Bibliometric Research Evaluation Study of the past performance of the Universidade Nova de Lisboa, Portugal (2006–2012/13)

Erik van Wijk

Center for Science and Technology Studies (CWTS)
Leiden University
PO Box 9555, 2300 RB Leiden
The Netherlands

December 09, 2014
Draft version

Contents

1 Introduction

The objective of the present study is to provide insight in important aspects of publication output and international citation impact of the Universidade Nova de Lisboa (UNL), Lisboa, Portugal. This report is an update of a previous study covering 2004 - 2010 (Costas-Comesaña, van Wijk, 2012).

In this report we present a 'Past Performance' analysis of the production of UNL as part of an every two years series of evaluations. The covered period of analysis is 2006 - 2012 for publications, with an extra year added for their citation period, so as to arrive at robust impact scores. The impact, as measured by citations, is compared to worldwide reference values. The study is based on a quantitative analysis of publications covered in journals and serials processed for the Web of Science (WoS) versions of the Science Citation Index and associated citation indices: the **Science Citation Index** (SCI), the **Social Science Citation Index** (SSCI), and the **Arts & Humanities Citation Index** (A&HCI); here the CWTS database containing these records as well as enhanced citation data is briefly indicated as **CI**.

Bibliometric Indicators

In this section, we discuss the methods underlying the bibliometric analyses presented in this report[1].

1.1 Database Structure

At CWTS, we calculate our indicators based on our in-house version of the Web of Science (WoS) database of Thomson Reuters. WoS is a bibliographic database that covers the publications of about 12,000 journals in the sciences, the social sciences, and the arts and humanities. Each journal in WoS is assigned to one or more subject categories. We note that our in-house version of the WoS database includes a number of improvements over the original WoS database. Most importantly, our database uses a more advanced citation matching algorithm and an extensive system for address unification. Our database also supports a hierarchically organized field classification system on top of the WoS subject categories.

To determine the appropriateness of our indicators for assessing a particular research entity, we often look at the internal WoS coverage of the entity. The internal WoS coverage of an entity is defined as the proportion of the references in its output that points to publications (also) covered by WoS (see section 3). The lower the internal WoS coverage of an entity's output, the more careful one should be in the interpretation of our indicators. The rest of this chapter provides an in-depth discussion of the main bibliometric indicators that we use in this report.

[1] We refer to Moed (2005) for a general introduction into the use of bibliometrics and citation analysis for research evaluation.

Overview of the bibliometric indicators.

Indicator	*Dimension*	*Definition*
P	Output	Total number of publications of a unit.
MCS	Impact	Average number of citations of the publications of a unit (self-citations not included).
TCS	Overall	Total number of citations.
MNCS	Impact	Average normalized number of citations of the publications of a unit (self-citations not included).
TNCS	Overall	Total average normalized number of citations.
MNJS	Journal Impact	Average normalized citation score of the journals in which a research group has published.
pp top 10%	Impact	Proportion of papers that belong to the top 10%.
pp uncited	Overall	Proportion of papers uncited.
Prop self cits	Overall	Proportion of self-citations.
pp collab	Collaboration	Percentage inter-institutional collaborative publications.
pp int collab	Collaboration	Percentage international collaborative publications.
Int_cov	Output	Internal coverage. Proxy of oeuvre being covered by Web of Science. Measured by the proportion of cited references in the oeuvre linking to other WoS publications.

1.2 Indicators of Output

To measure the total publication output of a unit, we use a very simple indicator. This is the number of publications indicator, denoted by P. This indicator is calculated by counting the total number of publications of a research unit. Only publications of the document types *article, letter* and *review* are taken into account. Other document types have not been considered for this study as they only make a smaller scientific contribution as compared to articles or reviews. Because of the sometimes erratic citation behavior of the document-type 'letter' these are counted for 25% both on the output side as on the impact side.

1.3 Coverage of CI Publications

To gain insight in the CI coverage of the publications included in the study, we studied the references of the papers included in the present study. To this end, references in UNL CI-covered publications (2006-2012) were matched with our extended CI publication database (1980 - 2013). In this way, we can estimate the importance of CI publications to the researchers participating in the study by determining to what extent they themselves cite CI Web of Science papers, and to what extent other, non-CI documents. The overall results (**table 0**) show that: overall internal coverage is quite sufficient to good. There are however exceptions:

Research Unit	Department	Internal Coverage
Faculty of Social Sciences and Humanities(FCSH)		30%
Faculty of Sciences and Technology(FCT)	Dept of Applied Social Sciences(DCSA)	35%
Faculty of Sciences and Technology(FCT)	Dept of Computer Sciences(DI)	43%
Faculty of Sciences and Technology(FCT)	Dept of Civil Engineering(DEC)	51%
Faculty of Sciences and Technology(FCT)	Dept of Mechanical & Industrial Engineering(DEMI)	52%
Faculty of Sciences and Technology(FCT)	Dept of Electronic Engineering(DEE)	54%
Faculty of Sciences and Technology(FCT)	Dept of Earth Sciences(DCT)	54%
Inst of Statistics & Information Management(ISEGI)		55%

Whenever the internal coverage drops beneath the 50% the outcome of the bibliometric analysis must be viewed with caution as the importance of Non-WoS material is then as high as or even higher than that of WoS material. Although the Faculty of Social Science and Humanities shows a low internal coverage, it is also fair to draw attention to the fact that it has been on the rise in this respect compared to previous reports. It had risen in the previous report from 10% to some 20% and again rises in this analysis to 30%, this against the backdrop of an increase in publications by a factor 3. Most of the entities that do not reach a sufficient internal coverage rate are departments within the faculty of science and technology, which in and by itself as a whole reaches 78%. The largest contributor to this faculty, the department of chemistry has a very high internal coverage of nearly 90% (table 0).

Table 0. Number of CI-covered UNL publications 2006 - 2012, by department.

Institute	Department	p	Internal Coverage
Faculty of Law(FD)		1.00	0%
Faculty of Medical Sciences(FCM)		591.75	89%
Faculty of Sciences and Technology(FCT)	Dept of Applied Social Sciences(DCSA)	26.00	35%
Faculty of Sciences and Technology(FCT)	Dept of Chemistry(DQ)	1110.50	89%
Faculty of Sciences and Technology(FCT)	Dept of Civil Engineering(DEC)	71.00	51%
Faculty of Sciences and Technology(FCT)	Dept of Computer Sciences(DI)	106.00	43%
Faculty of Sciences and Technology(FCT)	Dept of Conservation & Restoration(DCR)	85.00	67%
Faculty of Sciences and Technology(FCT)	Dept of Earth Sciences(DCT)	56.00	54%
Faculty of Sciences and Technology(FCT)	Dept of Electronic Engineering(DEE)	236.00	54%
Faculty of Sciences and Technology(FCT)	Dept of Environmental Sciences & Engineering(DCEA)	164.00	66%
Faculty of Sciences and Technology(FCT)	Dept of Life Sciences(DCV)	165.00	87%
Faculty of Sciences and Technology(FCT)	Dept of Material Sciences(DCM)	363.00	85%
Faculty of Sciences and Technology(FCT)	Dept of Mathematics(DM)	227.25	63%
Faculty of Sciences and Technology(FCT)	Dept of Mechanical & Industrial Engineering(DEMI)	97.00	52%
Faculty of Sciences and Technology(FCT)	Dept of Physics(DF)	270.50	84%
Faculty of Sciences and Technology(FCT)	Dept of Sciences & Technology of Biomass(DCTB)	53.00	77%
Faculty of Social Sciences and Humanities(FCSH)		160.50	30%
Inst of Chemical & Biological Technology(ITQB/IBET)		1505.75	91%
Inst of Hygiene & Tropical Medicine(IHMT)		397.25	84%
Inst of Statistics & Information Management(ISEGI)		104.25	55%
National School of Public Health(ENSP)		48.00	71%
Nova School of Business and Economics(NovaSBE)		228.00	61%
UNL		5584.25	81%

1.4 Indicators of Impact

A number of indicators are available for measuring the average scientific impact of the publications of a unit. These indicators are all based on the idea of counting the number of times the publications of a unit have been cited. Citations can be counted using either a fixed-length citation window or a variable-length citation window. In the case of a fixed-length citation window, only citations received within a fixed time period (e.g. for example four years) after the appearance of a publication are counted. In the case of a variable-length citation window, all citations received by a publication up to a fixed point in time are counted, which means that older publications have a longer citation window than more recent publications. An advantage of a variable-length window over a fixed-length window is that a variable-length window usually yields higher citation counts, which may be expected to lead to more reliable impact measurements.

