3	PM4	PM1	**PM6**	**<u>PM2</u>**	PM5	PM4	PM3	PM1	**PM6**	**<u>PM2</u>**	PM5	PM3	PM1	PM4
		0						0.50						0.50					
PHYS-C	PM5	**<u>PM5</u>**	PM4	PM6	PM2	PM3	PM1	**<u>PM5</u>**	PM6	PM2	PM4	PM3	PM1	**<u>PM5</u>**	PM6	PM4	PM2	PM3	PM1
		1						1						1					
PHYS- D	PM1	**<u>PM1</u>**	PM6	PM2	PM5	PM3	PM4	**<u>PM1</u>**	PM6	PM5	PM2	PM3	PM4	**<u>PM1</u>**	PM2	PM6	PM3	PM5	PM4
		1						1						1					
PHYS-E	PM4	PM6	PM5	PM2	**<u>PM4</u>**	PM3	PM1	PM6	PM5	PM2	PM1	**<u>PM4</u>**	PM3	PM5	PM6	PM2	PM3	PM1	**<u>PM4</u>**
		0						0						0					
PHYS- F	PM1	PM6	PM5	PM2	PM4	PM3	**<u>PM1</u>**	PM6	PM2	PM5	**<u>PM1</u>**	PM4	PM3	**<u>PM1</u>**	**PM3**	PM2	PM6	PM5	PM4
		0						0						0.50					
PHYS- G	PM4	**<u>PM4</u>**	PM5	PM6	PM3	PM2	PM1	**<u>PM4</u>**	PM5	PM6	PM2	PM3	PM1	**<u>PM4</u>**	PM5	PM3	PM6	PM2	PM1
		1						1						1					
PHYS- H	PM6	**<u>PM6</u>**	**PM2**	PM5	PM3	PM4	PM1	**<u>PM6</u>**	**PM2**	PM5	PM4	PM3	PM1	**PM2**	**<u>PM6</u>**	PM5	PM3	PM1	PM4
		0.50						1						0.50					
PHYS-I	PM3	PM5	PM4	PM6	**<u>PM3</u>**	PM2	PM1	PM5	PM6	PM2	**<u>PM3</u>**	PM4	PM1	**PM4**	**PM5**	**<u>PM3</u>**	**PM6**	**PM2**	PM1
		0						0						0.20					
	Score	**3.50**						**4.83**						**4.95**					

Table 106: Physics department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	**Main assessor**	**SAPV method in WoS SCs**						**WCS method in WoS SCs**						**Barycenter method in WoS SCs**					
PHYS-A	PM3	**PM 4**	**PM 3**	**PM 5**	**PM 6**	PM 2	PM 1	**PM3**	**PM5**	**PM4**	**PM6**	**PM2**	PM1	**PM 4**	**PM 3**	**PM 5**	**PM 6**	PM 1	PM 2
		0.25						0.20						0.25					
PHYS- B	PM2	PM 6	**PM 2**	PM 5	PM 4	PM 3	PM 1	**PM6**	**PM2**	PM5	PM4	PM3	PM1	PM 6	PM 5	**PM 2**	PM 1	PM 3	PM 4
		0						0.50						0					
PHYS-C	PM5	**PM 5**	PM 6	PM 2	PM 4	PM 3	PM 1	**PM5**	PM6	PM2	PM4	PM3	PM1	**PM 5**	PM 6	PM 2	PM 4	PM 3	PM 1
		1						1						0.50					
PHYS- D	PM1	**PM 1**	PM 3	PM 6	PM 4	PM 5	PM 2	**PM1**	PM6	PM3	PM5	PM4	PM2	**PM 1**	PM 3	PM 6	PM 5	PM 2	PM 4
		1						1						1					
PHYS-E	PM4	PM 6	PM 5	**PM 4**	PM 2	PM 3	PM 1	PM6	PM5	PM2	**PM4**	PM3	PM1	PM 5	PM 6	PM 3	PM 2	PM 1	**PM 4**
		0						0						0					
PHYS- F	PM1	PM 3	PM 6	PM 5	PM 2	PM 4	**PM 1**	PM6	PM3	PM2	**PM1**	PM5	PM4	**PM3**	**PM 1**	PM 5	PM 6	PM 4	PM 2
		0						0						0.50					
PHYS- G	PM4	**PM 4**	**PM 5**	**PM 6**	PM 2	PM 3	PM 1	**PM4**	PM5	PM6	PM2	PM3	PM1	**PM 4**	**PM 3**	**PM 5**	**PM 6**	PM 2	PM 1
		0.33						1						0.25					
PHYS- H	PM6	**PM 6**	**PM 2**	PM 5	PM 3	PM 4	PM 1	**PM6**	**PM2**	PM5	PM3	PM4	PM1	**PM 6**	**PM 5**	PM 2	PM 1	PM 3	PM 4
		0.50						0.50						0.50					
PHYS-I	PM3	**PM 4**	**PM 3**	**PM 5**	PM 6	PM 2	PM 1	**PM5**	**PM3**	**PM6**	**PM2**	**PM4**	PM1	**PM 4**	**PM 3**	**PM 5**	PM 1	PM 6	PM 2
		0.33						0.20						0.33					
	Score	**3.41**						**4.40**						**3.33**					

Table 107: Veterinary Science department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at journal level

Research groups	Main assessor	SAPV method in Journals				WCS method in Journals				Barycenter method in Journals			
VETE-A	PM4	PM2	PM1	**PM4**	PM3	PM2	**PM4**	PM1	PM3	PM2	PM1	**PM4**	PM3
		0				0				0			
VETE-B	PM2	**PM2**	PM1	PM4	PM3	**PM2**	**PM1**	**PM3**	PM4	**PM2**	**PM1**	**PM3**	PM4
		1				0.33				0.33			
VETE-C	PM1	**PM1**	PM3	PM2	PM4	**PM1**	PM3	PM2	PM4	**PM1**	**PM3**	**PM2**	PM4
		1				1				0.33			
	Score	**2**				**1.33**				**0.66**			

Table 108: Veterinary Science department's top ranked panel members according to Barycenter, SAPV and WCS methods (with confidence Intervals) at WoS SC level

Research groups	Main assessor	SAPV method in WoS SCs				WCS method in WoS SCs				Barycenter method in WoS SCs			
VETE-A	PM4	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM4**	**PM3**	**PM1**
		0.50				0.50				0.25			
VETE-B	PM2	**PM2**	**PM4**	PM1	PM3	**PM2**	**PM1**	**PM3**	PM4	**PM2**	**PM3**	**PM1**	**PM4**
		0.50				0.33				0.25			
VETE-C	PM1	**PM1**	PM3	PM2	PM4	**PM1**	PM3	PM2	PM4	**PM3**	**PM1**	**PM2**	PM4
		1				1				0.33			
	Score	**2**				**1.88**				**0.83**			

