WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

C. PCT 1398 November 4, 2013

Madam,
Sir,

Report on Characteristics of International Search Reports

This Circular is addressed to your Office in its capacity as an International Searching Authority under the Patent Cooperation Treaty (PCT). In that capacity, and as a member of
./. the quality subgroup of the Meeting of International Authorities under the PCT, your Office is invited to comment on the annexed draft report on characteristics of international search reports with a view to identifying indicators of what should be the focus of further work by the International Authorities.

Background

At the third informal meeting of the quality subgroup held in Munich in February 2013 prior to the twentieth session of the Meeting of International Authorities, the subgroup discussed a paper analyzing the responses from International Authorities to Circular C. PCT 1360, dated October 8, 2012, presenting various characteristics of international search reports. This discussion by the quality subgroup is summarized in paragraphs 20 to 24 of Annex II to document PCT/MIA/20/14. Paragraph 24 of that document outlines future work in this area recommended by the Subgroup as follows:

> "24. The Subgroup recommended that:
>
> "(a) the International Bureau should provide similar reports in coming years;

/...

34, chemin des Colombettes
1211 Geneva 20, Switzerland
T +4122 338 91 11 **F** +4122 733 54 28
www.wipo.int

"(b) the International Bureau should complete the assessment of which of the desired improvements were possible using existing data and make recommendations for changes which should be included in the next version;

"(c) the Subgroup should continue to analyze what improvements were desirable and make recommendations on the additional data which should be collected to support this."

Updated Report

./. Annex I to this Circular presents a report on characteristics of international search reports, updating the data presented in Circular C. PCT 1360 up to the end of the fourth quarter of 2011. The updated report takes into account the feedback provided by International Authorities to Circular C. PCT 1360 and that provided at the most recent informal meeting of the quality subgroup. Further explanation of the description of the data and definitions used
./. is included in Annex II to this Circular.

In addition to using more recent data as the basis for this year's report, a new indicator "total across all International Searching Authorities" has been added (where appropriate), all charts have been renumbered and the following new charts have been produced in response to requests from International Authorities:

1.7 – Percentage of PCT search reports with D citations

1.8 – Percentage of PCT search reports with O/T/L citations

2.4 – Average number of X/Y patent literature citations per search report

2.7 – Percentage of PCT search reports with at least one X/Y NPL citation.

In relation to section 3.1, showing percentage of patent citations in non-official languages (previously numbered section 1.1.18), the definition of an official language has remained unchanged from C. PCT 1360. However, it is recognized that the presently used definition produces results that differ widely in percentage terms between Offices where a majority of patent disclosures can be found in an official language and Offices where there are relatively few patent documents published in the official language and examiners are likely to have a high degree of proficiency in other languages. The International Bureau thus suggests to work towards finding a different definition for "non-official languages" of Offices to be used in future reports.

Issues for the Quality Subgroup

The production of updated information on characteristics of international search reports is an automated process using existing datasets (the PATSTAT database of the European Patent Office and the internal PCT database of the International Bureau). It is therefore a relatively straightforward exercise for the International Bureau. However, while the International Bureau is happy to continue processing these characteristics for the establishment of yearly reports, before making significant further improvements to their presentation, it is important to have a clear picture of how the characteristics are being used and any future intentions of International Authorities in this regard. Your Office as a member of the quality subgroup is therefore invited to respond to the following questions:

(a) How have the reports on the characteristics of international search reports helped in determining the direction of internal quality work of your Office and identifying matters to be discussed between International Authorities? Which metrics in the report have been of particular value in this respect?

(b) What practical use has your Office made of the information presented in the reports, especially as part of processes to improve the quality of international search? What further use of the information could be made in the future in quality-related work by International Authorities?

The updated reports present some, but not all of the data requested by International Authorities in their feedback on the report in Circular C. PCT 1360. Some of the data requested is not possible to extract from existing data sets and other data, while theoretically possible to present, would involve considerable effort making data from different sources fully compatible before processing. On the other hand, for some data, such as a breakdown of the charts by the 35 different technical fields, it is more straightforward to present graphically in a report, but the report would be an extremely long document, making it difficult to find the most relevant information. Furthermore, the data presented do not enable real-time active management of quality by an International Authority, given that the most recent figures available are about two years old.

The International Bureau is willing to continue to improve the international search report metrics presented in the future. However, this can only be possible where data can be obtained. Moreover, any improvements made would need to bring benefits that would be commensurate with the work involved. Therefore, with a view to considering further improvements to the report on characteristics of international search reports, your Office in its capacity as an International Authority is invited to respond to the following questions:

(c) Are there alternative ways of presenting the data in the report to improve accessibility to the information which would result in the data being more actively used by your Office?

(d) If it is desired to include new metrics using data that are currently not readily available to the International Bureau, to what extent is your Office willing to change its existing systems to provide these data in a structured format?

(e) Is there scope for making the data sources more up to date to enable active management of current issues rather than reviewing past performance?

Your Office in its capacity as an International Authority and a member of the quality subgroup is invited to provide comments to these questions, using the subgroup's electronic forum, by December 2, 2013.

Yours sincerely,

James Pooley
Deputy Director General

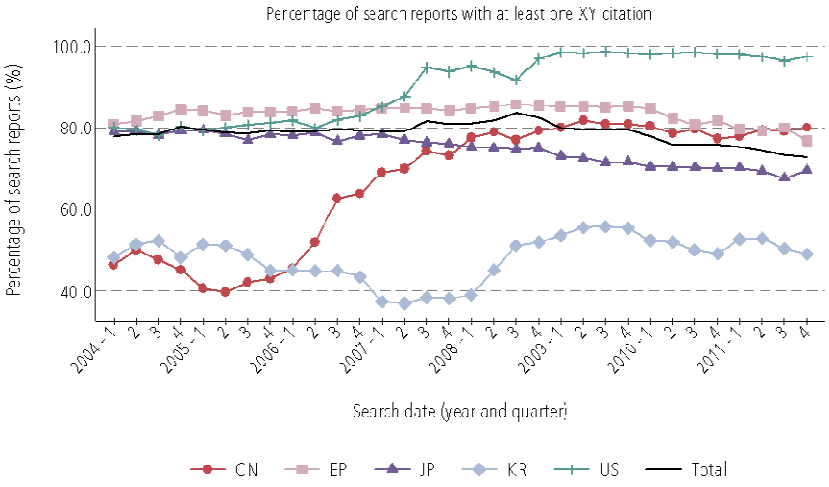
Enclosures: Annex I: Report on Characteristics of International Search Reports
Annex II: Description and Definitions

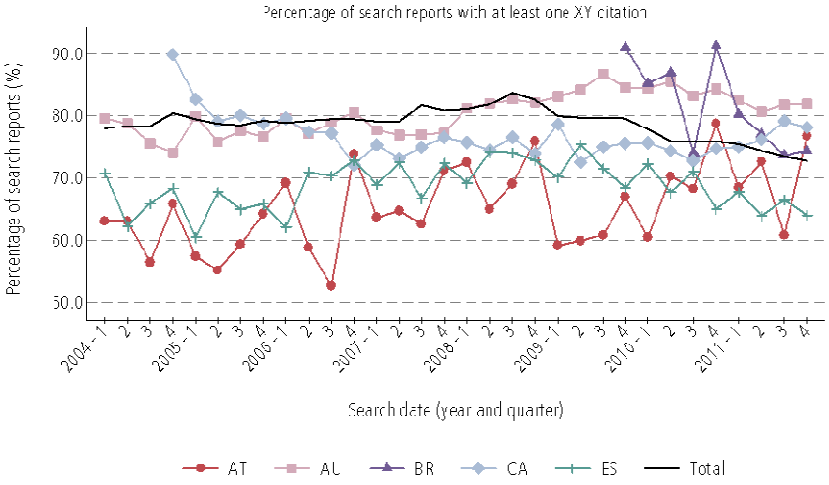
CHARACTERISTICS OF INTERNATIONAL SEARCH REPORTS

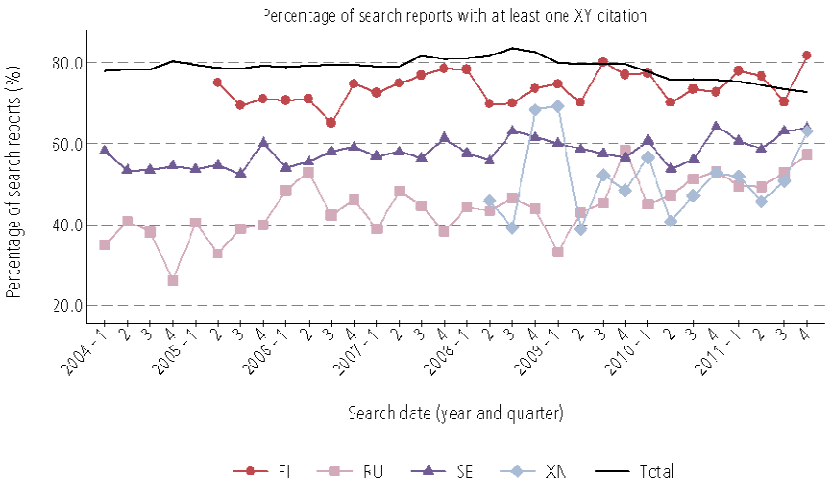
Table of Contents

1.1 – Percentage of PCT search reports with at least one X, Y or E citation 2
1.2 – Percentage of PCT search reports with A citations only 3
1.3 – Percentage of PCT search reports with P or E citations 4
1.4 – Percentage of citations in the category of P or E 5
1.5 – Percentage of citations in the category of P and X 6
1.6 – Percentage of PCT search reports with Y citations and without X citations 7
1.7 – Percentage of PCT search reports with D citations 8
1.8 – Percentage of PCT search reports with O/T/L citations 9
1.9 – P/E Citation Breakdown for PCT Search reports (2011) 10
1.10 – Percentage of PCT search reports with at least one X or Y citation by top applicant's origin (2011) 12
2.1 – Average number of citations per search report 13
2.2 – Average number of NPL citations per search report 14
2.3 – Average number of patent literature citations per search report 15
2.4 – Average number of X/Y patent literature citations per search report 16
2.5 – Percentage of NPL citations in all citations 17
2.6 – Percentage of NPL citations in the category of X or Y in all citations 18
2.7 – Percentage of PCT search reports with at least one X/Y NPL Citation 19
2.8 – Percentage of NPL citations in all citations 20
2.9 – Distribution of patent & non patent literature citations (2011) 21
3.1 – Percentage of patent citations in non-official languages 22
3.2 – Percentage of patent citations by top publication authorities (2011) 23
3.3 – Percentage of patent citations by processing authorities (i.e. publication authorities except WO) (2011) 24
3.4 – Percentage of PCT search reports by top applicant's origin (2011) 25
Data set description 1
Data issues 1
Definition of concepts 1
Technology breakdown 1
Applicant origin 2
XY rate (Searches with XY citations) 2
Citation category availability 2
A only rate 3
Y no X rate 3
Search date 3
Patent Literature/Non-Patent Literature 3
Non official language 4
Publication Authority (of citation) 4
Processing Authority (of citation) 4

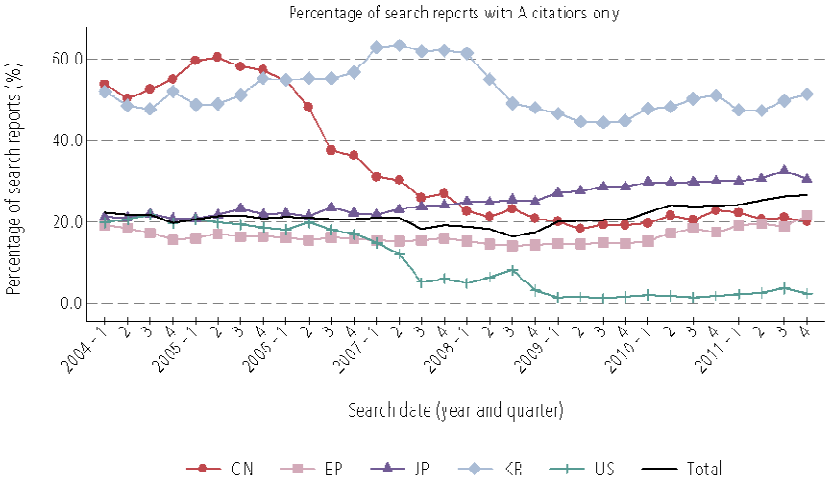
1.1 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X, Y OR E CITATION

