EDITORIAL

SPECIAL FOCUS: THE GLOBAL COVID-19 CHALLENGE

WOLFGANG GLÄNZEL1 & LIN ZHANG1,2
1 ECOOM and Dept. MSI, KU Leuven, Belgium
2 School of Information Management, Wuhan University, China

This number of the ISSI newsletter comes out with considerable delay. There were and still are severe reasons for that delay: The collection and preparation of the contributions for the present number was accompanied and overshadowed by the outbreak and spread of the COVID-19 pandemic. Therefore, we decided to put a relevant article – on modelling the diffusion patterns of COVID-19 and a scenario-driven thinking in forecasting the course of the pandemic – at the heart of the present issue of the Newsletter. The paper by Decock et al. (p. 2–6) is co-authored by our colleagues at KU Leuven. Extended versions and follow-ups of this contribution will be published in international scientific journals very soon.

In this context we would like to mention, that in a further article on COVID-19 (Zhang, L., Zhao, WJ., Sun, BB., Huang, Y, Glänzel, W. (2020), How scientific research reacts to international public health emergencies: a global analysis of response patterns, Scientometrics, in press), the two authors of this editorial jointly study with their co-workers how scientific research reacts to international public health emergencies. The article, which is to come in one of the next issues of the journal Scientometrics, provides the bibliometric characteristics and a comparative analysis of global academic response patterns to global public-health emergencies in previous pandemics/epidemics of this century with that experienced in the present situation of COVID-19.
SCENARIO-DRIVEN FORECASTING: LESSONS LEARNED FROM MODELING THE COVID-19 PANDEMIC

Department of Management, Strategy & Innovation, Faculty of Economics & Business, KU Leuven, Belgium
2 Flanders Business School, Belgium
3 KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Clinical and Epidemiological Virology, Institute for the Future, Leuven, Belgium;
4 Center for Global Health and Tropical Medicine, Unidade de Microbiologia, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal

KRIStOF DECOCK 1,2
KOENRAAD DEBACKERE 1
ANNE-MIEKE VANDAMME 3,4
BART VAN LOOY 1,2

SUMMARY: The Forecasting patterns and their implied end states remains cumbersome when few (stochastic) data points are available during the early stage of diffusion processes. Extrapolations based on compounded growth rates do not account for inflection points nor end-states. In order to remedy this situation, we advance a set of heuristics that blend forecasting and scenario thinking. Combined they provide an actionable decision space as short term predictions are accurate, while a portfolio of different end states informs the long view.

The creation of such a decision space requires temporal distance; only to the extent that one refrains from incorporating more recent data, more plausible end states become visible. As such, our contribution implies a plea for dynamically blending forecasting algorithms and scenario oriented thinking, rather than conceiving them as substitutes or complements.

ACKNOWLEDGMENT: This contribution benefited from useful input and reflections by Jorge Ricardo Blanco Nova (KU Leuven), Michela Bergamini (KU Leuven), Sien Luyten (Flanders Business School) and Xiaoyan Song (KU Leuven).
INTRODUCTION: MODELING DIFFUSION PATTERNS OF COVID-19 PANDEMIC

The current COVID-19 pandemic spurred efforts to model and forecast its diffusion patterns, either in terms of infections, people in need of medical assistance (hospitalization, ICU occupation) or casualties.

Forecasting the evolution of COVID-19 and related dynamics (e.g. ICU occupation) remains cumbersome when only few stochastic data points are available during the early stage of the outbreak. Extrapolation based on compounded growth rates could result in unstable predictions as they do not include inflection points (i.e. the peak of the net ICU increase) nor end states (i.e. the maximum ICU capacity). Several forecasting models suggest that the implied growth patterns follow sigmoid growth curves, like the logistic equation (Verhulst, 1838) or the Bass (1969) model. Predicting the evolution of these S-shaped curves involves an estimation of at least three parameters, related to the takeoff and the steepness of the curve as well as its saturation level. However, delineating robust parameter estimations – and consequently portraying the end-state – based on a limited time series results in quite rudimentary predictions. Recently, Decock, Debackere & Van Looy (2020) adapted the initial Bass (1969) model in order to quantitatively model different diffusion scenarios for Electrical Vehicles. The heuristics advanced start from the premise of multi-finality (Buckley, 1967): “similar initial conditions may lead to dis-similar end-states”. The proposed heuristics revolve around the development of a three-dimensional search space, reflecting the presence of three model parameters to be estimated.

In the context of the COVID-19 pandemic, it remains unclear at this stage which end states will occur with respect to e.g. casualties, capacity of ICU beds needed at its highest burden (in different regions/countries) as the available numbers for a wide range of countries suggest high spreads (both in terms of contamination and potential end states). This complicates the quest for accurate predictions, unless one considers the potential presence of multiple end states and different growth dynamics (reflected in the three parameters to gauge) simultaneously.