Pharmaceutical Sciences department

Research groups	Main assessor	SAPV method in Journals					WCS method in Journals					Barycenter method in Journals				
PHAR-A		PM5	PM3	PM4	PM1	PM2	PM5	PM1	PM4	PM2	PM3	PM1	PM5	PM4	PM3	PM2
PHAR-B		PM4	PM5	PM3	PM2	PM1	PM4	PM1	PM5	PM2	PM3	PM5	PM4	PM3	PM1	PM2
PHAR-C		PM2	PM3	PM4	PM5	PM1	PM2	PM4	PM5	PM3	PM1	PM2	PM3	PM4	PM5	PM1
PHAR-D		PM3	PM2	PM4	PM5	PM1	PM3	PM2	PM4	PM1	PM5	PM3	PM4	PM5	PM1	PM2
PHAR-E	Data is not	PM2	PM4	PM3	PM5	PM1	PM2	PM4	PM5	PM3	PM1	PM2	PM3	PM4	PM5	PM1
PHAR-F	available	PM4	PM5	PM1	PM3	PM2	PM1	PM4	PM5	PM2	PM3	PM5	PM4	PM3	PM1	PM2
PHAR-G		PM5	PM2	PM3	PM4	PM1	PM5	PM2	PM4	PM3	PM1	PM5	PM1	PM4	PM3	PM2
PHAR-H		PM4	PM5	PM3	PM2	PM1	PM4	PM1	PM2	PM5	PM3	PM3	PM4	PM5	PM2	PM1
PHAR-I		PM2	PM4	PM3	PM5	PM1	PM2	PM4	PM3	PM5	PM1	PM2	PM3	PM4	PM5	PM1
PHAR-J		PM2	PM4	PM3	PM5	PM1	PM2	PM5	PM4	PM3	PM1	PM2	PM3	PM4	PM5	PM1
Research groups	**Main assessor**	**SAPV method in WoS SCs**					**WCS method in WoS SCs**					**Barycenter method in WoS SCs**				
PHAR-A		PM5	PM1	PM4	PM3	PM2	PM5	PM1	PM3	PM4	PM2	PM1	PM5	PM3	PM4	PM2
PHAR-B		PM2	PM4	PM3	PM5	PM1	PM2	PM3	PM4	PM1	PM5	PM4	PM3	PM2	PM1	PM5
PHAR-C		PM2	PM3	PM4	PM5	PM1	PM2	PM3	PM4	PM5	PM1	PM2	PM4	PM3	PM1	PM5
PHAR-D		PM2	PM3	PM4	PM1	PM5	PM2	PM3	PM4	PM1	PM5	PM4	PM2	PM3	PM1	PM5
PHAR-E	Data is not	PM2	PM3	PM4	PM5	PM1	PM2	PM3	PM4	PM1	PM5	PM2	PM4	PM3	PM1	PM5
PHAR-F	available	PM4	PM3	PM2	PM1	PM5	PM5	PM2	PM3	PM1	PM4	PM4	PM3	PM1	PM5	PM2
PHAR-G		PM5	PM3	PM4	PM1	PM2	PM5	PM3	PM4	PM2	PM1	PM3	PM1	PM4	PM5	PM2
PHAR-H		PM2	PM4	PM3	PM5	PM1	PM2	PM3	PM4	PM1	PM5	PM2	PM4	PM3	PM1	PM5
PHAR-I		PM2	PM3	PM4	PM5	PM1	PM2	PM3	PM4	PM5	PM1	PM2	PM4	PM3	PM1	PM5
PHAR-J		PM2	PM3	PM4	PM5	PM1	PM2	PM3	PM4	PM5	PM1	PM2	PM4	PM3	PM1	PM5

Appendix F: List of scientific communications

List of original articles

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2017). Cognitive distances between evaluators and evaluees in research evaluation: a comparison between six informetric approaches. *Frontiers in Research Metrics & Analytics*, 2:6. doi: 10.3389/frma.2017.00006

Rousseau, R., Guns, R., Rahman, A. I. M. J., & Engels, T. C. E. (2017). Measuring cognitive distance between publication portfolios. *Journal of Informetrics*, *11*(2), 583-594

Rahman, A. I. M. J., Guns, R., Leydesdorff, L., & Engels, T. C. E. (2016). Measuring the match between evaluators and evaluees: cognitive distances between panel members and research groups at the journal level. *Scientometrics*, *109*(3), 1639–1663.

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2015). Is the expertise of evaluation panels congruent with the research interests of the research groups: a quantitative approach based on barycenters. *Journal of Informetrics*, *9*(4), 704–721.

Corrigendum

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2016). Corrigendum to "Is the expertise of evaluation panels congruent with the research interests of the research groups: a quantitative approach based on barycenters" (Journal of Informetrics (2015) 9(4) (704–721). *Journal of Informetrics*, 10(4), 1052-1054.

List of proceeding articles

Rahman, A.I.M.J., Guns, R., Rousseau, R., & Engels, T. C. E. (2015). Expertise overlap between an expert panel and research groups in global journal maps. In Albert Ali Salah, Yaşar Tonta, Alkım Almıla Akdağ Salah, Cassidy Sugimoto, & Umut Al (Eds.), *Proceedings of ISSI 2015 Istanbul: 15th International conference on scientometrics and informetrics, 29 June - 4 July 2015* (pp. 1035 – 1041). Istanbul: Boğaziçi University, Turkey.

Rahman, A.I.M.J., Guns, R., Rousseau, R., & Engels, T. C. E. (2014). Assessment of expertise overlap between an expert panel and research groups. In Ed Noyons (Ed.), *Context Counts: Pathways to Master Big and Little Data. Proceedings of the Science and Technology Indicators Conference 2014 Leiden* (pp. 295–301). Leiden: Leiden University.

List of presentations

"Cognitive distance between expert panels and units of assessment". A. I. M. Jakaria Rahman (presenter), Raf Guns, Ronald Rousseau, Loet Leydesdorff, Tim C. E. Engels - 20th Nordic Workshop on Bibliometrics and Research Policy - Oslo (Norway) - 1-2 October 2015.

"Expertise overlap between an expert panel and research groups in global journal maps". A. I. M. Jakaria Rahman (presenter), Raf Guns, Ronald Rousseau, Tim C. E. Engels – ISSI conference 2015 – Istanbul (Turkey) - 29 June to 4 July 2015.

"Assessment of expertise overlap between an expert panel and research groups". A. I. M. Jakaria Rahman (presenter), Raf Guns, Ronald Rousseau, Tim C. E. Engels - Science and technology indicators conference 2014 – Leiden (The Netherlands) – 3 to 5 September 2014.