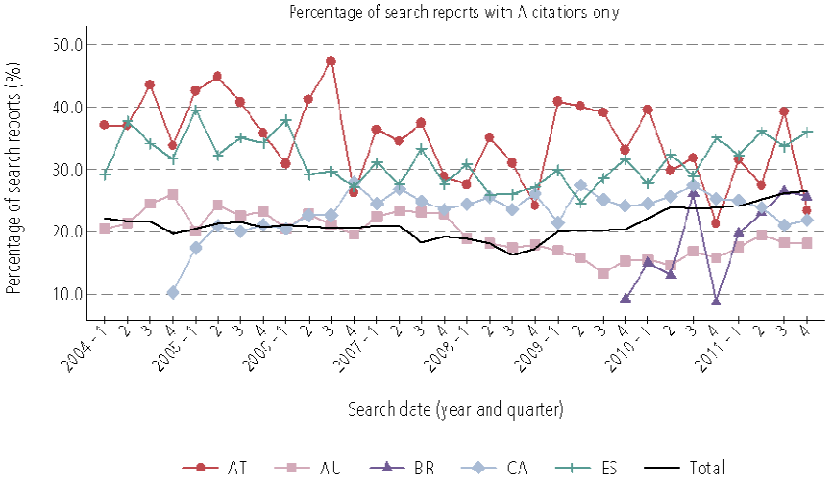


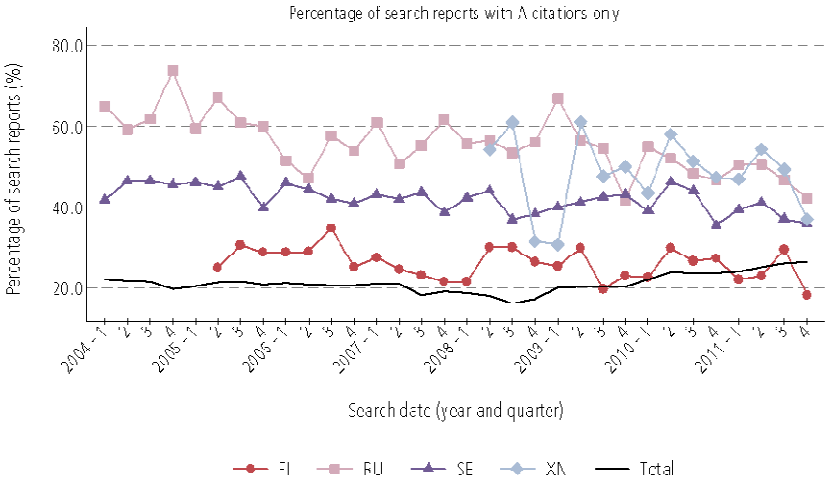




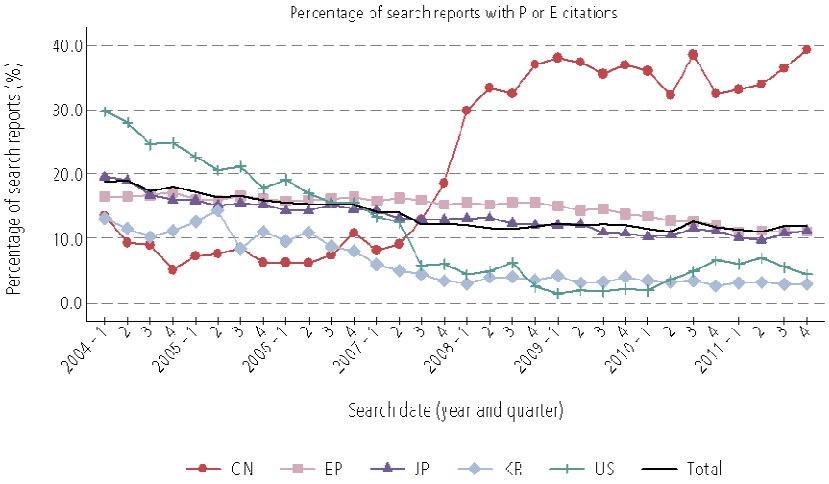
1.2 – PERCENTAGE OF PCT SEARCH REPORTS WITH A CITATIONS ONLY

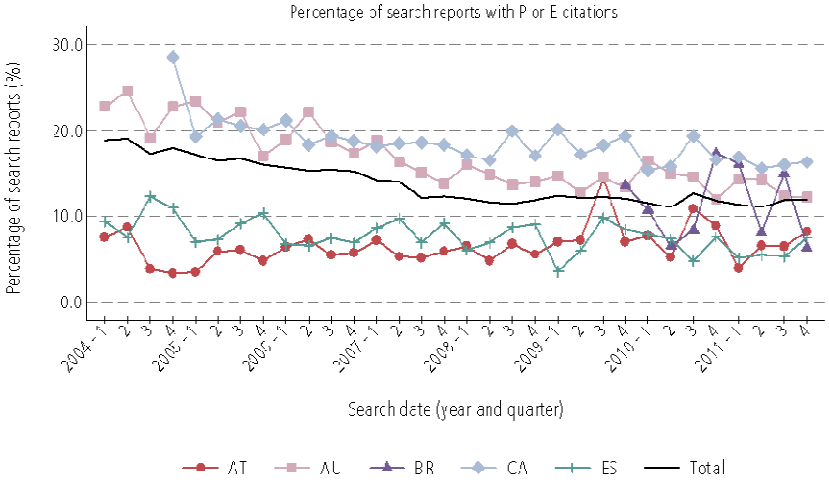


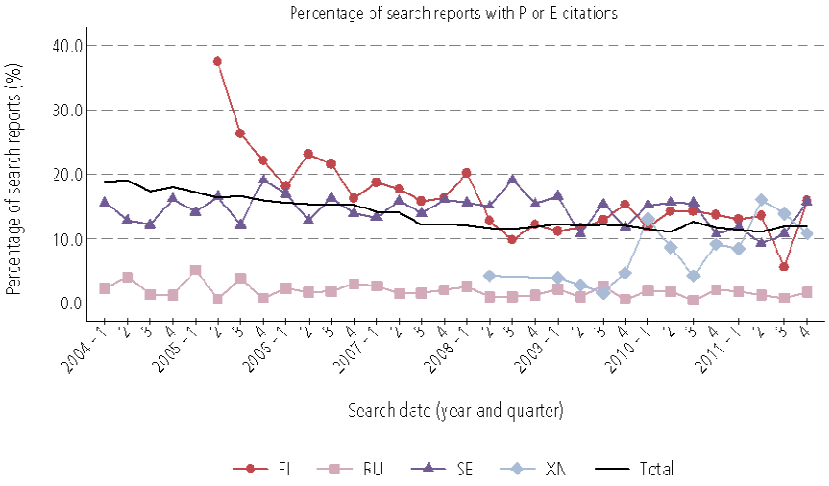




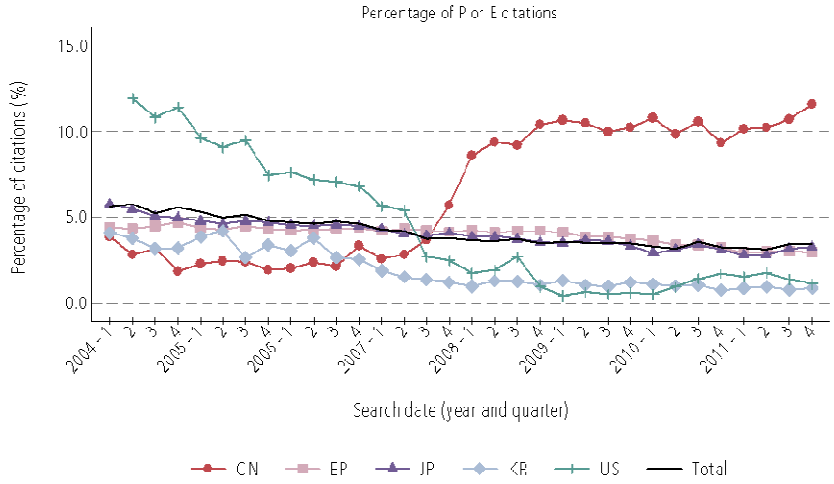
1.3 – PERCENTAGE OF PCT SEARCH REPORTS WITH P OR E CITATIONS

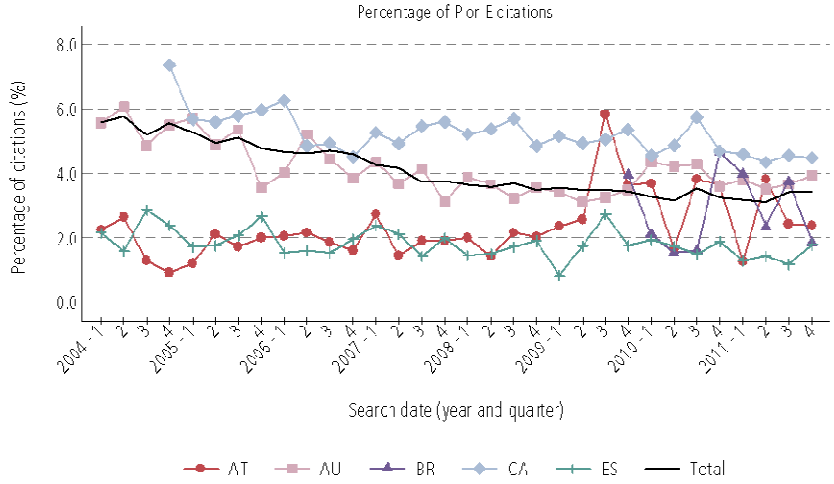


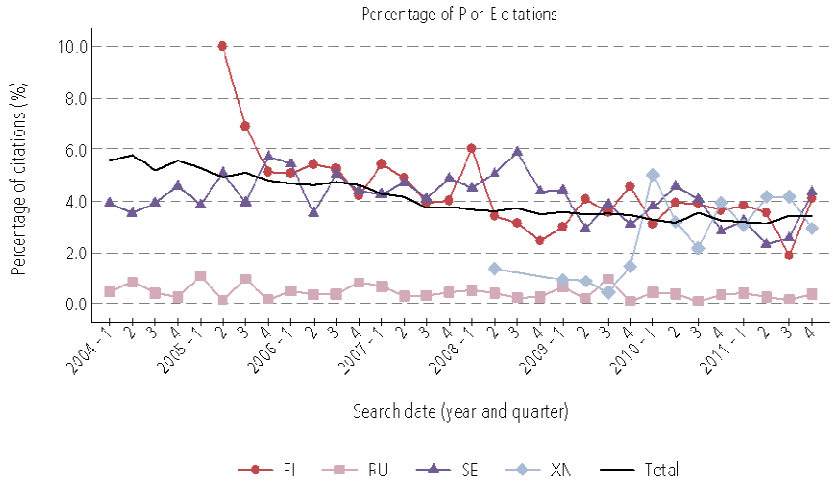




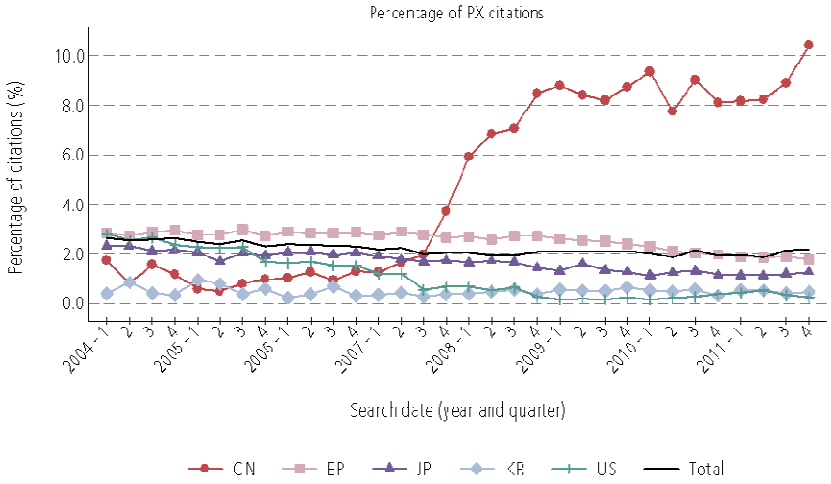
1.4 – PERCENTAGE OF CITATIONS IN THE CATEGORY OF P OR E

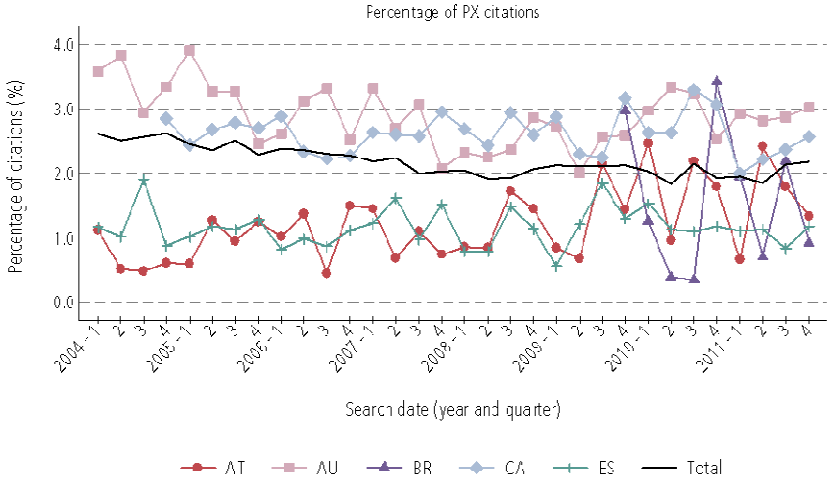


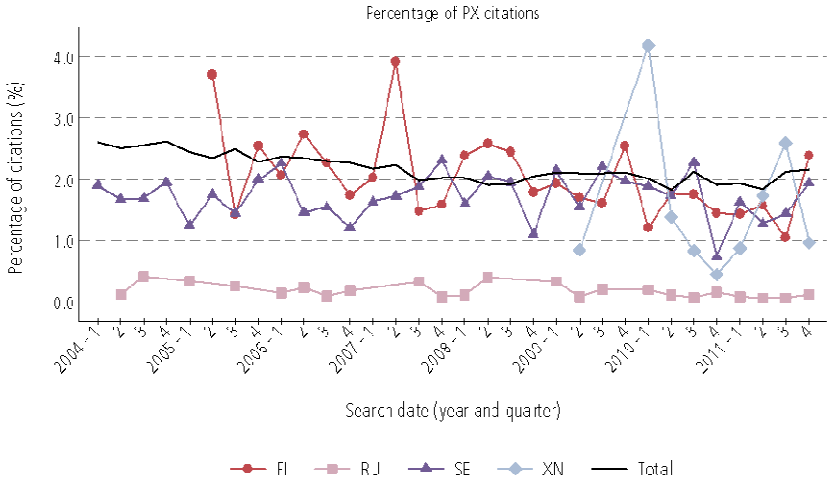




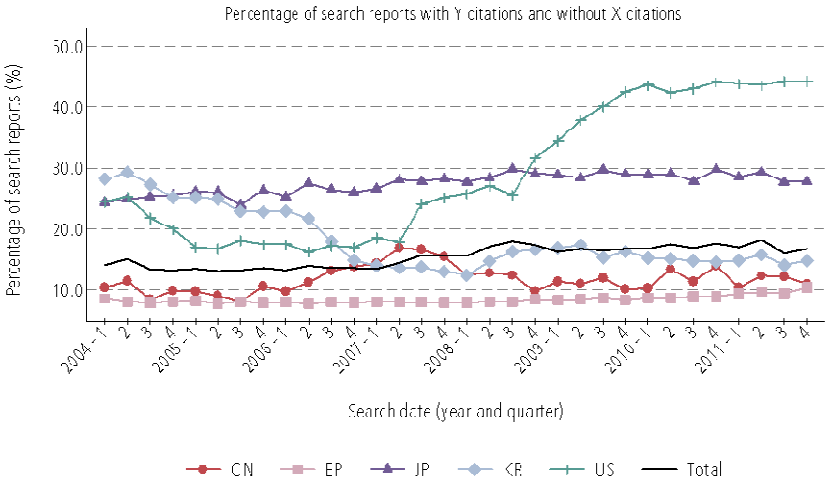
1.5 – PERCENTAGE OF CITATIONS IN THE CATEGORY OF P AND X

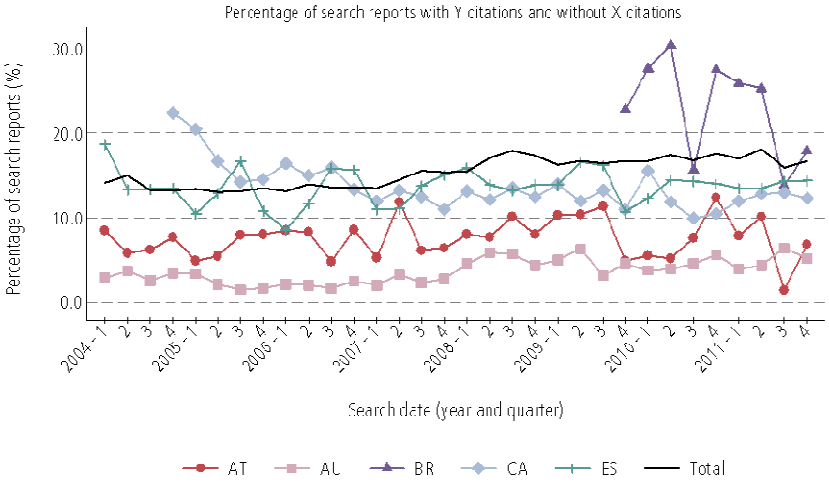


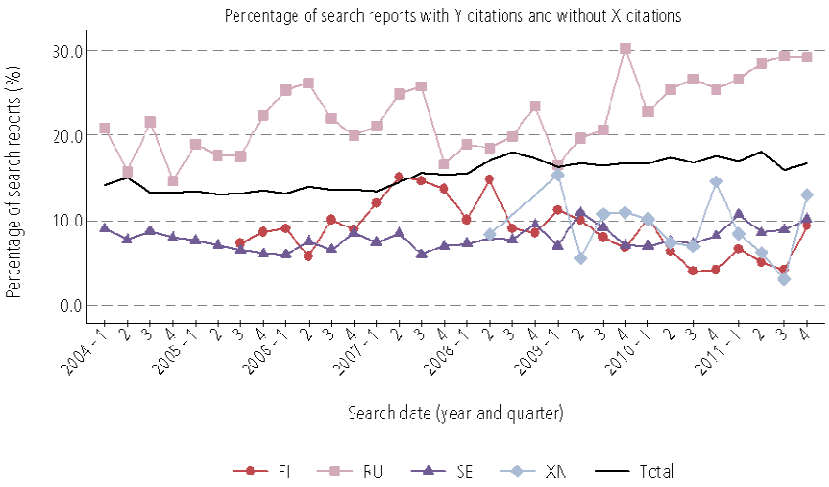