As such, the COVID-19 diffusion patterns for Belgium have been modeled by means of the heuristics proposed by Decock, Debackere and Van Looy to see whether they allow to arrive at a better informed decision space, both in the short (daily/weekly) and in the medium term (end states). Models have been developed both for deceased and for ICU occupation rates. In this contribution, we focus on the ICU models, due to space constraints (the insights are similar across indicators).

The considered search grid consists of 250,000 different parameter combinations (consisting of 10 different end-states scenarios multiplied with 250 values for the relevant infection parameter and 100 values for the relevant contamination parameter). By means of a loss function, all 250,000 pathways are assessed in terms of how well they explain the already available observations. In this exercise, the diffusion paths with a goodness-of-fit exceeding 99 percent have been considered as more likely scenarios and have been further analyzed in terms of growth dynamics and end states.

ICU OCCUPATION IN BELGIUM

In Belgium, a first COVID-19 deceased was registered on March 10. With an initial overall capacity around 1,900 ICU beds in the Belgian hospitals, policy makers decided to increase ICU capacity and allocated approximately 2,300 ICU beds exclusively for COVID-19 patients. The question that then becomes crucial: will this be enough for the coming weeks and months?

Table 1 depicts the number of patients in ICU (related to COVID-19) in Belgium, between March 12 and March 24:
Building on the observations until March 24 (i.e. time series of 13 data points), a multi-dimensional search space was assessed implying a potentially required ICU-capacity ranging from 1,000 to 10,000 beds. Figure 1 visualizes the initial observations (i.e., the green dotted line) combined with the stylized more likely scenarios (averaged by end-state).

Table 1 COVID-19 ICU occupation in Belgium between 12/03/2020 and 24/03/2020 (Source: Sciensano)

<table>
<thead>
<tr>
<th>DATE</th>
<th>ICU OCCUPATION</th>
<th>DATE</th>
<th>ICU OCCUPATION</th>
<th>DATE</th>
<th>ICU OCCUPATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/03/20</td>
<td>5</td>
<td>17/03/20</td>
<td>100</td>
<td>22/03/20</td>
<td>322</td>
</tr>
<tr>
<td>13/03/20</td>
<td>24</td>
<td>18/03/20</td>
<td>130</td>
<td>23/03/20</td>
<td>381</td>
</tr>
<tr>
<td>14/03/20</td>
<td>33</td>
<td>19/03/20</td>
<td>164</td>
<td>24/03/20</td>
<td>474</td>
</tr>
<tr>
<td>15/03/20</td>
<td>53</td>
<td>20/03/20</td>
<td>238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16/03/20</td>
<td>79</td>
<td>21/03/20</td>
<td>290</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 COVID-19 ICU occupation in Belgium between 25/03/2020 and 04/04/2020 (Source: Sciensano)

<table>
<thead>
<tr>
<th>DATE</th>
<th>ICU OCCUPATION</th>
<th>DATE</th>
<th>ICU OCCUPATION</th>
<th>DATE</th>
<th>ICU OCCUPATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>25/03/20</td>
<td>605</td>
<td>29/03/20</td>
<td>927</td>
<td>02/04/20</td>
<td>1,205</td>
</tr>
<tr>
<td>26/03/20</td>
<td>690</td>
<td>30/03/20</td>
<td>1,021</td>
<td>03/04/20</td>
<td>1,245</td>
</tr>
<tr>
<td>27/03/20</td>
<td>789</td>
<td>31/03/20</td>
<td>1,088</td>
<td>04/04/20</td>
<td>1,261</td>
</tr>
<tr>
<td>28/03/20</td>
<td>867</td>
<td>01/04/20</td>
<td>1,144</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig 2: Comparing the daily net increase in ICU occupation (i.e., dotted line) with the “more likely” scenarios ($R^2 > .99$) in Belgium.
As the data of the 11 days between March 25 and April 4 became available (see table 2 and visualized as red dots in figure 1 below), it becomes feasible to further qualify the likelihood of different pathways and their corresponding end-states.

As from March 28 onwards, it became clear that the end state of ICU occupation in Belgium was more likely to become situated in a range between 1,000 and 3,000. Consecutive updates and refinements in terms of feasible end states allowed us to predict the growth pattern and the peaks even more precisely, as depicted in figure 2. On April 2nd we informed the Belgian authorities that the growth peak has been reached and that the coming days, the growth rate for ICU capacity would decline markedly. This prediction materialized as shown by the red dots in figure 2.

DISCUSSION AND CONCLUSION

In this contribution we illustrate how blending forecasting models with scenario-oriented thinking yields novel insights, which inform decision makers in a number of distinctive ways. The heuristics implied, combine a broad range of end states with an assessment of more likely pathways. Not only do these models provide accurate predictions in the short term; when additional observations become available, they also signal plausible end-states in the scenario portfolio. The creation of such a decision space requires temporal distance: only to the extent that one refrains from incorporating immediately more recent data, more plausible end states become visible. At the same time, updating and re-calibrating the pathways seem to offer potential to start qualifying Knightian uncertainty (Knight, 1921).