List of technical reports

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Biology department (Technical report) (p. x, 77). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431560151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Biomedical Sciences department (Technical report) (p. xii, 93). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431570151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Chemistry department (Technical report) (p. x, 87). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431580151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Pharmaceutical Sciences department (Technical report) (p. x, 79). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431590151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Physics department (Technical report) (p. xii, 80). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431600151162165141

Rahman, A. I. M. J., & Guns, R. (2017). Determining cognitive distance between publication portfolios of evaluators and evaluees in research evaluation: A case study of Veterinary Sciences department (Technical report) (p. x, 68). Antwerp: University of Antwerp. http://hdl.handle.net/10067/1431610151162165141

Overview of authors of the included articles

A. I. M. Jakaria Rahman
Centre for Research & Developing Monitoring (ECOOM), Faculty of Social Sciences
University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium
jakaria.rahman@uantwerpen.be, jakaria@gmail.com
https://orcid.org/0000-0001-7876-4631

Raf Guns
Centre for Research & Developing Monitoring (ECOOM), Faculty of Social Sciences
University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium
raf.guns@uantwerpen.be
https://orcid.org/0000-0003-3129-0330

Ronald Rousseau
KU Leuven, Dept. of Mathematics, Celestijnenlaan 200B, 3001 Leuven, Belgium; and
Faculty of Social Sciences, University of Antwerp, 2000 Antwerp, Belgium
ronald.rousseau@kuleuven.be, ronald.rousseau@uantwerpen.be
http://orcid.org/0000-0002-3252-2538

Loet Leydesdorff
Amsterdam School of Communication Research (ASCoR), University of Amsterdam,
PO Box 15793, 1001 NG Amsterdam, The Netherlands
loet@leydesdorff.net
http://orcid.org/0000-0002-7835-3098

Tim C. E. Engels
Centre for Research & Developing Monitoring (ECOOM), Faculty of Social Sciences
University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium; and
Antwerp Maritime Academy, Noordkasteel Oost 6, B-2030 Antwerp, Belgium
tim.engels@uantwerpen.be
http://orcid.org/0000-0002-4869-7949

References

Abramo, G., & D’Angelo, C. A. (2011). Evaluating research: from informed peer review to bibliometrics. *Scientometrics*, *87*(3), 499–514.

Aggarwal, A., Imai, H., Katoh, N., & Suri, S. (1991). Finding k points with minimum diameter and related problems. *Journal of Algorithms*, *12*(1), 38–56.

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson’s correlation coefficient. *Journal of the Association for Information Science and Technology*, *54*(6), 550–560.

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research Evaluation*, *13*(1), 33–41.

Allen, L., Jones, C., Dolby, K., Lynn, D., & Walport, M. (2009). Looking for landmarks: the role of expert review and bibliometric analysis in evaluating scientific publication outputs. *PLOS ONE*, *4*(6), e5910.

Ankomah, P. K., & Crompton, J. L. (1992). Tourism cognitive distance: a set of research propositions. *Annals of Tourism Research*, *19*(2), 323–342.

Ankomah, P. K., Crompton, J. L., & Baker, D. (1996). Influence of cognitive distance in vacation choice. *Annals of Tourism Research*, *23*(1), 138–150.

Arnold, E. (2004). Evaluating research and innovation policy: a systems world needs systems evaluations. *Research Evaluation*, *13*(1), 3–17.

Baccini, A., & Nicolao, G. D. (2016). Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise. *Scientometrics*, *108*(3), 1651–1671.

Balázs, K., & Arnold, E. (1998). *Methods in the evaluation of publicly funded basic research: a review for OECD*. Brighton: Technopolis.

Barker, K. (2007). The UK research assessment exercise: the evolution of a national research evaluation system. *Research Evaluation*, *16*(1), 3–12.

Bazeley, P. (1998). Peer review and panel decisions in the assessment of Australian Research Council project grant applicants: what counts in a highly competitive context? *Higher Education*, *35*(4), 435–452.

Benda, W. G. G., & Engels, T. C. E. (2011). The predictive validity of peer review: a selective review of the judgmental forecasting qualities of peers, and implications for innovation in science. *International Journal of Forecasting*, *27*(1), 166–182.

Berendsen, R., de Rijke, M., Balog, K., Bogers, T., & Bosch, A. (2013). On the assessment of expertise profiles. *Journal of the American Society for Information Science and Technology*, *64*(10), 2024–2044.

Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., & Peracchi, F. (2015). Bibliometric evaluation vs. informed peer review: evidence from Italy. *Research Policy*, *44*(2), 451–466.

Borg, I., & Groenen, P. J. (2005). *Modern multidimensional scaling: theory and applications* (2nd ed.). New York: Springer.

Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, *36*(1), 2–72.

Börner, K. (2010). *Atlas of Science: visualizing what we know*. Massachusetts: MIT Press.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, *37*(1), 179–255.

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, *45*(1), 197–245.

Bornmann, L. (2014). Assigning publications to multiple subject categories for bibliometric analysis: an empirical case study based on percentiles. *Journal of Documentation*, *70*(1), 52–61.

Bornmann, L., & Daniel, H.-D. (2005). Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of board of trustees' decisions. *Scientometrics*, *63*(2), 297–320.

Bornmann, L., & Daniel, H.-D. (2006). Potential sources of bias in research fellowship assessments: effects of university prestige and field of study. *Research Evaluation*, *15*(3), 209–219.

Bornmann, L., & Leydesdorff, L. (2011). Which cities produce more excellent papers than can be expected? A new mapping approach, using Google Maps, based on statistical significance testing. *Journal of the Association for Information Science and Technology*, *62*(10), 1954–1962.

Bornmann, L., Mutz, R., Marx, W., Schier, H., & Daniel, H.-D. (2011). A multilevel modelling approach to investigating the predictive validity of editorial decisions: do the editors of a high profile journal select manuscripts that are highly cited after publication? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *174*(4), 857–879.

Borum, F., & Hansen, H. F. (2000). The local construction and enactment of standards for research evaluation: the case of the Copenhagen Business School. *Evaluation*, *6*(3), 281–299.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, *39*(1), 61–74.

Boyack, K. W. (2009). Using detailed maps of science to identify potential collaborations. *Scientometrics*, *79*(1), 27.

Boyack, K. W., Chen, M.-C., & Chacko, G. (2014). Characterization of the peer review network at the center for scientific review, National institutes of health. *PLoS ONE*, *9*(8), e104244.

Boyack, K. W., & Klavans, R. (2014a). Creation of a highly detailed, dynamic, global model and map of science. *Journal of the Association for Information Science and Technology*, *65*(4), 670–685.

Boyack, K. W., & Klavans, R. (2014b). Including cited non-source items in a large-scale map of science: what difference does it make? *Journal of Informetrics*, *8*(3), 569–580.

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, *64*(3), 351–374.

