BSH Background Paper #2 – Part One

"Thematic patterns of cross-border S&T cooperation based on co-publication analysis"

Authors: Hanna Scheck, Alexander Degelsegger, Carmen Heidenwolf, Katharina Büsel, Philipp Brugner, Johannes Simon, Judith Feichtinger, Ines Marinkovic (ed.)

Project Acronym:	BSH
Project full title:	BLACK SEA HORIZON
Project No:	H2020-INT-INCO-2014
Funding Scheme:	Horizon 2020
Coordinator:	ZSI GmbH – Centre for Social Innovation
Project start date:	February 1, 2015
Project duration:	36 months

Abstract	The study underlying this deliverable (D1.2) scrutinises the co-publications of Black Sea (BS) countries for the years of 2003-2013. The actual analysis of the publication data was carried out along several dimensions: overall co-publication output numbers per country to provide an overview, the internationalisation of publications, main scientific research fields, and finally some highlights regarding scientific impact. The target group of this study involves all interested relevant stakeholders in the field of international STI cooperation from the EU and the non-EU BS countries.

BLACK SEA HORIZON POLICY BRIEFS	The BSH Policy Briefs and related Background Papers are available for download from the Black Sea Horizon Website: https://blacksea-horizon.eu/policy-dialogue/briefs **Policy Brief #1:** "Obstacles, drivers and opportunities to enhance EU-Black Sea STI cooperation" – **BSH Policy Paper #1** (D1.1) "Obstacles, drivers and opportunities to enhance EU-Black Sea STI cooperation – **BSH Background Paper #1**" (D1.1) **Policy Brief #2:** "Thematic patterns of cross-border S&T cooperation based on co-publication and co-patent analysis" – **BSH Policy Paper #2** (D1.2) "Thematic patterns of cross-border S&T cooperation based on co-publication analysis" – **BSH Background Paper Part One"** (D1.2) "Thematic patterns of cross-border S&T cooperation based on co-patent analysis" – **BSH Background Paper Part Two"** (D1.2) For further information please visit https://blacksea-horizon.eu/ or contact office@blacksea-horizon.eu.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AC – Associated Countries to Horizon 2020 (Iceland, Norway, Albania, Bosnia and Herzegovina, the former Yugoslav Republic of Macedonia, Montenegro, Serbia, Turkey, Israel, Moldova, Switzerland, Faroe Islands, Ukraine)

BS – Black Sea

BSC – Black Sea Countries

BSH – Black Sea Horizon

EU – European Union

RTDI – Research, Technology, Development and Innovation

STI – Science, Technology and Innovation

EXECUTIVE SUMMARY

The study underlying this deliverable scrutinises the co-publications of Black Sea (BS) countries for the years of 2003-2013.

The methodology of the co-publication analysis is based on the two main academic citation databases, namely Web of Science and Scopus. The unification of the data of those two sources is fairly unique and entails a number of complex operations to ensure the quality and compatibility of the data, which goes hand in hand with a considerable data normalisation effort. The gain achieved by this unification both in terms of quantity (a roughly 25 % higher publication coverage) and quality (each data source can be taken as a quality check for the other or as an additional source for missing information) is well worth the effort.

The actual analysis of the publication data was carried out along several dimensions: overall co-publication output numbers per country to provide an overview, the internationalisation of publications, main scientific research fields, and finally some highlights regarding scientific impact.

The main outcomes of the analysis will inform the "Policy brief on thematic patterns of cross-border S&T cooperation based on co-publication and co-patent analysis".

1. Introduction

1.1. Short description of the EU project Black Sea Horizon (aims)

The BLACK SEA HORIZON (BSH) EU project within H2020, which started in February 2015 with a duration of 3 years, aims to support the EU's external relations with the Black Sea region by significantly contributing to ongoing bi-regional and regional Science, Technology and Innovation (STI) policy dialogues, and by increasing the knowledge base about the EU's external environment. It also tries to stimulate bi-regional STI cooperation and to strengthen the EU's economic competitiveness as well as to contribute to the establishment of supportive framework conditions by facilitating the pooling of resources and by identifying challenging thematic areas for mutual STI cooperation.

1.2. Methodology

This chapter describes the main processes involved in the study underlying this deliverable.

This report uses the term BSCs (Black Sea Countries) when referring to the non EU-countries Armenia, Azerbaijan, Georgia, Moldova, Russia and Turkey. This definition takes also into consideration the Easter Partnership countries Armenia, Azerbaijan and Moldova that are not directly riparian Black Sea countries, but can be considered as part of the wider BS-region and that are represented with partners in the BSH project.

The analysis of Black Sea co-publication output in the years 2003 to 2013 is based on the two best known and most comprehensive multidisciplinary academic citation data bases:

- Elsevier's *Scopus*
- Thomson Reuter's *Web of Science* (short: WoS; at present containing the following databases: Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index)

To understand and adequately interpret the results presented in this deliverable, a few basic terms need to be defined. More details are provided in the Key definitions for co-publication analysis (see 41 page and following).

An 'affiliation' links an author to her/his institution(s). As these can be more than one and also located in different BS countries, several affiliations are counted and also included as international co-publications. The analysis uses 'categories' and 'main categories' that are basically thematic keywords to classify the scientific literature. Those categories are based on the Science Metrix Ontology for journal classification.

The study consolidates different document types from the databases and used articles, conference papers, meeting abstracts, reviews, editorials, letters, and others to describe the units of analysis. As a specific sub-chapter of the report deals with the 'impact' of publications, readers should keep in mind that the impact given below is just a snapshot: since there is a lag between the publication of a work and the occurrence of references to it in later works, the most recent works will typically show no or few citations.

With the term 'record' we refer to an entry in our database containing the meta-data of a uniquely identified publication. So, as soon as the same publication is identified in both data sources, it is treated as one record.

The study analyses all publications from both databases that featured any affiliation to one of the countries of the BS region in order to be able to draw conclusions on the differences between "all publications" and the "co-publications". The exceptions are Russia and Turkey, in whose case the sheer number of publications would have by far exhausted the resources available for this study; consequently, only Russia's and Turkey's co-publications with EU/ACs countries are covered in this report. And to get a comparative overview of the total publication output we used data from Scimago which are based on data from Scopus.

The study does not make any discrimination regarding document types, meaning that scientific articles are taken into account the same way as conference proceedings, academic letters, and other document types that were tracked by the both data sources (see Annex A -Key definitions for co-publication analysis).

The reason for this procedure is the idea that jointly published conference papers can indicate international cooperation activity, which is of prime interest to the Black Sea Horizon project. The data is available only with partly different field names and different quality (depending on the data source). A bundle of software tools was especially developed to assure (1) that the formats of the data allow unification and (2) the rise of quality of metadata of publications tracked in both sources after unification. The steps involved are described in detail in Annex B - Data cleaning, consolidation of data sources and thematic areas

There are a few things to be kept in mind when interpreting the results and data presented:

- First of all, the sample was huge. Especially with regards to Russia and Turkey, as noted above, the study had to limit itself to cover only those co-publications that involve at least one other EU/AC-country.

- We have put a lot of effort in data cleaning and processing. Depending on the type of analysis (overall figures, subject areas, impact data, etc.), a rough analysis of possible error points to an error probability of 2-8 %. This may become especially important for those results, which are based on only a small number of publications.

- Impact data is a snapshot at a given point in time. While the number of publications in the two databases is stable approximately half a year after the end of the year of publication, the times cited counts are constantly being updated in the future as new publications refer to already recorded ones. In addition, older publications had more time to get cited than recent publications, i.e. the citation count for the latter is typically lower.

2. Results of the bibliometric analysis

This chapter presents the analysis results. The chapter is further divided into the following sections:

(1) overall numbers that provide an overview of the publications and especially co-publications of the EU/AC and BSCs, in the time span from 2003 to 2013;

(2) the internationalisation of publications as well as short profiles of the individual BSCs;

(3) a more detailed view in terms of research fields, and

(4) as a means to measure the quality of the recorded publications, their calculated impact is presented in the last section of this chapter.

2.1. Overall numbers – descriptive statistics

This study bases its analysis on both Thomson Reuter's Web of Science and Elsevier's Scopus databases. The number of involved records amount to 273,917 records from 2003 – 2013, which constitutes the overall publication output of BS countries from 2003 – 2013, with the exception of Russia and Turkey, for which only co-publications with at least EU/AC-country were taken into account.

The following pie chart (see Figure 1) shows the share of Web of Science and Scopus records within the total amount of covered datasets. One can immediately see, that the highest amount of covered data is overlapping, meaning it was included both in Scopus and WoS.
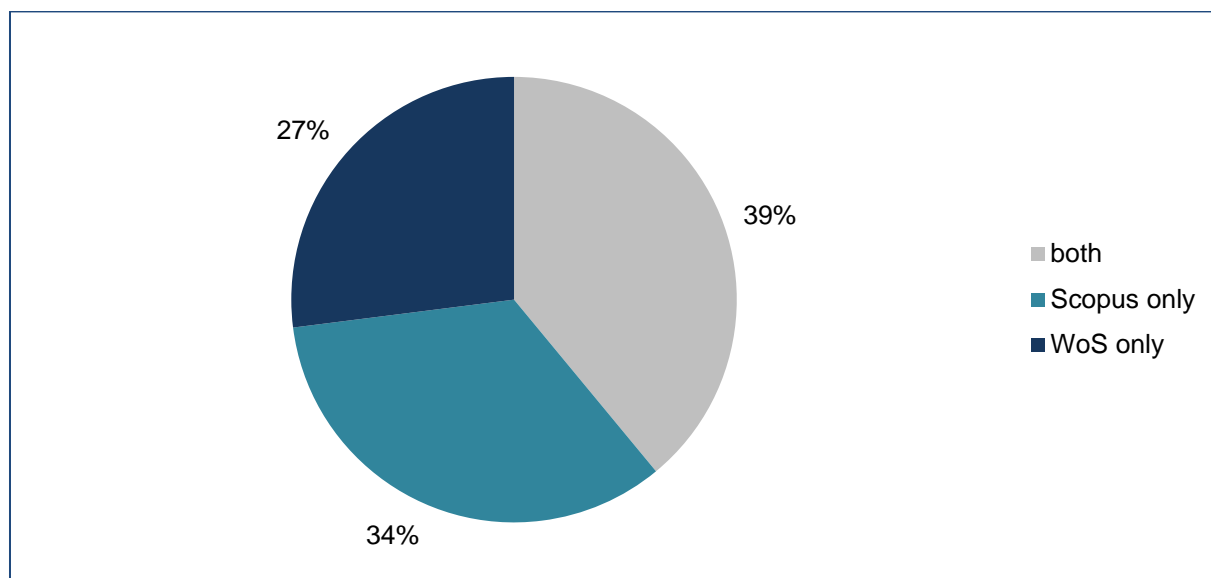
FIGURE 1: Data coverage of the citation databases WoS and Scopus for the BLACK SEA Region, 2003 – 2013 (Source: WoS+Scopus)

This bibliometric analysis covers the co-publication patterns of the individual BS countries. A BS co-publication refers to an international co-publication, i.e. a co-publication between at least one BS member state and at least one other state outside the BS region.

Of the 273,917 records that were analysed, each has, on average, been authored by 15.67 researchers affiliated in 2.41 countries and been cited about 8.44 times.

On country level, the data covered by WoS and Scopus include:

	Total publications	co-publications	co-publications with EU/AC
ARM	9,408	4,527	3,129
AZE	7,775	2,816	922
GEO	8,205	4,217	2,694
MDA	4,243	2,421	1,788
RUS	-	-	99,840
TUR	-	-	29,905
UKR	95,254	32,796	22,100

TABLE 1: Data coverage of the citation database WoS and Scopus for the BS Region 2003 – 2013

2.2. Publication output over time

The Black Sea region is a diverse geographical area, where Turkey, Russia and Ukraine are bigger countries, whereas Armenia, Azerbaijan, Georgia and Moldova are smaller countries. In terms of the development of national research and innovation systems, these countries are quite diverse as well. This difference, resulting from more or less spending on RTDI system affects also the publication output of its scientific and research institutions.