As such, our contribution implies a plea for combining forecasting algorithms with a foresight-oriented methodology and vice versa. This dynamic blending of forecast and foresight has the potential to inform policy makers in situations of urgent decision needs conditioned by profound uncertainty. At such critical moments during an unfolding crisis, the use of large amounts of data cannot inform decision-making, as those data are largely absent in such instances. Though, it is the judicious use of the limited data available combined with various scenarios that may unfold. Our combination of forecast and foresight hence does not signal yet another case in evidence-based policy, but it rather illustrates the good governance of evidence as advocated by Parkhurst (2016).

REFERENCES


CONFERENCE REPORT ON ISSI2019

CINZIA DARAIO
Sapienza University of Rome, Italy

The 17th International Conference of the International Society for Scientometrics and Informetrics (ISSI2019) took place at Sapienza University of Rome on 2-5 September 2019. For the first time, Italy hosted an ISSI Conference and the city of Rome, the “eternal city”, offered a beautiful frame to the Conference. As we wrote in the last email to all conference participants: “we waited for our conference for years (it was discussed and approved by the ISSI Board in 2015 and confirmed in 2017) but it passed like a flash, remaining in our mind as a dream...” Anyway, we have the gadgets of the conference to remind us that it happened...

The conference has been a major event attended by more than 650 participants from 44 countries from all the world. The conference included a special STI Indicators Conference Track organized in collaboration with the European Network of Indicator Designers (ENID).

The program of the four days was very dense, although we tried to organize the parallel, the plenary and the posters sessions, in such a way to balance the different themes at the benefit of a variegated audience.

The first day of the conference, on Monday 2 September 2019 we had several interesting workshops and sponsored sessions, ranging from Altmetrics and open citations, to impact assessment and application of Big Data in Scientometrics, including dedicated platforms (CADRE and NETSCITY), collective datasets (ETER), in-
vestigation on scholarly communication in China and the study of review articles. The program of the first conference day was completed by a rich Doctoral Forum, and a light welcome dinner preceded the Jazz concert performed by Musa Jazz, one of the Sapienza’s Orchestras, in the Aula Magna.

All posters were categorized using the same classification into themes as that used to group the papers. Next, they were aggregated into four main poster groups: i) Indicators & databases, ii) Social context, innovation, policy, iii) Analytical methods and iv) Application domains.

The poster session around the poster boards was preceded by a session in which all posters in a particular group were presented in short presentations of about one minute, in a conference room.

The Conference Proceedings of more than 2800 pages organized in two volumes contain all the contributions presented at the conference: around 260 papers and more than 150 posters.

All the topics presented during the plenary sessions, although addressed by scholars with different backgrounds and analysing different aspects, were around the multidimensionality and the complexity of the assessment in Scientometrics and Informetrics.

The first keynote was delivered by Prof. John Carson from the University of Michigan.

Table 1  All accepted papers were grouped into the following eighteen themes and were grouped according to their theme into eight parallel lines:

<table>
<thead>
<tr>
<th>Altmetrics</th>
<th>Higher Education Systems</th>
<th>Patent analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated Text Processing</td>
<td>Humanities and Social Sciences</td>
<td>Research evaluation models</td>
</tr>
<tr>
<td>Databases</td>
<td>Informetric distributions</td>
<td>Research field mapping</td>
</tr>
<tr>
<td>Funding and determinants</td>
<td>Library science aspects</td>
<td>Scholarly publishing</td>
</tr>
<tr>
<td>Gender studies</td>
<td>National R&amp;D systems</td>
<td>Social context of R&amp;D</td>
</tr>
<tr>
<td>Globalization &amp; collaboration</td>
<td>Novel bibliometric indicators</td>
<td>Technological innovation</td>
</tr>
</tbody>
</table>
(USA), entitled “Quantification, its affordances and limits”. John Carson is an expert in the history of quantification and large part of his research has been devoted to the quantification and measurement of intelligence. We recommend his very interesting, deep and inspiring book on “The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750-1940”
In his plenary talk, Carson pointed out some crucial aspects connected to the evaluation that every scholar in our field should consider in our opinion. He started his talk by considering quantification as “moral endeavor”.