Broström, A., & McKelvey, M. (2016). *Knowledge transfer at the science policy interface: how cognitive distance and the degree of expert autonomy shapes the outcome* (CESIS electronic working paper series No. 441). The Royal institute of technology: center of excellence for science and innovation studies (CESIS). Retrieved from https://static.sys.kth.se/itm/wp/cesis/cesiswp441.pdf

Buckley, H. L., Sciligo, A. R., Adair, K. L., Case, B. S., & Monks, J. M. (2014). Is there gender bias in reviewer selection and publication success rates for the New Zealand Journal of Ecology? *New Zealand Journal of Ecology*, *38*(2), 335–339.

Buter, R. K., Noyons, E. C. M., & Van Raan, A. F. J. (2004). A combination of quantitative and qualitative maps in an evaluative bibliometric context. In E. Banissi, K. Borner, C. Chen, M. Dastbaz, G. Clapworthy, A. Faiola, … J. J. Zhang (Eds.), *Eighth International Conference on Information Visualisation, Proceedings* (pp. 978–982). London, UK: IEEE computer society.

Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy, 34*(8), 565–574.

Campanario, J. M. (1998a). Peer review for journals as it stands today—Part 1. *Science Communication*, *19*(3), 181–211.

Campanario, J. M. (1998b). Peer review for journals as it stands today—Part 2. *Science Communication*, *19*(4), 277–306.

Casey-Campbell, M., & Martens, M. L. (2009). Sticking it all together: a critical assessment of the group cohesion–performance literature. *International Journal of Management Reviews*, *11*(2), 223–246.

Cattaneo, M., Meoli, M., & Signori, A. (2016). Performance-based funding and university research productivity: the moderating effect of university legitimacy. *The Journal of Technology Transfer*, *41*(1), 85–104.

Chen, C. (2006a). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*(3), 359–377.

Chen, C. (2006b). *Information visualization: beyond the horizon* (2nd edition). London: Springer-Verlag.

Chen, C. (2013). *Mapping scientific frontiers: the quest for knowledge visualization* (2nd ed.). London: Springer-Verlag.

Chen, S., Arsenault, C., Gingras, Y., & Lariviere, V. (2015). Exploring the interdisciplinary evolution of a discipline: the case of Biochemistry and Molecular Biology. *Scientometrics*, *102*(2), 1307–1323.

Cicchetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: a cross-disciplinary investigation. *Behavioral and Brain Sciences*, *14*(1), 119–135.

Clarivate Analytics. (2016). Web of knowledge - real facts. Retrieved January 10, 2017, from http://wokinfo.com/citationconnection/realfacts/#regional

Clarivate Analytics. (2017). Web of Science core collection help. Retrieved from https://images.webofknowledge.com/images/help/WOS/hp_subject_category_terms_tasca.html

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: review, analysis, and cooperative study among tools. *Journal of the Association for Information Science and Technology*, *62*(7), 1382–1402.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. *The Economic Journal*, *99*(397), 569–596.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Cook, W. D., Golany, B., Kress, M., Penn, M., & Raviv, T. (2005). Optimal allocation of proposals to reviewers to facilitate effective ranking. *Management Science*, *51*(4), 655–661.

Coryn, C. L. S., & Scriven, M. (2008). Editor's notes. *Reforming the evaluation of research: new directions for evaluation, 118,* 1-5.

Creswell, J. W. (2014). *Educational research: planning, conducting, and evaluating quantitative and qualitative Research* (5th edition). New Jersey: Pearson.

Cronin, B. (2005). *The hand of science: academic writing and its rewards*. Lanham, MD: Scarecrow Press.

Dahler-Larsen, P. (2012). Evaluation as a situational or a universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look like in practice, and why it is not fashionable. *Scandinavian Journal of Public Administration, 16*(3), 29–46.

Daniel, H., Mittag, S., & Bornman, L. (2007). The potential and problems of peer evaluation in higher education and research. In A. Cavalli (Ed.), *Quality Assessment for Higher Education in Europe* (pp. 71–82). London, UK: Portland Press.

de Nooy, W., Mrvar, A., & Batagelj, V. (2012). *Exploratory social network analysis with Pajek* (2nd edition). New York: Cambridge University Press.

Demicheli, V., & Pietrantonj, C. D. (2007). Peer review for improving the quality of grant applications. *Ochrane Database of Systematic Reviews*, Issue 2, Art. No.: MR00000.

Dumais, S. T., & Nielsen, J. (1992). Automating the assignment of submitted manuscripts to reviewers. In *SIGIR '92 Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 233–244). Copenhagen, Denmark: ACM Press.

Efron, B., & Tibshirani, R. J. (1998). *An introduction to the bootstrap*. London: Chapman & Hall/CRC.

Egghe, L., & Rousseau, R. (1990). *Introduction to Informetrics*. Elsevier Science Publishers. Retrieved from https://uhdspace.uhasselt.be/dspace/handle/1942/587

Eisenhart, M. (2002). The Paradox of peer review: admitting too much or allowing too little? *Research in Science Education*, *32*(2), 241–255.

Elsevier. (2016). *Scopus: content coverage guide*. Retrieved from https://www.elsevier.com/solutions/scopus/content

Engels, T. C. E., Goos, P., Dexters, N., & Spruyt, E. H. J. (2013). Group size, h-index, and efficiency in publishing in top journals explain expert panel assessments of research group quality and productivity. *Research Evaluation*, *22*(4), 224–236.

Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics*, *93*(2), 373–390.

ESF. (2011). *European peer review guide: integrating policies and practices into coherent procedures*. Strasbourg: European Science Foundation.

Fedderke, J. W. (2013). The objectivity of national research foundation peer review in South Africa assessed against bibliometric indexes. *Scientometrics*, *97*(2), 177–206.

Fields, C. (2015). How small is the center of science? Short cross-disciplinary cycles in co-authorship graphs. *Scientometrics*, *102*(2), 1287–1306.

Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: a bibliometric perspective. *Journal of Informetrics*, *5*(2), 275–291.

Franklin, J. J., & Johnston, R. (1988). Co-citation bibliometric modeling as a tool for S&T policy and R&D management: issues, applications, and developments. In A. F. J. Van Raan (Ed.), *Handbook of quantitative science and technology research* (Vols. 1–325-389). Amsterdam: Elsevier Science Publishers.

Garfield, E., Pudovkin, A. I., & Istomin, V. S. (2003). Why do we need algorithmic historiography? *Journal of the Association for Information Science and Technology*, *54*(5), 400–412.

Garfield, E., Sher, I. H., & Torpie, R. J. (1964). *The use of citation data in writing the history of science*. Institute for Scientific Infomration Inc Philadelphia.

Geuna, A., & Martin, B. R. (2003). University research evaluation and funding: an international comparison. *Minerva*, *41*(4), 277–304.

Godin, B. (2006). On the origins of bibliometrics. *Scientometrics*, *68*(1), 109–133.

