

Paper presented at the European Alliance for Innovation, EAI's international conference  
on *Technology, R&D, Education & Economy for Africa*  
– TREE for Africa 2018 –, Abidjan, Côte d'Ivoire, 21<sup>st</sup>–22<sup>nd</sup> March 2018

## What drives regional scientific integration in Africa? Evidence from publications

Lorenzo Cassi<sup>1</sup>, Mafini Dosso<sup>2</sup>, Wilfriedo Mescheba<sup>3</sup>

[lorenzo.cassi@univ-paris1.fr](mailto:lorenzo.cassi@univ-paris1.fr)<sup>1</sup>, [mafini.dosso@ec.europa.eu](mailto:mafini.dosso@ec.europa.eu)<sup>2</sup>, [wilfriedo.mescheba@hceres.fr](mailto:wilfriedo.mescheba@hceres.fr)<sup>3</sup>

Paris School of Economics University of Paris 1 and HCERES-OST, Paris, France<sup>1</sup>  
European Commission, Joint Research Centre Seville, Spain<sup>2</sup>  
HCERES-OST, Paris, France<sup>3</sup>

Regional scientific integration is considered as fundamental for the development of an African system of research and innovation (AOSTI 2014, NEPAD 2006). This study exploits co-publications data on the period 2000-2015 to discuss the factors driving regional scientific integration in Africa overtime. It argues that a better understanding of the bottom up and regional factors driving the evolution of African scientific collaborations is key to design evidence-informed policies in a globalizing knowledge society.

**Keywords:** *scientific collaborations, African countries, publications, bibliometrics, gravity model*

### 1 Introduction

The strengthening of domestic research capacities and cooperation are pivotal for the transition towards the globalizing knowledge society. However, as many emerging and developing economies, African countries, which are often characterised by weak or fragmented national innovation and science systems, are constrained in their ability to generate scientific knowledge domestically (Arocena & Sutz, 2001). This contrasts with the important role of scientific knowledge in the process of economic and technological development of countries (Mowery, 1992; Cusmano et al. 2010; Mazzoleni & Nelson, 2007).

Besides the low scientific production of local universities and public research centres, the potential synergies and complementarities between the individual systems also remain very limited. At the same time, the African international collaborations are still greatly influenced by the global thematic agenda. In this perspective, Africa could benefit both from boosting international scientific cooperation that represents an important channel to tap into frontier scientific knowledge and in turn to build national scientific capabilities.

Awareness of the importance of scientific cooperation for socio-economic development has led since the years 2000s to the adoption of several regional and national policy instruments and frameworks in Africa.<sup>1</sup> Developed in the context of the New Partnership for Africa's Development (NEPAD) and endorsed in 2006, the African Science and Technology Consolidated Plan of Action (CPA) constitutes such a framework. The CPA intended to support, among other, the strengthening of links between science systems and the development of networks of centres of excellence through the thematic flagship programmes (NEPAD 2006). So far, few progresses have been observed outside the agricultural research field, mainly due to the lack of funding, of effective national policy prioritization as well as limited steering capabilities of African Union's (AU) implementation bodies (DIE, 2010). Nonetheless, improvements can be observed both in the quantity and quality of Africa's overall research output as well as in the regional cooperation (DIE 2010, World Bank 2014, Dosso et al 2017).

Understanding the existing degree of regional scientific integration in Africa is a key step to design policies aimed at building a common research area in this region, and it represents the main research interest of this study. This is even more relevant as research collaborations in Africa present distinctive features that should be mapped in order to design evidence-informed policies and policy practices (Adam et al 2014, World Bank 2014). In addition, it can reasonably be assumed that the last two decades, marked by fast economic growth and enhanced support to S&T-related activities in many African countries, may have generated new opportunities and dynamics, and thus modified Africa's scientific collaborations, i.e. the patterns, bottom-up drivers as well as the role of regional 'zones' and hubs (Guns and Wang, 2017, for a study on the networks of African collaborations).

In this framework, the aim of this paper is twofold: it analyses the determinants of scientific collaborations among African countries; it investigates the role of regional factors (e.g. regional 'communities') together with the influence of historical ties (e.g. Great Britain vs. France) in these collaborations and how they changed overtime. In line with earlier studies (Hoekman et al. 2010), we use bibliometric data on co-authorships to identify the patterns of scientific collaborations (Katz and Martin 1997, for a critical assessment of this proxy) and we trace their evolution overtime.