Quantification can never stand alone, Carson said, it is as much human as it is mechanical, subjective as well as objective. “Why quantify? The world is messy and complex and full of noise; combining bits of information and analyzing them rigorously is often the only way to detect patterns and truths concealed in the flood of information with which we are inundated. Every act of seeing, however, is also an act of occlusion, of not-seeing. To attend to one part of the visual spectrum one must pay much less attention to other parts; to count one kind of thing, one cannot include other kinds.” Quantification is not just a technical pursuit, but a moral one; the reductions and simplifications it relies on can bring great and important clarity, but always at a cost. The moral questions for the practitioner are first, is the cost justified, and second, and even more critically, who will pay it? The powerful or the powerless? The established or those fighting for a place? After a description of the problems of archives in history that are almost the same as those that apply to data selection and treatment in our field, John Carson called for “epistemic modesty”, that is for scholars in all fields to be aware of the limits of their work, its partialness, and the particular standpoint from which they are coming. By applying his suggestion, we should then try to (1) take care to make the framing of problems to be explored expansive and inclusive; (2) pay particular attention to the most vulnerable; (3) insure that our work is distributed to a wide network of those who might be affected by it so that they have a chance to respond; and (4) consider whatever we come up with to be provisional and open to revision as we learn more. “None of these techniques can ensure that we don’t make mistakes, but they should at least remind us to think hard about the silences our archive will contain and to think carefully about who might bear the costs of whatever we have rendered invisible” he finally concluded.

The second keynote was given by Prof. Mike Thelwall, from the University of Wolverhampton (UK), very well known in our field and recipient of the prestigious Derek de Solla Price Award, with the title “New
developments in scientometric and informetric research”. In his talk, Thelwall discussed three recent developments that are changing and challenging scientometric practice: altmetrics, full text mining, and impact assessment. After decades of research into altmetrics and webometrics, alternative indicators have emerged as a standard part of scholarly communication infrastructure. This can be seen in the availability of Altmetric.com scores in many publisher websites and the informal use of altmetrics supporting research evaluation narratives. Altmetrics are well enough understood that we can now recommend appropriate uses, and are ready to fully exploit them. Of particular interest for the readers is Mike Thelwall’s nice book “Web Indicators for Research Evaluation: A Practical Guide”. This book describes how to analyse altmetric and webometric data for research evaluation purposes, including appropriate methods for field normalising it and reporting it for use in evaluations.

Full text mining is a second major development presented by Mike Thelwall, which is taking advantage of the increasing availability of collections of open access documents, such as the PubMed Central Open Access Subset, to get fine-grained information about citation contexts. This approach has the potential to identify important types of citation to get more precise evidence of the type of impact reflected by citation counts. The impact agenda in the UK Research Excellence Framework (REF) was the object of the last part of Thelwall’s talk and was presented as the requirement for researchers to produce evidence-based narratives describing how their research has non-academic societal benefits. This has led to changes within UK universities to promote societal impact, such as through the appointment of impact officers. Altmetric and Webometric indicators are used, Thelwall said, as part of the evidence base of some of these narratives, which provide a tough challenge for scientometricians. His talk ended with a few considerations in line with John Carson’s talk, with a few words about the importance of using research indicators responsibly. Practitioners should always be aware of potential biasing and systemic effects that can lead to unintended consequences.

The plenary sessions were enriched by the celebration of the Eugene Garfield Award, the celebration of the Derek de Solla Price Award and that of the Best ISSI Paper Award. In addition, within the plenary sessions, five invited lectures were presented. These lectures were selected among the submitted papers at the conference highly appreciated by the reviewers, and that in the view of the Program Chairs transcend the boundaries of their specialized topic, by putting this topic in a broader perspective, and/or by combining methods or findings from different specialized domains. The list of the invited lectures is reported in the followings:

- Invited lecture 1: Frederique Bone and Daniele Rotolo. Text-Mining Historical Sources to Trace Technical Change: The Case of Mass Production
- Invited lecture 2: Grit Laudel. Studying the embeddedness of researchers’ careers: Can bibliometric methods help?
- Invited lecture 3: Tao Han and Xiaoyu Cai. Analysis of Division of Labor in High Quality Life Science Research of China
- Invited lecture 4: Francois Van Schalkwyk, Jonathan Dudek and Rodrigo Costas. Using altmetrics to study social movements and cognitive bridges in the communication of science in the social media: The case of the anti-vaccination movement on Twitter
- Invited lecture 5: Matthias Held and Theresa Velden. How to interpret algorithmically constructed topical structures of research specialties? A case study comparing an internal and an external mapping of the topical structure of invasion biology
Palazzo Doria Pamphilj: welcome cocktail, mirrors hall and cloister just before the conference dinner, Wednesday 4 September 2019. Photo: © Maurizio Perciballi
The Conference Dinner was held at Palazzo Doria Pamphilj, at the center of the center of Rome, on Wednesday 4 September, 2019. Some of the photos reported here were taken during the Welcome Cocktail, in the cloister where the dinner was held, after the guided tours of the art galleries, including the extraordinary hall of mirrors. We take the opportunity here to thank very much all our Sponsors including Elsevier, Digital Science, Data for Impact, Clarivate Analytics, RISIS, Frontiers, Journal of

Special Plenary session “Memorial of Judit Bar-Illan” Thursday 5 September 2019. Photo: © Maurizio Perciballi
Data and Information Science and Quantitative Science Studies, for their generous contributions that allowed us to organize the conference and all its social events in a comfortable way.