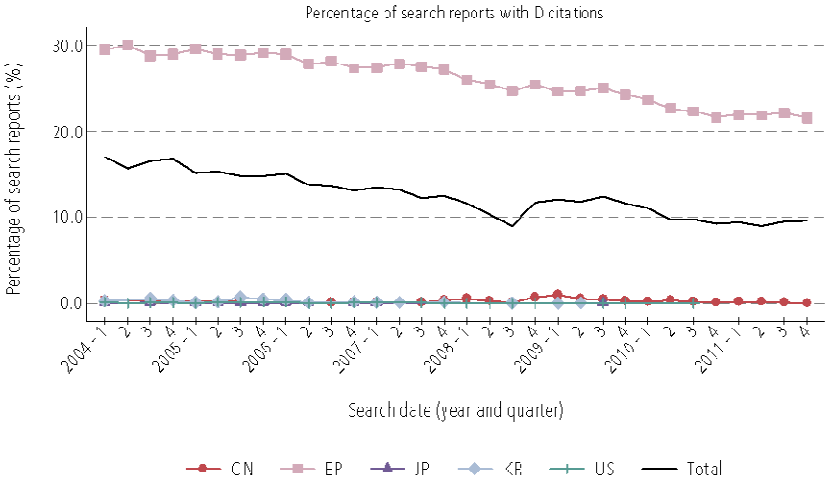
1.6 – PERCENTAGE OF PCT SEARCH REPORTS WITH Y CITATIONS AND WITHOUT X CITATIONS

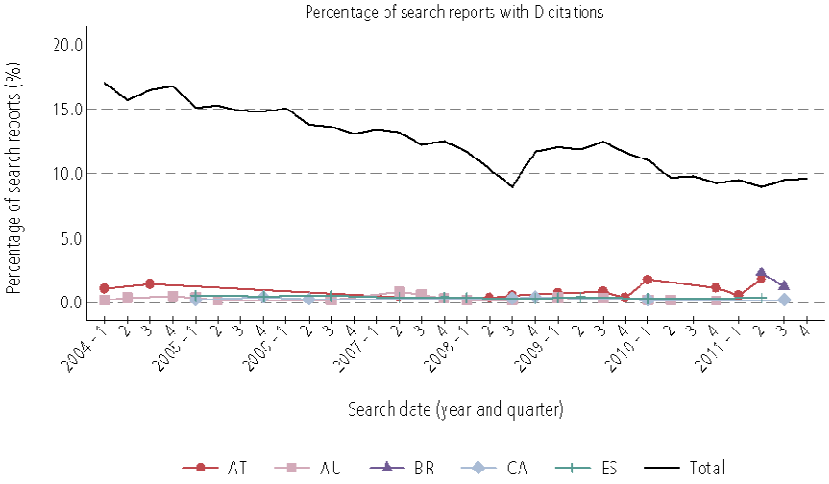


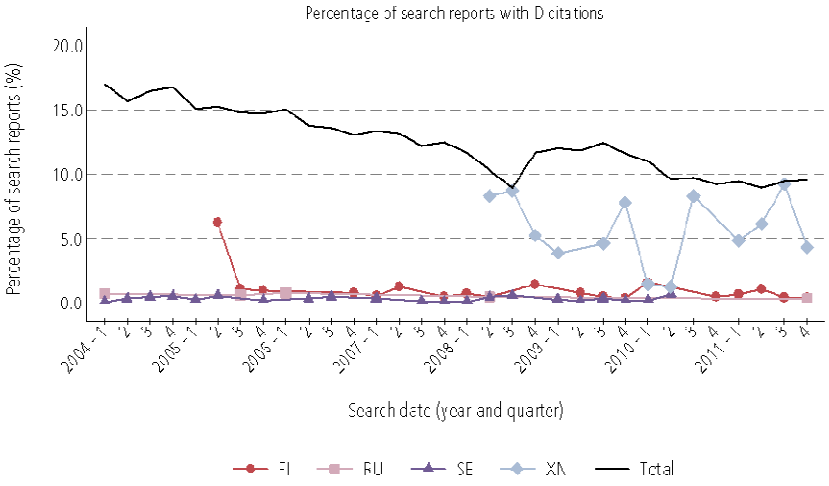




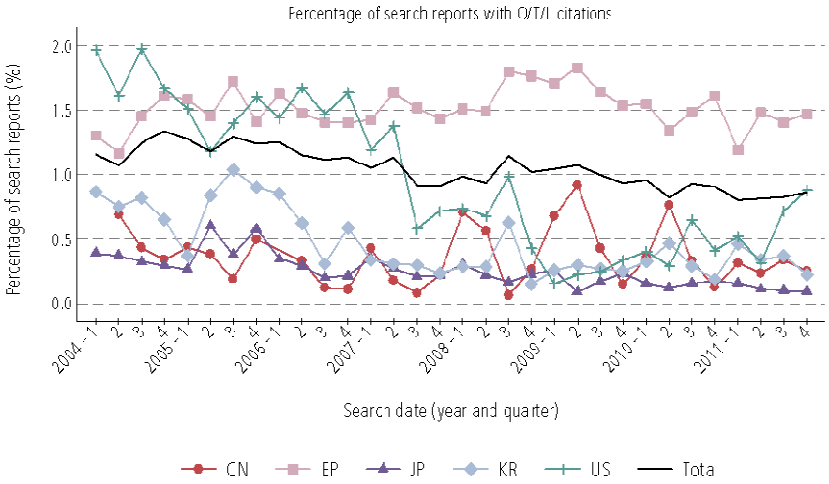
1.7 – PERCENTAGE OF PCT SEARCH REPORTS WITH D CITATIONS

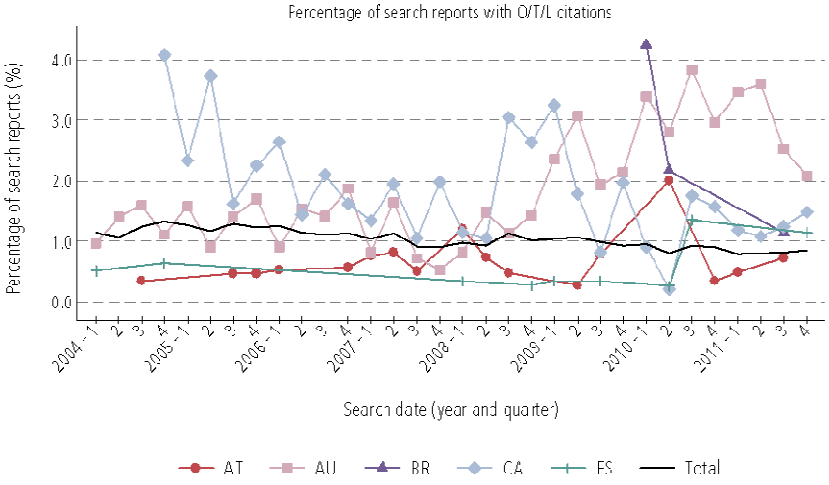


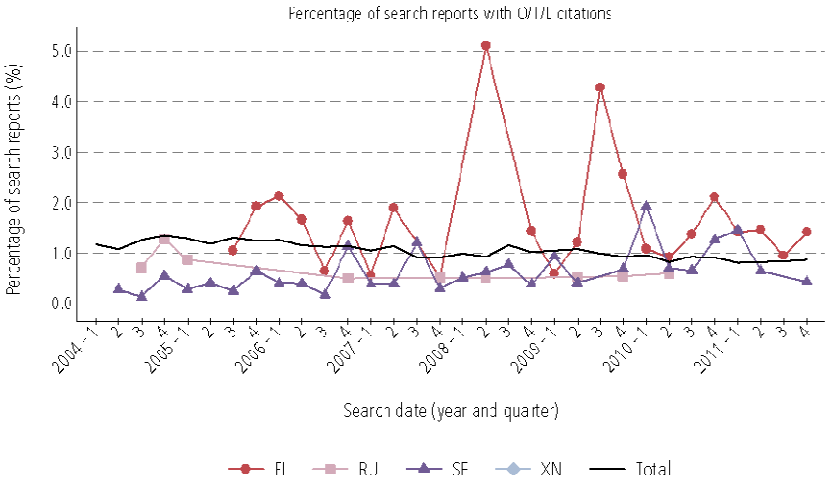




1.8 – PERCENTAGE OF PCT SEARCH REPORTS WITH O/T/L CITATIONS

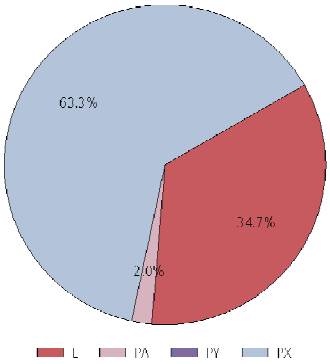




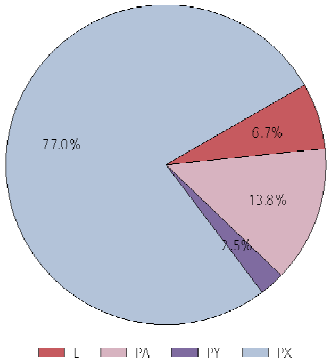


1.9 – P/E CITATION BREAKDOWN FOR PCT SEARCH REPORTS (2011)

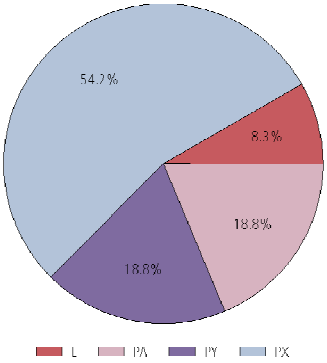
P/E citations breakdown (2011): ISA – AT



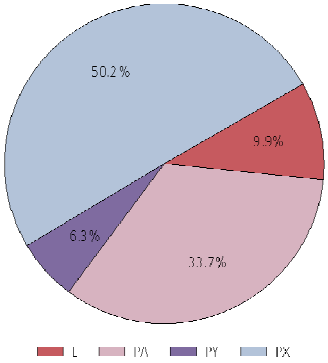
P/E citations breakdown (2011): ISA – AU



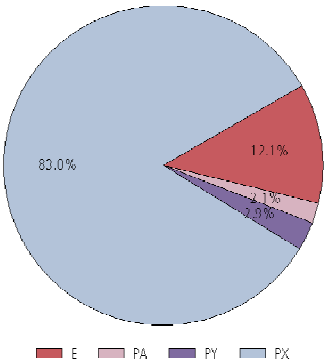
P/E citations breakdown (2011): ISA – BR