The remainder of the paper is organized as follows. Section 2 focuses on the drivers of scientific collaboration put forward by the studies in the *spatial scientometrics* tradition. Then, it presents the main research questions departing, among other, from earlier studies on the patterns of African scientific collaborations. Section 3 describes our methodological approach and provides the main features of our dataset and variables for the empirical analysis. Section 4 presents the main results and the outcomes of further explorations on the effects of geographical distance and 'common partner' over time. Section 5 concludes.

---

<sup>1</sup> In addition of continental plans, the Economic Community of West African States (ECOWAS) and the Southern African Development Community (SADC) have adopted regional strategies for STI. More recently, the Science, Technology and Innovation Strategy for Africa (STISA-2024) has been adopted as the successor of the CPA (AUC, 2014) and in 2015 the establishment of the pan-African platform Alliance for Accelerating Excellence in Science in Africa, (AESA <http://www.nepad.org/programme/alliance-accelerating-excellence-science-africa-aesa>).

## 2 Background literature and main hypotheses

In order to understand which factors shape scientific collaborations across countries in Africa, we draw from a stream of studies in *spatial scientometrics* (Frenken et al. 2009) that has set out a comprehensive analytical framework to analyse the geographical distribution of scientific activities. This framework builds on the proximity approach developed in regional studies (Boschma 2005), which claims that personal and organisational interactions, in particular those involving the exchange of intangible goods, such as scientific knowledge, are influenced by the proximity between actors. The literature has focused on five main types of proximities: geographical, institutional, cognitive, organisational and social. In a nutshell, it has been argued that effect of geographical proximity can be attenuated or substituted by the other forms of proximities (Boschma, 2005; Hardeman, et al. 2015). *Geographical* proximity favours serendipitous encounters with potential partners and reduces the costs of collaboration at distance (e.g. travelling). These face-to-face contacts are important because they favour the exchange of tacit knowledge. *Institutional* proximity captures the difference in the “rules of the game”, therefore similar institutional contexts should favour collaborations, since partners are aware of the prevailing formal and informal rules and more easily align their incentive systems. On the contrary, dissimilar institutional contexts can generate misunderstanding and misalignment of the incentive systems, which in turn discourage collaboration. *Cognitive* proximity refers to the distance in the mind-set and knowledge domain of actors. Actors might need a certain degree of cognitive similarity in order to understand each other and benefit from their interaction, though also dissimilarity might be required since actors often look for complementary knowledge when searching for a partner. *Organisational* proximity refers to the belonging to either similar or dissimilar organisations. If they belong to the same organisation, actors might have stronger incentive and lower costs in establishing collaboration than if they would belong to different organisations. However, organisational dissimilarity might be also a prerequisite of the collaboration if actors need to access assets that are not available in their own organisation. *Social* proximity is regarded a positive determinant of collaboration, since actors that share some social traits tend to trust each other and avoid opportunistic behaviours. However, it is also argued that too close social ties might be irrelevant since actors will access only to redundant knowledge, which overall reduces the benefit of collaborating.

Several empirical studies using co-publications as a proxy for research collaborations have adopted the proximity approach to identify the patterns of scientific and technological collaboration across cities, countries and regions both in advanced (Cassi et al. 2015; Hardeman et al. 2015; Hoekman et al. 2010; Ponds et al. 2007; Picci, 2010) and emerging economies (Montobbio and Sterzi, 2013; Scherngell and Hu, 2011).

Since the early 2000s, the scientific works dedicated to mapping and measuring the research output and impact of Africa using bibliometric data have flourished (among other, Arvanitis et al. 2000, Tijssen 2007, Pouris and Pouris 2009; Boshoff, 2009, 2010, Irikefe et al 2011, Toivanen and Ponomariov 2011, Mègnigbeto 2013, Pouris and Ho 2014, Confraria and Godinho 2015, Sooryamoorthy, 2018). The current patterns and outputs of scientific collaborations in Africa clearly suggest that African countries are still far from exploiting the potential for synergies and complementarities between their research systems. Nevertheless, as the total scientific outcomes of the continent, research collaborations are slightly increasing in Africa and some signs of (local) growing self-reliance and changing local dynamics can be observed (Adams et al 2014, Guns and Wang 2017). Also, as underlined by the AOSTI (2014), more than 36 AU Member States have recorded increases in their intra-continental

cooperation in the past decade. Earlier statistical and networks analyses suggest that such scientific activities rely upon common language, history and geographical “proximity” as key dimensions for collaboration (Mêgnibeto 2013 for an overview of West African collaborations, Adams et al., 2014). Yet, as underlined by Adams et al. (ibid.), more research is needed to properly understand the complexity of African collaborations and the bottom-up and local drivers behind the Africa’s research and scientific collaboration outcomes. Such evidence is needed for the design and implementation of policies targeting further regional scientific integration in Africa.