The special plenary session organized by Gali Halevi on the 5 September 2019 and devoted to the Memorial of Judit Bar-Ilan, who passed away on July 16, 2019, was particularly touching.

Another special plenary session entitled “Evaluative informetrics– the art of metrics-based research assessment”, including the conferral of a Doctorate Honoris Causa in Research Assessment Methodologies awarded by the Rector of Sapienza University of Rome to Prof. Henk F. Moed took place in the second part of the plenary session of the 5 September 2019. During this session, Prof. Wolfgang Glänzel introduced the content of the forthcoming book by Springer Nature on “Evaluative informetrics: the art of metrics-based research assessment – Fest-schrift in honour of Henk F. Moed” edited by Cinzia Daraio and Wolfgang Glänzel and including several original contributions by well-known scholars.

In the last part of this special session, Henk Moed delivered his Lectio Magistralis on “The Application Context of Research Assessment Methodologies”, just after the presentation of the paper by Yong Zhao, Jiayan Han, Jian Du and Yishan Wu on: “Origin and Impact: A Study on the Intellectual Transfer of Professor Henk Moed”. In his Lectio Magistralis, Moed discussed a complex and controversial topic that is the use of bibliometric or informetric indicators in the assessment of research performance. His aims were to spread light on the application context of these indicators. He proposed an application model for the use of indicators in the assessment of academic research, highlighting a series of analytical distinctions and useful building blocks, to contribute in this way to enlightening and further developing the search for best practices in research assessment. His proposals was organized in three sections. After an introduction on the use of indicators in research assessment, based on his recent monograph “Applied Evaluative Informetrics” published by the author in 2017 that we highly recommend to the readers, he proposed a series of useful analytical distinctions. These analytical distinctions were useful to draw conclusions on the validity and usefulness of the most
common indicators of the performance of three principal units of assessment: individual researchers, research institutions and scientific-scholarly journals. Finally, an application model for the use of indicators in academic research was presented, and an extended version can be found by the readers in his new journal Scholarly Assessment Reports (Moed, H.F., 2020. Appropriate Use of Metrics in Research Assessment of Autonomous Academic Institutions. Scholarly Assessment Reports, 2(1), p.i. DOI: http://doi.org/10.29024/sar.8) open now for submissions at https://www.scholarlyassessmentreports.org/

Finally, we like to end this conference report with a photo and the lightness and the view towards the blue sky of a song. The photo, reported below, was taken at the end of the closing session of the conference, and the song is “Volare (to fly)” by Domenico Modugno, a classical Italian song played during the concert of the Musa Jazz of Sapienza Orchestra the first day of the conference.
“Volare”

by Domenico Modugno

Penso che un sogno così non ritorni mai più
mi dipingevano le mani e la faccia di blu
poi d'improvviso venivo dal vento rapito
e incominciai a volare nel cielo infinito

Volare oh, oh,
cantare oh, oh,
nel blu dipinto di blu,
felice di stare lassù
e volavo volavo felice più in alto del sole
ed ancora più su
mentre il mondo pian piano spariva lontano lassù
una musica dolce suonava soltanto per me.

Volare oh, oh,
cantare oh, oh,
nel blu degli occhi tue blu,
felice di stare quaggiù
e continuai a volare felice più in alto del sole
ed ancora più su
mentre il mondo pian piano si cancellava lentamente
la tua voce è una musica dolce che suona solo per me.

I think that dream like that will never return.
I painted my hands and face blue.
Then all of sudden I was abducted by the wind.
And I started to fly in the infinite sky.

To fly ho ho
To sing ho ho ho ho,
In the blue, painted blue.
Happy to stay up there.
And I flew and flew happy, higher than the sun.
And yet more up.
While the world slowly disappeared down below.
A sweet music playing just for me.

To fly ho ho
To sing ho ho ho ho,
In the blue, painted blue.
Happy to stay up there.
But all dreams vanish at dawn because.
When the moon sets, it takes dreams with itself.
But I keep dreaming in your beautiful eyes.
That are blue as the sky spangled with stars.

To fly, ho ho
To sing ho ho ho ho,
In the blue of your blue eyes.
Happy to be down here.
And I keep flying, happy happy, higher than the sun.
And yet more up.
While the world is slowly vanishing in your blue eyes.
Your voice is a soft music that plays for me.

To fly ho ho
To sing ho ho ho ho,
In the blue of your blue eyes.
Happy to be down here.
In the blue of your blue eyes.
Happy to be down here.
With you.

(English translation: https://lyricstranslate.com)
INTERDISCIPLINARITY OF AN ARTICLE AS MEASURED THROUGH REFERENCES: IS SCALE-INVARiance REALLY NEEDED?

RONALD ROUSSEAU
University of Antwerp, Faculty of Social Sciences, Antwerpen, Belgium
ronald.rousseau@uantwerpen.be
Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Leuven, Belgium
ronald.rousseau@kuleuven.be

ABSTRACT: I argue that scale-invariance may not be a necessary condition for measuring interdisciplinarity through the diversity of references.