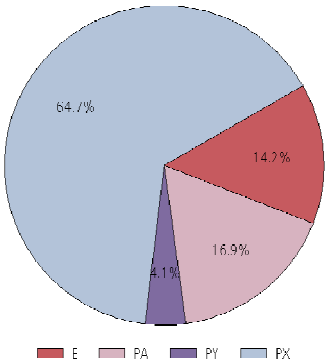
P/E citations breakdown (2011): ISA – CA



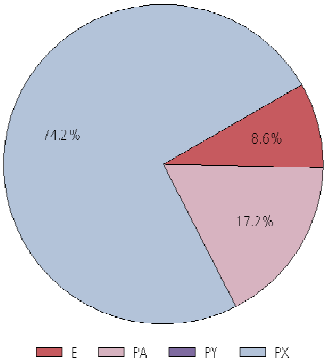
P/E citations breakdown (2011): ISA – CN



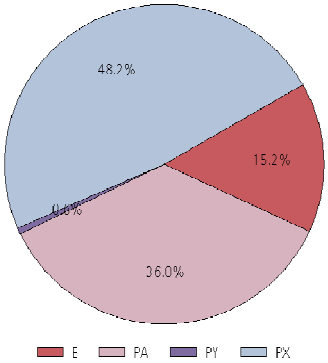
P/E citations breakdown (2011): ISA – EP



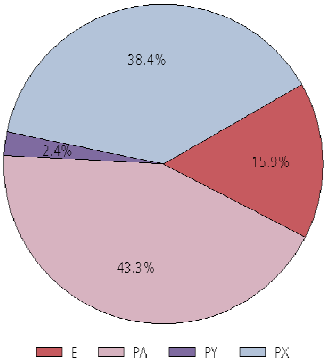
P/E citations breakdown (2011): ISA – ES



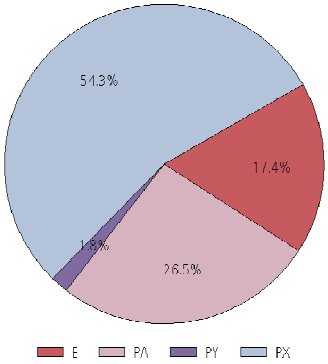
P/E citations breakdown (2011): ISA – FI



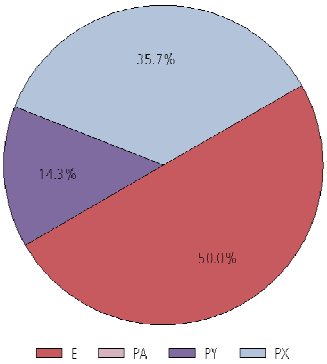
P/E citations breakdown (2011): ISA = JP



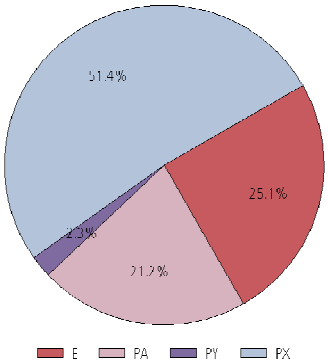
P/E citations breakdown (2011): ISA = KR



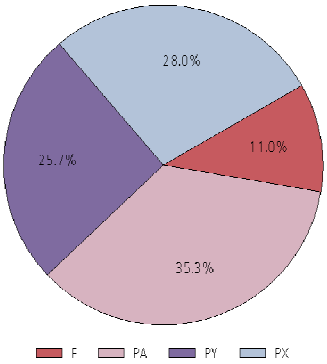
P/E citations breakdown (2011): ISA = RU



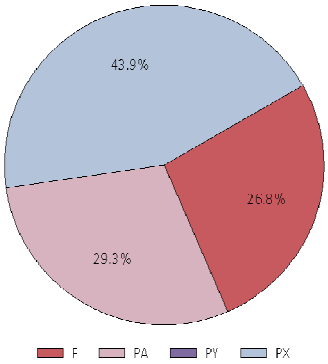
P/E citations breakdown (2011): ISA = SE



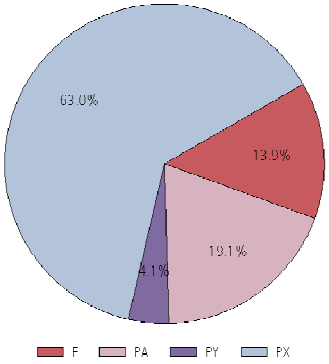
P/E citations breakdown (2011): ISA = US



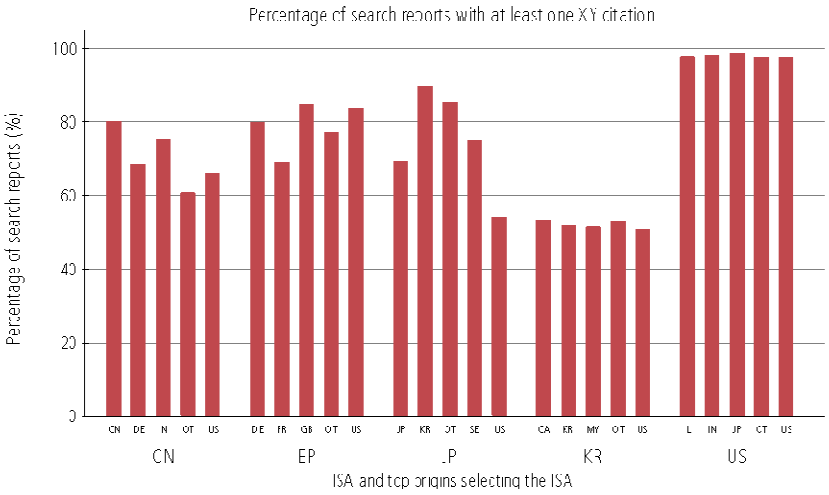
P/E citations breakdown (2011): ISA = XN

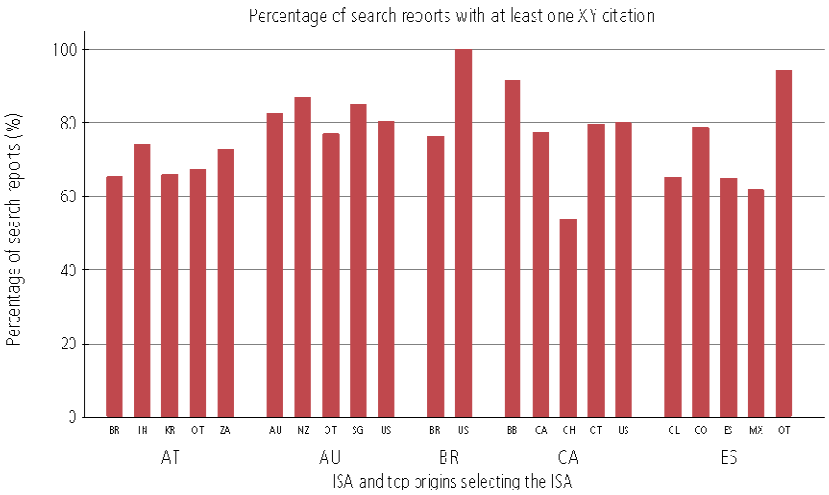


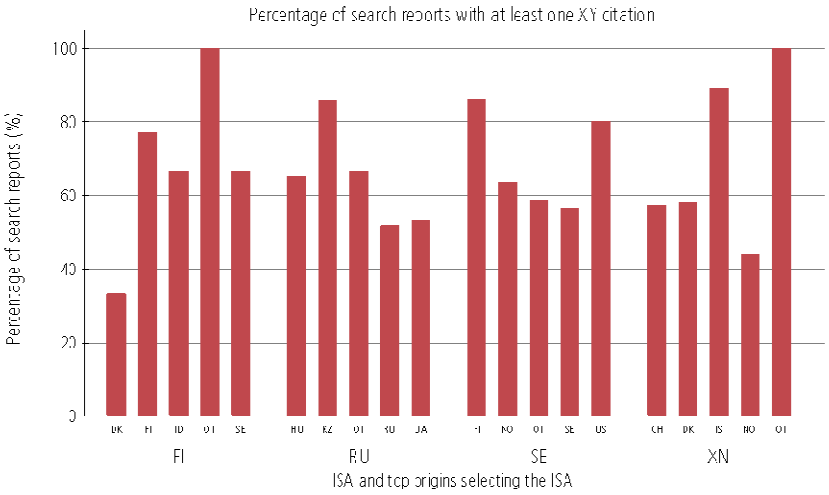
P/E citations breakdown (2011): Total



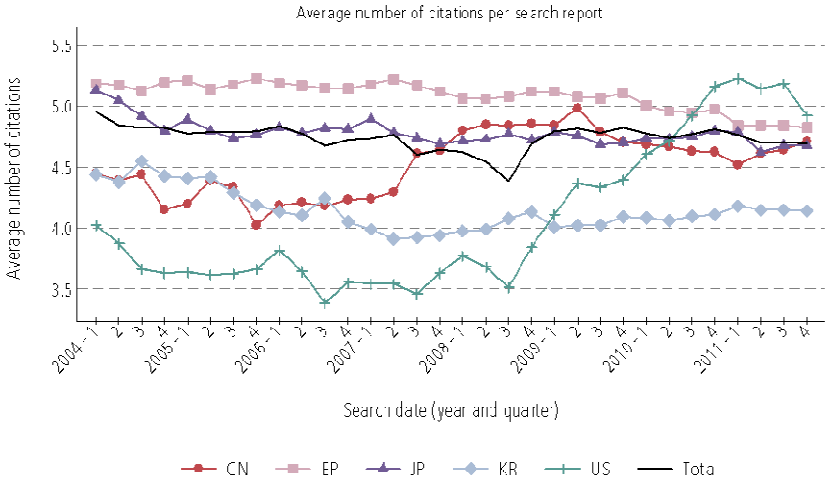
1.10 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X OR Y CITATION BY TOP APPLICANT'S ORIGIN (2011)