In this study, we have used a variable length window, so all citations up to 2013 received by all publications are counted, for the whole period of analysis 2006-2012. In the calculation of our

impact indicators, we disregard self-citations. We classify a citation as a self-citation if the citing publication and the cited publication have at least one author name (i.e., last name and initials) in common. Self-citations are disregarded because they are of a somewhat different nature than other citations. Many self-citations are given for good reasons in particular to indicate how different publications of a researcher build on each other. Particularly in narrow or emerging fields where the number of active scholars is low and the number of collaborations is high, the relevance of self-citations should be considered more carefully. However self-citations can also serve as a mechanism for self-promotion rather than as a mechanism for indicating relevant related work. This is why we consider it preferable to exclude self-citations from the calculation of our impact indicators. By disregarding self-citations, the sensitivity of our impact indicators to manipulation is reduced. Disregarding self-citations means that our impact indicators focus on measuring the impact of the work of a researcher on other members of the scientific community. The impact of the work of a researcher on his own work is ignored[2].

As we have mentioned previously each journal in WoS is assigned to one or more subject categories. These subject categories can be interpreted as scientific fields. There are about 250 subject categories in WoS. Publications in multidisciplinary journals such as *Nature*, *Proceedings of the National Academy of Sciences* and *Science* were individually allocated, if it was possible, to subject fields on the basis of their references. The reassignment was done proportionally to the number of references pointing to a subject category. It is important to highlight that the impact indicators are calculated based on this assignment. Each publication in WoS has a document type. The most frequently occurring document types are article, book review, correction, editorial material, letter, meeting abstract, news item, and review. In the calculation of bibliometric indicators, we only take into account publications of document types: article, letter and review. Publications of other document types usually do not make a significant scientific contribution.

Our most straight forward impact indicator is the mean citation score indicator, denoted by MCS. This indicator equals the average number of citations per publication. Only citations within the relevant citation window are counted, and self-citations are excluded. Also, only citations to publications of the document types: article, letter and review are taken into account.

[2] The main reason to exclude self-citations is because they somehow capture a different idea of 'scientific impact' than it is intended to be measured through the citations given by authors different from those of the original publication.

A major shortcoming of the MCS indicator is that it cannot be used to make comparisons between scientific fields. This is because different fields have very different citation characteristics. It makes no sense to make comparisons between fields using the MCS indicator. Furthermore, when a variable-length citation window is used, the MCS indicator also cannot be used to make comparisons between publications of different ages. In the case of a variable-length citation window, the MCS indicator favours older publications over more recent ones because older publications tend to have higher citation counts.

Our mean normalized citation score indicator, denoted by MNCS, provides a more sophisticated alternative to the MCS indicator. The MNCS indicator is similar to the MCS indicator except that it performs a normalization that aims to correct for differences in citation characteristics between publications from different scientific fields and different ages (in the case of a variable-length citation window)[3]. To calculate the MNCS indicator for a unit, we first calculate the normalized citation score of each publication of the unit.

The normalized citation score of a publication equals the ratio of the actual and the expected number of citations of the publication, where the expected number of citations is defined as the average number of citations of all publications in WoS that belong to the same field and that have the same publication year. The field (or the fields) to which a publication belongs is determined by the WoS subject categories of the journal in which the publication has appeared. The MNCS indicator is obtained by averaging the normalized citation scores of all publications of a unit. If a unit has an MNCS indicator of one, this means that on average the actual number of citations of the publications of the unit equals the expected number of citations. In other words, on average the publications of the unit have been cited equally frequently as publications that are similar in terms of field and publication year. An MNCS indicator of, for instance, two means that on average the publications of a unit have been cited twice as frequently as would be expected based on their field, publication year, and document type. We refer to Appendix I and to Waltman, Van Eck, Van Leeuwen, Visser, and Van Raan (2011a and b) for more details on the MNCS indicator.

[3] We do not normalize for document types *article* and *review*. One of the main criteria used by WoS to distinguish between these two document types is the number of references of a publication. In general, a publication with fewer than 100 references is classified as *article* while a publication with at least 100 references is classified as *review*. It is clear that this criterion does not yield a very accurate distinction between ordinary articles and review articles; therefore we have opted for considering them as the same document type.

In addition to the MNCS indicator, we have another important impact indicator. This is the *proportion top 10% publications indicator*, denoted by $PP_{top\ 10\%}$. For each publication of a research group, this indicator determines whether based on its number of citations the publication belongs to the top 10% of all WoS publications in the same field (i.e., the same WoS subject category) and the same publication year. The $PP_{top\ 10\%}$ indicator equals the proportion of the publications of a research group that belong to the top 10%. If a research group has a $PP_{top\ 10\%}$ indicator of 10%, this means that the actual number of top 10% publications of the group equals the expected number. A $PP_{top\ 10\%}$ indicator of, for instance, 20% means that a group has twice as many top 10% publications as expected. Of course, the choice to focus on top 10% publications is somewhat arbitrary. Instead of the $PP_{top\ 10\%}$ indicator, we can also calculate for instance a $PP_{top\ 1\%}$, $PP_{top\ 5\%}$, or $PP_{top\ 20\%}$ indicator. In this study, however, we focus on the $PP_{top\ 10\%}$ indicator. On the one hand this indicator has a clear focus on high impact publications, while on the other hand the indicator is more stable than for instance the $PP_{top\ 1\%}$ indicator. However we will also touch upon more percentage indicators highly cited publications.

To assess the impact of the publications of a unit, our general recommendation is to rely on a combination of the MNCS indicator and the $PP_{top\ 10\%}$ indicator. The MCS indicator does not correct for field differences and should therefore be used only for comparisons of groups that are active in the same field. An important weakness of the MNCS indicator is its strong sensitivity to publications with a very large number of citations. If a unit has one very highly cited publication, this is usually sufficient for a high score on the MNCS indicator, even if the other publications of the group have received only a small number of citations. Because of this, the MNCS indicator may sometimes seem to significantly overestimate the actual scientific impact of the publications of a unit. The $PP_{top\ 10\%}$ indicator is much less sensitive to publications with a very large number of citations, and it therefore does not suffer from the same problem as the MNCS indicator. A disadvantage of the $PP_{top\ 10\%}$ indicator is the artificial dichotomy it creates between publications that belong to the top 10% and publications that do not belong to the top 10%. A publication whose number of citations is just below the top 10% threshold does not contribute to the $PP_{top\ 10\%}$ indicator, while a publication with one or two additional citations does contribute to the indicator. Because the MNCS indicator and the $PP_{top\ 10\%}$ indicator have more or less opposite strengths and weaknesses, the indicators are strongly complementary to each other. This is why we recommend taking into account both indicators when assessing the impact of a unit's publications.

It is important to emphasize that the correction for field differences that is performed by the MNCS and $PP_{top\ 10\%}$ indicators is only a partial correction. As already mentioned, the field

definitions on which these indicators are based on the WoS subject categories. It is clear that, unlike these subject categories, fields in reality do not have well-defined boundaries. The boundaries of fields tend to be fuzzy, fields may be partly overlapping, and fields may consist of multiple subfields that each have their own characteristics. From the point of view of citation analysis, the most important shortcoming of the WoS subject categories seems to be their heterogeneity in terms of citation characteristics. Many subject categories consist of research areas that differ substantially in their density of citations. For instance, within a single subject category, the average number of citations per publication may be 50% larger in one research area than in another. The MNCS and $PP_{top\ 10\%}$ indicators do not correct for this within-subject-category heterogeneity. This can be a problem especially when using these indicators at lower levels of aggregation, for instance at the level of departments or individuals.