Keywords: scale-invariance; Shorrocks’ generalized Lorenz curve; references; interdisciplinarity

1. INTRODUCTION: INTERDISCIPLINARITY

When considering the term interdisciplinarity I first like to point out that defining a concept and measuring it are two different things. Moreover, the notion of interdisciplinarity includes another one, namely that of a discipline. I think it is safe to say that the notion of a discipline cannot be formulated in a way that covers all possible situations and applications. In practice, the notion of a discipline or knowledge domain is operationalized in three broad ways (Zitt et al., 2019): using ready-made classifications, using an algorithmic approach and determining the knowledge domain of an article through the domains to which its authors belong. In this note, I will not go deeper into this aspect and assume that an appropriate choice has been made.

Next, I focus on one way to determine the interdisciplinarity of an article, namely through the interdisciplinarity of its refer-
ences, or of the set of references expanded with the set of references of references, i.e. using one or two backward generations (Rafols & Meyer, 2010; Hu et al., 2011). Further on in this note I will use the term references when referring to one or more generations of references. For the next step, there are again many alternatives, but I focus on one often applied method, namely applying a diversity measure on the set of disciplines covered/used in these references.

This note is further subdivided as follows: first I give a short introduction about how interdisciplinarity can be measured. Next, I come to the main part of this note, namely introducing the question if scale-invariance is really a requirement when studying interdisciplinarity through the diversity of used references. One suggestion is offered to measure diversity in a way that is not scale-invariant. Finally, a short list is offered of problems that must be solved if one wants to continue in the direction of non-scale-invariance for studying interdisciplinarity.

2. MEASURING DIVERSITY

Traditionally diversity consists of two concepts: variety and balance. Variety refers to the number of cells, i.e., the number of nominal classes such as species in biology and disciplines in informetrics studies, while evenness or balance is defined as the relative apportionment of abundances among cells, actually present, or assumed to be possibly present (Rousseau & Van Hecke, 1999). It has been shown that evenness on its own is best represented by a classical Lorenz curve (Lorenz, 1905; Nijssen et al., 1998). Yet, it has been convincingly argued that, besides variety and evenness, a third notion must be taken into account when measuring diversity, namely the disparity (or its opposite: proximity) between cells. Stirling (2007) points out that species in ecology or WoS categories in informetrics are not independent entities, but they are shaped by patterns of common developments or ancestry leading to proximities between units, see also (Leinster & Cobbold, 2012; Zhang et al., 2016). Finally, Jost (2006, 2009) pointed out that one should use measures that make it possible to discuss percentage changes in a meaningful way, see also (Leinster & Cobbold, 2012; Zhang et al., 2016). For a review on interdisciplinarity I refer to (Wagner et al., 2011; Rousseau et al., 2019).

Next, I introduce some notation. Let $X = [x_1, \ldots, x_N]$ be an array of $N$ sources or cells (N a natural number larger than 1), ranked according to the number of items (a non-negative value) in each cell. In this note, I will refer to such arrays as diversity arrays. Let $f$ be a diversity measure and $f(X)$ the value of $f$ for the array $X$. This measure $f$ is said to be scale-invariant if, for each $c > 0$, $f(cX) = f(X)$. In most cases being scale-invariant is considered a necessary property for diversity (and concentration) measures. The idea behind this requirement is that diversity or concentration should not depend on monetary units (dollars, euros, Chinese yuan) in the case of income studies, or weight units (kg, g, pound) in case of biomass. All measures based on the Lorenz curve, such as the Gini index or the coefficient of variation, are by definition scale-invariant as the Lorenz curve is constructed using relative values (see further), not absolute ones.

3. IS SCALE-INVARIANCE REALLY NEEDED?

I return now to the specific case I want to study, namely measuring interdisciplinarity of an article through the diversity of its references. If article A1 is based on four disciplines occurring 10, 6, 2 and 1 times, while article A2 is also based on the same four disciplines occurring 20, 12, 4 and 2 times, which one reflects the most interdisciplinarity? In this hypothetical case, I assume that there are no spurious references and no redundancies: each reference is ‘really’ needed. Then classical approaches consider these two articles to be equally interdisciplinary.

Yet, because each article, theoretically, makes new and different knowledge claims, it is clear that the second article used the
knowledge embedded in these four disciplines more thoroughly than article A1. For this reason, one may say that A2 bears testimony of more interdisciplinarity than A1. Consequently, I claim that depending on the purpose of the investigation, scale-invariance is not a necessary property.

4. HOW TO REPRESENT “NON-SCALE-INVARIENCE”? 

How can the above observation be taken into account? Already in 1983 Shorrocks when studying income inequality, observed that, if for a fixed number of cells, the partial order introduced by the Lorenz curve, which is actually the dominance order as studied e.g. in (Hardy et al., 1934), is often inconclusive among two situations, as Lorenz curves tend to cross and for this reason, are intrinsically incomparable (Shorrocks, 1983). For this reason, he proposed the generalized Lorenz curve, constructed by scaling up the ordinary Lorenz curve by the mean of the distribution. This construction reduced the number of crossing curves.