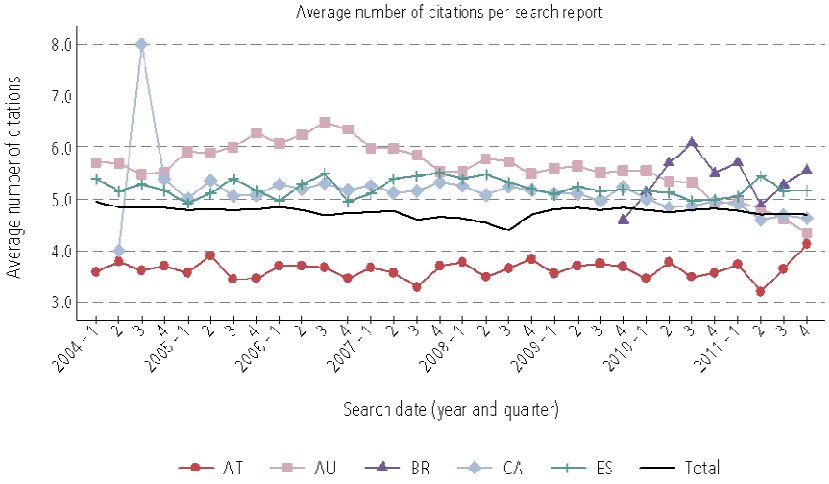


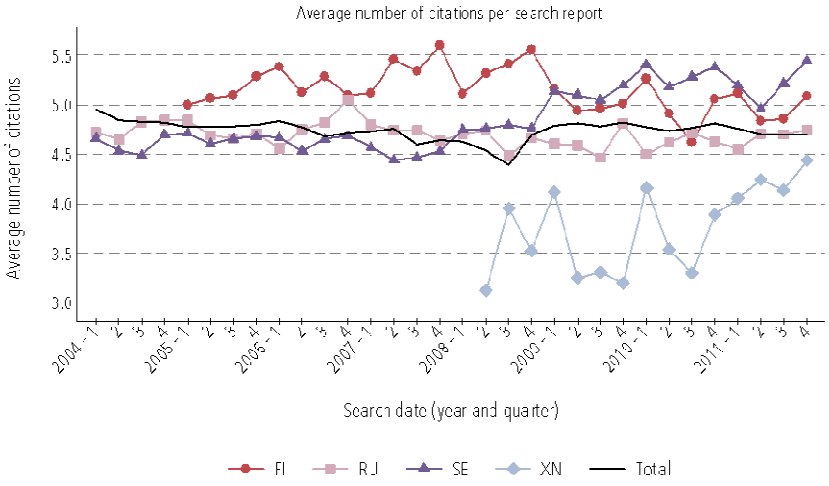




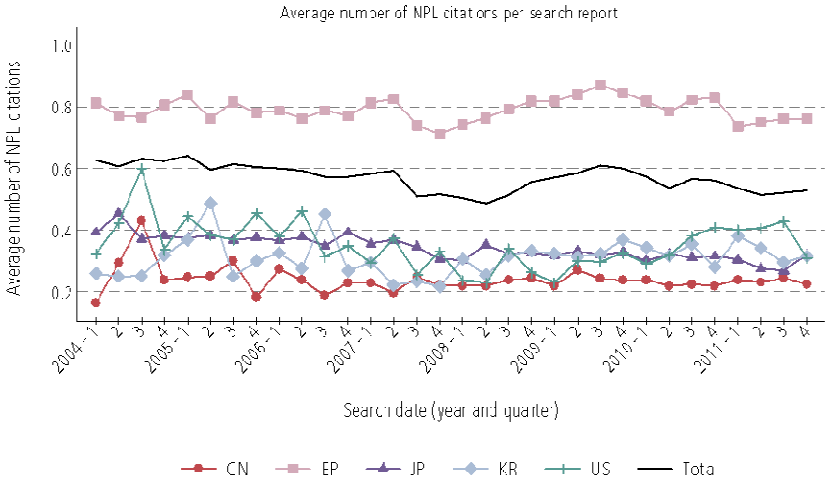
2.1 – AVERAGE NUMBER OF CITATIONS PER SEARCH REPORT

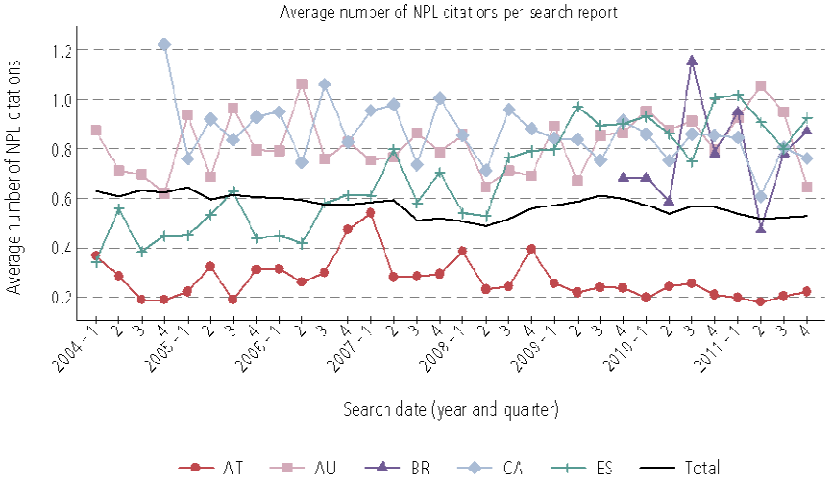


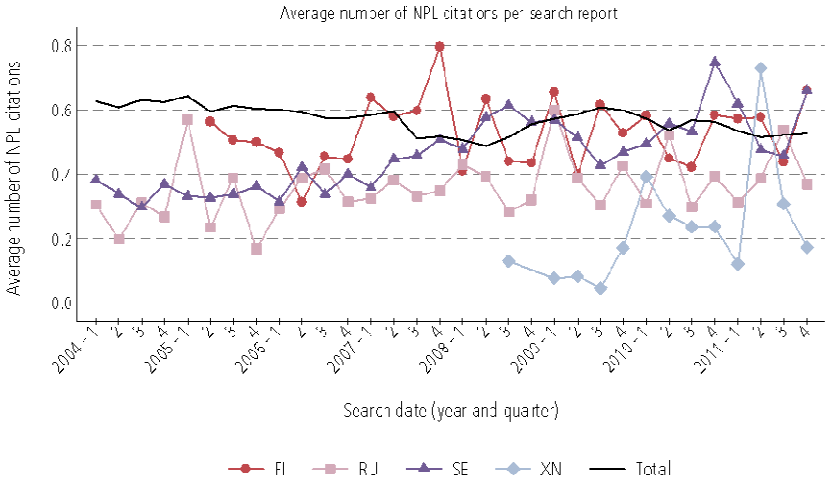




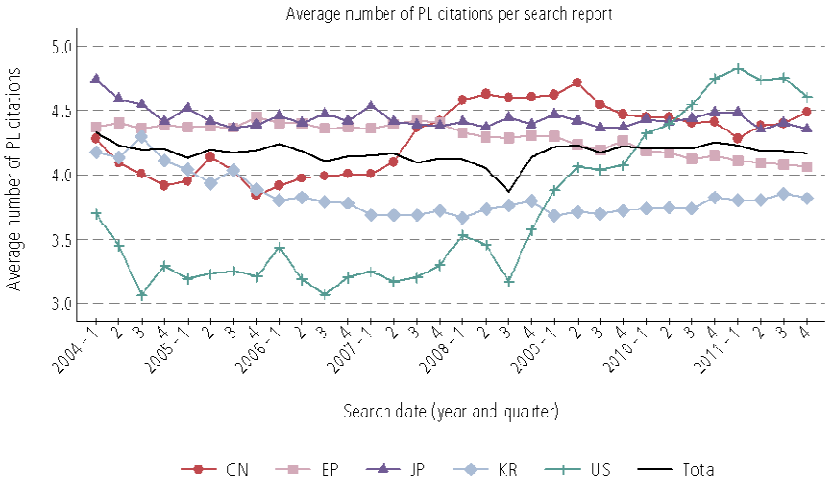
2.2 – AVERAGE NUMBER OF NPL CITATIONS PER SEARCH REPORT

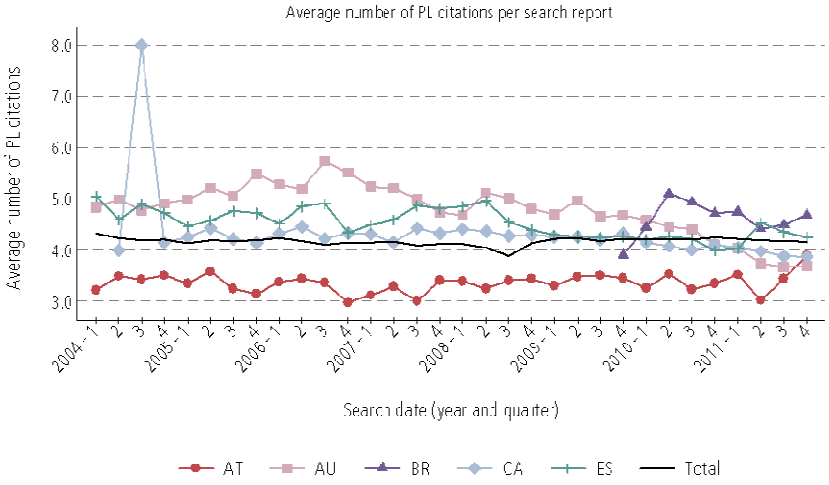


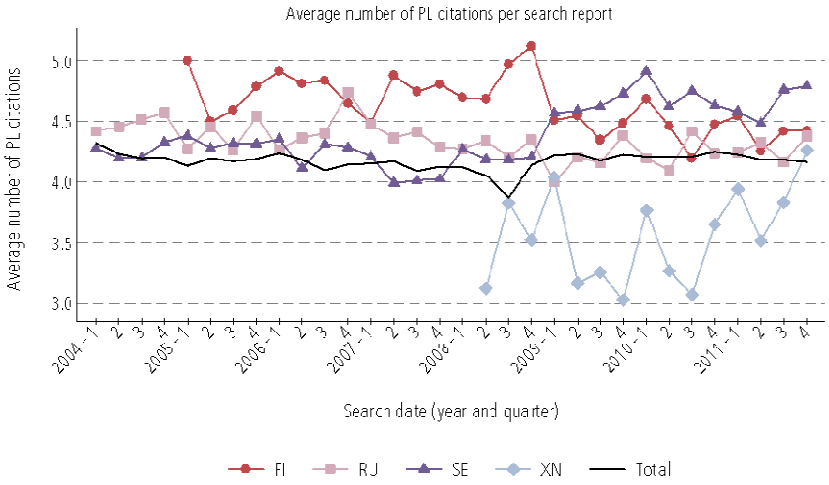




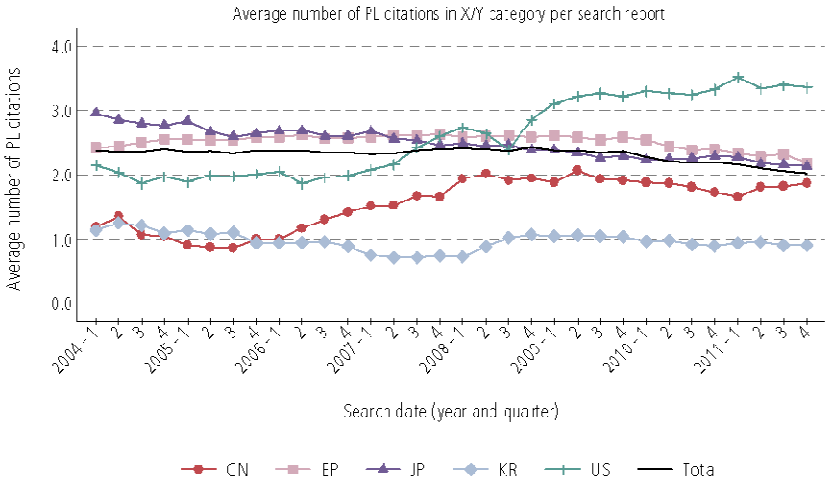
2.3 – AVERAGE NUMBER OF PATENT LITERATURE CITATIONS PER SEARCH REPORT

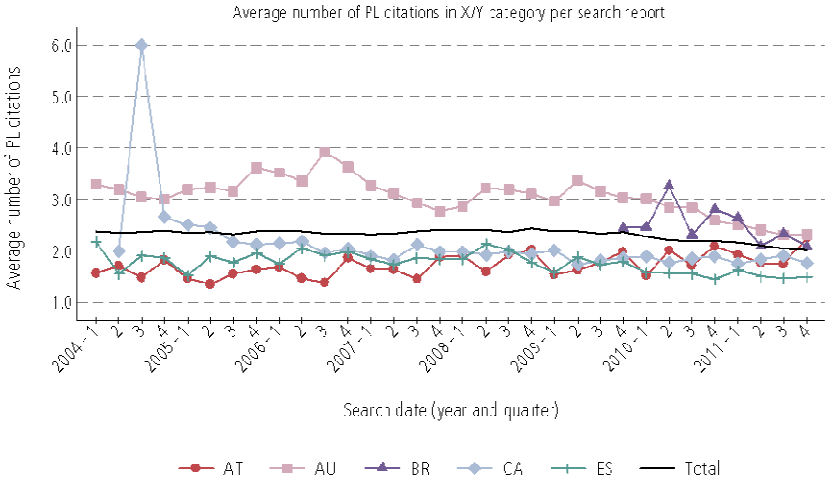


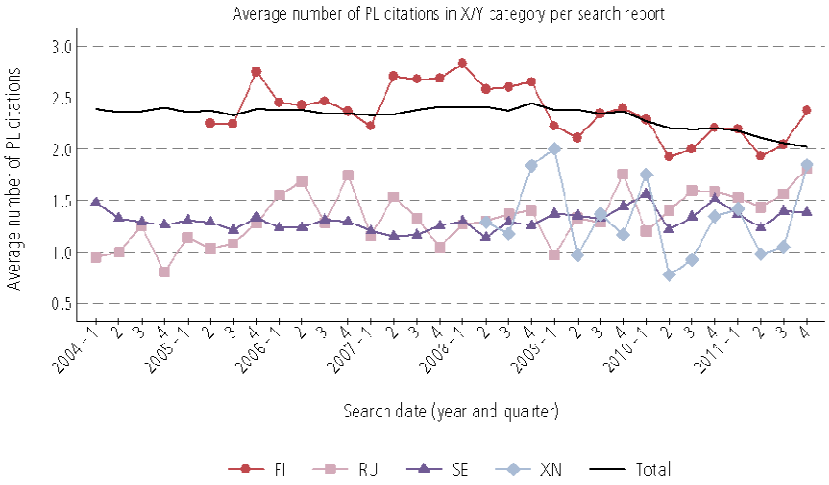




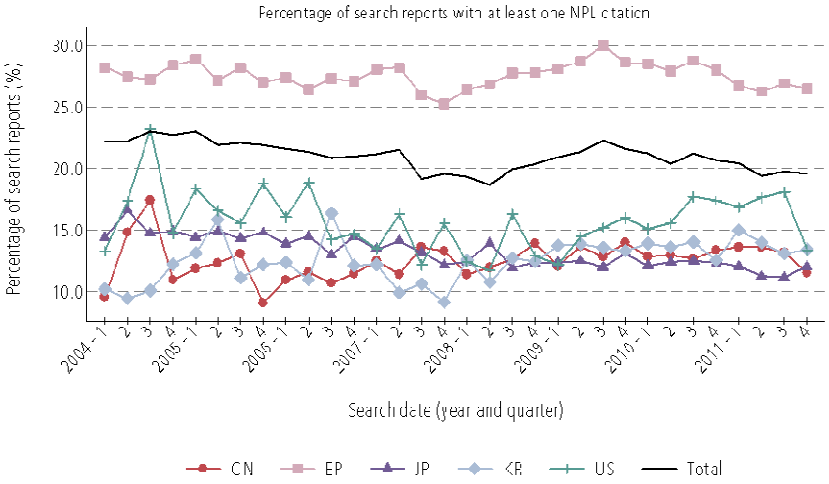
2.4 – AVERAGE NUMBER OF X/Y PATENT LITERATURE CITATIONS PER SEARCH REPORT