Accordingly, our analysis aims to assess and explain the existing degree of scientific integration of African countries. To do that, we perform an econometric analysis using essentially bibliometric to capture scientific collaborations. More specifically we test if:

1) *Geographical distance is decreasing over time.* In an integrated area, geographical distance should not in principle matter (Hoekman et al., 2010). Other determinants should prevail. So, a decreasing importance of the geographical distance effect in shaping scientific collaboration between African countries can be interpreted as a first evidence of an increasing integration or at least an intensification of collaboration network. In addition, the effect of geographical distance is also expected to be attenuated due, for instance, to improved (air) mobility especially within regional zones, a greater facilitation of visa (upon arrival policies) in a few countries as well as the rise of internet on the continent and the related information-sharing opportunities.

2) *Legacy of common past colonial ties is decreasing over time.* The historical legacy of past ties, especially the difference between Francophone and Anglophone countries, still matters in determining scientific collaboration between countries (Adams et al 2010, Boshoff, 2009, 2010 on central and southern Africa, Mêgnibeto 2013 for West Africa). However, we should expect that this effect is decreasing over time as the integration effort is increasing.

3) *The role of non-African countries is decreasing.* Usually, international collaboration with non-African countries can be interpreted as index of scientific quality, however a too high share of ‘outside’ collaboration “may denote a situation of dependence” (AOSTI, 2014, p. xvi) on, for instance, external resources (e.g. financial supporting funds). Again, the emergence of African countries should imply a decreasing importance of external partners in shaping local collaborations.

4) *Regional ‘Community’ effect is increasing.* Mainly established to facilitate the regional and continental economic integration, the Regional Economic Communities (RECs) are closely integrated with the AU’s work and have been increasingly involved in the coordination of AU’s Member States in different areas. Thus and depending on the RECs, we can expect that they might have played a structuring and integrative role among the member countries’ research systems.

### **3 Methodology and data**

#### *3.1 Econometric approach*

In line with previous empirical studies on scientific and technological collaborations, we model the co-publication patterns between any two countries using a gravity model (Ponds et

al. 2007; Hoekman et al. 2010; Montobbio and Sterzi, 2013). The gravity approach, which has been initially exploited in the trade literature to explain inter-country trade flows (Bergstrand, 1985), takes inspiration from the Newtonian law of gravitation, where the attraction between two objects depends positively on their masses and negatively on their distance. In our conceptualisation, the model implies that the intensity of co-publication activity between two countries increases with their number of publications, which proxy their mass, and decreases with their distances (e.g. geographical). Our gravity model can be represented by the following equation:

$$\ln(Copub)_{ijt} = \beta_0 + \beta_1 A_{it} + \beta_2 A_{jt} + \beta_3 PROX_{ijt} + \varepsilon_{ijt} \quad (1).$$

where  $Copub_{ijt}$  is the natural log of the number of co-publications between countries  $i$  and  $j$ , at time  $t$ .  $A_{it}$  and  $A_{jt}$  respectively capture the masses (e.g. number of publications) of countries  $i$  and  $j$  at time  $t$ . We take the inverse of the distance (i.e.  $PROX$ ) to capture the different types of proximities between organisations: geographical, scientific, social, institutional.

Our model is estimated using a negative binomial to tackle the issues of skewed distribution, excess of zeros and over-dispersion in the co-publication data (Hilbe, 2011, Cameron and Trivedi, 2013). We mainly performed two sets of regressions. In the first set, the determinants of co-publications are explored mainly according to the hypotheses suggested in Section 2 (baseline model). The second set of regressions aims at exploring the changes in the effects of the geographical distance and of ‘having specific common partners’, i.e. whether they belong to the political North or subsets or only to Africa. In this latter framework, twelve regressions are performed with dependent variables defined over three years, i.e.  $t/t-2$ ; where  $t=2004\dots2015$ .