Figure 2 shows the distribution of publications over time from 1994 – 2014. For this overview data from Scimago, and thus only the data of Scopus were used, to allow a comparison of the total publication outputs.

Turkey shows an extraordinary increase in publications between 1996 and 2013. Also Russia's publications show a strong growth of publication output, however starting from a higher level than Turkey. The publication output of Ukraine has only slightly increased.
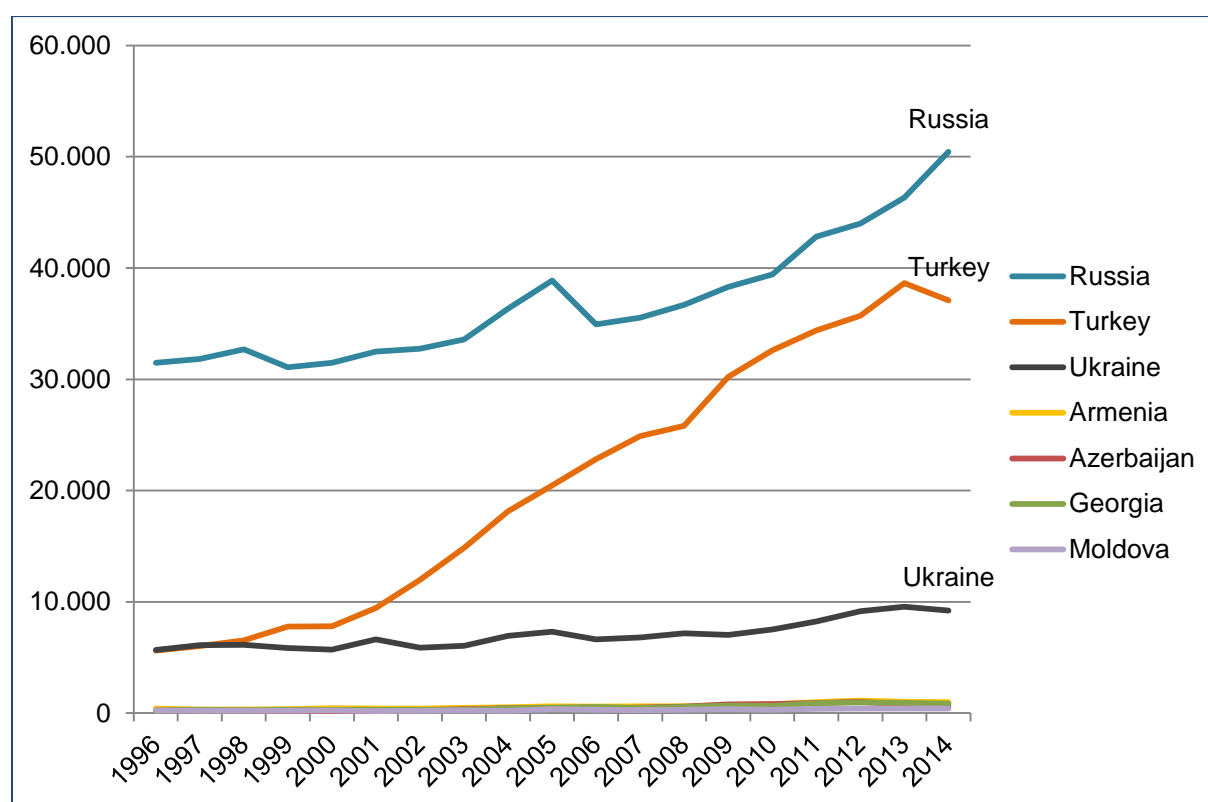
FIGURE 2: Distribution of Publications FROM 1996 - 2014; (Source: Scimago)

To have a closer look at the development of countries with less publication output, in the figure below Russia, Turkey and Ukraine are excluded.
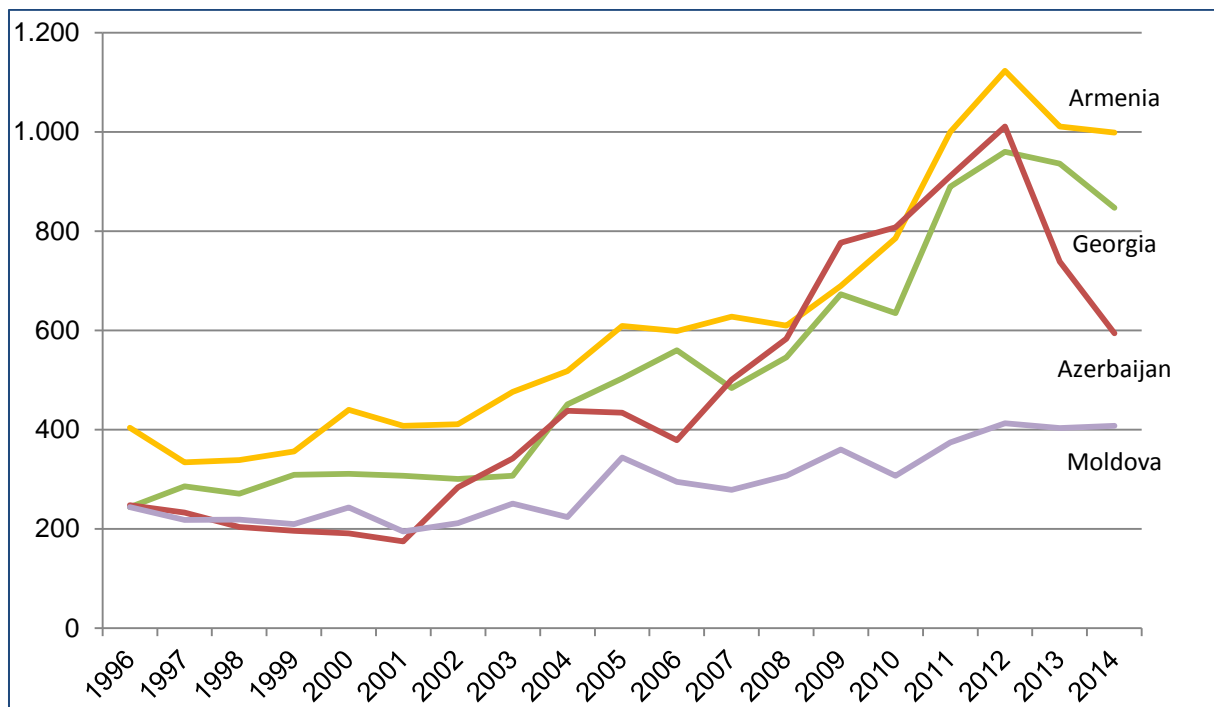
FIGURE 3: Distribution of Publications FROM 1996 - 2014; (Source: Scimago)

Over the period 1996 to 2012 the amount of publications rose significantly in Armenia, Azerbaijan and Georgia with a drop in 2012. Over the time, Armenia mostly kept its fore-runner role, with an exception from 2008 to 2010 where Azerbaijan was stronger. Moldova´s publication output increased only slowly over time.

2.3. Internationalisation of publications

Glänzel and Schubert (2005) have made an important observation with regards to the relation between the size and the publication output of a country: "Big countries have […] lower shares of international co-publications than medium-sized or small countries have. Nevertheless, the growth of the share of international co-publications can be observed independently of the country's size. The increase is thus a global law."
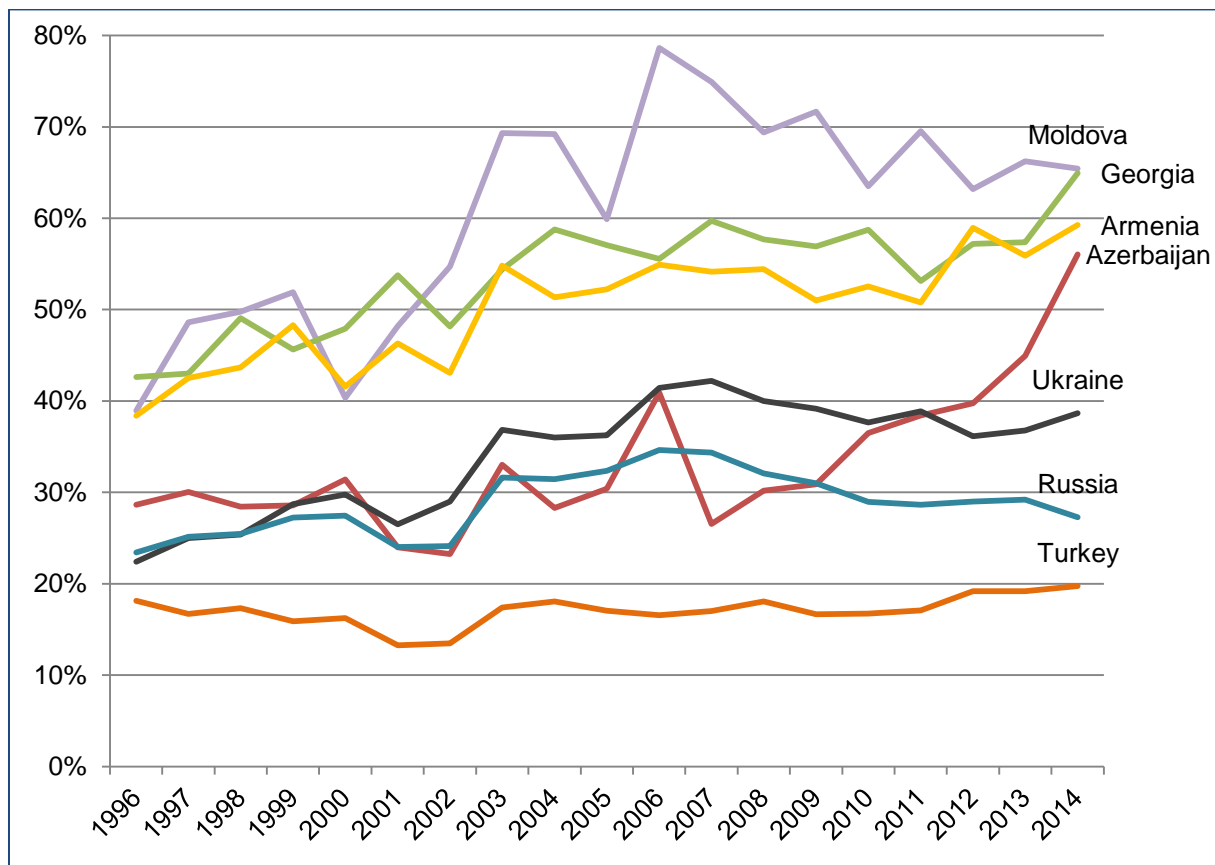
FIGURE 4: Percentage of co-Publications over time; (Source: Scimago)

Also for the comparison of co-publication shares we used the data of Scimago, which would not have been possible with our combined WoS and Scopus data, as the total publication output of Turkey and Russia is not available.

In the observed countries, the share of internationally co-authored publications has a wide range. Generally, between 2006 and 2014 the percentage of co-publications has risen in all countries. Looking at 2014, two groups of countries can be identified: on the one hand, smaller countries with a co-publication rate between 55% and 65% and on the other hand bigger countries between 20% and roughly 40%. In that sense, the general estimation that smaller research communities are, due to limited national capacities, better internationally connected seems to be valid. As Glänzel and Schubert (2005, p. 271) point out, this "international ambition" in producing publications has another positive consequence: International co-publications are more likely to appear in high-impact journals and have a better chance to be cited, than "domestically" produced papers.