2 Data Collection

2.1 Data selection

In general, bibliometric past performance analyses can be conducted in two ways. One approach is to collect the publications produced by a research unit in the past on the basis of address affiliation. The other approach is to collect data on the basis of publications provided by the institute itself. The 'Past performance' data collection in this case was on the basis of publications that the institute itself provided CWTS with.

The analysis focuses on publications from the period 2006–2012. Only WoS indexed publications are considered. This means that books, book chapters, journal publications not indexed in WoS (as well as conference proceedings publications, working papers, etc.), although deemed important in this particular field of research, are not included in the analysis. Each publication in WoS has a document type, such as 'article', 'book review', 'editorial material', 'letter', or 'review'. In our analysis, we only take into account publications of the document types 'article', 'letter' and 'review'. In general, these three document types cover the most significant publications and are referred to as 'citable items' since these publication types are the most commonly cited.

The publications of the researchers affiliated with one of the departments were collected as follows. UNL provided CWTS with a list of unique ISI publication keys. There were 5613 article, letter and reviews papers provided. Because letters are treated differently (they contribute only 0.25 in) the actual total comes to 5584.25. In **table 0** we have presented the breakdown of the publication numbers by department and the level at which an analysis within the CI can be considered representative of overall performance.

3 Overall results

3.1 Aggregated publication output and impact

Table 1 includes the overall bibliometric indicators for UNL for the period 2006 to 2012. In terms of the citation analysis, we have used a variable length window; so all citations received by all publications up to 2012 are counted. This means that for publications from each of the publication years (2009 – 2012), citations are counted up to and including 2013. A similar method has been applied to the four-year periods between 2006 - 2009 and 2009 - 2012. To facilitate comparison between periods, citations were counted for the same number of years plus one. If we take the 2006 - 2009 four-year period as an example, this means that, for publications from 2006 citations are counted during 2006 - 2010, for publications from 2007 citations are counted in 2007 - 2010, for 2008 publications citations from 2008 - 2010 are taken into account and the last publication year 2009 will harvest citations from 2009 and 2010.

The indicator of output (P) for the period 2006-2012 shows that: 5584.25 publications were used in the analysis, which is an increase of some 20% with respect to the report of 2004 - 2010. We see an increase from first year-block to last year-block of nearly 26%. For UNL as a whole its output is cited 49780.25 (Tcs) times excluding self-citations, on average this results in an Mcs of 8.91. The Mean Normalized Citation Score (MNCS) value for the whole period is 1.16, this means that UNL performs above (16%) world average. This is a significant rise in comparison to the previous report which presented an MNCS of 1.02. The publications appear in journals (MNJS) with an impact value slightly over world average (1.06) in which the MNCS is higher than that of the journals they publish in. In terms of the $PP_{top\ 10\%}$ indicator, **Table 1** shows that the publications of UNL keep up with the world level in the top 10% of most highly cited papers. The MNCS doesn't follow a well-defined pattern; it rises, drops and rises again but the share of top 10% best cited papers moves onto a slightly higher level. Against this backdrop it is not surprising that the level of uncited papers and proportion self-citations slightly decreases.

Table 1. Main Bibliometric indicators for UNL for the period (2006-2012/13)

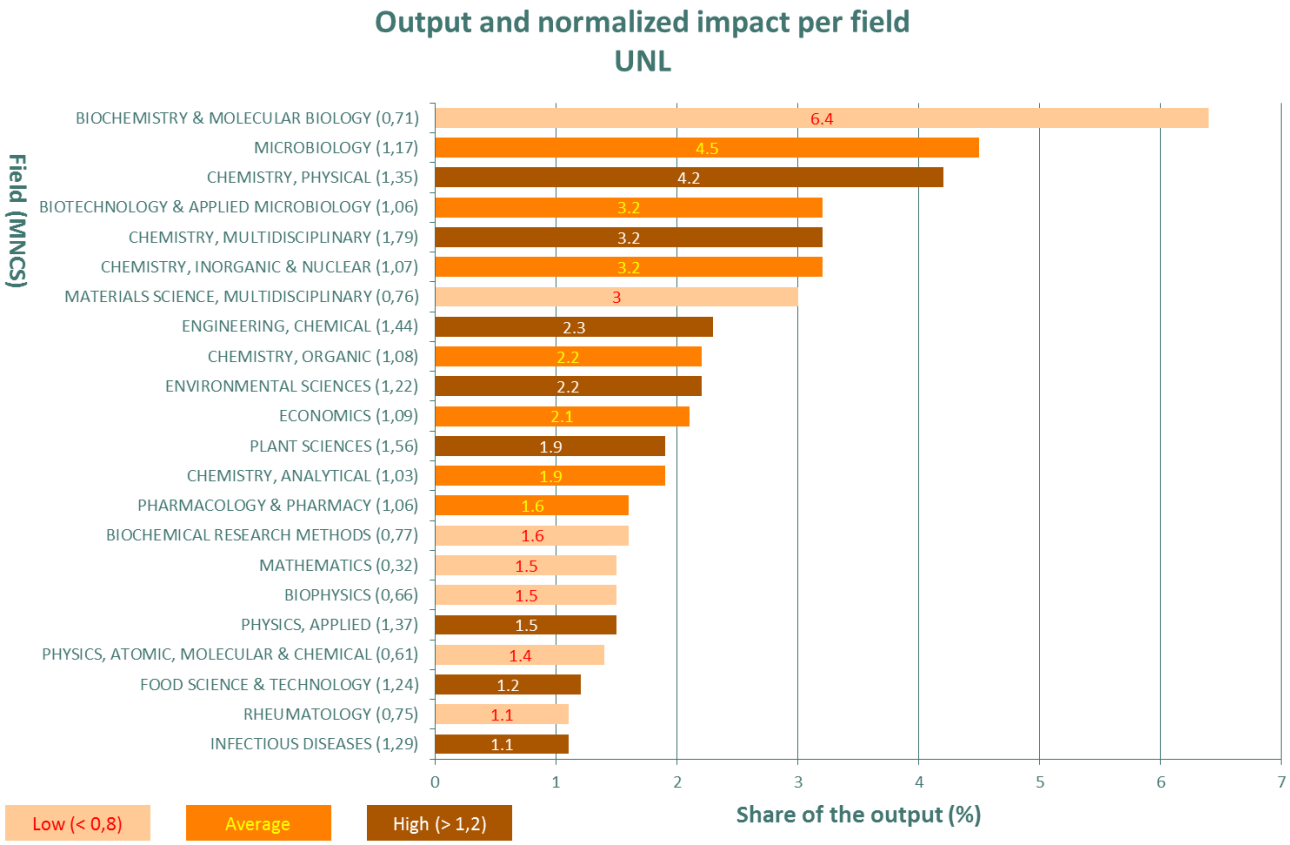
Year	Unit	P	MCS	TCS	MNCS	MNJS	TNCS	PP (top 10%)	PP (uncited)	Proportion self citations	PP (collab)	PP (int collab)	Internal coverage
2006 - 2012	UNL	5584.25	8.91	49780.25	1.15	1.06	6425.81	11%	20%	25%	81%	49%	81%
2006 - 2009	UNL	2819.00	5.18	14593.50	1.03	1.02	2906.98	10%	28%	28%	79%	48%	81%
2007 - 2010	UNL	2966.75	5.52	16378.25	1.09	1.04	3219.32	10%	25%	27%	80%	48%	81%
2008 - 2011	UNL	3282.75	5.41	17776.00	1.07	1.06	3520.96	11%	26%	28%	82%	49%	81%
2009 - 2012	UNL	3544.25	5.95	21104.00	1.19	1.08	4209.39	11%	24%	27%	82%	48%	81%

3.2 Research profiles.

Usually, scientists publish not only in journals within their specialty, but also in journals outside their field. Often research is of a multidisciplinary nature. An analysis of the publication output according to CI subject categories shows in which subfields the different research units are cited above or below the world subfield average. One should keep in mind that a subject category, for example 'Microbiology', refers only to a combination of journals, and not to an institutional or departmental affiliation. As a consequence, it is quite common that publications in one subject category have been contributed to by members from several research units within one institute. Scientific categories in the report have been limited to those who contributed more than 1% to the output total.