### 3.2 Data and variables

The main source of data<sup>2</sup> is the HCERES-OST scientific publication dataset, an enriched<sup>3</sup> version of Thompson Reuters Web of Science - SCIENCE CITATION INDEX EXPANDED (henceforth WoS). Even if we have used an enriched version of WoS, our data suffers of the same drawbacks of the original data in terms of coverage<sup>4</sup>. First, in terms of scientific fields: WoS has a bias towards hard science and social sciences and humanities are highly underrepresented. Second, de facto mainly scientific publications in English are reported and many national journals are not considered, and this is especially true for developing countries (Rafols et al., 2016).

---

<sup>2</sup> Other source is the CEPII gravity dataset, cepii.fr.

<sup>3</sup> For instance, the geolocalisation of authors.

<sup>4</sup> A comprehensive comparison of WoS and Scopus (Mongeon and Paul-Hus, 2016) shows that Scopus is subject to the same biases as WoS, in spite of its larger journal coverage. This bias is hard to avoid for international comparative studies but should be kept in mind when interpreting the results.

We focus on the subset of publications published in eleven scientific fields<sup>5</sup> between 2000 and 2015, reporting among the authors' affiliation located in 52 African countries.<sup>6</sup> This means that our sample concerns all publications involving at least an author located in Africa and that we have information on collaboration with authors are located out of the continent.

Based on this, we can define a series of variables adapted to a gravity equation approach (Table 1). Time-variant explanatory variables are integrated with a 3-years lag, i.e.  $t-1$  to  $t-3$ . Table 2 gives the related descriptive statistics.

**Table 1 – List of variables and their definition**

<b>Variable</b>	<b>Definition</b>
<b>Copub<sub>ijt</sub></b>	Number of scientific papers co-authored by residents of countries $i$ and $j$ at time $t$
<b>lang_off<sub>ij</sub></b>	$i$ and $j$ share a common language (dummy)
<b>comcol<sub>ij</sub></b>	$i$ and $j$ have ever had a colonial link (dummy)
<b>logdistw<sub>ij</sub></b>	$i$ and $j$ geographical distance weighted by population of main agglomeration (logarithm)
<b>logpastcopubijt</b>	Stock of co-publication between country $i$ and $j$ , time $t-1$ to $t-3$ (logarithm)
<b>logpub<sub>it</sub></b>	Stock of publication of country $i$ , time $t-1$ to $t-3$ (logarithm)
<b>rec<sub>ij</sub></b>	$i$ and $j$ belong to a same REC (dummy)
<b>CEN_SAD<sub>ij</sub> (etc.)</b>	$i$ and $j$ belong to a specific REC (dummy)
<b>dist_tech<sub>ijt</sub></b>	The complement to 1 of cosine similarity of the scientific profiles of country's $i$ and $j$ , time $t-1$ to $t-3$ . The scientific profile is proxied by the shares of publications by field.
<b>lognorth<sub>ijt</sub></b>	Number of common developed countries scientific partners between countries $i$ and $j$ , time $t-1$ and $t-3$ (logarithm)
<b>logcolony<sub>ijt</sub></b>	Number of common past colonial (GB, FR, NL, PT) scientific partners between countries $i$ and $j$ , time $t-1$ to $t-3$ (logarithm)
<b>logafrique<sub>ijt</sub></b>	Number of common African scientific partners between countries $i$ and $j$ , time $t-1$ to $t-3$ (logarithm)
<b>logonlyafrique<sub>ijt</sub></b>	Number of common African scientific partners between countries $i$ and $j$ , time $t-1$ and $t-3$ (logarithm) taken into account only publications involving exclusively African countries.

For each African country  $i$  at year  $t$ , we calculate the stock of publications in the previous three years. Moreover, based on this stock, we map the scientific profile of the country in terms of the eleven fields. The same is done for each country for each of the eleven fields, i.e. we have calculated its stock and mapped its profile in terms of subfields. On this basis, the scientific proximity between two countries  $i$  and  $j$  (i.e.  $science\_prox_{ijt}$ ) is calculated as 1 minus the cosine similarity between the scientific profiles of the two countries.