Since 2002, Moldova has kept its position with the highest share of co-publications. Azerbaijan has an outstanding position in two ways: First it is the only country having a steady increase since 2007 and second; it doubled almost the share of co-publication output since 1996.

In the group of the bigger countries, Turkey´s co-publication rate is steady between 15% and 20%. Russia ranges between 21% and 34%.
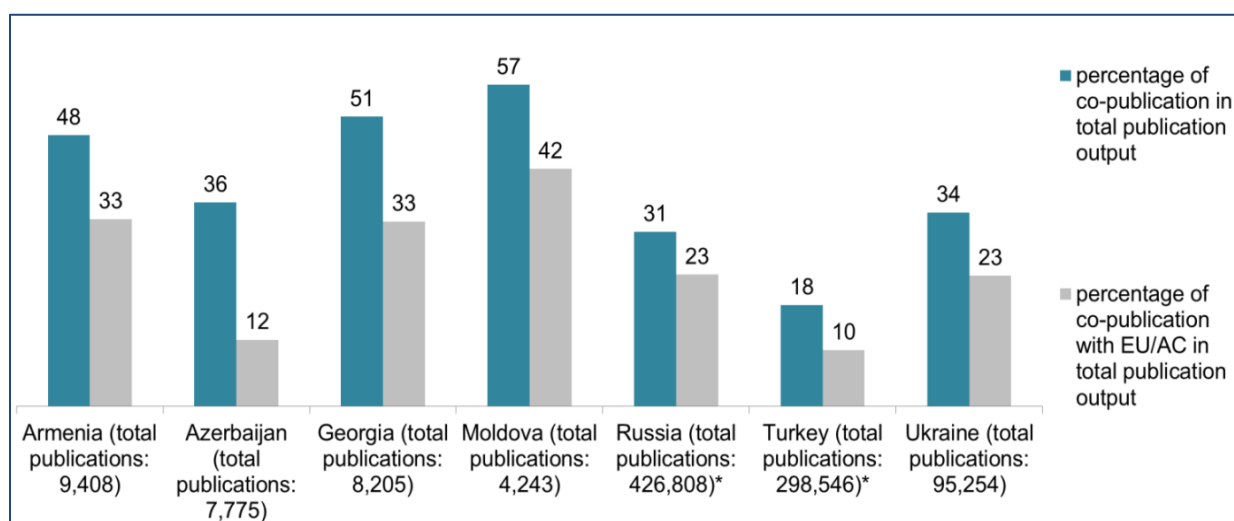
FIGURE 5: Percentage of co-publications in total publications and percentage of publications with EU/AC; 2003 – 2013, (Source: WoS+Scopus); *for Russia and Turkey the data regarding total publications and percentage of co-publications from Scimago were used.

For Figure 5 we also had to refer to Scimago for the total publication output of Turkey and Russia, which means that we have to compare them another data level of the other countries where we used both Scopus and WoS. However, it gives an indicated picture of the share of co-publications in total and co-publications with the EU/AC.

In Figure 5 in blue, we see the percentage of co-publication in total publication output, in red, we see the percentage of co-publication with EU/AC of the total publication output for the years 2003 - 2013.

In the group of the smaller countries, Moldova reaches as high as 57% of co-publications, followed by Georgia (51%), Armenia (48%) and Azerbaijan (36%). In the group of the bigger countries, Ukraine (34%) and Russia (31%) are almost equal, followed by Turkey (18%).

The share of co-publication with EU/AC in total publication output in descending order is: Moldova (42%), Georgia and Armenia (33%), Russia and Ukraine (23%), Azerbaijan (12%) and Turkey (10%).

The comparison of percentage of co-publication with EU/AC and with other countries shows an interesting picture: Azerbaijan has a co-publication share of 12% with EU/AC and 24% with other countries. This means, only 1/3 of co-publication shares are with EU/AC. Azerbaijan is an exception because all other countries have a higher share of co-publications with EU/AC than with other countries.

Georgia shows 19% co-publication share with other countries, followed by Moldova and Armenia (15%), Ukraine (11%) and Russia and Turkey (each 8%)

2.4. Strongest co-publication linkages with the EU/AC

When comparing the most important co-publication partner countries within the EU/AC we can clearly see that for almost all countries Germany is the most important partner, with one exception, Azerbaijan, where it is Turkey. Having a look at the following position the picture is rather diverse. Table 2 below shows the top 10 co-publication countries for the BS countries.

Armenia		Azerbaijan		Georgia		Moldova	
Germany	1,716	Turkey	1,321	Germany	1,288	Germany	489
France	1,507	Germany	483	Italy	985	Romania	292
Italy	1,444	United Kingdom	430	United Kingdom	944	Poland	245
United Kingdom	1,292	France	371	France	863	France	234
Poland	1,090	Italy	361	Spain	828	Spain	212
Switzerland	1,043	Switzerland	334	Switzerland	821	Italy	202
Czech Republic	888	Portugal	329	Poland	798	Ukraine	168
Spain	859	Poland	328	Turkey	714	Belgium	98
Greece	827	Spain	325	Austria	706	United Kingdom	95
Serbia	729	Netherlands	314	Greece	694	Switzerland	94

Russia		Turkey		Ukraine	
Germany	37,659	Germany	8,033	Germany	6,956
France	20,609	United Kingdom	7,298	Poland	5,416
United Kingdom	16,004	Italy	5,296	France	3,837
Italy	13,311	France	4,816	United Kingdom	3,231
Poland	8,699	Spain	3,368	Italy	2,401
Spain	8,473	Netherlands	3,268	Spain	1,961
Switzerland	8,075	Switzerland	2,621	Switzerland	1,340
Ukraine	7,311	Greece	2,351	Czech Republic	1,221
Netherlands	7,280	Belgium	2,124	Austria	1,178
Sweden	6,925	Austria	2,040	Sweden	1,156

TABLE 2: Development of co-publications with 10 most important partners in EU/AC 2003 – 2013 (Source: WoS+Scopus)