Figure 1 shows the research profile for UNL as a whole. The share of the output (%) and the fields, together with the impact (MNCS) by field is displayed. The most important subject categories for UNL are 'Biochemistry & Molecular Biology', 'Microbiology', 'Chemistry, Physical', all of them with a share in total output of more than 4%. All of them also have an impact that's ample over world average except for the largest category 'Biochemistry & Molecular Biology' that is slipping to 0.71 where it was 0.83 in the previous study. The reason for this decline is not all together clear. There are three subject fields which are at the same importance level at just over 3% share in total output; two chemistry categories of which one has one of the highest impacts of all the categories presented and 'Biotechnology & applied microbiology' which has an average impact.

Figure 1. Research Profile for UNL per scientific field



3.3 Knowledge users

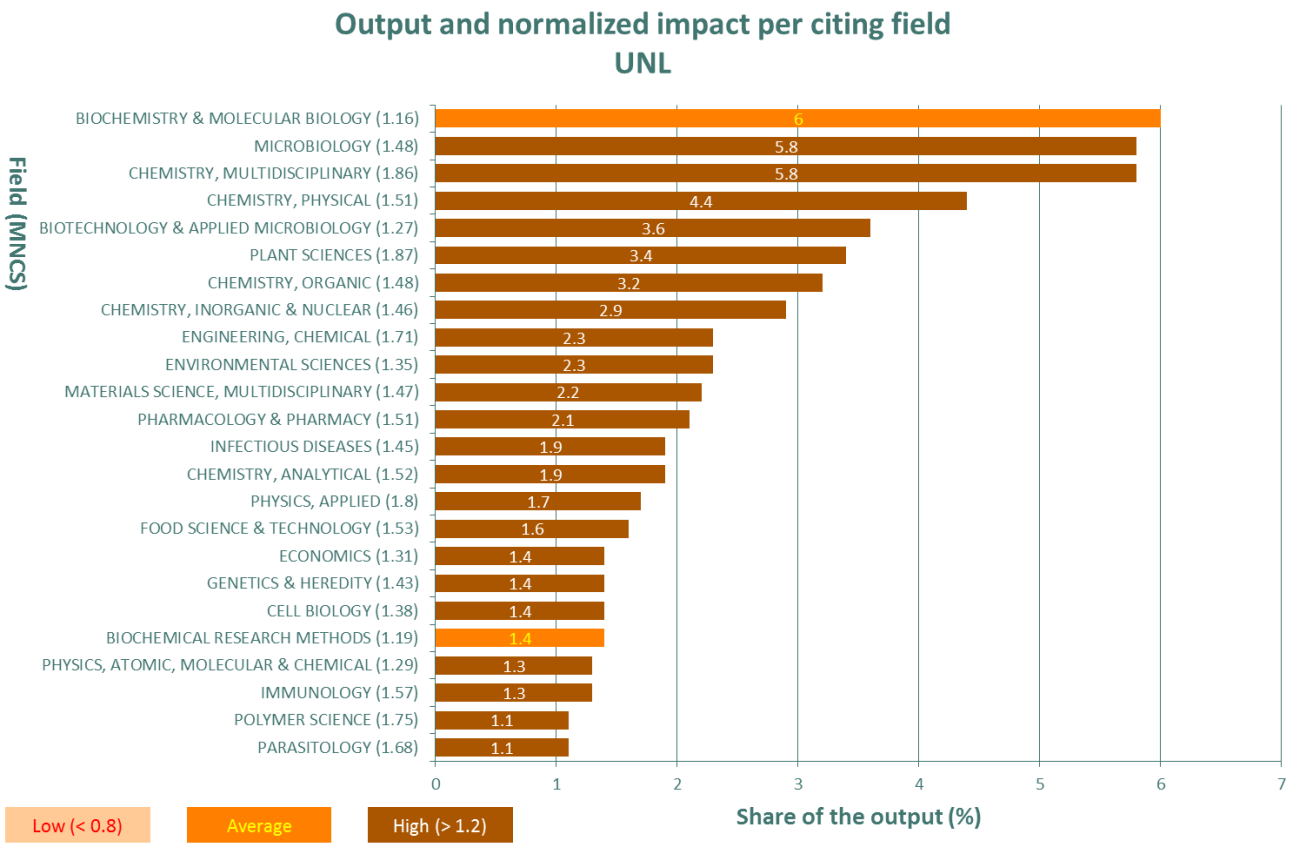
To identify users of published UNL knowledge, an 'impact profile' is calculated for UNL. A knowledge user impact profile is a breakdown of the publications citing UNL papers into entities of interest. First we categorize these citing publications into subfields of science. A citing publication is categorized only once, even if it cites more than one UNL paper. Author self-citations are excluded from the analysis.

Figure 2 presents the impact profile for UNL knowledge users for the period 2006 - 2012. It lists the most frequently occurring subfields of citing publications. Thus, **Figure 2** shows that among the top fields citing UNL publications we find the fields they are themselves most active in. Most notably: 'Biochemistry & Molecular Biology', 'Microbiology' and 'Chemistry Physical' which all have a citation share of over 4% and all have a high to very high impact (be it that strictly speaking : 'Biochemistry & Molecular Biology' doesn't qualify for the 'High' impact label by 0.04 index points).

The impact profile of the users of UNL knowledge differs from the cognitive profile (see **Figure 1**) of UNL output in that the citation density within categories is higher with UNL users than it is

with UNL output itself. Three fields are in the top of both lists: 'Biochemistry & Molecular Biology', 'Microbiology' and 'Chemistry Physical'. On the knowledge user side papers have a high- to very high impact.