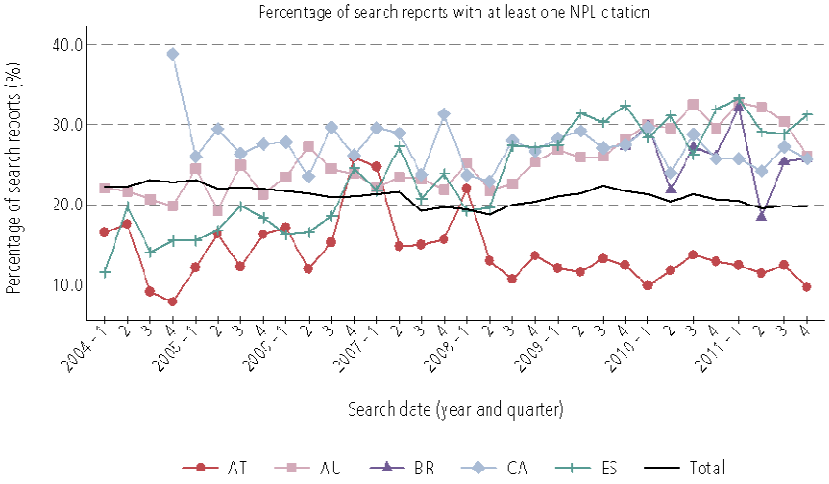


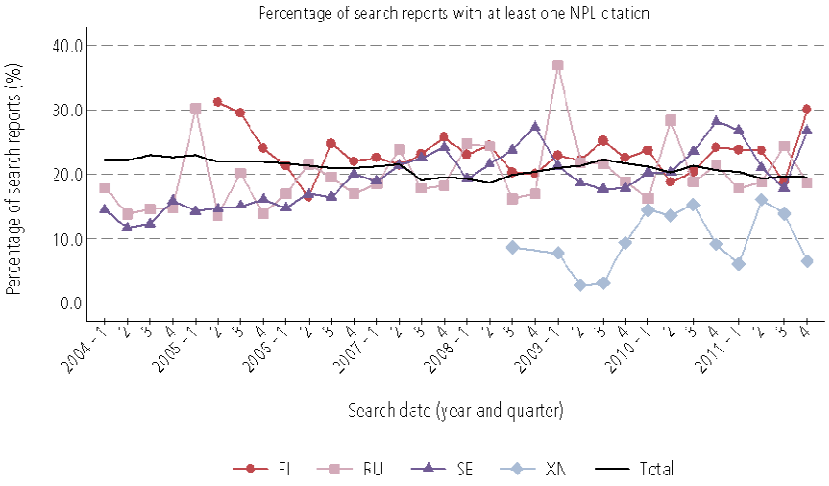




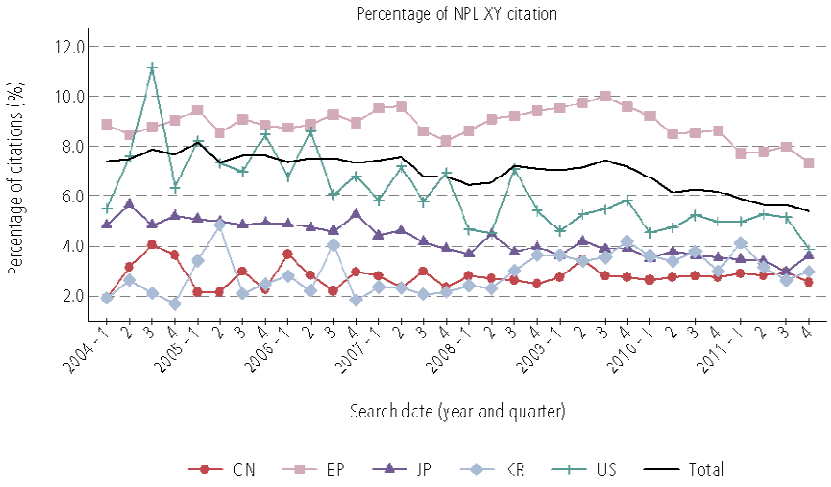
2.5 – PERCENTAGE OF NPL CITATIONS IN ALL CITATIONS

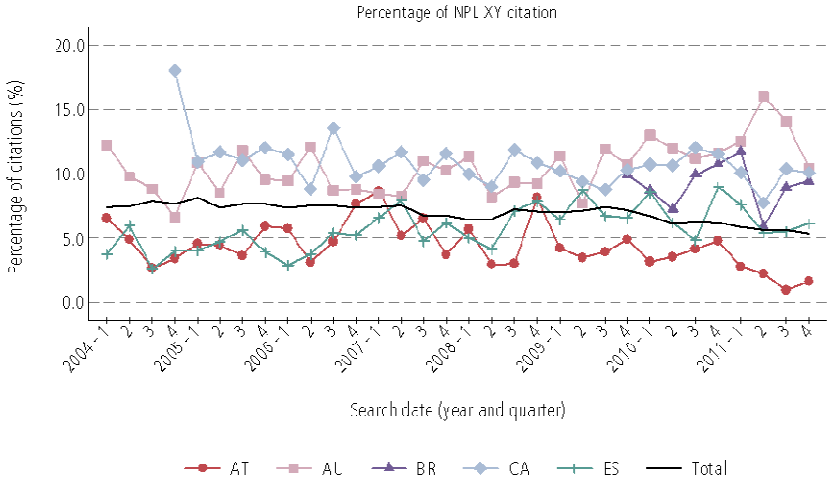


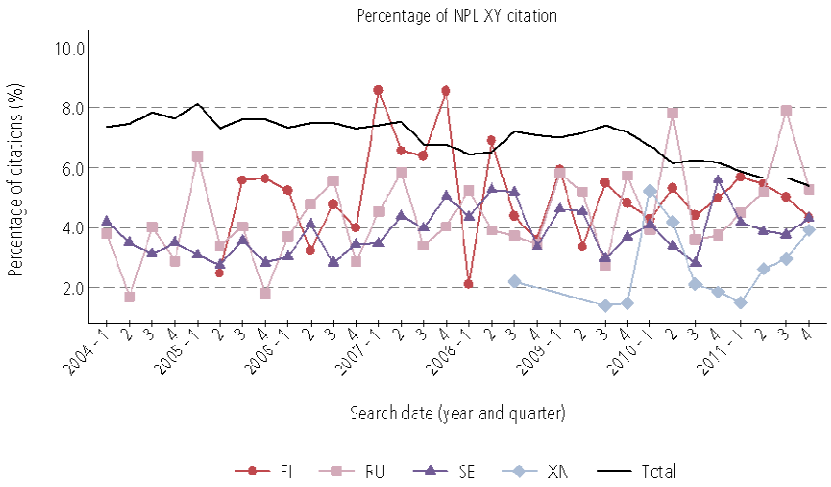




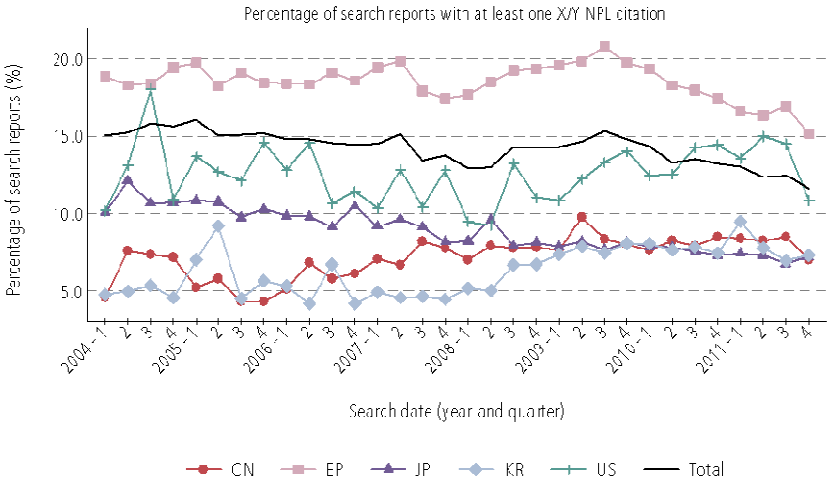
2.6 – PERCENTAGE OF NPL CITATIONS IN THE CATEGORY OF X OR Y IN ALL CITATIONS

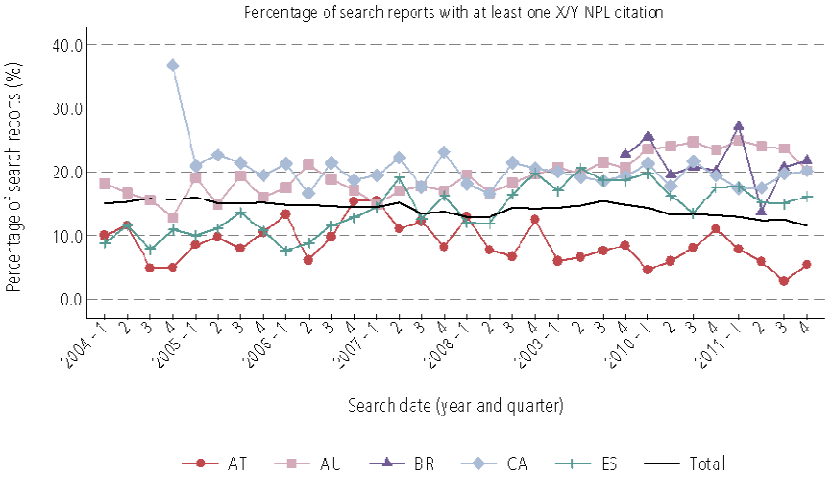


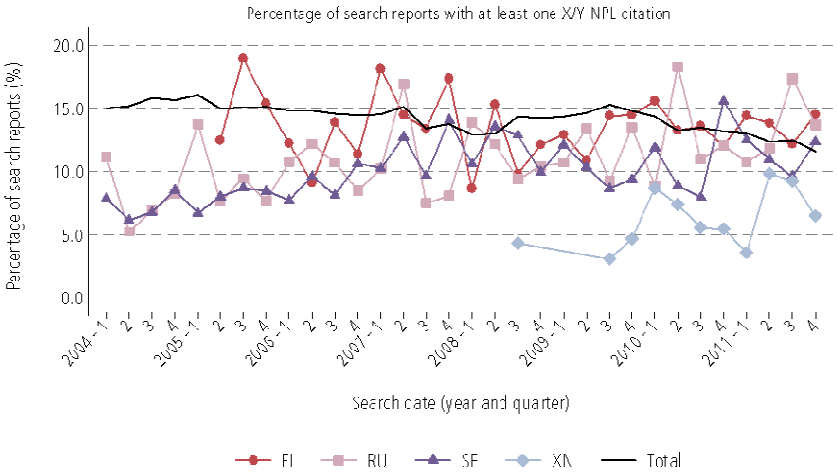




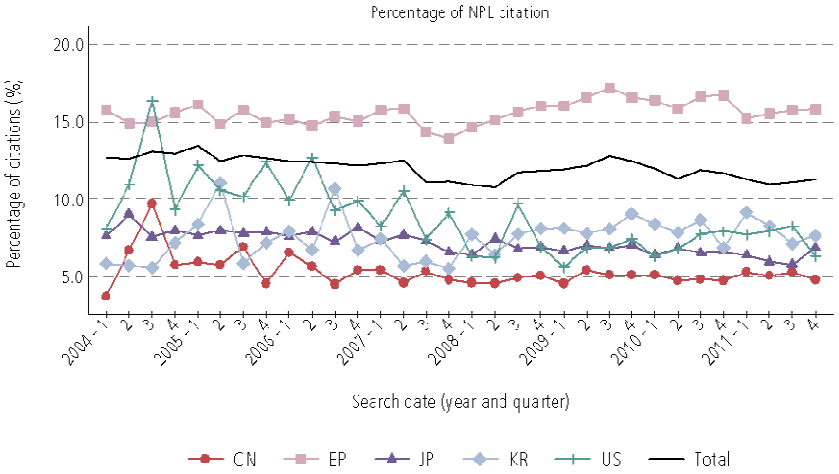
2.7 – PERCENTAGE OF PCT SEARCH REPORTS WITH AT LEAST ONE X/Y NPL CITATION

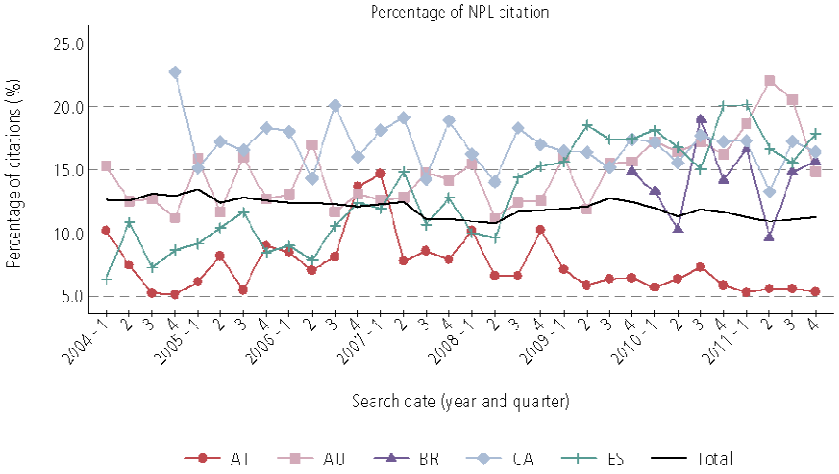


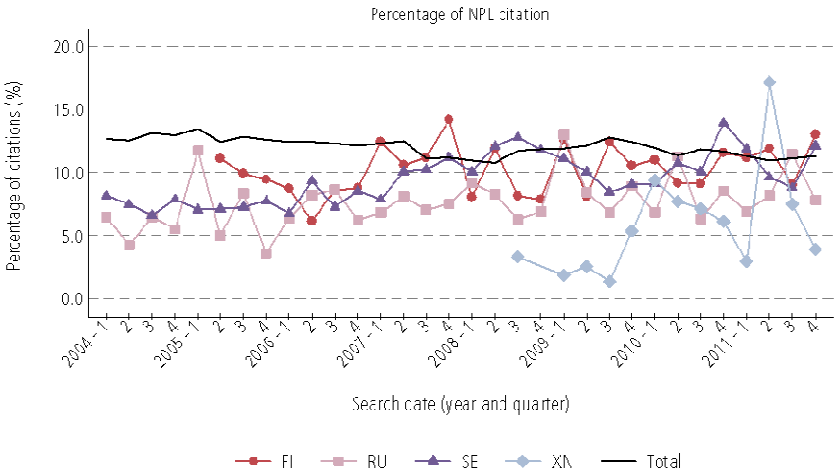




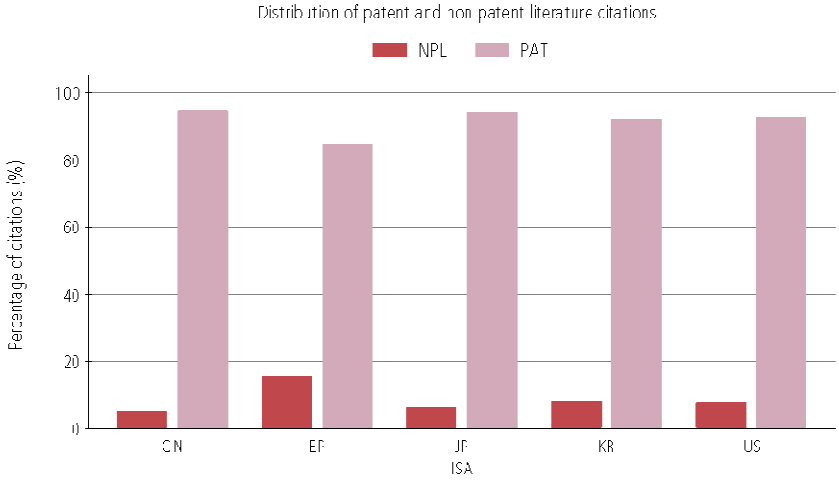
2.8 – PERCENTAGE OF NPL CITATIONS IN ALL CITATIONS