Figure 6 to Figure 12 show the development of the co-publications with the most important partners over time. It can be observed that the relation to the EU/AC countries developed more or less synchronously and is also strongly related to the total publication output. Only Moldova is different. Due to the limited number of publications no clear trend can be observed.
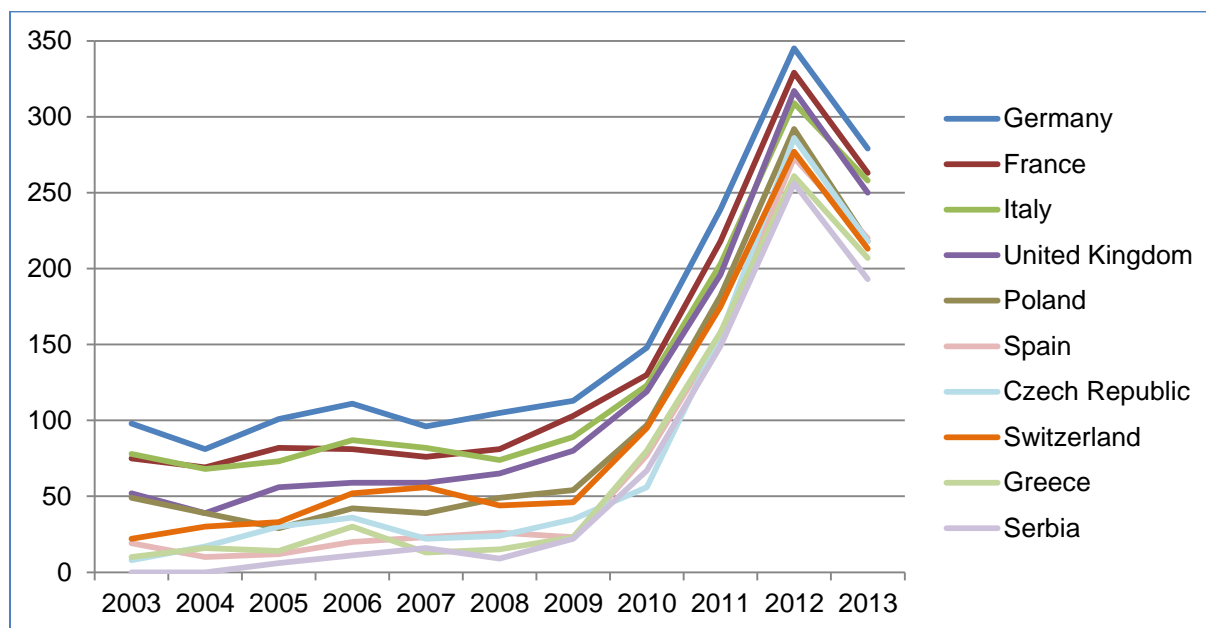
FIGURE 6: Armenia; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
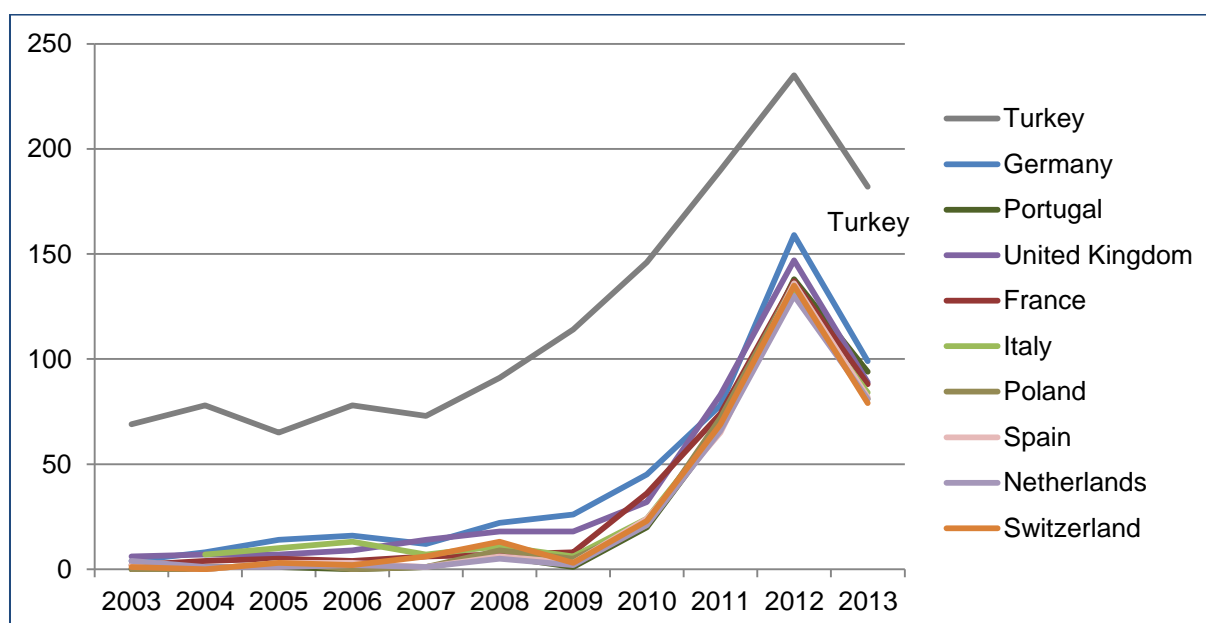
FIGURE 7: Azerbaijan; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
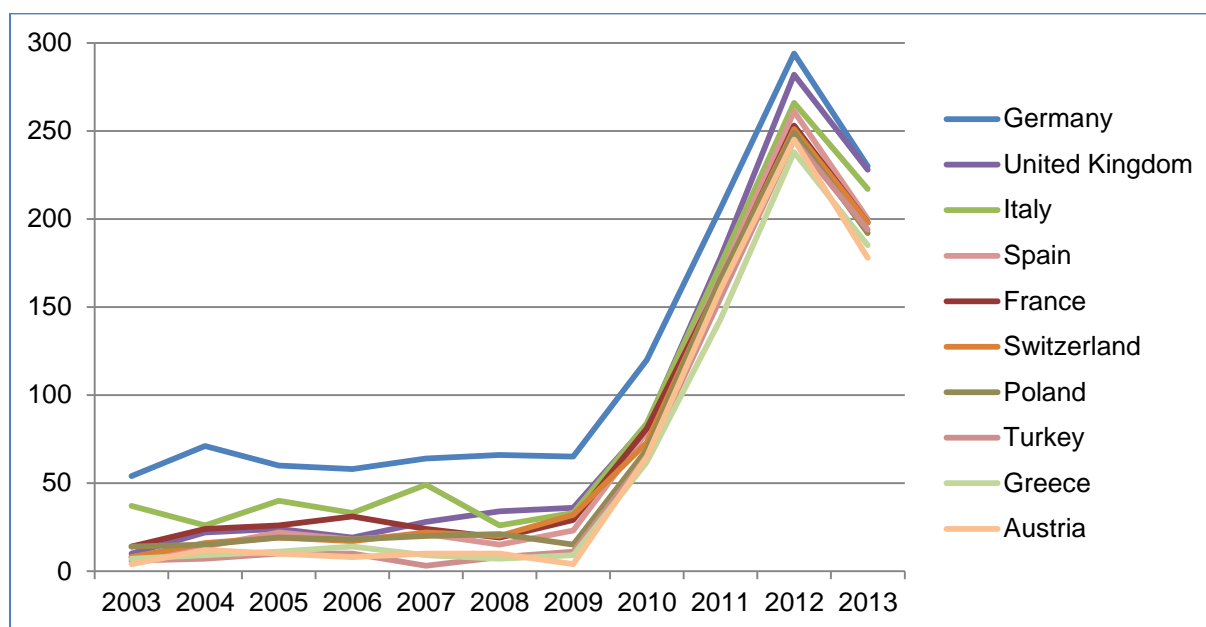
FIGURE 8: Georgia; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
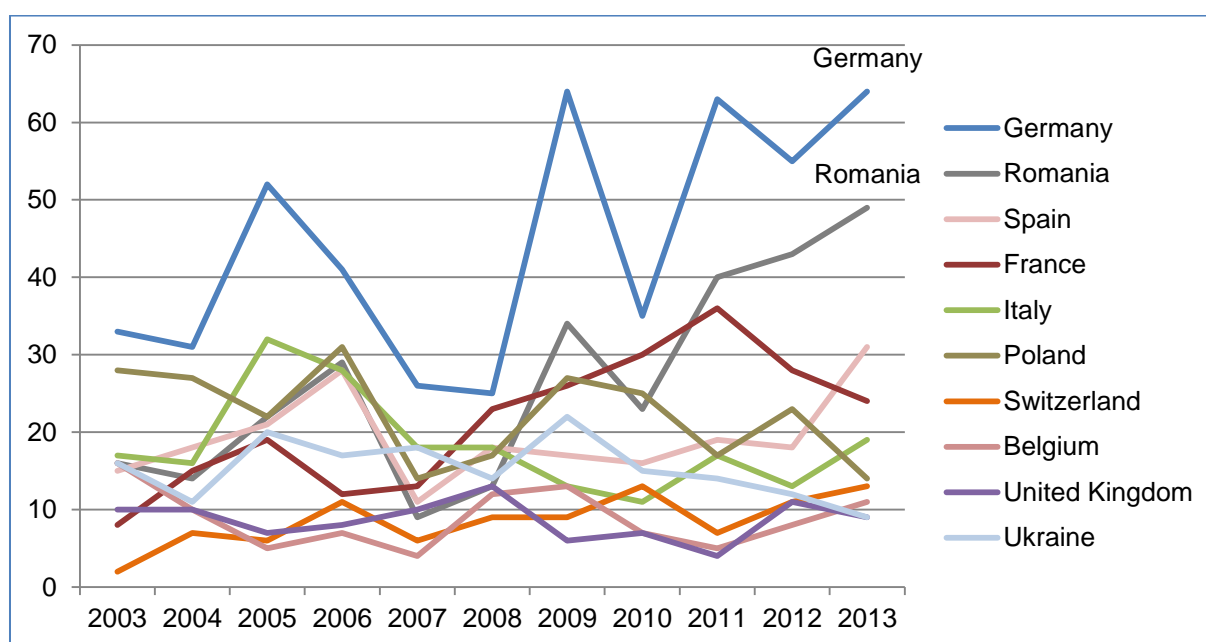
FIGURE 9: Moldova; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
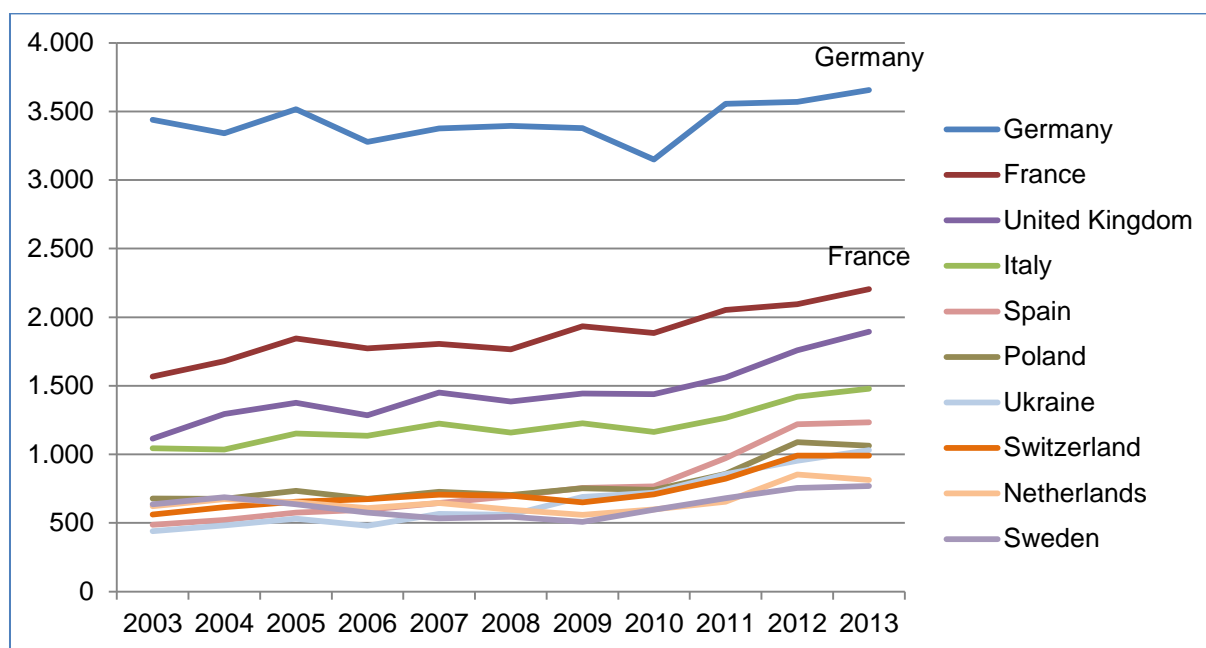
FIGURE 10: Russia; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
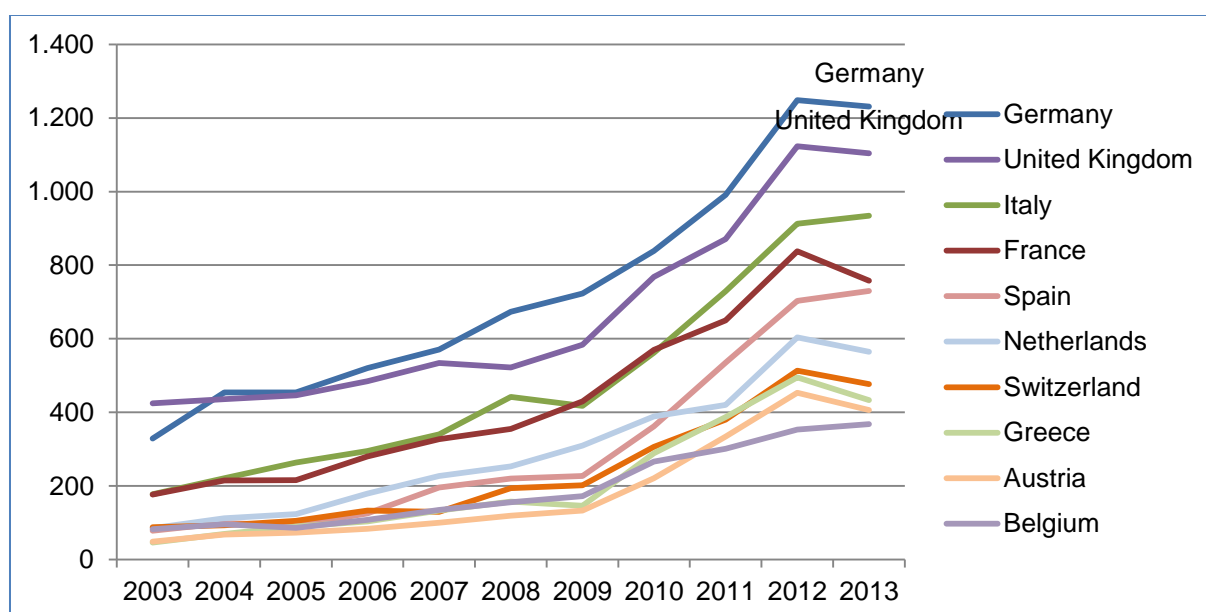
FIGURE 11: Turkey; Development of co-publications with 10 most important partners in EU/AC, 2003 – 2013 (Source: WoS+Scopus)
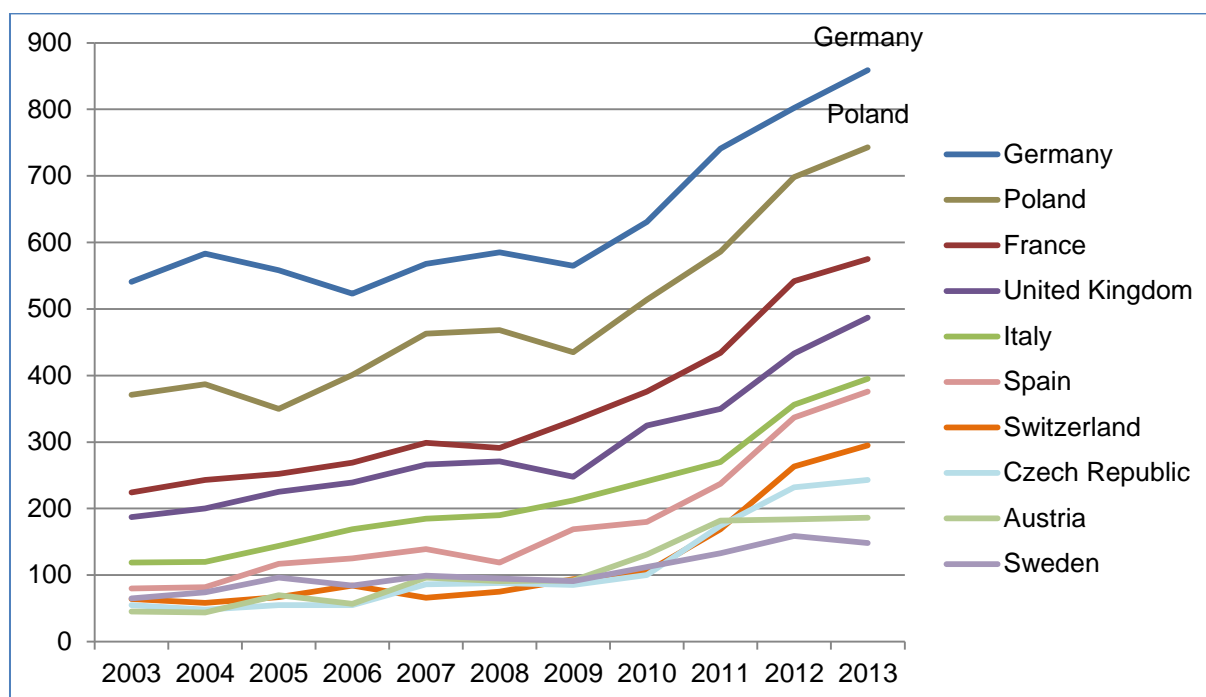
FIGURE 12: Ukraine; Development of co-publications with 10 most important partners in EU/AC (Source: WoS+Scopus)

2.5. Scientific research fields

To examine the scientific research topics in the Black Sea region co-publication output in the following section, we use the Science-Metrix Ontology of Science classification[1] (mostly on the area of the research field and only those scientific (co-)publications, which are citable[2]).