Figure 2. Knowledge User Research Profile for UNL per scientific field



Identification of users

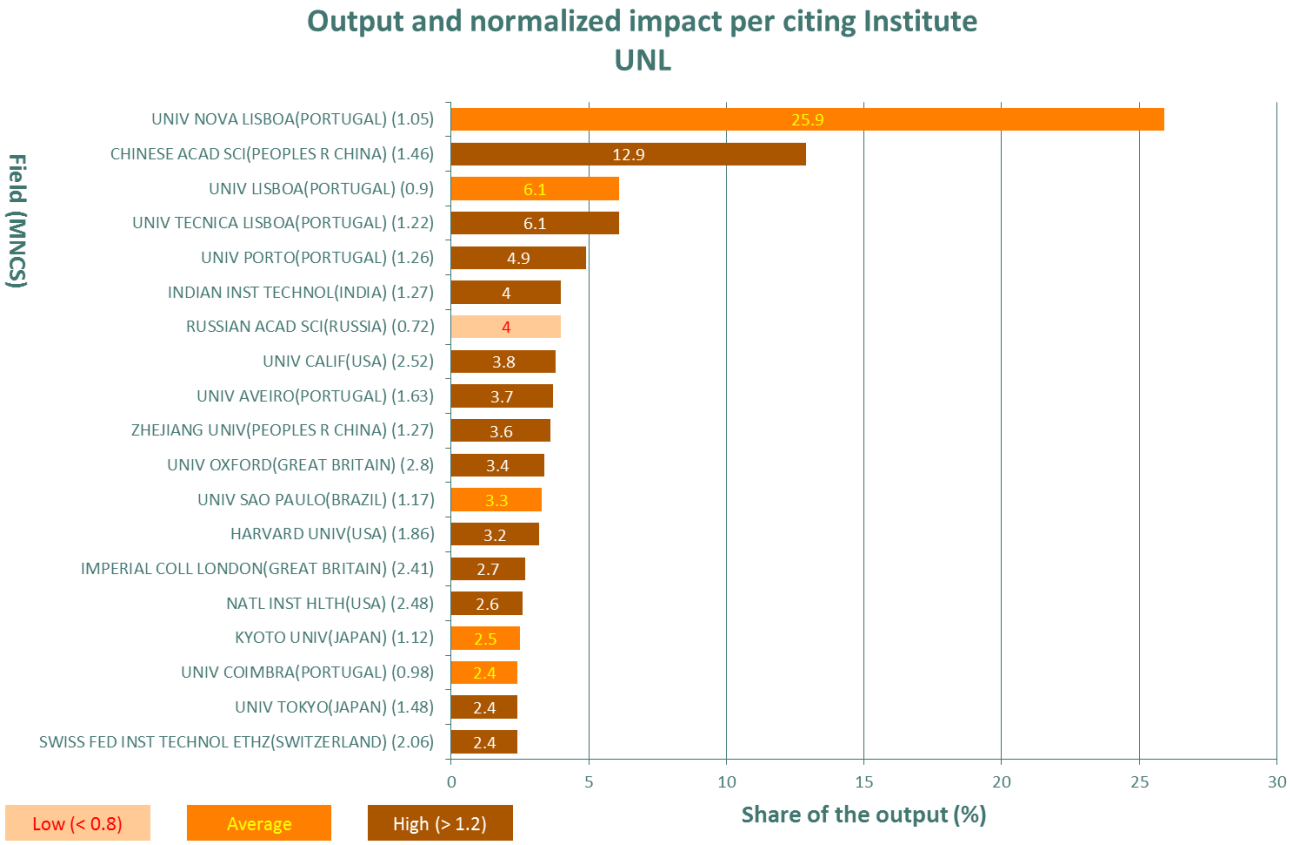
Figure 3 shows the whereabouts of the main users of published UNL knowledge from the period 2006 – 2012 in a knowledge user breakdown categorized by citing institute. Because of the shear data-volume in this analysis (some 23000 citing Institutes) we brought the number of institutes included in the analysis down to the 19 most important citing institutes, which is about all you can fit in a graph and the number of citations involved drops significantly below 100 after that. The cut-off point was at 0.25% share in total number of citing institutes. As a result the percentage share depicted is the share within these 19 most important knowledge users.

Although author self-citations were excluded, UNL heads the list. The list is highly dominated by universities from Portugal. Of the first 5 highest scoring usage institutes no less than 4 are from Portugal. The fifth (second largest user in the list) is the Chinese academy of Science from not

surprisingly China. Here the universities from the Lisbon area have an impact round world average whereas the other citing institutes from Portugal have a high impact. If we take the analysis further down the list to the first 12 institutes that have a share of over 0.4% we still find 40% of the usage from inside Portugal. Further down the line more important users from all over the world start to appear: Oxford University from England, University of California from the USA, the University of Sao Paolo from Brazil and Harvard University from the US.

Users of UNL knowledge tend to be very highly cited themselves. This indicates that UNL work is used by research of high impact, at the edge of the research frontier. Be it that contrary to the previous report from April 2012 the users now tend to originate in Portugal itself more.

Figure 3. Knowledge User Research Profile for UNL per Institute



4 Indicators for faculties, institutes and research departments

4.1 Block indicators and trend analysis for faculties/institutes

Block indicators and trend analysis for UNL faculties, institutes and school during 2006 – 2012/13 are presented in **Table 2**. There are large differences among these research units in terms of number of CI papers and citation impact. When indicators are based on very low volumes of both papers and citations bibliometric indicators may be less robust and should be interpreted with caution.

Focusing on the field-normalized citation impact (***MNCS***) during 2006 - 2012, **Table 2** shows that FCSH is cited well above average (+86%) and rising over time, although both the low numbers for this faculty as well as the low coverage (30%) (cfr. **Table 0**) suggest that these values must be observed with caution. ITQB/IBET, IHTM , ISEGI and FCT have an impact that is comfortably above the world average. FCM is a fraction low when compared to the world. The citation impact of the school NovaSBE has improved over the years and is now a good deal above average. In the case of the other school: The National School of Public Health (ENSP), the impact is very clearly lagging behind world measure although a slight improvement over the years may be developing. The Institute of Statistics & Information Management(ISEGI) is developing very well, although CWTS must warn up front that the statistics for this institute are based on a limited CI internal coverage level. Nevertheless it is seen to rise rather spectacularly in impact and output. Impact improves from some 30% under world level to nearly 30% over world level and at the same time the output increases more than 100%.

The output trend is clearly not wildly fluctuating with a slight upward slant for most faculties and institutes for the time period 2009 - 2012 as compared to 2006 - 2009. The two schools are an exception in this respect; they are more or less stationary. Two other further exceptions, FCSH and ISEGI, rise considerably in output.

Much the same development can be seen for the citation impact of publications (MNCS) and the journal mix published in (MNJS). No wild fluctuations but with a slight upward trend. FCSH shows a somewhat more pronounced wave pattern. Both MNCS and MNJS go up and down in an opposite movement. NovaSBE is marked by a rising MNCS and is publishing in a journal-mix that has an increase in impact with every passing year block.

Table 2. Bibliometric statistics for UNL faculties and institutes 2006 – 2012/13

Year	P	MCS	TCS	MNCS	MNJS	TNCS	PP (top 10%)	PP (uncited)	Proportion self citations	PP (collab)	PP (int collab)	Internal coverage
Faculty of Medical Sciences(FCM)												
2006 - 2012	**591.75**	**9.87**	**5840.25**	**0.92**	**0.95**	**563.07**	**9%**	**16%**	**19%**	**86%**	**43%**	**89%**
2006 - 2009	280.25	6.20	1736.75	0.86	1.01	283.65	7%	23%	20%	80%	41%	89%
2007 - 2009	344.25	6.95	2391.25	0.95	1.00	342.78	8%	23%	20%	81%	39%	89%
2008 - 2011	368.25	5.94	2185.75	0.86	0.91	335.16	8%	19%	21%	85%	40%	88%
2009 - 2012	407.00	6.01	2447.25	0.88	0.88	359.97	9%	21%	19%	89%	42%	89%
Faculty of Sciences and Technology(FCT)												
2006 - 2012	**2800.25**	**7.46**	**20876.50**	**1.09**	**1.06**	**2956.77**	**9%**	**23%**	**28%**	**79%**	**45%**	**78%**
2006 - 2009	1511.75	4.05	6128.50	0.96	1.00	1505.03	9%	33%	32%	77%	45%	78%
2007 - 2009	1507.50	4.52	6819.75	1.00	1.04	1571.61	10%	28%	31%	77%	45%	78%
2008 - 2011	1612.00	4.22	6810.25	0.94	1.06	1705.44	9%	29%	32%	80%	46%	78%
2009 - 2012	1672.75	5.05	8448.25	1.14	1.10	1846.49	9%	28%	30%	80%	46%	78%
Faculty of Social Sciences and Humanities(FCSH)												
2006 - 2012	**160.50**	**1.86**	**298.50**	**1.27**	**0.95**	**153.24**	**13%**	**64%**	**27%**	**41%**	**25%**	**30%**
2006 - 2009	32.00	0.78	25.00	1.02	0.79	25.25	12%	63%	31%	34%	19%	18%
2007 - 2009	46.25	1.29	59.75	1.57	0.86	39.65	13%	58%	26%	44%	26%	29%
2008 - 2011	89.25	1.70	151.50	0.84	1.04	92.54	12%	69%	28%	47%	29%	33%
2009 - 2012	142.50	1.76	251.50	1.22	0.95	135.94	13%	66%	29%	42%	27%	33%
Inst of Hygiene & Tropical Medicine(IHMT)												
2006 - 2012	**397.25**	**8.52**	**3385.75**	**1.08**	**0.99**	**392.91**	**11%**	**16%**	**30%**	**87%**	**64%**	**84%**
2006 - 2009	165.75	5.42	898.75	1.09	0.99	164.45	11%	20%	32%	82%	59%	88%
2007 - 2009	188.75	5.94	1121.25	1.08	1.03	193.87	11%	20%	32%	83%	63%	87%
2008 - 2011	248.50	6.13	1522.50	1.16	1.07	266.60	11%	23%	34%	86%	65%	84%
2009 - 2012	281.75	6.18	1741.00	1.05	0.99	278.72	11%	21%	32%	90%	66%	83%
Inst of Chemical & Biological Technology(ITQB/IBET)												
2006 - 2012	**1505.75**	**13.18**	**19842.75**	**1.35**	**1.14**	**1718.48**	**14%**	**10%**	**23%**	**87%**	**55%**	**91%**
2006 - 2009	770.50	8.06	6207.50	1.30	1.09	843.53	13%	16%	26%	88%	58%	90%
2007 - 2009	785.25	8.03	6301.75	1.32	1.08	849.20	14%	14%	25%	88%	56%	90%
2008 - 2011	857.25	8.36	7163.00	1.40	1.13	965.49	14%	15%	25%	88%	56%	91%
2009 - 2012	927.25	8.51	7888.00	1.39	1.18	1095.81	15%	12%	26%	87%	52%	92%
Inst of Statistics & Information Management(ISEGI)												
2006 - 2012	**104.25**	**5.07**	**529.00**	**1.18**	**0.85**	**89.04**	**15%**	**27%**	**15%**	**88%**	**43%**	**55%**
2006 - 2009	40.25	1.32	53.00	0.71	0.70	28.35	4%	43%	20%	92%	35%	59%
2007 - 2009	54.25	2.23	121.00	1.11	0.87	47.33	10%	35%	20%	92%	41%	55%
2008 - 2011	64.00	3.78	242.00	1.16	0.88	56.07	13%	31%	19%	91%	42%	54%
2009 - 2012	91.00	5.01	456.00	1.27	0.85	77.50	16%	29%	16%	89%	46%	53%