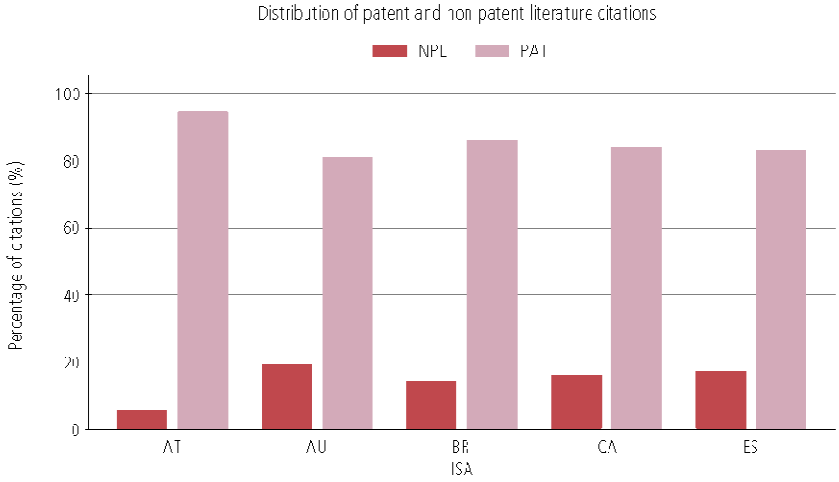


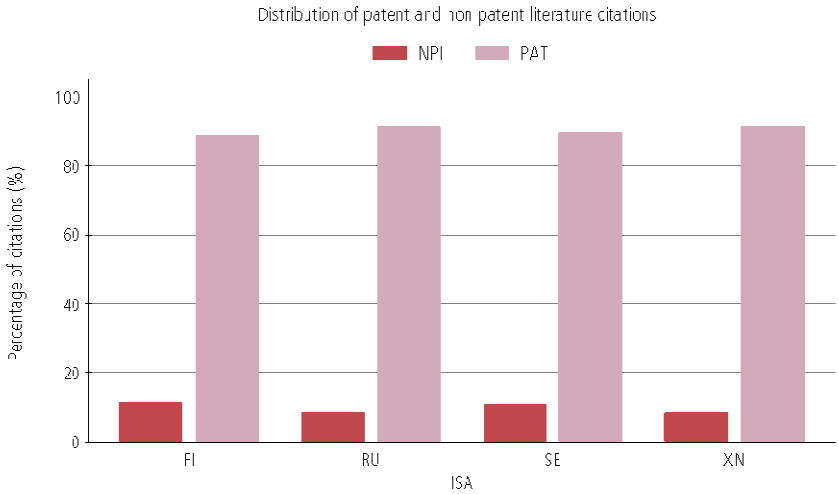




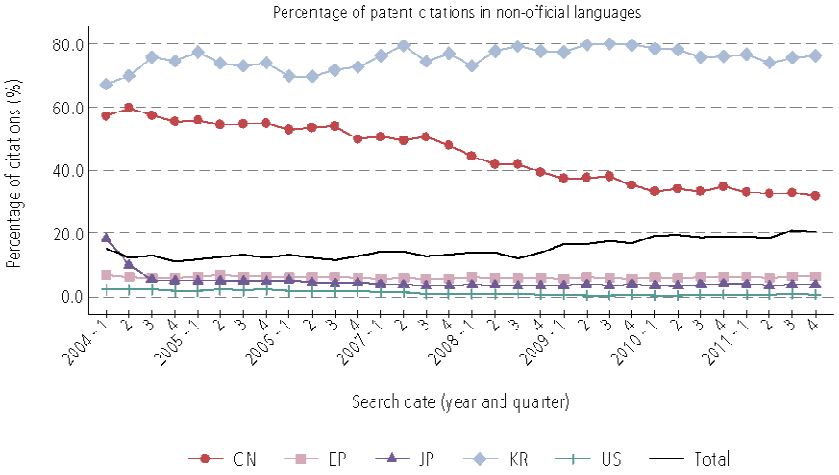
2.9 – DISTRIBUTION OF PATENT & NON PATENT LITERATURE CITATIONS (2011)

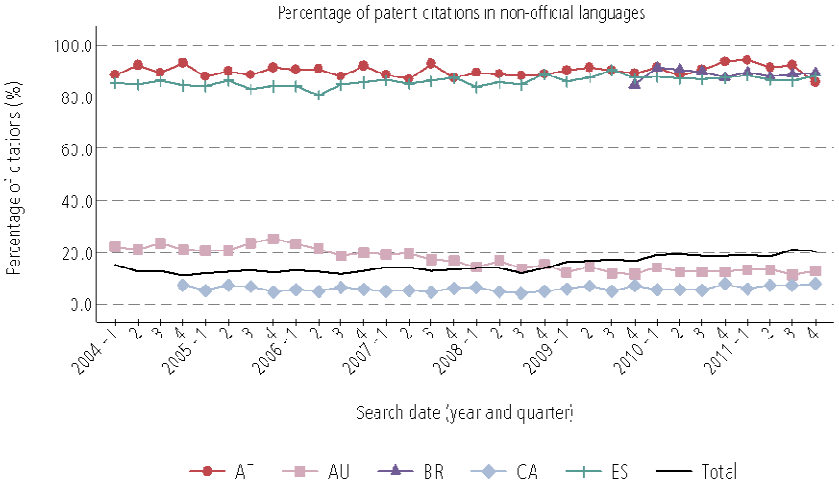


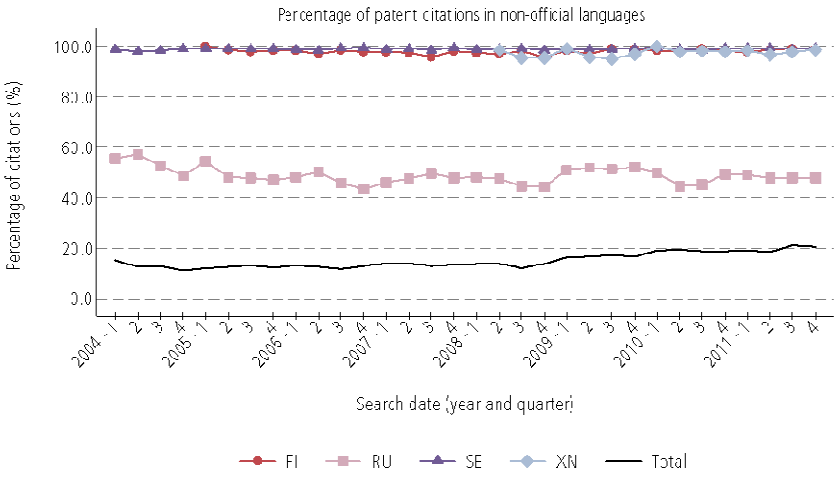




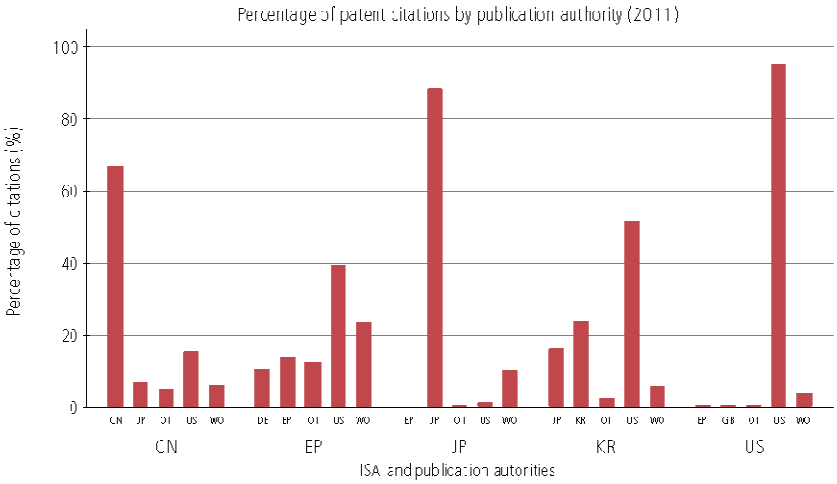
3.1 – PERCENTAGE OF PATENT CITATIONS IN NON-OFFICIAL LANGUAGES

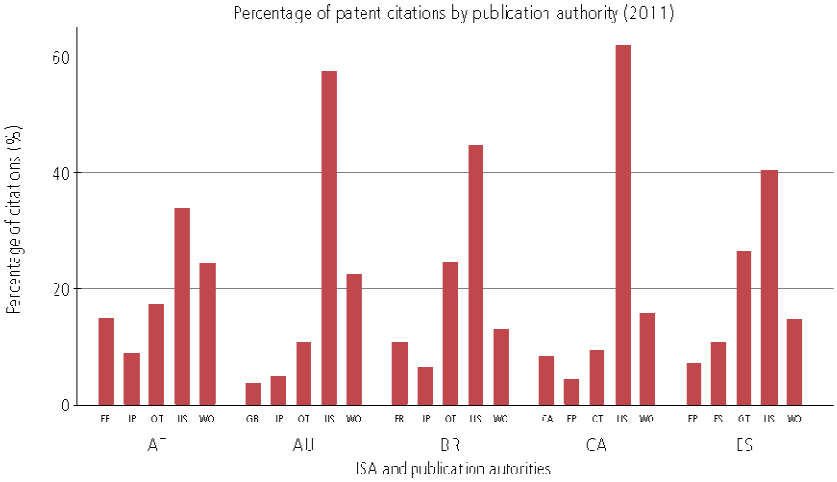


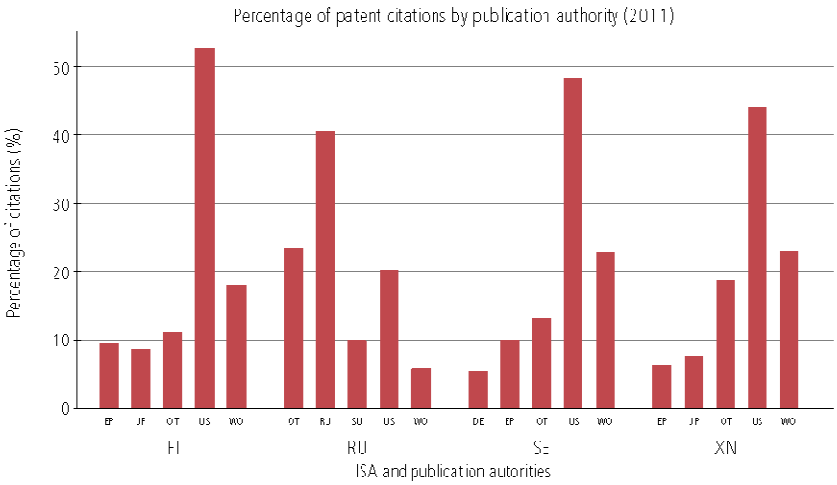




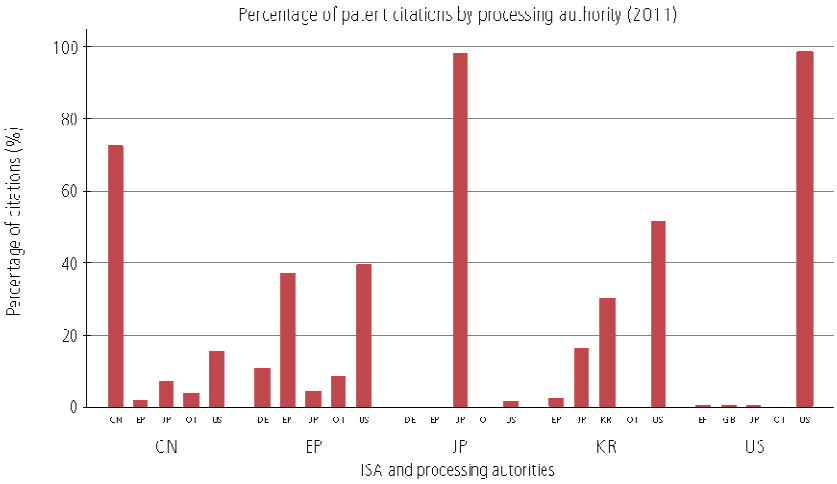
3.2 – PERCENTAGE OF PATENT CITATIONS BY TOP PUBLICATION AUTHORITIES (2011)

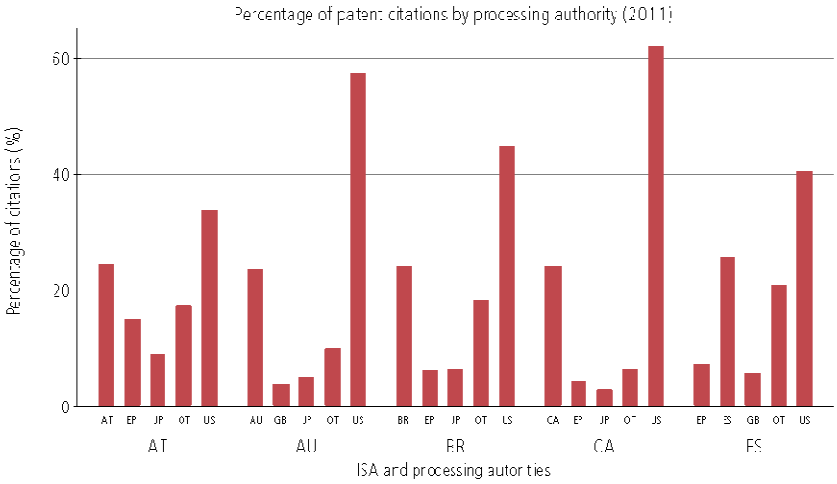


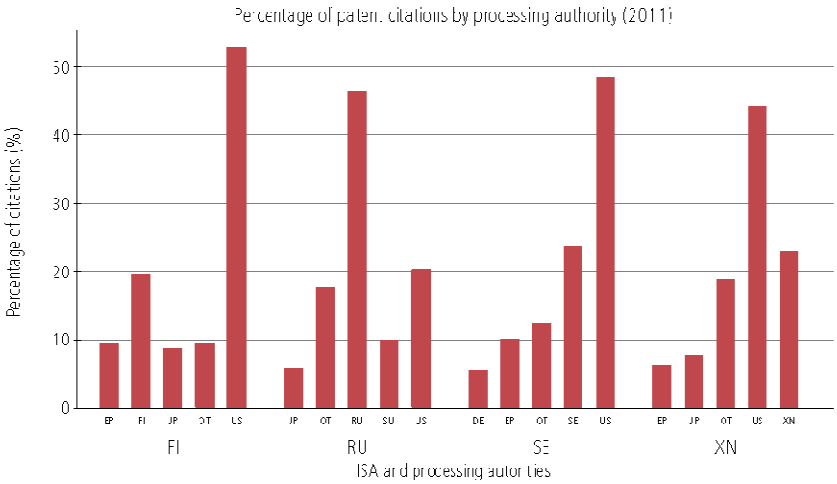




3.3 – PERCENTAGE OF PATENT CITATIONS BY PROCESSING AUTHORITIES (I.E. PUBLICATION AUTHORITIES EXCEPT WO) (2011)