---

[1] Science Metrix, a Canada-based company, developed a multi-lingual three-level journal subject classification system: the Science Metrix Ontology of Science. It builds on comprehensive work on standardisation and classification of journals, partly financed by the European Commission. The main difference between the Science Metrix Ontology and classification systems used by Scopus and Web of Science is the disjunct classification, i.e. each journal is attributed to one (not one or more) subject category.

[2] Here we are only using (co-)publications which are citable, which means we are using only (co-)publications which are published in citable document types: Article, Conference/Proceedings Paper, Letter and Review (and don't count normally not cited documents like Editorials).
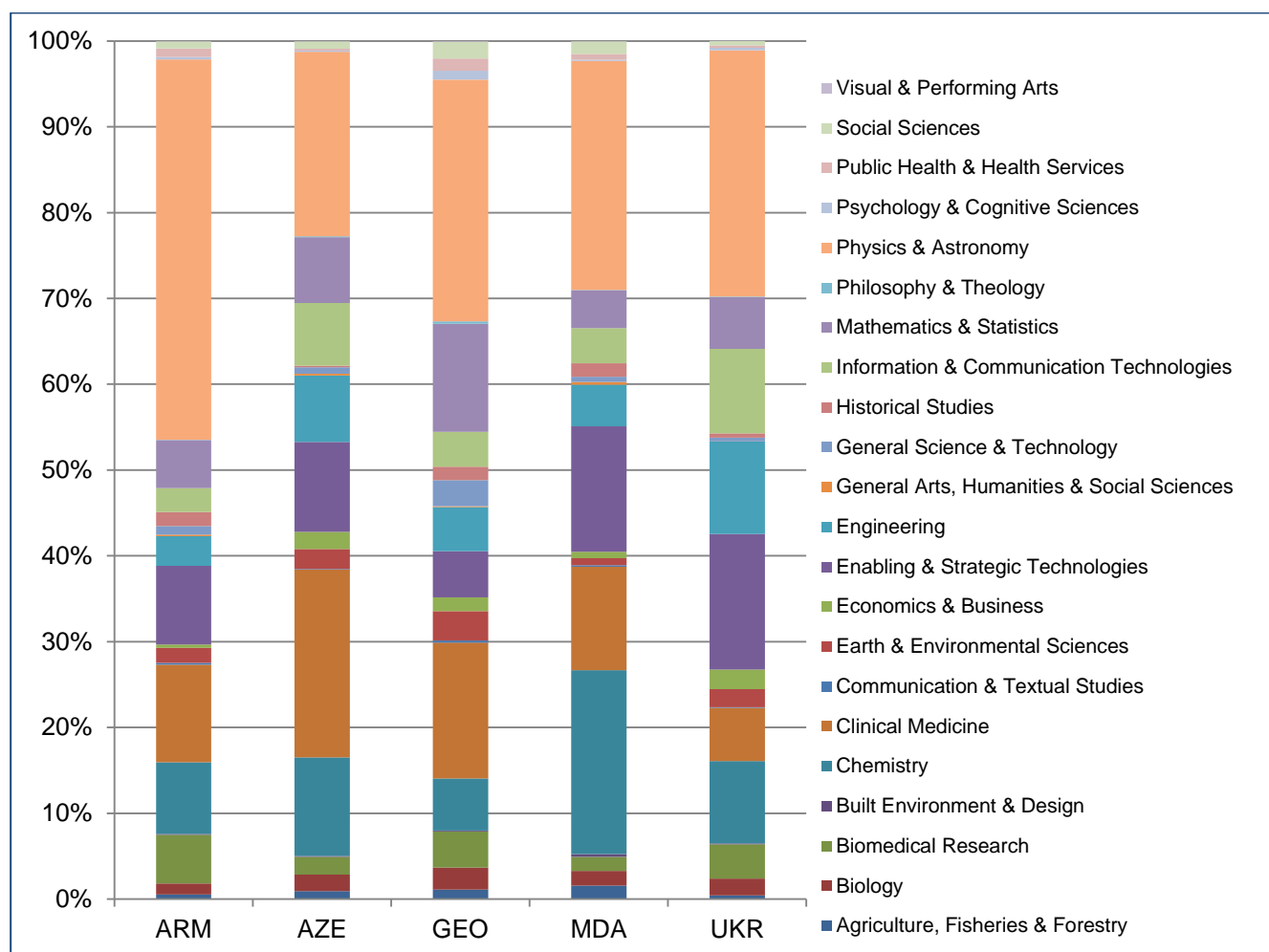
The comparison of scientific fields distribution in total publications[3] 2003-2013 (Figure 13) shows that we find the highest share of publication output in six scientific fields, namely Physics & Astronomy, Clinical Medicine, Chemistry, and Enabling and Strategic Technologies. All countries have the same field with the highest share of publication output in common: Physics & Astronomy which is one of the special characteristics of the BS region. Armenia takes an outstanding position, with a share in Physics & Astronomy as high as 44%. In the field of Clinical Medicine, we see the second biggest share in Armenia (11%), Azerbaijan (11%) and Georgia (15%). The second biggest share in Moldova lies in Chemistry (21%) and in Ukraine in the field of Enabling and Strategic Technologies (15%).

---

[3] As the data for Russia and Turkey are not available in our data set, they do not appear in this figure.
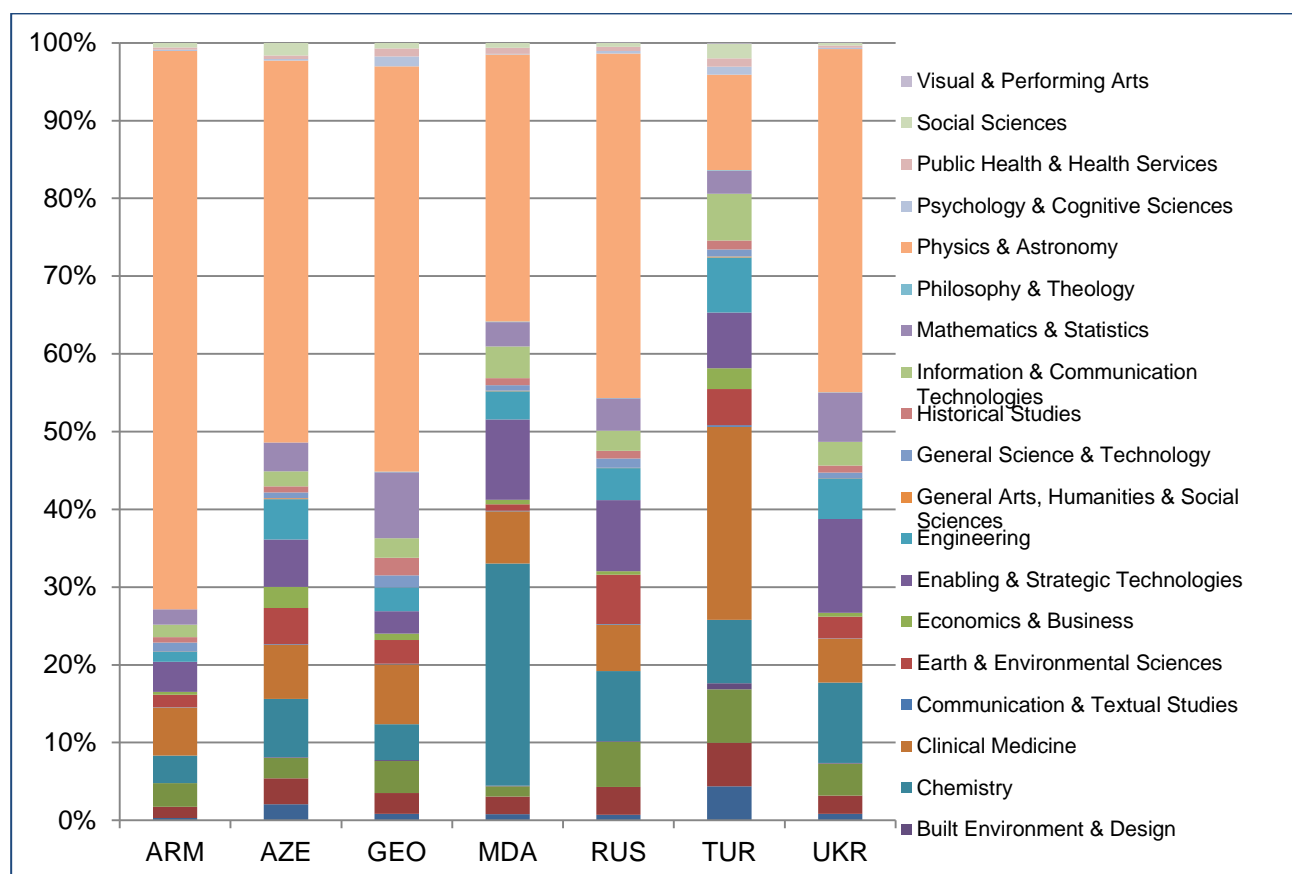
FIGURE 14: Comparison of scientific fields distribution in Co-Publications with EU/AC, 2003 – 2013 (Source: WoS+Scopus)

Figure 14 illustrates a comparison in co-publications with EU/AC countries (2003-2013). If we compare this figure to Figure 13 we can see that the field Physics & Astronomy is much stronger in the EU/AC cooperation than on national level. This is especially true for Armenia, Azerbaijan and Georgia. Other fields like Clinical Medicine (in Armenia, Azerbaijan, Georgia and Moldova), Enabling & Strategic Technologies (in Armenia, Azerbaijan and Ukraine) and Information and Communication Technologies (in Azerbaijan and Ukraine) are less internationalised and show significantly stronger performance in comparison on national level.

In the next sub-sections, the thematic portfolios of each Black Sea region country are discussed. Armenia, Azerbaijan, Georgia, Moldova and Ukraine are analysed and compared based on the total publication output. Russia and Turkey are analysed based on the co-publication output with the EU/AC, as for these countries the total

publication output is not available on the level of scientific fields. Thus, a direct comparison of all countries is not possible on all levels.

Firstly, each country´s overall publication output (including EU/AC co-publication links as well as further international links of each country beyond the EU/AC region) is compared to the country's international co-publications (BS links) and secondly, the Black Sea region co-publication activity of each country is compared to the thematic distribution of the country's overall publications and international co-publications. In particular and to go a little bit into more detail, the 14 scientific research fields with the most Black Sea region publication output were examined for this comparison at country level.