Year	P	MCS	TCS	MNCS	MNJS	TNCS	PP (top 10%)	PP (uncited)	Proportion self citations	PP (collab)	PP (int collab)	Internal coverage
National School of Public Health(ENSP)												
2006 - 2012	**48.00**	**5.81**	**279.00**	**0.63**	**0.91**	**43.89**	**1%**	**19%**	**11%**	**85%**	**35%**	**71%**
2006 - 2009	25.00	3.64	91.00	0.52	0.93	23.36	0%	32%	9%	76%	32%	64%
2007 - 2009	20.00	4.45	89.00	0.69	1.10	21.98	0%	15%	9%	75%	40%	65%
2008 - 2011	21.00	2.67	56.00	0.66	1.08	22.65	0%	19%	20%	86%	43%	73%
2009 - 2012	27.00	3.74	101.00	0.72	0.95	25.60	2%	26%	18%	89%	41%	74%
Nova School of Business and Economics(NovaSBE)												
2006 - 2012	**228.00**	**4.99**	**1138.00**	**1.17**	**1.21**	**276.45**	**13%**	**22%**	**15%**	**78%**	**54%**	**61%**
2006 - 2009	128.00	1.88	241.00	0.76	1.08	138.72	6%	46%	23%	78%	52%	57%
2007 - 2009	146.00	2.17	317.00	0.80	1.10	160.22	8%	40%	24%	76%	50%	59%
2008 - 2011	156.00	3.03	472.00	1.10	1.19	185.31	10%	33%	22%	78%	54%	60%
2009 - 2012	144.00	3.69	532.00	1.35	1.25	180.33	16%	28%	17%	78%	56%	63%

4.2 Most frequently cited publications for UNL faculties/institutes

Table 3. Highly cited papers 2006 – 2010/13

Unit	P 2006-2010	P 1%	pp top 1%	P 2%	pp top 2%	P 5%	pp top 5%	P 10%	pp top 10%	p 20%	pp top 20%
Faculty of Medical Sciences(FCM)	383.75	4.00	1.04%	9.00	2.35%	17.32	4.51%	32.36	8.43%	63.65	16.59%
Faculty of Sciences and Technology(FCT)	1900.75	16.08	0.85%	33.52	1.76%	77.48	4.08%	170.44	8.97%	371.80	19.56%
Faculty of Social Sciences and Humanities(FCSH)	51.25	0.02	0.04%	1.39	2.72%	6.57	12.82%	8.91	17.39%	11.75	22.93%
Inst of Chemical & Biological Technology(ITQB/IBET)	991.50	14.49	1.46%	30.10	3.04%	68.15	6.87%	136.43	13.76%	251.00	25.31%
Inst of Hygiene & Tropical Medicine(IHMT)	224.75	2.83	1.26%	5.72	2.55%	15.87	7.06%	21.30	9.48%	46.81	20.83%
Inst of Statistics & Information Management(ISEGI)	59.25	1.00	1.69%	3.00	5.06%	5.58	9.42%	12.41	20.94%	14.84	25.04%
National School of Public Health(ENSP)	27.00	0.00	0.00%	0.00	0.00%	0.00	0.00%	0.00	0.00%	3.71	13.75%
Nova School of Business and Economics(NovaSBE)	164.00	1.00	0.61%	1.00	0.61%	7.51	4.58%	13.92	8.49%	28.66	17.48%
UNL	3639.75	38.08	1.05%	81.05	2.23%	192.43	5.29%	380.46	10.45%	753.91	20.71%