Goldfinch, S., & Yamamoto, K. (2012). *Prometheus assessed?: research measurement, peer review, and citation analysis*. Oxford: Chandos Publishing.

Golledge, R. G. (1987). Environmental cognition. In D. Stokols & I. Altman (Eds.), *Handbook of Environmental Psychology* (Vol. 1, pp. 131–174). New York: Wiley.

Golub, G. H., & Van Loan, C. F. (1996). *Matrix computations* (3rd ed). Baltimore, Md: Johns Hopkins University Press.

Gorjiara, T., & Baldock, C. (2014). Nanoscience and nanotechnology research publications: a comparison between Australia and the rest of the world. *Scientometrics*, *100*(1), 121–148.

Gould, T. H. P. (2013). *Do we still need peer review?: An argument for change*. UK: Scarecrow Press.

Grabher, G., & Ibert, O. (2013). Distance as asset? Knowledge collaboration in hybrid virtual communities. *Journal of Economic Geography*.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Granovetter, M. S. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory*, *1*, 201–233.

Grauwin, S., & Jensen, P. (2011). Mapping scientific institutions. *Scientometrics, 89*(3), 943–954.

Griffith, B. C., Small, H. G., Stonehill, J. A., & Dey, S. (1974). The Structure of scientific literatures II: toward a macro- and microstructure for science. *Science Studies*, *4*, 339-365.

Guerrero-Bote, V. P., Zapico-Alonso, F., Espinosa-Calvo, M. E., Gómez-Crisóstomo, R., & Moya-Anegón, F. de. (2007). Import-export of knowledge between scientific subject categories: the iceberg hypothesis. *Scientometrics*, *71*(3), 423–441.

Guns, R. (2016a). Bootstrapping confidence intervals for the distances between barycenters. Retrieved from http://nbviewer.jupyter.org/gist/rafguns/6fa3460677741e356538337003692389

Guns, R. (2016b). Confidence intervals for weighted cosine similarity. Retrieved from http://nbviewer.jupyter.org/gist/rafguns/faff8dc090b67a783b85d488f88952ba

Guston, D. H. (2003). The expanding role of peer review processes in the United States. In P. Shapira & S. Kuhlmann (Eds.), *Learning from science and technology policy evaluation: experiences from the United States and Europe* (pp. 81–97). Cheltenham, UK: Edward Elgar Publishing.

Hammarfelt, B., & de Rijcke, S. (2015). Accountability in context: effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. *Research Evaluation*, *24*(1), 63–77.

Hansson, F. (2010). Dialogue in or with the peer review? Evaluating research organizations in order to promote organizational learning. *Science and Public Policy*, *37*(4), 239–251.

Harrison-Hill, T. (2001). How far is a long way? contrasting two cultures' perspectives of travel distance. *Asia Pacific Journal of Marketing and Logistics*, *13*(3), 3–17.

Harzing, A.-W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804.

Hashemi, S. H., Neshati, M., & Beigy, H. (2013). Expertise retrieval in bibliographic network: a topic dominance learning approach. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management* (pp. 1117–1126). San Francisco, US: ACM.

Hautala, J. (2013). Cognitive proximity in international research groups. *Journal of Knowledge Management*, *15*(4), 601–624.

Healey, P., Rothman, H., & Hoch, P. K. (1986). An experiment in science mapping for research planning. *Research Policy*, *15*(5), 233–251.

HEFCE. (2015). *The Metric tide: correlation analysis of REF2014 scores and metrics (Supplementary report II to the independent review of the role of metrics in research assessment and management)*. Higher education funding council for England (HEFCE).

Hemlin, S. (1996). Research on research evaluation. *Social Epistemology*, *10*(2), 209–250.

Hemlin, S., & Rasmussen, S. B. (2006). The Shift in academic quality control. *Science, Technology & Human Values*, *31*(2), 173–198.

Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems* (pp. 473–496). Dordrecht: Kluwer Academic Publishers.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, *41*(2), 251–261.

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. D., & Rafols, I. (2015). Bibliometrics: the Leiden manifesto for research metrics. *Nature*. 520 (7548) : 429-31.

Hofmann, K., Balog, K., Bogers, T., & de Rijke, M. (2010). Contextual factors for finding similar experts. *Journal of the American Society for Information Science and Technology*, *61*(5), 994–1014.

Holbrook, J. B. (2010). Peer review. In R. Frodeman (Ed.), *The Oxford handbook of interdisciplinarity* (pp. 321–332). New York: Oxford University Press.

Howlett, M. (2009). Policy analytical capacity and evidence-based policy-making: Lessons from Canada. *Canadian Public Administration, 52*(2), 153–175.

Huson, L. W. (2007). Performance of some correlation coefficients when applied to zero-clustered data. *Journal of Modern Applied Statistical Methods*, *6*(2), 530–536.

Iman, R. L., & Conover, W. J. (1987). A measure of Top-Down correlation. *Technometrics*, *29*(3), 351–357.

Janak, S. L., Taylor, M. S., Floudas, C. A., Burka, M., & Mountziaris, T. J. (2006). Novel and effective integer optimization approach for the NSF panel-assignment problem: a multiresource and preference-constrained generalized assignment problem. *Industrial & Engineering Chemistry Research*, *45*(1), 258–265.

Jin, B., & Rousseau, R. (2001). An introduction to the barycenter method with an application to China's mean centre of publication. *Libri*, *51*, 225–233.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, *31*(1), 7–15.

Kay, L., Newman, N., Youtie, J., Porter, A. L., & Rafols, I. (2014). Patent overlay mapping: visualizing technological distance. *Journal of the Association for Information Science and Technology*, *65*(12), 2432–2443.

Kington, J. (2014). Balanced cross sections, shortening estimates, and the magnitude of out-of-sequence thrusting in the Nankai trough accretionary prism, Japan. *Figshare*.

Klavans, R., & Boyack, K. (2007). Is there a convergent structure of science? A comparison of maps using the ISI and Scopus databases. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of ISSI 2007* (vol. 1, pp. 437–448). Madrid, Spain: Center for Information and Scientific Documentation (CINDOC) and Spanish National Research Council (CSIC).

Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics*, *68*(3), 475–499.

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the Association for Information Science and Technology*, *60*(3), 455–476.

Kogan, M. (1989). The Evaluation of higher education: an introductory note. In *Evaluating Higher Education*. London: Jessica Kingsley Publisher.

Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010a). The iceberg hypothesis revisited. *Scientometrics*, *85*(2), 443–461.

Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010b). What lies behind the averages and significance of citation indicators in different disciplines? *Journal of Information Science*, *36*(3), 371–382.

Langfeldt, L. (2004). Expert panels evaluating research: decision-making and sources of bias. *Research Evaluation*, *13*(1), 51–62.