2.5.1. Armenia

Science Metrix fields	Overall publications		Co-publications		Co-publications with EU/AC		Differences in shares between co-publications and co-publications with EU/AC
	Number	share in %	Number	share in %	Number	share in %	
Agriculture, Fisheries & Forestry	51	0.54	38	0.84	9	0.29	-0.55
Biology	122	1.30	93	2.05	46	1.47	-0.58
Biomedical Research	533	5.67	154	3.40	95	3.04	-0.37
Built Environment & Design	8	0.09	6	0.13		0.00	-0.13
Chemistry	786	8.35	222	4.90	111	3.55	-1.36
Clinical Medicine	1,070	11.37	327	7.22	193	6.17	-1.06
Communication & Textual Studies	20	0.21	3	0.07	1	0.03	-0.03
Earth & Environmental Sciences	165	1.75	94	2.08	51	1.63	-0.45
Economics & Business	35	0.37	21	0.46	11	0.35	-0.11
Enabling & Strategic Technologies	865	9.19	279	6.16	121	3.87	-2.30
Engineering	324	3.44	100	2.21	41	1.31	-0.90
General Arts, Humanities & Social Sciences	16	0.17	6	0.13	1	0.03	-0.10
General Science & Technology	94	1.00	61	1.35	35	1.12	-0.23
Historical Studies	154	1.64	42	0.93	22	0.70	-0.22
Information & Communication Technologies	264	2.81	90	1.99	51	1.63	-0.36
Mathematics & Statistics	524	5.57	110	2.43	61	1.95	-0.48
Philosophy & Theology	4	0.04	2	0.04		0.00	-0.04
Physics & Astronomy	4,172	44.35	2,793	61.70	2,247	71.81	10.12
Psychology & Cognitive Sciences	24	0.26	7	0.15	5	0.16	0.01
Public Health & Health Services	95	1.01	47	1.04	10	0.32	-0.72
Social Sciences	78	0.83	32	0.71	18	0.58	-0.13
Visual & Performing Arts	4	0.04		0.00		0.00	0.00
Total	9,408	100.00	4,527	100.00	3,129	100.00	

TABLE 3: Science Metrix fields in Armenia (co-)publications, 2003-2013 (Source: WoS+Scopus)

Armenia´s specialisation in the area of Physics & Astronomy is clearly visible with 44% of the total publication output. Even though in all of the other countries of the BS region, Physics & Astronomy shows the highest share in total publication as well, no other country reaches such a high percentage. The second most important field for Armenia is Clinical Medicine with 14% of all publications followed by Chemistry and Enabling & Strategic Technologies with 11% each.

Relative strengths in overall publications compared to the BS region are in Physics & Astronomy. Relative weaknesses compared to the BS region are in Enabling & Strategic Technologies, Engineering, Chemistry and Clinical Medicine.

Looking at the co-publication shares, we see that Armenia´s co-publication rate in Physics & Astronomy of 61% is exceptionally high and more than ¾ of the co-publications is with EU/AC. Another strong co-publication field with EU/AC is Clinical Medicine.

In comparison to the BS region, the share of co-publications in Chemistry is rather low.

2.5.2. Azerbaijan

Science Metrix fields	Overall publications		co-publications		Co-publications with EU/AC		Differences in shares between co-publications and co-publications with EU/AC
	Number	share in %	Number	share in %	Number	share in %	
Agriculture, Fisheries & Forestry	73	0.94	47	1.67	19	2.06	0.39
Biology	150	1.93	96	3.41	31	3.36	-0.05
Biomedical Research	159	2.05	64	2.27	24	2.60	0.33
Built Environment & Design	8	0.10	5	0.18	1	0.11	-0.07
Chemistry	893	11.49	315	11.19	69	7.48	-3.70
Clinical Medicine	1,702	21.89	191	6.78	64	6.94	0.16
Communication & Textual Studies	8	0.10	4	0.14	1	0.11	-0.03
Earth & Environmental Sciences	180	2.32	80	2.84	43	4.66	1.82
Economics & Business	155	1.99	91	3.23	25	2.71	-0.52
Enabling & Strategic Technologies	814	10.47	263	9.34	56	6.07	-3.27
Engineering	603	7.76	245	8.70	48	5.21	-3.49
General Arts, Humanities & Social Sciences	15	0.19	4	0.14	1	0.11	-0.03
General Science & Technology	58	0.75	35	1.24	7	0.76	-0.48
Historical Studies	16	0.21	9	0.32	7	0.76	0.44
Information & Communication Technologies	568	7.31	92	3.27	18	1.95	-1.31
Mathematics & Statistics	595	7.65	230	8.17	34	3.69	-4.48
Philosophy & Theology	9	0.12	1	0.04		0.00	-0.04
Physics & Astronomy	1,671	21.49	1,000	35.51	453	49.13	13.62
Psychology & Cognitive Sciences	10	0.13	3	0.11	2	0.22	0.11
Public Health & Health Services	19	0.24	11	0.39	4	0.43	0.04
Social Sciences	64	0.82	30	1.07	15	1.63	0.56
Visual & Performing Arts	5	0.06		0.00		0.00	0.00
Total	7,775	100.00	2,816	100.00	922	100.00	

TABLE 4: Science Metrix fields in Azerbaijan (co-)publications, 2003-2013 (Source: WoS+Scopus)

Azerbaijan has specialisations in the scientific field of Physics & Astronomy and Clinical Medicine each accounting about 22% of the overall publication output. Whereas all other countries have a higher publication output in Physics & Astronomy no other country in the Black Sea region has a higher share of publication output in Clinical Medicine than Azerbaijan. Further strong scientific fields are Chemistry (11%) and Enabling & Strategic Technologies (10%).

Azerbaijan´s relative strengths in shares of publication output in the scientific fields compared to the other BS countries are in Clinical Medicine while the relative weaknesses lie within Physics & Astronomy and Enabling and Strategic Technologies.

Also in Azerbaijan, Physics & Astronomy is the most internationalised field with 36%, although in comparison to the BS region significantly more national than international. Chemistry (11%), Enabling & Strategic Technologies and Engineering (both 9%) are also high in co-publication shares. Almost half of the co-publications in Physics & Astronomy are with EU/AC. Other strong co-publication fields with EU/AC are Chemistry, Clinical Medicine and Enabling & Strategic Technologies.

2.5.3.   Georgia

Science Metrix fields	Overall publications		co-publications		Co-publications with EU/AC		Differences in shares between co-publications and co-publications with EU/AC
	Number	share in %	Number	share in %	Number	share in %	
Agriculture, Fisheries & Forestry	90	1.10	55	1.30	23	0.85	-0.45
Biology	212	2.58	136	3.23	72	2.67	-0.55
Biomedical Research	346	4.22	217	5.15	111	4.12	-1.03
Built Environment & Design	8	0.10	6	0.14	3	0.11	-0.03
Chemistry	497	6.06	251	5.95	124	4.60	-1.35
Clinical Medicine	1,298	15.82	456	10.81	208	7.72	-3.09
Communication & Textual Studies	21	0.26	5	0.12	2	0.07	-0.04
Earth & Environmental Sciences	281	3.42	147	3.49	82	3.04	-0.44
Economics & Business	133	1.62	61	1.45	22	0.82	-0.63
Enabling & Strategic Technologies	441	5.37	165	3.91	78	2.90	-1.02
Engineering	422	5.14	163	3.87	82	3.04	-0.82
General Arts, Humanities & Social Sciences	8	0.10	1	0.02		0.00	-0.02
General Science & Technology	249	3.03	65	1.54	42	1.56	0.02
Historical Studies	127	1.55	76	1.80	61	2.26	0.46
Information & Communication Technologies	336	4.10	125	2.96	68	2.52	-0.44
Mathematics & Statistics	1,031	12.57	347	8.23	228	8.46	0.23
Philosophy & Theology	24	0.29	6	0.14	2	0.07	-0.07
Physics & Astronomy	2,310	28.15	1,726	40.93	1,405	52.15	11.22
Psychology & Cognitive Sciences	85	1.04	53	1.26	34	1.26	0.01
Public Health & Health Services	119	1.45	85	2.02	28	1.04	-0.98

Social Sciences	160	1.95	68	1.61	19	0.71	-0.91
Visual & Performing Arts	7	0.09	3	0.07		0.00	-0.07
Total	8,205	100.00	4,217	100.00	2,694	100.00	

TABLE 5: Science Metrix fields in Georgia (co-)publications, 2003-2013 (Source: WoS+Scopus)

Georgia´s specialisation in the area of Physics & Astronomy is clearly visible with 28% of the total publication output, thus having together with Ukraine the second biggest share in Physics & Astronomy after Armenia. The second important field for Georgia is Clinical Medicine with 16% of overall publications. In the field of Mathematics & Statistics with 13%, Georgia has the highest share of overall publications in the BS region.

When we compare the shares of scientific fields with the other BS countries, Georgia´s strengths are in Mathematics & Statistics and General Science & Technologies while its weaknesses lie in Enabling & Strategic Technologies, Physics & Astronomy and Chemistry.

The co-publication share in Physics & Astronomy is 41% and in Clinical Medicine 11%. Relative weakness in international co-publication is Chemistry and Enabling & Strategic Technologies. Slightly over average is Clinical Medicine and Mathematics & Statistics.

Georgia´s highest share of co-publications with EU/AC are in Physics & Astronomy, Mathematics & Statistics and Clinical Medicine.

2.5.4. Moldova

Science Metrix fields	Overall publications		co-publications		Co-publications with EU/AC		Differences in shares between co-publications and co-publications with EU/AC
	Number	share in %	Number	share in %	Number	share in %	
Agriculture, Fisheries & Forestry	67	1.58	21	0.87	14	0.78	-0.08
Biology	72	1.70	52	2.15	41	2.29	0.15
Biomedical Research	70	1.65	45	1.86	23	1.29	-0.57
Built Environment & Design	13	0.31	2	0.08	2	0.11	0.03
Chemistry	910	21.45	637	26.31	511	28.58	2.27
Clinical Medicine	512	12.07	163	6.73	120	6.71	-0.02
Communication & Textual Studies	7	0.16	1	0.04	1	0.06	0.01
Earth & Environmental Sciences	35	0.82	24	0.99	15	0.84	-0.15
Economics & Business	32	0.75	10	0.41	10	0.56	0.15
Enabling & Strategic Technologies	620	14.61	303	12.52	185	10.35	-2.17
Engineering	206	4.86	101	4.17	65	3.64	-0.54
General Arts, Humanities & Social Sciences	13	0.31	1	0.04	1	0.06	0.01
General Science & Technology	26	0.61	15	0.62	13	0.73	0.11
Historical Studies	66	1.56	19	0.78	16	0.89	0.11
Information & Communication Technologies	175	4.12	97	4.01	73	4.08	0.08
Mathematics & Statistics	187	4.41	86	3.55	56	3.13	-0.42
Philosophy & Theology	2	0.05	1	0.04	1	0.06	0.01
Physics & Astronomy	1,131	26.66	801	33.09	614	34.34	1.25
Psychology & Cognitive Sciences	6	0.14	4	0.17	2	0.11	-0.05
Public Health & Health Services	29	0.68	22	0.91	14	0.78	-0.13
Social Sciences	62	1.46	15	0.62	11	0.62	0.00
Visual & Performing Arts	2	0.05	1	0.04		0.00	-0.04
Total	4,243	100.00	2,421	100.00	1,788	100.00	

TABLE 6: Science Metrix fields in Moldova (co-)publications, 2003-2013 (Source: WoS+Scopus)

Also in Moldova the specialisation is clearly visible in the area of Physics & Astronomy with 27% of the total publication output. The second most important field for Moldova is Chemistry with 22%, thus having the highest share of all the Black Sea region countries. The third is Enabling and Strategic Technologies with 15%. This is also a relative strength of Moldova compared in the BS region lies in Chemistry.

Relative weaknesses are in Physics & Astronomy, Mathematics and Statistics and Engineering.