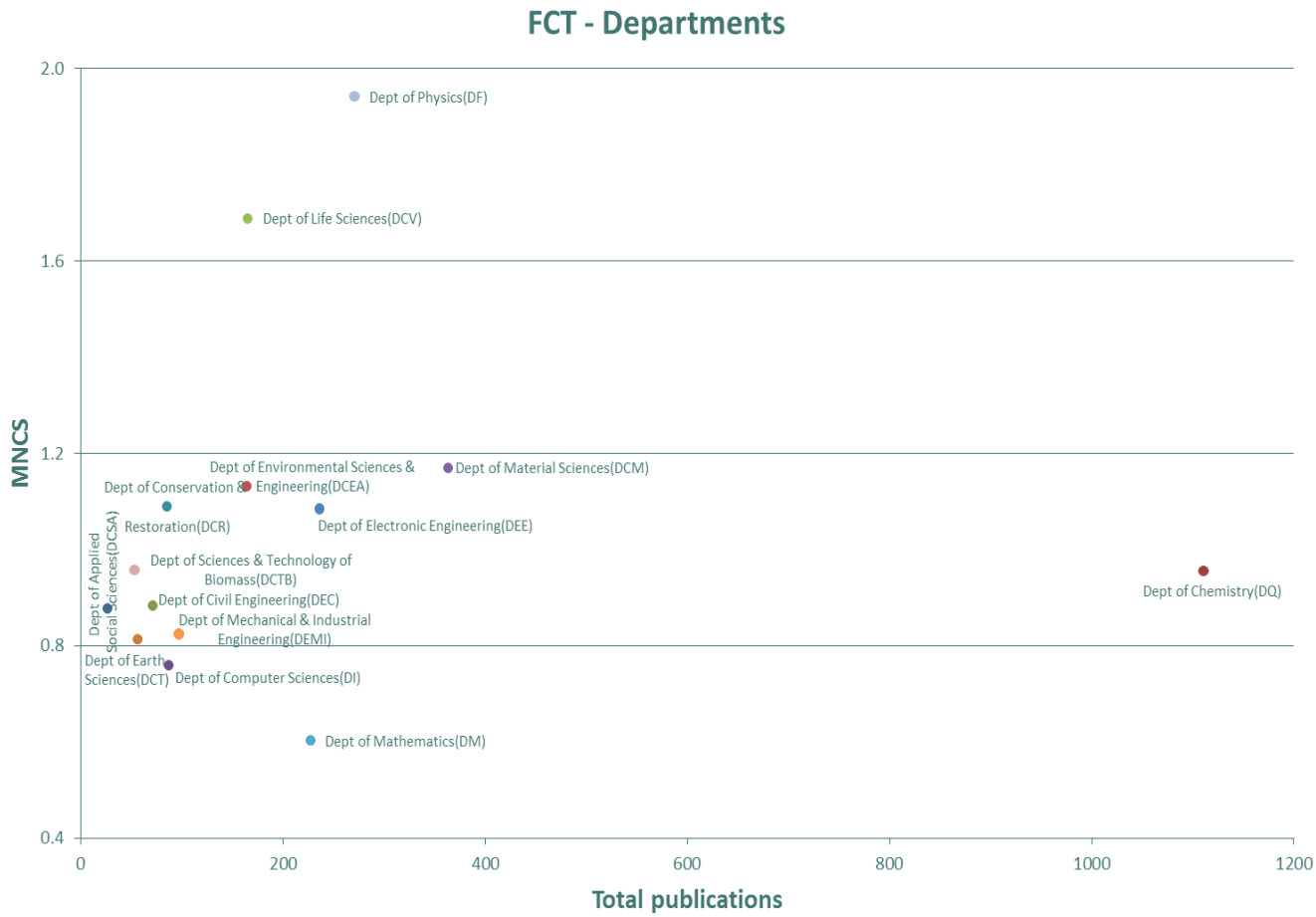
Four units in the most frequently cited analysis (almost) constantly show a citation visibilty above average. These are 'Faculty of Social Sciences and Humanities (FCSH)', 'Institute of Chemical & Biological Technology (ITQB/IBET)', 'Institute of Hygiene & Tropical Medicine (IHMT)' and 'Institute of Statistics & Information Management (ISEGI)'. It must be noted here again however that the 'Faculty of Social Sciences and Humanities (FCSH)' has a limited CI coverage. UNL itself is keeping perfect pace with the highly cited world levels. 'Faculty of Medical Sciences (FCM)' is in line with world average highly cited in the pp top 1- and 2% but on the more modest pp top levels they undercut the standard slightly. 'Faculty of Sciences and Technology (FCT)' and 'Nova School of Business and Economics (NovaSBE)' underperform somewhat over the entire spectrum of values. We can expect this for NovaSBE to be a result from the past rather than the present as their results as shown previously are on the rise in this period of analysis. 'National School of Public Health(ENSP)' hardly registers in this analysis.

4.3 A survey of output and impact results for research departments

We made an analysis of the relation between output (in terms of CI-publications) and citation impact on the level of departments and sections of the Faculty of Sciences and Technology (FCT). **Figure 4** combines output figures (P) and field-normalized impact results (***MNCS***) for 2006 – 2012/13. Some departments recorded both a low number of CI papers and a low citation impact which frequently coincides with a low CI coverage. In these cases, indicators may be less robust and not adequate for inferential analysis.

It clearly shows that three layers are present in the graphical representation of the departments and their P versus MNCS: a higher level, an average to high level and a low to average level. The departments position themselves markedly in one of the three different levels. Worthy of note is that 'Department of Physics' and 'Department of Life Sciences' dominate on the impact dimension and 'Department of Chemistry' dominates on the output dimension. The departments all fall into three well delineated comparison levels, in between 0.4 and 0.8, from 0.5 to 1.2 and between 1.2 and 2.0. So compliant to the 'high', 'average' and 'low' impact levels used throughout this report. The department of Computer Sciences edges in the low impact realm towards the boundary of the average layer.

Figure 4. Impact compared to world average.



5 Collaboration Analysis

In this section the main collaborative patterns of UNL's output is analyzed. In **table 1** it was already shown that 81% of UNL publications are done in collaboration and 49% in international collaboration, thus indicating that the latter is the most important collaboration typology of the publications.

Three types of scientific collaboration were distinguished. Publications with only one address were assigned to '*No collaboration'* or '*Single institute*'. Publications with multiple addresses, all from the same country, were assigned to '*National collaboration'*. Finally, all publications with at least one address outside Portugal (since we may assume that all publication addresses share at least that country) were marked with the collaboration type '*International collaboration*'.

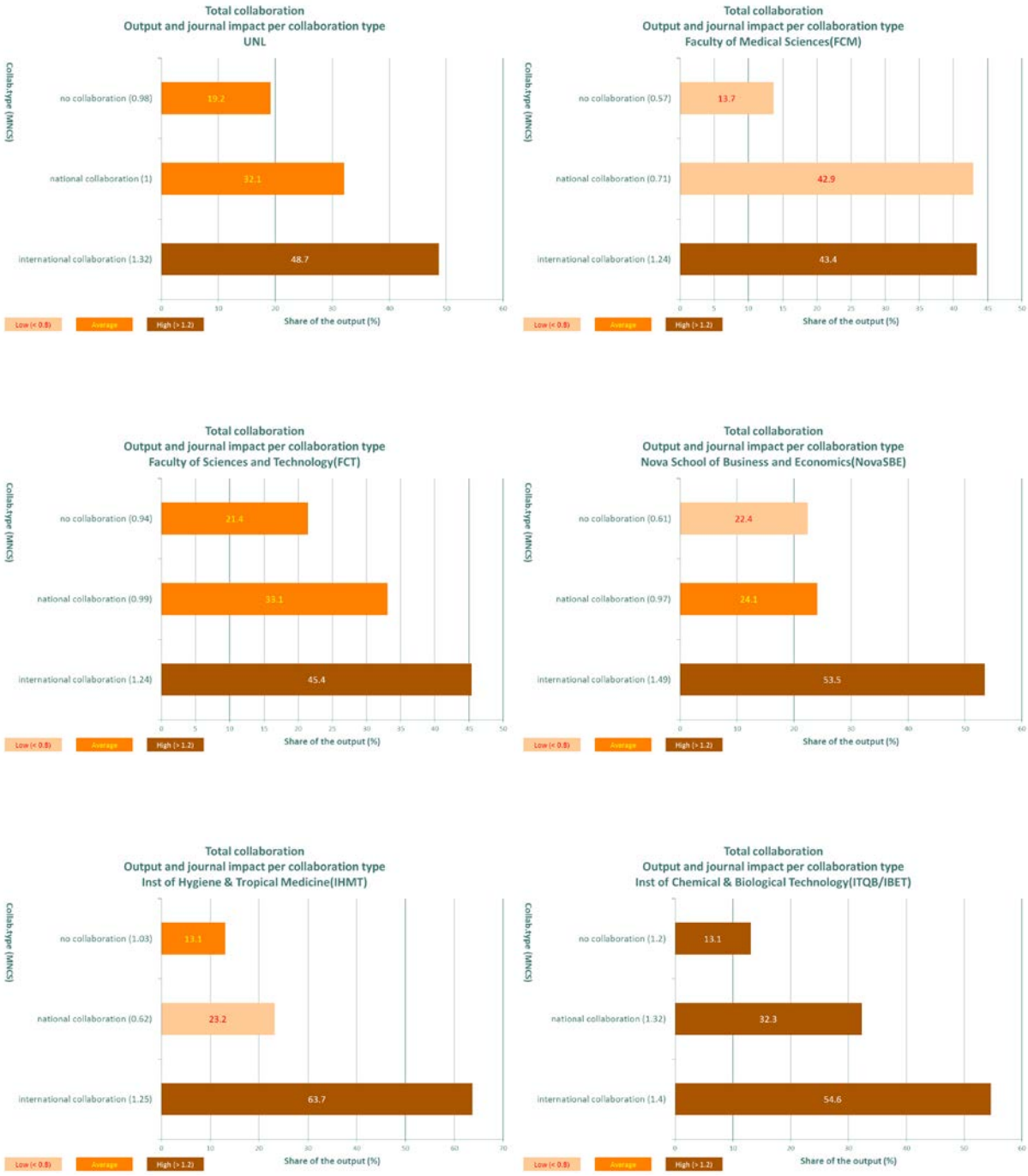
Figure 5 shows the percentage of the total output (P) represented by each of the three types of collaboration for UNL and its major faculties during 2006 - 2012. It has been indicated whether the impact compared to the world subfield average (MNCS) is 'low' (<0.80), 'average' (0.80 -

1.20) or 'high' (>1.20). Only the most important units with at least 200 CI papers (the top 5 output producing departments/institutes) were included in the figure.