I next recall the construction of the ordinary Lorenz curve (Lorenz, 1905; Rousseau et al., 2018) and then build on this to describe Shorrocks’ generalized Lorenz curve. Given the diversity array \( X = [x_1, \ldots, x_N] \), the ordinary Lorenz curve (Lorenz, 1905) is constructed as follows: for \( j = 1, \ldots, N \), let 

\[
S_j = \sum_{i=1}^j x_i
\]

be the \( j \)-th partial sum and hence \( S_N = \text{TOT} \) the total number of items. Moreover, \( S_0 \) is set equal to 0. Now, take the points with coordinates 

\[
\left( \frac{k}{N}, \frac{S_k}{N} \right)
\]

\( k = 0, \ldots, N \) are connected by straight line segments.

Just as the Lorenz curves introduce a partial order in the set of diversity arrays, also Shorrocks’ generalized Lorenz curves introduce a – finer – partial order. Acceptable diversity measures must then respect this partial order.

5. CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

In this note, I just want to put forward the idea that scale-invariance may not be a necessary requirement when using diversity measures to study interdisciplinarity via references. This leads to at least three questions to be studied in further investigations:

1. Is the Shorrocks curve an acceptable way to study interdisciplinarity in a way that is not scale-invariant? Or should it be adapted?

2. How to include the other diversity requirements into this study? A first step could be to take disparity “seriously” as proposed in (Rousseau, 2018), while adapting the Leydesdorff-Wagner-Bornmann approach (Leydesdorff et al., 2019) or intrinsic diversity profiles (Patil & Taillie, 1979; Rousseau et al., 1999) are other alternatives.

I hope to be able to answer these questions in further investigations.
ACKNOWLEDGEMENT

I thank Lin Zhang (Wuhan University and KU Leuven) for helpful discussions.

REFERENCES


A MULTILAYER NETWORK ANALYSIS OF MOBILITY AND COLLABORATION IN EUROPEAN UNIVERSITY RESEARCH

PEI-SHAN CHI

ZSOLT KOSZTYÁN

LÁSZLO GADÁR

VIVIAN CSÁNYI

BART THIJS

ANDRÁS TELCS

WOLFGANG GLÄNZEL

ABSTRACT: The present report gives an overview of a recent joint Hungarian-Flemish project funded by the Research Foundation - Flanders (FWO). The project aimed to study European university research from the viewpoint of scientists’ mobility and collaboration patterns of the most important universities of selected countries.
INTRODUCTION

A two-year joint research project between the ECOOM team at KU Leuven and the Budapest Ranking Research Group at Hungarian Academy of Sciences were conducted during 2018 and 2019. This project aims to provide a high level of analytic and visualization system to explore the scientific contribution of European universities. Based on its capacities the parties studied and developed fine-tuned, targeted university multilayer networks. In short, the partners created a database which is basically a multi-labelled graph in which vertices are institutions and edges are citations. Institutions carry some feature describing labels while edges carry labels describing the citation relation in details (publication time, citation time, subject classification etc.). Unlike other endeavours, this project did not aim to conduct any ranking exercise, which provided us the liberty to choose appropriate data sources and to select European countries and universities for a more complex multilayer exercise. In particular, we have selected the following data sources to highlight on essential aspects in researchers’ mobility and collaboration.

The Erasmus activity network served as the main source of tracking mobility and analysing European networks from this perspective. The student and teacher mobility patterns based on 2013/2014 Erasmus data are somewhat different (see Figure 1; top: students, bottom: teachers). This is to a certain extent due to the different motivations for and durations of their stays abroad. Generally Spanish universities attract more students and academic teachers compared to their outgoing ones. By contrast, Turkish and Polish universities have a disproportionate number of outgoing teachers as compared with the incoming ones, whereas German and Italian universities have much more outgoing students than incoming ones. Figure 1 also illustrates the active participation of Spanish universities in the Erasmus project.

In order to gain a deeper insight into the Erasmus student mobility, activities in three different subject areas were studied (cf. Figure 2). We have selected applied fields in the sciences along with the humanities based on Erasmus project classification. In particular, in Social science,
Figure 1. Scatter plots of 2013/2014 Erasmus incoming vs. outgoing staff. Top: students; bottom: academic teachers.
Business and Law, Spanish universities attract more students from other universities than their own students who leave for staying at other institutions abroad. In Science, Mathematics and Computing, Nordic and some Belgian and Dutch universities show their advantage to attract more students from other universities than sending out their own students. In Health and Welfare, a few universities from Italy, Czech Republic, Sweden, Denmark and Hungary show their receiving role for unproportionally more incoming than outgoing students. The observed subject-related difference revealed in Figure 2 points us a new direction to dig in from the mobility data. We further looked at the spatial multiplex network of the student mobility based on the determined GPS coordinates in the three main subject areas.