The co-publication share in Physics & Astronomy is 33% which is in comparison to the BS region relatively low. In the field of Chemistry, the co-publication rate is relatively high with 26% compared to the BS countries. Enabling & Strategic Technologies have a 13% co-publication rate which is the average in the BS region. The most prominent co-publication fields with EU/AC are Physics & Astronomy, Chemistry and Enabling & Strategic Technologies.

2.5.5. Russia

Science Metrix fields	Co-publications with EU/AC	
	Number	share in %
Agriculture, Fisheries & Forestry	729	0.73
Biology	3,548	3.55
Biomedical Research	5,800	5.81
Built Environment & Design	93	0.09
Chemistry	9,005	9.02
Clinical Medicine	5,950	5.96
Communication & Textual Studies	83	0.08
Earth & Environmental Sciences	6,317	6.33
Economics & Business	468	0.47
Enabling & Strategic Technologies	9,141	9.16
Engineering	4,100	4.11
General Arts, Humanities & Social Sciences	33	0.03
General Science & Technology	1,179	1.18
Historical Studies	1,017	1.02
Information & Communication Technologies	2,596	2.60
Mathematics & Statistics	4,141	4.15
Philosophy & Theology	25	0.03
Physics & Astronomy	44,253	44.32
Psychology & Cognitive Sciences	324	0.32
Public Health & Health Services	537	0.54
Social Sciences	497	0.50
Visual & Performing Arts	4	0.00
Total	99,840	100.00

TABLE 7: Science Metrix fields in Russia co-publications with EU/AC, 2003-2013 (Source: WoS+Scopus)

As mentioned already, for Russia we can only analyse the data available for the co-publications with the EU/AC.

Russia´s scientific field with the highest co-publication with EU/AC is Physics & Astronomy with 44% which is relatively low compared to the shares of the BS region. Russia´s co-publication shares with EU/AC in the field of Chemistry and Enabling & Strategic Technologies (9%) are around average in the BS region.

2.5.6. Turkey

Science Metrix fields	Co-publications with EU/AC	
	Number	share in %
Agriculture, Fisheries & Forestry	1,310	4.38
Biology	1,671	5.59
Biomedical Research	2,061	6.89
Built Environment & Design	228	0.76
Chemistry	2,444	8.17
Clinical Medicine	7,424	24.83
Communication & Textual Studies	58	0.19
Earth & Environmental Sciences	1,391	4.65
Economics & Business	804	2.69
Enabling & Strategic Technologies	2,136	7.14
Engineering	2,130	7.12
General Arts, Humanities & Social Sciences	25	0.08
General Science & Technology	273	0.91
Historical Studies	343	1.15
Information & Communication Technologies	1,805	6.04
Mathematics & Statistics	883	2.95
Philosophy & Theology	26	0.09
Physics & Astronomy	3,674	12.29
Psychology & Cognitive Sciences	307	1.03
Public Health & Health Services	321	1.07
Social Sciences	565	1.89
Visual & Performing Arts	26	0.09
Total	29,905	100.00

TABLE 8: Science Metrix Fields in Turkey, co-publications with EU/AC, 2003 – 2013 (Source: WoW+Scopus)

Also for Turkey we have only the co-publication data with the EU/AC available.

Turkey´s scientific field with the highest co-publication with EU/AC is Clinical Medicine with a share of 25% which is a relatively high share within the BS region. Whereas the co-publication share in the field of Physics & Astronomy with the EU/AC countries is in all other countries much higher than in Turkey (12%). In the field of Chemistry, the co-publication share with EU/AC of 8% is average compared to the BS region.

2.5.7. Ukraine

Science Metrix fields	Overall publications		co-publications		Co-publications with EU/AC		Differences in shares between co-publications and co-publications with EU/AC
	Number	share in %	Number	share in %	Number	share in %	
Agriculture, Fisheries & Forestry	423	0.44	262	0.80	189	0.86	0.06
Biology	1,844	1.94	770	2.35	517	2.34	-0.01
Biomedical Research	3,829	4.02	1,416	4.32	909	4.11	-0.20
Built Environment & Design	81	0.09	24	0.07	14	0.06	-0.01
Chemistry	9,134	9.59	3,495	10.66	2,290	10.36	-0.29
Clinical Medicine	5,929	6.22	1,857	5.66	1,246	5.64	-0.02
Communication & Textual Studies	55	0.06	7	0.02	4	0.02	0.00
Earth & Environmental Sciences	2,014	2.11	1,088	3.32	625	2.83	-0.49
Economics & Business	2,163	2.27	148	0.45	104	0.47	0.02
Enabling & Strategic Technologies	15,052	15.80	4,204	12.82	2,664	12.05	-0.76
Engineering	10,289	10.80	2,036	6.21	1,161	5.25	-0.95
General Arts, Humanities & Social Sciences	18	0.02	5	0.02	1	0.00	-0.01
General Science & Technology	408	0.43	228	0.70	160	0.72	0.03
Historical Studies	449	0.47	243	0.74	198	0.90	0.15
Information & Communication Technologies	9,384	9.85	1,183	3.61	673	3.05	-0.56
Mathematics & Statistics	5,759	6.05	1,849	5.64	1,403	6.35	0.71
Philosophy & Theology	47	0.05	13	0.04	8	0.04	0.00
Physics & Astronomy	27,360	28.72	13,662	41.66	9,761	44.17	2.51
Psychology & Cognitive Sciences	245	0.26	57	0.17	29	0.13	-0.04
Public Health & Health Services	271	0.28	127	0.39	72	0.33	-0.06
Social Sciences	491	0.52	120	0.37	71	0.32	-0.04
Visual & Performing Arts	9	0.01	2	0.01	1	0.00	0.00
Total	95,254	100.00	32,796	100.00	22,100	100.00	

TABLE 9: Science Metrix fields in Ukraine (co-)publications, 2003-2013 (Source: WoS+Scopus)

Similar to the other BS countries, Ukraine´s specialisation is in the field of Physics & Astronomy (28% of the total publication output). The second most important field for Ukraine is Enabling & Strategic Technologies with 16% which is the highest amount in the Black Sea region, only Moldova has a closely high as 15%.

Relative strengths are in Information & Communication Technologies and in Engineering. Relative weaknesses are in Clinical Medicine and in Chemistry.

The co-publication share in Physics & Astronomy is 42% which is relatively weak compared to the shares of other BS countries. Enabling & Strategic Technologies have a co-publication share of 13% which is a little bit under the average. And Chemistry with 11% co-publication rate is as well a bit under the average. Almost 2/3 of all co-publications in the field of Physics & Astronomy is with the EU/AC, which represents the strongest co-publication field with EU/AC.

2.6. Impact highlights

In bibliometrics, impact is regarded as the impact on the research community. One fairly straightforward means to measure such an impact is citations per publication or – simply put – citation counts. This section highlights simply and concisely the most noteworthy observations.

Armenia, Azerbaijan, Georgia, Moldova and Ukraine are analysed based on the total publication output (see Table 10), Russia and Turkey are analysed based on the co-publication output with the EU/AD (see Table 11), as for these countries the total publication output is not available. As co-publications tend to have more citations than publications from a single country, a direct comparison is not possible between the two levels.

As it can be seen from the Table 10, Armenia has the most citations in average per publication (7.94), followed by Georgia (6.31), Moldova (5.94), Ukraine (3.72) and Azerbaijan (3.07).

The fields with the most citations are: General Science & Technology, Physics & Astronomy, Psychology & Cognitive Sciences, Earth & Environmental Sciences, Biomedical Research, Chemistry and Biology.

To identify country strengths, we have analysed where countries have on average more citations than the others. The average citations of these fields are marked in the Table 10 in grey.

The country-strengths are for:

- Armenia: General Science & Technology, Communication & Textual Studies and Physics & Astronomy
- Azerbaijan: Agriculture, Fisheries & Forestry and Biomedical Research
- Georgia: General Arts, Humanities & Social Sciences, Historical Studies, Social Sciences, Economics & Business, Information & Communication Technologies, Built Environment & Design, Communication & Textual Studies, Clinical Medicine and Biology
- Moldova: Psychology & Cognitive Sciences, Chemistry, Enabling & Strategic Technologies and Information & Communication Technologies
- Ukraine: Visual & Performing Arts, Philosophy & Theology, Built Environment & Design, Agriculture, Fisheries & Forestry, Clinical Medicine and Historical Studies

	Armenia	Azerbaijan	Georgia	Moldova	Ukraine
Applied Sciences	**1.97**	**1.75**	**2.79**	**3.13**	**1.77**
Agriculture, Fisheries & Forestry	1.63	4.59	2.09	1.63	5.20
Built Environment & Design	1.88	0.00	2.51	0.39	2.57
Enabling & Strategic Technologies	2.17	1.30	3.19	4.05	2.50
Engineering	2.13	2.70	2.54	1.77	1.71
Information & Communication Technologies	1.17	1.03	2.78	2.26	0.51
Arts & Humanities	**1.57**	**0.56**	**4.76**	**1.62**	**3.40**
Communication & Textual Studies	1.70	0.76	1.48	0.00	0.69
Historical Studies	1.62	0.88	6.33	1.88	3.99
Philosophy & Theology	0.25	0.11	0.71	0.50	1.51
Visual & Performing Arts	0.00	0.00	0.00	0.00	0.11
Economic & Social Sciences	**0.61**	**1.50**	**2.36**	**0.97**	**0.57**
Economics & Business	0.85	1.70	2.98	1.88	0.51
Social Sciences	0.50	1.03	1.85	0.50	0.84
General	**28.36**	**1.59**	**4.92**	**8.81**	**10.47**
General Arts, Humanities & Social Sciences	0.69	0.07	5.50	0.15	1.08
General Science & Technology	33.07	1.98	4.90	13.14	10.89
Health Sciences	**3.55**	**1.90**	**5.20**	**2.41**	**5.41**
Biomedical Research	4.39	8.18	6.76	4.42	5.52
Clinical Medicine	3.11	1.29	4.78	1.87	5.51
Psychology & Cognitive Sciences	4.92	4.70	5.93	17.83	2.20
Public Health & Health Services	3.41	2.58	4.80	4.04	4.46
Natural Sciences	**10.77**	**4.63**	**8.24**	**8.47**	**4.95**
Biology	4.15	3.09	6.23	5.22	3.60
Chemistry	2.88	1.86	5.72	10.38	4.72
Earth & Environmental Sciences	6.30	8.05	7.17	5.15	4.61
Mathematics & Statistics	1.11	2.16	2.09	2.04	2.05
Physics & Astronomy	13.84	6.76	11.85	8.31	5.75
TOTAL	**7.94**	**3.07**	**6.31**	**5.94**	**3.72**

TABLE 10: Average citations from total publications (Source: WoS+Scopus)

As already mentioned, we can compare Russia and Turkey only on the level of co-publications with the EU/AC.

In terms of impact, Turkey is especially strong in General Science & Technology (it has the highest value with 40.31 citations in average), Biomedical Research and Clinical Medicine. For Russia the fields with the highest impacts are General Science & Technology (68.78 citations), Clinical Medicine and Biomedical Research in the cooperation with the EU/AC.

When comparing the differences of impact, Russia shows a higher impact in General Science & Technology, Historical Studies and Clinical Medicine, while Turkey has a stronger performance in Engineering, Philosophy & Theology and General Arts, Humanities & Social Sciences. The related numbers are marked in grey in the Table 11.