Lawrenz, F., Thao, M., & Johnson, K. (2012). Expert panel reviews of research centers: the site visit process. *Evaluation and Program Planning*, *35*(3), 390–397.

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the Association for Information Science and Technology*, *64*(1), 2–17.

Leydesdorff, L. (1987). Various methods for the mapping of science. *Scientometrics*, *11*(5–6), 295–324.

Leydesdorff, L. (2004). Top-down decomposition of the journal citation reportof the Social Science citation index: graph- and factor-analytical approaches. *Scientometrics*, *60*(2), 159–180.

Leydesdorff, L. (2006). Can scientific journals be classified in terms of aggregated journal-journal citation relations using the journal citation reports? *Journal of the American Society for Information Science and Technology*, *57*(5), 601–613.

Leydesdorff, L., & Bornmann, L. (2012). Mapping (USPTO) patent data using overlays to Google Maps. *Journal of the American Society for Information Science and Technology*, *63*(7), 1442–1458.

Leydesdorff, L., & Bornmann, L. (2016). The operationalization of "fields" as WoS subject categories (WCs) in evaluative bibliometrics: The cases of "library and information science" and "science & technology studies." *Journal of the Association for Information Science and Technology*, *67*(3), 707–714.

Leydesdorff, L., Bornmann, L., & Zhou, P. (2016). Construction of a pragmatic base line for journal classifications and maps based on aggregated journal-journal citation relations. *Journal of Informetrics*, *10*(4), 902–918.

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new web-of-science categories. *Scientometrics*, *94*(2), 589–593.

Leydesdorff, L., de Moya-Anegón, F., & de Nooy, W. (2016). Aggregated journal–journal citation relations in scopus and web of science matched and compared in terms of networks, maps, and interactive overlays. *Journal of the Association for Information Science and Technology*, *67*(9), 2194–2211.

Leydesdorff, L., de Moya-Anegón, F., & Guerrero-Bote, V. P. (2010). Journal maps on the basis of Scopus data: a comparison with the journal citation reports of the ISI. *Journal of the Association for Information Science and Technology*, *61*(2), 352–369.

Leydesdorff, L., & de Nooy, W. (2016). Can “Hot Spots” in the Sciences be mapped using the dynamics of aggregated journal-journal citation relations? *Journal of the Association for Information Science and Technology*, *68*(21), 197-213.

Leydesdorff, L., Heimeriks, G., & Rotolo, D. (2016). Journal portfolio analysis for countries, cities, and organizations: maps and comparisons. *Journal of the Association for Information Science and Technology*, *67* (3), 741–748.

Leydesdorff, L., Kushnir, D., & Rafols, I. (2014). Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC). *Scientometrics*, *98*(3), 1583–1599.

Leydesdorff, L., Moya-Anegón, F., & Guerrero-Bote, V. P. (2015). Journal maps, interactive overlays, and the measurement of interdisciplinarity on the basis of Scopus data (1996–2012). *Journal of the Association for Information Science and Technology*, *66*(5), 1001–1016.

Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: distribution patterns and networks of relations among cities and institutes. *Journal of the Association for Information Science and Technology*, *61*(8), 1622–1634.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the Association for Information Science and Technology*, *60*(2), 348–362.

Leydesdorff, L., & Rafols, I. (2012). Interactive overlays: a new method for generating global journal maps from Web-of-Science data. *Journal of Informetrics*, *6*(2), 318–332.

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal–journal citations. *Journal of the American Society for Information Science and Technology*, *64*(12), 2573–2586.

Li, D., & Agha, L. (2015). Big names or big ideas: Do peer-review panels select the best science proposals? *Science*, *348*(6233), 434–438.

Li, J., Sanderson, M., Willett, P., Norris, M., & Oppenheim, C. (2010). Ranking of library and information science researchers: comparison of data sources for correlating citation data, and expert judgments. *Journal of Informetrics*, *4*(4), 554–563.

Lovegrove, B. G., & Johnson, S. D. (2008). Assessment of research performance in Biology: how well do peer review and bibliometry correlate? *BioScience*, *58*(2), 160–164.

Massy, W. F. (Ed.). (1996). *Resource allocation in higher education*. Michigan: University of Michigan Press.

McCullough, J. (1989). First comprehensive survey of NSF applicants focuses on their concerns about proposal review. *Science, Technology, & Human Values*, *14*(1), 78–88.

McDavid, J. C., Huse, I., & Ingleson, L. R. L. (2012). *Program evaluation and performance measurement: an introduction to practice* (2nd edition). Los Angeles: SAGE.

McKenna, H. P. (2015). Research assessment: the impact of impact. *International Journal of Nursing Studies*, *52*(1), 1–3.

Milat, A. J., Bauman, A. E., & Redman, S. (2015). A narrative review of research impact assessment models and methods. *Health Research Policy and Systems, 13(*18), 1-7.

Mimno, D., & McCallum, A. (2007). Expertise modeling for matching papers with reviewers. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 500–509). California: ACM.

Moed, H. F. (2002). The impact-factors debate: the ISI's uses and limits. *Nature*, *415*(6873), 731–732.

Moed, H. F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy*, *34*(8), 575–583.

Molas-Gallart, J. (2012). Research governance and the role of evaluation: a comparative study. *American Journal of Evaluation*, *33*(4), 583–598.

Molas-Gallart, J., D'Este, P., Llopis, O., & Rafols, I. (2015). Towards an alternative framework for the evaluation of translational research initiatives. *Research Evaluation*, *25*(3), 235–243.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, *106*(1), 213–228.

Montello, D. R. (1991). The measurement of cognitive distance: methods and construct validity. *Journal of Environmental Psychology*, *11*, 101–122.

Morris, S. A., & Martens, B. V. der V. (2008). Mapping research specialties. *Annual Review of Information Science and Technology*, *42*(1), 213–295.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., & Munoz-Fernández, F. J. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics*, *61*(1), 129–145.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: a Review. *Scientometrics*. *66*(1), 81-100.

Nederhof, A. J., & van Raan, A. F. J. (1993). A bibliometric analysis of six economics research groups: a comparison with peer review. *Research Policy*, *22*(4), 353–368.

Nedeva, M., Georghiou, L., Loveridge, D., & Cameron, H. (1996). The use of co-nomination to identify expert participants for technology foresight. *R&D Management*, *26*(2), 155–168.

Neshati, M., Beigy, H., & Hiemstra, D. (2012). Multi-aspect group formation using facility location analysis. In *Proceedings of the Seventeenth Australasian Document Computing Symposium* (pp. 62–71). New York, USA: ACM.

Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R., … Levine, K. (2015). Peer review: still king in the digital age. *Learned Publishing*, *28*(1), 15–21.

Nooteboom, B. (1999). *Inter-firm alliances: analysis and design*. London: Routledge.