Figure 3 – inspired by Gadár et al. (2019) – linked the data from the Erasmus project, the European Tertiary Education Register (ETER) and the Global Research Identifier Database (GRID) to convert a spatial network on geographical map. It shows different connections in the three subject areas. In Social science, Business and Law, the connections between the UK and European countries are strikingly strong. Universities in Spain, France, Germany, Belgium and Italy are also important contributors in the Erasmus student mobility network. In the field Science, Mathematics and Computing, Nordic universities play a more important role in the network. Finally, in Health and Welfare, the connections are more diversified.

Figure 2. Scatter plots of 2013/2014 Erasmus incoming VS. outgoing students in different fields. Top: Social science, Business and Law; Middle: Science, Mathematics and Computing; Bottom: Health and Welfare.
Welfare sciences, The Netherlands is sending a large number of students to Denmark. Italy, Spain and Czech Republic are central nodes of student mobility in this field.

A further perspective to focus on the countries we are especially interested in, was taken and presented in Figure 4. This new dimension was added on the basis of scholarly collaboration measured by co-authorship of scientific publications indexed in Clarivate Analytics Web of Science Core Collection (WoS). Although the time frames and actors are different in these dimensions, i.e., we have not studied the publication patterns of the Erasmus students but the complete research staff affiliated with the selected universities, at this level of aggregation this approach may add a new dimension not directly linked to the network created by previous actors. To zoom into targeted countries, the student and teacher mobility and scholarly publication collaboration networks were limited into the international connections among Belgium, Germany and Hungary. Regarding the student mobility, universities in Berlin are highly connected with Flemish universities and universities in Budapest. Universities in Cologne area and Munich area received and sent out many students from or to universities in Budapest and Flanders. The teacher mobility was more equally distributed among universities. The connections from Hungary are not as strong as those among Belgium and Germany. In terms of WoS publications in 2013, Belgian universities strongly collaborated with the other two countries, especially with German universities.

DISCUSSION AND FUTURE RESEARCH

The joint project has offered a new perspective. The usual macro and meso bibliometric study of research collaboration and the socio-bibliometric analysis of students’ and researchers’ mobility could be

Figure 3. Mobility networks with different subject dimensions in the 2013-2014 academic year similarly to Gadár et al. (2019). The size and colour intensity of the nodes are proportional to the number of incoming students. The size and colour intensity of the edges are proportional to the number of students that travel between the connected HEIs. Top: Social science, Business and Law; Middle: Science, Mathematics and Computing; Bottom: Health and Welfare
combined into a first multi-layer framework, even if only a small number of institutions could be taken into account. Overlapping network layers indicate how correlated (or not) mobility and research collaboration systems are with each other. Based on our preliminary results, there does not appear to be a significant overlap between mobility and research collaborations, due to the different motivations and driving forces of different collaboration.

In the present project we could much benefit from our local knowledge and experience, which was all the more important because we had to clean a large amount of data on the bibliometric part for the integration into the framework. The main objective and achievement of the project was therefore rather the creation of the framework and the elaboration of the possible components of this multi-layer network. Figure 4 has shown that different patterns can be attributed to individual components. Beyond the programmes enabling student and staff mobility, there are different individual interests influencing and shaping the patterns of student mobility particular making destinations attractive on the long-term (education, career) compared to rather short-term teacher mobil-

Figure 4. Mobility and scholarly collaboration international networks between Belgium, Germany and Hungary. Top: students; Middle: academic teachers; Bottom: WoS publications
ity. Collaboration patterns are, by contrast, influenced to a lesser extent by individual interests as those are part of larger, mostly subject-specific, European and even global collaboration frameworks. These patterns can be recognised in Figure 4. The sparser mobility networks top and notably in the centre is contrasted by the denser collaboration networks that are often based multilateral projects. This also implies the involvement of more universities and the higher density of networks. At the future works, subnetworks of subjects can be compared. In addition, further layers can be integrated to the multilayer structure such as co-authorship network of publications and the patents. In this way the mobility and collaboration activities of HEIs can be deeper analysed.

A further recognition in the project is that we found different network patterns in subject area layers, which is mostly related to the characteristics of the institutions. In case of real sciences, the attractiveness of Northern European universities for student mobility is apparent, while in other subject areas Southern European institutions receive the most students. The characteristics of institutions depend on the representation of disciplines moreover the educational and/or research and/or innovation focus of the institution. Different characteristics result different collaboration systems resulting different network patterns. Future research directions could be the examination of correlations between the strength of institutions, the position of higher education rankings and the characteristics of the emerged collaboration networks.

ACKNOWLEDGMENTS

The project was partially funded by “Analysis of Scientific Performance of Universities” – bilateral mobility project of HAS – FWO 2018-19. A bilateral collaboration programme of the Research Foundation – Flanders (FWO).

REFERENCES