	Russia	Turkey
Applied Sciences	**6.97**	**8.42**
Agriculture, Fisheries & Forestry	11.40	10.81
Built Environment & Design	4.52	6.45
Enabling & Strategic Technologies	8.28	8.61
Engineering	5.31	9.03
Information & Communication Technologies	3.81	6.04
Arts & Humanities	**10.90**	**6.18**
Communication & Textual Studies	2.56	4.58
Historical Studies	11.88	6.99
Philosophy & Theology	0.60	5.45
Visual & Performing Arts	0.28	0.15
Economic & Social Sciences	**4.92**	**6.38**
Economics & Business	5.96	8.01
Social Sciences	4.00	4.09
General	**66.79**	**38.37**
General Arts, Humanities & Social Sciences	1.09	13.90
General Science & Technology	68.78	40.31
Health Sciences	**18.80**	**16.79**
Biomedical Research	17.78	18.04
Clinical Medicine	20.79	16.87
Psychology & Cognitive Sciences	9.88	12.72
Public Health & Health Services	12.70	10.02
Natural Sciences	**12.96**	**10.66**
Biology	11.44	9.19
Chemistry	12.28	12.06
Earth & Environmental Sciences	13.14	13.10
Mathematics & Statistics	4.94	5.60
Physics & Astronomy	13.97	11.20
TOTAL	**13.18**	**12.04**

TABLE 11: Average citations from total co-publications with EU/AC (Source: WoS+Scopus)

REFERENCES

Archibugi, Daniele / Iammarino, Simona (1999): The policy implications of the globalisation of innovation, in: Research Policy, 28, 317-336.

Archibugi, Daniele / Pianta, Mario (1996): Measuring Technological Change through patents and innovation surveys, in: Technovation, 16(9), 451-468.

Bergek, Anna / Bruzelius, Maria (2010): Are patents with multiple inventors from different countries a good indicator of international R&D collaboration? The case of ABB. Research Policy 39, 1321-1334

De Prato, Giuditta / Nepelski, Daniel (2014): Global technological collaboration network: network analysis of international co-inventions, in: Journal of Technology Transfer, 39, 358-375.

De Rassenfosse, Gaétan et al. (2013): The worldwide count of priority patents: A new indicator of inventive performance, in: Research Policy, 42(3), 720-737.

Gambardella, Alfonso et al. (2012): Report on empirical results and policy, Deliverable 8.1 of the project 'Innovative S&T indicators combining patent data and surveys: Empirical models and policy analyses', Milano: InnoS&T.

Giuri, Paola et al. (2007): Inventors and invention processes in Europe: Results from the PatVal-EU survey, in: Research Policy, 36, 1107-1127.

Glänzel, Wolfgang / Schubert, Andras (2005): Analyzing scientific networks through co-authorship, in: Handbook of Quantitative Science and Technology Research, p. 257-276, Springer Netherlands, p. 265

Guan, Jiancheng / Chen, Zifeng (2012): Patent collaboration and international knowledge flow, in: Information Processing and Management, 48, 170-181.

Guellec, Dominique / van Pottelsberghe de la Potterie, Bruno (2001): The internationalisation of technology analysed with patent data, in: Research Policy, 30, 1253-1266.

Hagedoorn, John (2003): Sharing intellectual property rights – an exploratory study of joint patenting amongst companies, in: Industrial and Corporate Change, 12(5), 1035-1050.

Kuemmerle, Walter (1997): Building Effective Capabilities Abroad, in: Harvard Business Review, March-April, 61-70.

Meyer, Martin (2006): Academic Inventiveness and Entrepreneurship: On the Importance of Start-Up Companies in Commercializing Academic Patents, in: Journal of Technology Transfer, 31, 501-510.

Meyer, Martin / Bhattacharya, Sujit (2004): Commonalities and differences between scholarly and technical collaboration, in: Scientometrics, 61(3), 443-456.

Morescalchi, Andrea et al. (2015): The evolution of networks of innovators within and across borders: Evidence from patent data, in: Research Policy, 44, 651-668.

Niosi, Jorge (1999): The internationalisation of industrial R&D: from technology transfer to the learning organisation, in: Research Policy, 28, 107-117.

OECD (2013): Glossary of Statistical Terms. Patent, online at: https://stats.oecd.org/glossary/detail.asp?ID=2023, last accessed: 22 April 2015.

Pakes, Ariel / Griliches, Zvi (1984): Patents and R&D at the Firm Level: A First Look, in: Griliches, Zvi (ed.): R&D, Patents, and Productivity, Chicago: University of Chicago Press, 55-72.

Patel, Parimal (1995): Localised production of technology for global markets, in: Cambridge Journal of Economics, 19(1), 141-154.

Penner-Hahn, Joan / Shaver, J. Myles (2005): Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms, in: Strategic Management Journal, 26, 121-140.

Picci, Lucio (2010): The Internationalization of Inventive Activity: A Gravity Model using Patent Data, in: Research Policy, 39(8), 1070-1081.

Picci, Lucio / Savorelli, Luca (2012): The Structural Changes of Internationalized R&D Activities: An Analysis of Patent Data, Working Paper, SSRN Electronic Journal, 12.

Pohl, Hans / Warnan, Guillaume / BaasBaas, Jeroen (2014): Level the playing field in scientific international collaboration with the use of a new indicator: Field-Weighted Internationalization Score, in Research Trends, Issue 39, December 2014.

URL: http://www.researchtrends.com/issue-39-december-2014/field-weighted-internationalization-score/ (last accessed: 27.02.2015)

Scherer, Frederic M. and Harhoff, Dietmar (2000): Technology policy for a world of skew-distributed incomes, in: Research Policy, 29, 559-566.

Song, Jaeyong / Asakawa, Kazuhiro / Chu, Youngeun (2011): What determines knowledge sourcing from host locations of overseas R&D operations? A study of global R&D activities of Japanese multinationals, in: Research Policy, 40, 380-390.

Swiss Federal Institute of Intellectual Property (2014): Patents, online at: https://www.ige.ch/en/ip4all/patents/patents.html?type=kxftokfxvjku, last accessed: 22 April 2015.Webster, Elizabeth / Jensen, Paul H. (2011): Do Patents Matter for Commercialization?, in: Journal of Law and Economics, 54(2), 431-453.

ANNEXES

Annex A. Key definitions for co-publication analysis

Affiliation By affiliation we refer to a unique author-institution combination related to one record. The same author can be affiliated with several institutions within one single record. If this is the case, we consequently count several affiliations. Therefore, publications with one author, but two affiliations, one in one country of the Danube Region and one in another country, are included in the analysis and considered a co-publication. The number of affiliations in the Danube Region co-publications therefore shall not be confused with the number of authors.

BibTex BibTex on the one hand is a software package for creating literature references and indices in TeX or LaTeX documents (TeX is a typesetting system with integrated macro language, LaTeX is a variant of TeX). On the other hand we use the term in context of BibTeX exports from our data sources. In this case we refer to the BibTeX format which makes literature database entries available, coded in a particular way. The BibTeX format was the common denominator present to receive data from both different source databases with the same format, though slightly different in detail features.

Categories and main categories The two scientific literature databases used in this study assign the recorded books or periodicals to one or more thematic key words based on a classification system. In Elsevier's Scopus we have around 340 of these thematic keywords and around 250 in the case of Thomson Reuter's Web of Science (as listed in the annex). Only a small percentage of the scientific works is classified independently of the general classification of the periodical. To remove potential ambiguities, this study has used the Science Metrix Ontology that classifies journals on three levels of granularity: the domain, the field, and the sub-field.

Co-publication In the context of this study we refer to international scientific publications, indexed in literature databases, with the participation of at least two institutions/organisations in at least two different countries. For this study the term co-publication therefore is only used for international co-publications, unless explicitly stated otherwise.

Document types Each of the data sources used assigns a certain document type to the tracked publications to better describe them. These types reach from articles over abstracts and conference papers to editorials, errata and even music, movie or soft-ware reviews. To have comparable document types available we consolidated the two document type sets of our data sources to the following list: article, conference paper, meeting abstract, review, editorial, letter, other.

FRASCATI Manual The FRASCATI Manual is a standard methodology developed by the OECD to gather data on research and technology development activity of countries and contains a classification system of topic areas.

Impact Talking about impact in the framework of this study, we refer to the passive citations per record, i.e., the number of cases in which the respective publication was cited by a different younger publication. The data can only be punctual snapshots (summer/autumn 2014 in the case of this study). Citation counts for publications from very recent years are to be treated differently from the ones of very old publications and therefore, of course, comparison only makes sense for citation data from 3 or more years in the past. Publications that are tracked in both data sources tend to be assigned with different passive citation counts. Internally, we work with various algorithms to level this bias (e.g. the weight factor for citation counts from Web of Science or the preferential usage of the higher citation count).

Institute/Organisation Because the scientific literature databases used in this study relate authors to different organisational entities (i.e.: in one case the university as a whole is named, in another case we have detailed description of the institute or even the research group, etc.), we agreed on the usage of the label "institute" for the more detailed, subordinate level often called "organisational unit" (university institute, department, laboratory, sub entity of a company or international organisation) and the term "organisation" as the bigger entity, for example university, academy or intergovernmental organisation, etc.

Overlap factor The overlap factor is a measure we used to numerically express the intersection's size of the sets of journals listed in one ASJC category in comparison to a WoS category.

Record With record we refer to an entry in our database containing the meta data of a uniquely identified publication. In case the same publication appears in both data sources (Scopus and Web of Science), it is still dealt with as one record.

Annex B. Data cleaning, consolidation of data sources and thematic areas

The process starts with database-specific tables, into which parsed BibTeX data are inserted. The resulting tables contain records and affiliations for Scopus and WoS separately; they are subsequently unified into one record table and an affiliation table.

On the basis of raw data tables, we created a unified data set using a series of processing steps:

- Unification of journal names: the number and set of journals that are registered by Scopus and Web of Science are different. Many records appear in both databases, but with different spelling, institution or author notation, etc. The first unification step normalises syntax and spelling of journal names detected as identical (e.g. with differing capitalisation). In a next step we use Document Object Identifiers (DOIs) of all records in our database, which are unique (disregarding typing errors in the original databases, whose rate of occurrence lies at roughly 1%) for any registered publication worldwide (but unfortunately often are missing), to identify identical journals (in different notations). If one record is available with the same DOI in both databases, the journals linked to this record must as well be identical. Remaining journal names are examined for their similarity and are suggested as merging candidates, which then are controlled and manually assigned.
- Removal of duplicates in both record tables: Of course, publications that are registered in both databases must not appear twice in our unified data set. The identification of records from both sources describing the same publication is led through by searching for conformities in the following variables:
  - DOI
  - title, year, begin page
  - ISBN and begin page
  - journal ID or ISSN and begin page, year and author , title or volume
  - begin page and author-keywords
- Unification of journal names, second round: the results of the record unification can now be used to run through another round of journal name unification; a procedure to enhance data quality once more.
- Based on the previous steps a unified record-table can be established and filled according to the queries of interest.
- A similar data cleaning procedure takes place for the affiliations (author-institution combinations) – details below. After these data cleaning steps it can be shown which benefits the consultation of both data sources can offer for the present analysis: of the 1,026,556 observed Danube Region publications, 815,812 are listed by Scopus and 779,024 are listed by Web of Science.

Each cleaned record not only contains keywords given by the author(s) but has also been assigned with the journal subject categories of the respective source database(s). Unfortunately, the two thematic classification systems of Web of Science and of Scopus not only distinguish themselves in the way of assignment, but also in the set of the used categories. Each database

classifies each listed journal with one or more journal subject categories (249 in Web of Science) or with the help of All Science Journal Classification numbers (ASJC; 334 categories in Scopus).

A third classification scheme, the Science Metrix ontology, offers the advantage of a clear attribution of a journal to a single category called sub-field. Sub-fields are aggregated into fields which again are aggregated into domains. The ZSI developed a semi-automatic system to connect the two different category systems. Web of Science categories and Scopus ASJC categories are compared and rated for their overlap in the especially designed web-interface (see screenshot below).


FIGURE 15: Web interface for the assignment journal categories, showing an exemplary assignment of Web of Science subject areas to Scopus ASJC categories, Centre for Social Innovation, 2011-2015