Nooteboom, B. (2000). Learning by interaction: absorptive capacity, cognitive distance and governance. *Journal of Management and Governance*, *4*(1–2), 69–92.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., & van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, *36*(7), 1016–1034.

Norris, M., & Oppenheim, C. (2003). Citation counts and the research assessment exercise V: archaeology and the 2001 RAE. *Journal of Documentation*, *59*(6), 709–730.

Noyons, E. C. M. (2001). Bibliometric mapping of science in a policy context. *Scientometrics*, *50*(1), 83–98.

Noyons, E. C. M. (2004). Science maps within a science policy context: improving the utility of science and domain maps within a science policy and research management context. In H. F. Moed, W. Glänzel, & Ulrich Schmoch (Eds.), *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems* (pp. 237–255). Netherlands: Springer.

OECD. (1997). *The evaluation of scientific research: selected experiences* (pp. 1–112). Paris: OECD. Retrieved from http://www.oecd.org/science/sci-tech/2754549.pdf

OECD. (2015). *Frascati manual 2015: guidelines for collecting and reporting data on research and experimental development*. Paris: OECD Publishing.

Oleinik, A. (2014). Conflict(s) of interest in peer review: its origins and possible solutions. *Science and Engineering Ethics*, *20*(1), 55–75.

Oppenheim, C. (1997). The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology. *Journal of Documentation*, *53*(5), 477–487.

Over, R. (1996). Perceptions of the Australian research council large grants scheme: differences between successful and unsuccessful applicants. *The Australian Educational Researcher*, *23*(2), 17–36.

Papponetti, V., & Bucchi, M. (2007). *Research evaluation as a policy design tool: mapping approaches across a set of case studies* (FEEM (Fondazione Eni Enrico Mattei) Research paper series No. 75). Trento: University of Trento.

Patterson, M. S., & Harris, S. (2009). The relationship between reviewers' quality-scores and number of citations for papers published in the journal Physics in Medicine and Biology from 2003–2005. *Scientometrics*, *80*(2), 343–349.

Pina, D. G., Hren, D., & Marušić, A. (2015). Peer review evaluation process of Marie Curie actions under EU's seventh framework programme for research. *PLoS ONE*, *10*(6), e0130753.

Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, *53*(13), 1113–1119.

RAE2001. (2002). Overview of the research assessment exercise. Retrieved August 19, 2016, from http://www.rae.ac.uk/2001/pubs/2_99/section1.htm

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, *61*(9), 1871–1887.

Rahm, E. (2008). Comparing the scientific impact of conference and journal publications in computer science. *Information Services and Use*, *28*(2), 127–128.

Rahman, A. I. M. J., Guns, R., Leydesdorff, L., & Engels, T. C. E. (2016). Measuring the match between evaluators and evaluees: cognitive distances between panel members and research groups at the journal level. *Scientometrics*, *109*(3), 1639–1663.

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2014). Assessment of expertise overlap between an expert panel and research groups. In Ed Noyons (Ed.), *Context counts: Pathways to master big and little data. Proceedings of the science and technology indicators conference 2014 Leiden* (pp. 295–301). Leiden: Universiteit Leiden.

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2015). Is the expertise of evaluation panels congruent with the research interests of the research groups: a quantitative approach based on barycenters. *Journal of Informetrics*, *9*(4), 704–721.

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2016). Corrigendum to “Is the expertise of evaluation panels congruent with the research interests of the research groups: a quantitative approach based on barycenters” (Journal of Informetrics (2015) 9(4) (704–721). *Journal of Informetrics*, *10*(4), 1052-1054.

Rahman, A. I. M. J., Guns, R., Rousseau, R., & Engels, T. C. E. (2017). Cognitive distances between evaluators and evaluees in research evaluation: a comparison between three informetric methods at the journal and subject category aggregation level. *Frontiers in Research Metrics and Analytics*, *6*(2).

Ramos, A., & Sarrico, C. S. (2016). Past performance does not guarantee future results: lessons from the evaluation of research units in Portugal. *Research Evaluation*, *25*(1), 94–106.

REF2014. (2014). Research excellence framework. Retrieved from http://www.ref.ac.uk/

Rehn, C., Kronman, U., Gornitzki, C., Larsson, A., & Wadskog, D. (2014). *Bibliometric handbook for Karolinska Institutet*. Stockholm, Sweden: Karolinska Institute.

Reinhart, M. (2009). Peer review of grant applications in biology and medicine. Reliability, fairness, and validity. *Scientometrics*, *81*(3), 789–809.

Rinia, E. J., van Leeuwen, T. N., van Vuren, H. G., & van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria: evaluation of condensed matter physics in the Netherlands. *Research Policy*, *27*(1), 95–107.

Riopelle, K., Leydesdorff, L., & Jie, L. (2014). *How to create an overlay map of science using the Web of Science* (pp. 1–38). Retrieved from https://www.leydesdorff.net /overlaytoolkit/manual.riopelle.pdf

Rip, A., & Meulen, B. J. R. van der. (1995). The patchwork of the Dutch evaluation system. *Research Evaluation*, *5*(1), 45–53.

Rodriguez, M. A., & Bollen, J. (2008). An algorithm to determine peer-reviewers. In *Proceedings of the 17th ACM conference on Information and knowledge management* (pp. 319–328). California: ACM.

Rons, N., Bruyn, A. D., & Cornelis, J. (2008). Research evaluation per discipline: a peer-review method and its outcomes. *Research Evaluation*, *17*(1), 45–57.

Rousseau, R. (1989a). Kinematical statistics of scientific output. Part I: geographical approach. *Revue Française de Bibliométrie*, *4*, 50–64.

Rousseau, R. (1989b). Kinematical statistics of scientific output. Part II: geographical approach. *Revue Française de Bibliométrie*, *4*, 65–77.

Rousseau, R. (2008). Triad or Tetrad: another representation. *ISSI Newsletter*, *4*(1), 5–7.

Rousseau, R., Guns, R., Rahman, A. I. M. J., & Engels, T. C. E. (2017). Measuring cognitive distance between publication portfolios. *Journal of Informetrics*, *11*(2), 583–594.

Rousseau, R., Rahman, A. I. M. J., Guns, R., & Engels, T. C. E. (2016). A note and a correction on measuring cognitive distance in multiple dimensions. *ArXiv:1602.05183 [Cs]*. Retrieved from http://arxiv.org/abs/1602.05183

Rybak, J., Balog, K., & Nørvåg, K. (2014). ExperTime: tracking expertise over time. In *Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval* (pp. 1273–1274). Broadbeach, Australia: ACM.

Salas, E., Grossman, R., Hughes, A. M., & Coultas, C. W. (2015). Measuring team cohesion: observations from the Science. *Human Factors*, *57*(3), 365–374.