Figure 5 also shows the importance of collaboration for UNL researchers: 'internally' produced publications represent less than 20% of the total output. Both 'national' and in particular 'international collaboration' are of considerable importance. For UNL as a whole, publications stemming from all types of collaboration are cited competitively with world average. However, **Figure 5** also shows that several faculties have somewhat divergent signatures in scientific collaboration. It is then clearly visible that publications produced in international collaborations are the most fruitful endeavours. These publications have invariably the highest impact, between 25% and up to even 50% higher than world average.

The results show that UNL researchers as a whole tend to contribute as substantially to international scientific networks as they do to other types of collaboration and that international co-authored publications are the most visible publication types, except for the institute of Hygiene & Tropical Medicine which has a higher representation in international collaboration than the other two combined. The citation pattern of the 'Institute of 'Chemical & Biological Technology (ITQB/IBET)' is high, not only in international publications but overall.

Figure 5. Impact analysis for type of collaboration 2006 – 2012/13



6 Conclusion

In this study, the research performance of UNL research during 2006 – 2012 has been compared with a number of international citation reference values. In this study we find that from 2006 – 2009 to 2009 – 2012 the output again increased by almost 26% with a total increase over the entire period that's 20% higher than in the previous report. MNCS appears to be higher than in the previous study and presents a score that differs sufficiently from previous findings in the report of 2004 – 2010 to call this a significant improvement. What's beyond doubt important is that in this study also, the UNL performs consistently above world average and that UNL papers outperform the journals they publish in, which are in their turn also competitive on the world level. The most important scientific field in which UNL is active 'Biochemistry & Molecular Biology' has decreased noticeably in impact and the most frequent knowledge users tend to be situated more in Portugal itself. When working in international cooperation, the papers UNL puts out have a high impact, for the UNL as a whole and for the 5 highest output producing departments/institutes as well.

On the level of faculties and institutes, the field-normalized impact of FCSH that increased spectacular over 2002 – 2008 and still was 50% over world average in 2004 – 2010, seems to top off in 2007 – 2010, goes into decline and regains some of its impact at the end of the period. As in the previous study the impact of ITQB/IBET is very comfortably above average and the same goes for both ISEGI and to a lesser extent IHTM, for which the impact seems to top in the year-block 2008 - 2011. The citation impact for most of the other faculties/Institutes is close to world average. ENSP is however severely below average. The trend analysis shows that the citation impact for NovaSBE has gone through a performance mark-up and the same goes for Inst of Statistics & Information Management (ISEGI). Starting at a low impact in 2006 – 2009, they have been increasing with some 0.60 MNCS index points to arrive at a value over world average at the end of the period.

The citation impact of UNL researchers is competitive with the world average in four out of its five top output subfields, however the largest scientific field UNL publishes in 'Biochemistry & Molecular Biology' (more than 50% of total output) does not hold its own on the world level. Citation results for this scientific field in the previous study were also below average by some 20% but it has slipped since to arrive at 30% under world average. This is an unexpected turn of events within the overall picture of a well performing institute that could warrant further investigation.

Acknowledgments

The author acknowledges the critical reading of this report by Mark Neijssel en Ed Noyons from CWTS – Leiden University.

References

Borgatti, S. P. (2002). *Netdraw: Network Visualization Software*. Harvard: Analytic Technologies. http://www.analytictech.com/downloadnd.htm.

De Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory Social Network Analysis with Pajek.* New York: Cambridge University Press.

Moed, H.F. (2005). *Citation analysis in research evaluation*. The Netherlands: Springer.

Reijnhoudt, L., Costas, R., Noyons, E., Boerner, K., & Scharnhorst, A. (2013). "Seed+Expand": A validated methodology for creating high quality publication oeuvres of individual researchers. *arXiv*. Retrieved from http://arxiv.org/abs/1301.5177

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E., Tijssen, R., van Eck, N.J., van Leeuwen, T., van Raan, A., Visser, M., & Wouters P. (2012) The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, 63, 12, 2419-2432.

Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S., & van Raan, A.F.J. (2011a). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5, 1, 37-47.

Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S., & van Raan, A.F.J. (2011b). Towards a new crown indicator: An empirical analysis. *Scientometrics*, 87, 3, 467-481.

Appendix I

To illustrate the calculation of the MNCS indicator, we consider a hypothetical research group that has only five publications. **Table** provides some bibliometric data for these five publications. For each publication, the table shows the scientific field, to which the publication belongs, the year in which the publication appeared, and the actual and the expected number of citations of the publication. (For the moment, the last column of the table can be ignored.) The five publications are all of them document type article.

Citations have been counted using a variable-length citation window. As can be seen in the table, publications 1 and 2 have the same expected number of citations. This is because these two publications belong to the same field and have the same publication year and the same document type. Publication 5 also belongs to the same field and has the same document type. However, this publication has a more recent publication year, and it therefore has a smaller expected number of citations. It can further be seen that publications 3 and 4 have the same publication year and the same document type. The fact that publication 4 has a larger expected number of citations than publication 3 indicates that publication 4 belongs to a field with a higher citation density than the field in which publication 3 was published. The MNCS indicator equals the average of the ratios of actual and expected citation scores of the five publications. Based on Table 1, we obtain

$$\text{MNCS} = \frac{1}{5}\left(\frac{7}{6.13} + \frac{37}{6.13} + \frac{4}{5.66} + \frac{23}{9.10} + \frac{0}{1.80}\right) = 2.08$$

Hence, on average the publications of our hypothetical research group have been cited more than twice as frequently as would be expected based on their field, publication year, and document type.

Table X: Bibliometric data for the publications of a hypothetical research group.

Publication	*Field*	*Year*	*Actual citations*	*Expected citations*	*Top 10% threshold*
1	Surgery	2007	7	6.13	15
2	Surgery	2007	37	6.13	15
3	Clinical neurology	2008	4	5.66	13
4	Hematology	2008	23	9.10	21
5	Surgery	2009	0	1.80	5

To illustrate the calculation of the $PP_{top\ 10\%}$ indicator, we use the same example as we did for the MNCS indicator. **Table** shows the bibliometric data for the five publications of the hypothetical

research group that we consider. The last column of the table indicates for each publication the minimum number of citations needed to belong to the top 10% of all publications in the same field and the same publication year and of the same document type.[4] Of the five publications, there are two (i.e., publications 2 and 4) whose number of citations is above the top 10% threshold. These two publications are top 10% publications. It follows that the $PP_{top\ 10\%}$ indicator equals

$$PP_{top\ 10\%} = \frac{2}{5} = 0.4 = 40\%$$

In other words, top 10% publications are four times overrepresented in the set of publications of our hypothetical research group.

Appendix II

List of underlying data to the analyses in this report and sent with the report in Excel format:

1. 10 most highly cited papers.xlsx
2. Cognitive Orientation UNL.xlsx
3. Collaboration.xlsx
4. Departments.xlsx
5. Highly Cited.xlsx
6. ISI Share.xlsx
7. Knowledge User Institutes.xlsx
8. Knowledge User Scientific Fields.xlsx
9. Trend Faculties and Institutes.xlsx
10. Trend.xlsx

[4] If the number of citations of a publication is exactly equal to the top 10% threshold, the publication is partly classified as a top 10% publication and partly classified as a non-top-10% publication. This is done in order to ensure that for each combination of a field, a publication year, and a document type we end up with exactly 10% top 10% publications.