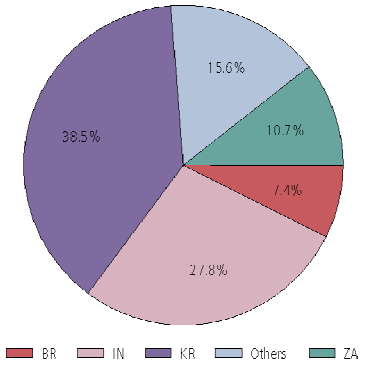




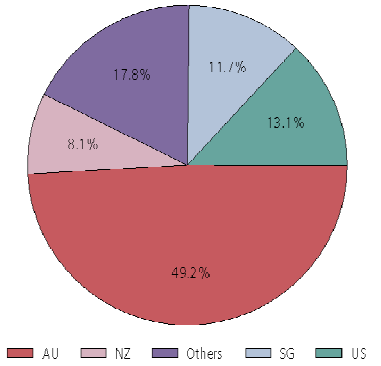


3.4 – PERCENTAGE OF PCT SEARCH REPORTS BY TOP APPLICANT'S ORIGIN (2011)

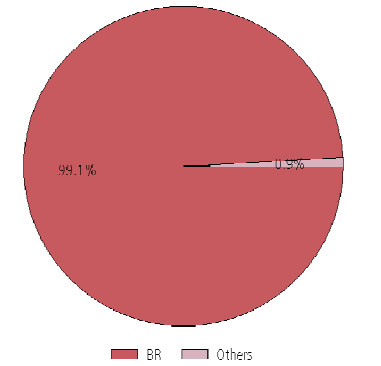
Percentage of search reports by applicants' origin (2011): ISA = AT



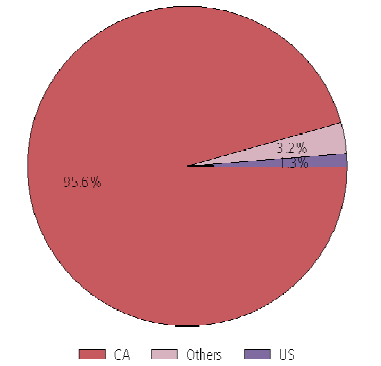
Percentage of search reports by applicants' origin (2011): ISA = AU



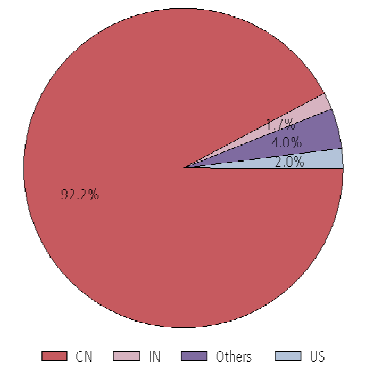
Percentage of search reports by applicants' origin (2011): ISA = BR



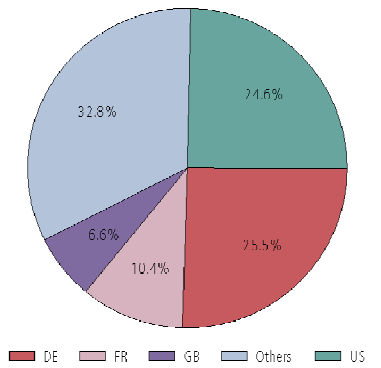
Percentage of search reports by applicants' origin (2011): ISA = CA



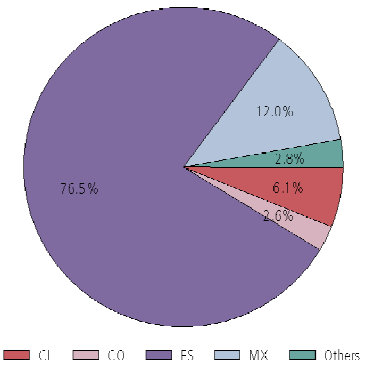
Percentage of search reports by applicants' origin (2011): ISA = CN



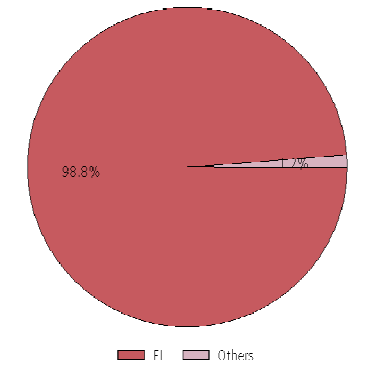
Percentage of search reports by applicants' origin (2011): ISA = EP



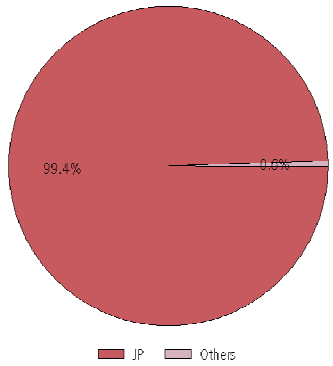
Percentage of search reports by applicants' origin (2011): ISA = ES



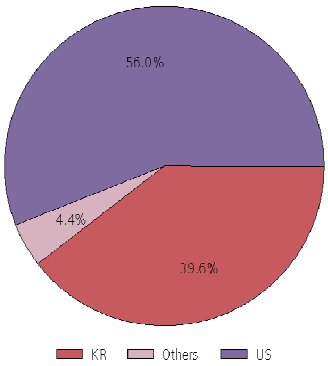
Percentage of search reports by applicants' origin (2011): ISA = FI



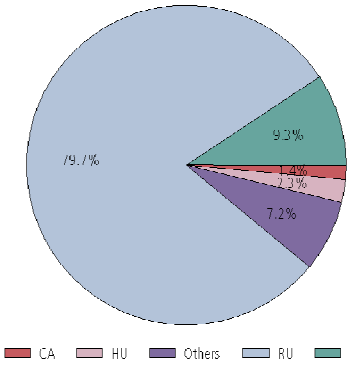
Percentage of search reports by applicants' origin (2011): ISA – JP



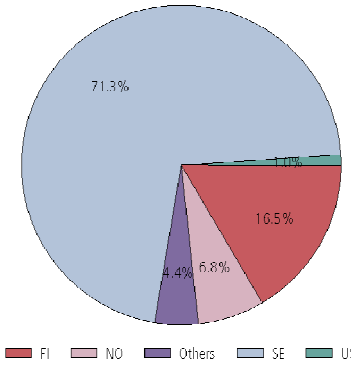
Percentage of search reports by applicants' origin (2011): ISA – KR



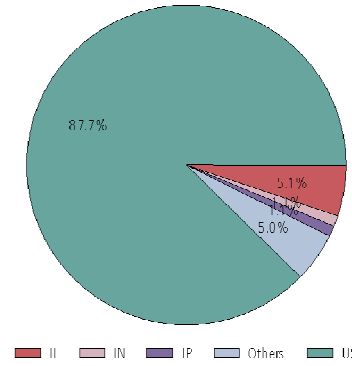
Percentage of search reports by applicants' origin (2011): ISA = RU



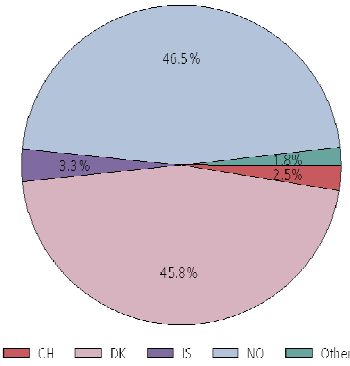
Percentage of search reports by applicants' origin (2011): ISA = SE



Percentage of search reports by applicants' origin (2011): ISA = US



Percentage of search reports by applicants' origin (2011): ISA = XN



[Annex II follows]

Descriptions and Definitions

DATA SET DESCRIPTION

- The data source is the European Patent Office's PATSTAT database for all citation information. Bibliographic information for international applications is taken mainly from the PATSTAT database, supplemented by information from WIPO internal databases where information could be provided which was not available from PATSTAT.
- The data provided is based on published PCT searches.
- Statistics are presented by search date up to 2011 Q4, meaning the date on which an international search report was transmitted to the International Bureau (since this information is available more consistently than the actual date of search).
- The date ranges for statistics take into account data availability. This is constrained by procedural latency such as time to publication, as well as cut-off dates for database extracts.
- No filing date constraint is applied.

DATA ISSUES

- Applications with no citation recorded are removed, as this generally means that no meaningful international search was carried out for these applications.
- A small number of patent citations are without category codes.
- In case of citn_origin = 5 (documents cited during international search), those citations are considered; otherwise, citations with citn_origin = 0 (documents cited during search) are considered. Citations with other citn_origin codes are removed.
- NPL citations with no category assigned and with ID >= 900000000 are removed, as they don't seem to be in the original search reports.
- All citation category codes recorded in the database for the valid citations are considered.
- Citation language codes for national patent documents are those recorded in the Patstat database, citation language codes for PCT documents are assigned using WIPO's PCT database as they are more reliable. The language codes are further cleaned up according to information of the authorities who publish those documents.
- No attempt has been made to determine the language of publication of non-patent literature documents.

DEFINITION OF CONCEPTS

Technology breakdown

- Technology sector and field are derived from the IPC classes assigned in the international phase search report or publication.
- The grouping into technology sector and field is based on a concordance provided by WIPO. (http://www.wipo.int/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf).

- This technology breakdown includes 8 technology sectors (Electrical engineering, Energy technology, Instruments, Mechanical engineering, Micro-structural and nano-technology, Other fields and Semiconductors), which are further broken down into 35 technology fields.

- Multiple IPC classes are often assigned to applications. For the present statistics, fractional counting method is applied, that is, an international application and all citations in its search report are evenly distributed to multiple technological fields when multiple fields are associated with it.

- IPC class information is not available for approximately 1% of applications.

Applicant origin

- In general this is the State in which the first-named applicant is resident (overall, this gives a more useful indication of origin of the application than the receiving Office because the International Bureau and regional Offices work for many States, whereas some States do not themselves operate a receiving Office).

- "Unknown" code is used for a small percentage of applications.

XY rate (Searches with XY citations)

- XY rate refers to share of search reports where at least one citation is in the category of X or Y.

- In addition the use of an E citation is counted as XY if it can be assumed that the E citation is prejudicial to novelty. This is the case unless the E category is assigned in combination with A.

Citation category availability

- PATSTAT does not contain all citation categories for each citation. The database contains one citation category per group of categories for each citation. The category groups are defined as follows:

Group 1	X Y A
Group 2	P E
Group 3	D
Group 4	O T L

- Only one category from the same group is selected. The category selected is determined from the order in the table above. In this way a citation will be categorized as X if the citation categories in the search report are XY for this citation. Priority for selecting the letter is according to the ranking of categories left to right within the groups above, rather than the order of their appearance within the citation in the international search report (that is, X will be shown even if the search report lists Y category claims first).

- A maximum of 3 categories is recorded.

- Citation Category Examples:

Search report citation	*Citation categories present in PATSTAT*
X, Y, A, P	X, P
Y, A, P, E	Y, P
Y, X, O, T	X, O
X, P, D, O	X, P, D

- This means for example in row 1 above the Y nor the A is not stored in PATSTAT.
- In practice it is therefore possible to determine whether a search has at least one X or Y citation. It is also possible to correctly count the number of X citations.
- In approximately 20% of cases it is not possible to correctly count the number of Y categories used, although it is possible to count the use of Y without an X.
- EPO data has been refined with an additional internal data source.

A only rate

- A-only rate refers to the share of search reports where no citation is in the category of X, Y or E.

Y no X rate

- Y no X rate refers to the share of search reports where at least one citation is in the category of Y and there is no X citation.

Search date

- The date when the search report is transmitted to WIPO (the actual date of search is not available in all cases).

Patent Literature/Non-Patent Literature

- Citations in PATSTAT are categorized into patent literature and non-patent literature.
- A citation is considered patent literature if it relates to patent abstracts provided by various providers.
- Less information is available for NPL citations. For example, the language of a NPL citation is not available.

Non official language

- This is used for counting patent citations that are not in an official language of the respective ISA:

ISA	*Official language*
AT	German
AU	English
BR	Portuguese
CA	English
CA	French
CN	Chinese
EP	German
EP	English
EP	French
ES	Spanish
FI	Finnish
JP	Japanese
KR	Korean
RU	Russian
SE	Swedish
US	English
XN	Danish
XN	Icelandic
XN	Norwegian

- The statistics are based on the actual official languages of the Office, but can easily be redefined to reflect any set of core languages which an Offices considers to be useful in assessing how effective its processes may be at discovering prior art beyond those languages.

Publication Authority (of citation)

- This is the patent organization who published a citation document.

- It is normally a national patent office, a regional office such as the EPO, or WIPO.

Processing Authority (of citation)

- Generally processing authority is assigned from the publication authority of the citation.

- For WO publications, the international search authority is chosen to indicate which office processed the cited patent publication. This gives an indication of the nature of the publication which will be more useful for some purposes than simply the number of WO citations, which may be in any of 10 languages.

[End of Annex II and of Circular]