Salton, G., & McGill, M. J. (1986). *Introduction to modern information retrieval*. New York: McGraw-Hill, Inc.

Shiffrin, R. M., & Börner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences*, *101*(suppl 1), 5183–5185.

Simon, D., & Knie, A. (2013). Can evaluation contribute to the organizational development of academic institutions? An international comparison. *Evaluation*, *19*(4), 402–418.

Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, *107*(2), 357–368.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science*, *50*(9), 799–813.

Small, H., & Garfield, E. (1985). The geography of science: disciplinary and national mappings. *Journal of Information Science*, *11*, 147–159.

Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, *99*(4), 178–182.

Sobkowicz, P. (2015). Innovation suppression and clique evolution in peer-review-based, competitive research funding systems: an agent-based model. *Journal of Artificial Societies and Social Simulation*, *18*(2), article number: UNSP 13.

Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, *38 (Part II)* (20), 1409–1438.

Sombatsompop, N., & Markpin, T. (2005). Making an equality of ISI impact factors for different subject fields. *Journal of the Association for Information Science and Technology*, *56*(7), 676–683.

Soós, S., & Kampis, G. (2012). Beyond the basemap of science: mapping multiple structures in research portfolios: evidence from Hungary. *Scientometrics*, *93*(3), 869–891.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society, Interface*, *4*(15), 707–719.

Suter, L. E. (1997). United States: the experience of the NSF's education and human resources directorate. In OECD (Ed.), *The Evaluation of Scientific Research: Selected Experiences* (pp. 107–112). Paris: OECD.

Taylor, J. (2011). The assessment of research quality in UK universities: peer review or metrics? *British Journal of Management*, *22*(2), 202–217.

Tian, Q., Ma, J., & Liu, O. (2002). A hybrid knowledge and model system for R&D project selection. *Expert Systems with Applications*, *23*(3), 265–271.

Tseng, Y. H., & Tsay, M. Y. (2013). Journal clustering of library and information science for subfield delineation using the bibliometric analysis toolkit: CATAR. *Scientometrics*, *95*(2), 503–528.

Tu, W. (2006). Zero-inflated data. In Abdel H. El-Shaarawi & Walter W. Piegorsch (Eds.), *Encyclopedia of Environmetrics* (Vol. 4, pp. 2387–2391). Chichester: John Wiley & Sons.

van den Besselaar, P., & Leydesdorff, L. (2009). Past performance, peer review and project selection: a case study in the social and behavioral sciences. *Research Evaluation*, *18*(4), 273–288.

van Eck, N. J. (2011). *Methodological advances in bibliometric mapping of science* (PhD Thesis). Erasmus Research Institute of Management, Rotterdam. Retrieved from http://repub.eur.nl/pub/26509

van Eck, N. J., & Waltman, L. (2007). VOS: a new method for visualizing similarities between objects. In R. Decker & H.-J. Lenz (Eds.), *Advances in Data Analysis: Proceedings of the 30th annual conference of the German Classification Society advances in data analysis* (pp. 299–306). Heidelberg: Springer.

van Eck, N. J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, *60*(8), 1635–1651.

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact* (pp. 285–320). London: Springer International Publishing.

van Eck, N. J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, *61*(12), 2405–2416.

van Leeuwen, T. N., & Medina, C. C. (2012). Redefining the field of economics: Improving field normalization for the application of bibliometric techniques in the field of economics. *Research Evaluation*, *21*(1), 61–70.

van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. J. (2003). The Holy grail of science policy: exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, *57*(2), 257–280.

van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics*, *36*(3), 397–420.

van Raan, A. F. J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, *62*(1), 133–143.

van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.

van Steen, J., & Eijffinger, M. (1998). Evaluation practices of scientific research in the Netherlands. *Research Evaluation*, *7*(2), 113–122.

Verleysen, F. T., & Engels, T. C. E. (2014a). Barycenter representation of book publishing internationalization in the Social Sciences and Humanities. *Journal of Informetrics*, *8*(1), 234–240.

Verleysen, F. T., & Engels, T. C. E. (2014b). Internationalization of peer reviewed and non-peer reviewed book publications in the Social Sciences and Humanities. *Scientometrics*, *101*(2), 1431–1444.

VSNU. (2003). *Standard evaluation protocol 2003-2009 for public research organisations*. Utrecht/den Haag/Amsterdam: VSNU, NWO and KNAW.

VSNU. (2009). *Standard evaluation protocol 2009-2015: protocol for research assessment in The Netherlands*. Utrecht/den Haag/Amsterdam: VSNU, NWO and KNAW.

VSNU, KNAW, & NWO. (2014). *Standard evaluation protocol 2015-2021: protocol for research assessment in The Netherlands*. Utrecht/den Haag/Amsterdam: VSNU, NWO and KNAW.

Wainer, J., & Vieira, P. (2013). Correlations between bibliometrics and peer evaluation for all disciplines: the evaluation of Brazilian scientists. *Scientometrics*, *96*(2), 395–410.

Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*(12), 2378–2392.

Wang, F., Zhou, S., & Shi, N. (2013). Group-to-group reviewer assignment problem. *Computers & Operations Research*, *40*(5), 1351–1362.

Wang, Q., & Sandström, U. (2015). Defining the role of cognitive distance in the peer review process with an explorative study of a grant scheme in infection biology. *Research Evaluation*, *24*(3), 271–281.

Warner, J. (2003). Citation analysis and research assessment in the United Kingdom. *Bulletin of the American Society for Information Science and Technology*, *30*(1), 26–27.

Wessely, S. (1998). Peer review of grant applications: what do we know? *The Lancet*, *352*(9124), 301–305.

Whitley, R. (2007). Changing governance of the public sciences. In R. Whitley & J. Gläser (Eds.), *The Changing governance of the sciences: The advent of research evaluation systems* (Vol. 26, pp. 3–27). Netherlands: Springer.

Wood, F. Q., & Wesseley, S. (2003). Peer review of grant applications: a systematic review. In F. Godlef & T. Jefferson (Eds.), *Peer Review in Health Sciences* (pp. 14–44). London, UK: BMJ Books.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., Rijcke, S. de, Rushforth, A., & Franssen, T. (2015). *The Metric tide: literature review (supplementary report I to the independent review of the role of metrics in research assessment and management)*. Higher education funding council for England (HEFCE).

Wuyts, S., Colombo, M. G., Dutta, S., & Nooteboom, B. (2005). Empirical tests of optimal cognitive distance. *Journal of Economic Behavior & Organization*, *58*(2), 277–302.

Zhou, Q., Rousseau, R., Yang, L., Yue, T., & Yang, G. (2012). A general framework for describing diversity within systems and similarity between systems with applications in informetrics. *Scientometrics*, *93*(3), 787–812.

Ziman, J. (2002). *Real science. what it is, and what it means*. Cambridge, UK: Cambridge University Press.