<sup>5</sup> Fundamental biology, Medical research, Applied biology, Chemistry, Physics, Earth Sciences, Astronomy and Astrophysics, Engineering and Computer Sciences, and Mathematics, Social sciences and Humanities, Multidisciplinary. They are defined as an aggregation of Thomson Reuters Scientific Categories.

<sup>6</sup> Gambia and South Sudan are excluded. Gambia's country code was not available at the time of the download and only 5 years are available for South Sudan. The country obtained its independence (from the North) since 2011.

For each couple of countries,  $i$  and  $j$ , we have some relational variables, which are not varying over time, such as geographical distance from the capital cities (i.e.  $distance_{geo_{ij}}$ ). The geographical distance has been transformed in proximity considering its inverse (i.e.  $geo\_prox_{ij}$ ).<sup>7</sup> Moreover, we define a dummy if the two countries share a common past colonial tie, if they speak the same language and if they belong to a same or specific Regional Economic Community ( $rec_{ij}$  and  $REC\_name_{ij}$ ). RECs are not exclusive, i.e. the same country can belong to one or more communities. The African union recognizes eight RECs including the Arab Maghreb Union (AMU, 5 countries), the Common Market for Eastern and Southern Africa (COMESA, 19), the Community of Sahel-Saharan States (CEN-SAD, 29), the East African Community (EAC, 5), the Economic Community of Central African States (ECCAS, 10), the Economic Community of West African States (ECOWAS, 15), the Intergovernmental Authority on Development (IGAD, 8) and the Southern Africa Development Community (SADC, 15).

**Regional Economic Communities** - The African Union (AU) recognizes 8 Regional Economic Communities (RECs, number of members, see also at: <https://au.int/en/organs/recs>, access Feb 2018):

**Arab Maghreb Union (AMU, 5):** Algeria, Libya, Mauritania, Morocco, Tunisia

**Common Market for Eastern and Southern Africa (COMESA, 19):** Burundi, Comoros, DR Congo, Djibouti, Egypt, Eritrea, Ethiopia, Kenya, Libya, Madagascar, Malawi, Mauritius, Rwanda, Seychelles, Sudan, Swaziland, Uganda, Zambia, Zimbabwe

**Community of Sahel-Saharan States (CEN-SAD, 29):** Benin, Burkina Faso, Cabo Verde, Central African Republic, Chad, Comoros, Côte d'Ivoire, Djibouti, Egypt, Eritrea, Gambia, Ghana, Guinea, Guinea Bissau, Kenya, Liberia, Libya, Mali, Mauritania, Morocco, Niger, Nigeria, São Tomé and Príncipe, Senegal, Sierra Leone, Somalia, Sudan, Togo, Tunisia

**East African Community (EAC, 5):** Burundi, Kenya, Rwanda, Uganda, UR of Tanzania

**Economic Community of Central African States (ECCAS, 10):** Angola, Burundi, Cameroon, Central African Republic, Chad, Congo, DR Congo, Equatorial Guinea, Gabon, São Tomé and Príncipe

**Economic Community of West African States (ECOWAS, 15):** Benin, Burkina Faso, Cabo Verde, Côte d'Ivoire, Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone, Togo

**Intergovernmental Authority on Development (IGAD):** Djibouti, *Eritrea (check status on website)*, Ethiopia, Kenya, Somalia, South Sudan, Sudan, Uganda

**Southern Africa Development Community (SADC, 15):** Angola, Botswana, DR Congo, Lesotho, Madagascar, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, UR of Tanzania, Zambia, Zimbabwe.

Moreover, for each dyad we calculate variables that can vary over time. The co-publication is the count of the publications co-authored by two countries at time  $t$  (i.e.  $copub_{ijt}$ ) and past co-publications is the amount of publication co-authored by the two countries in the previous three years (i.e.  $pastcopub_{ijt}$  and its logarithm:  $\log pastcopub_{ijt}$ ). Besides, we calculate for the past co-publications the share involving a non-African country. Finally, we count separately the number of past common partners from developed countries or political North ( $\lognorth_{ijt}$ ), of past common partners from colonial North ( $\logcolony_{ijt}$  - France, United Kingdom, the Netherlands and Portugal) and (only) from Africa ( $\logafrique_{ijt}$ ).

---

<sup>7</sup> If the distance is zero then proximity is assumed to be equal to 1.

Table 2 – Descriptive statistics

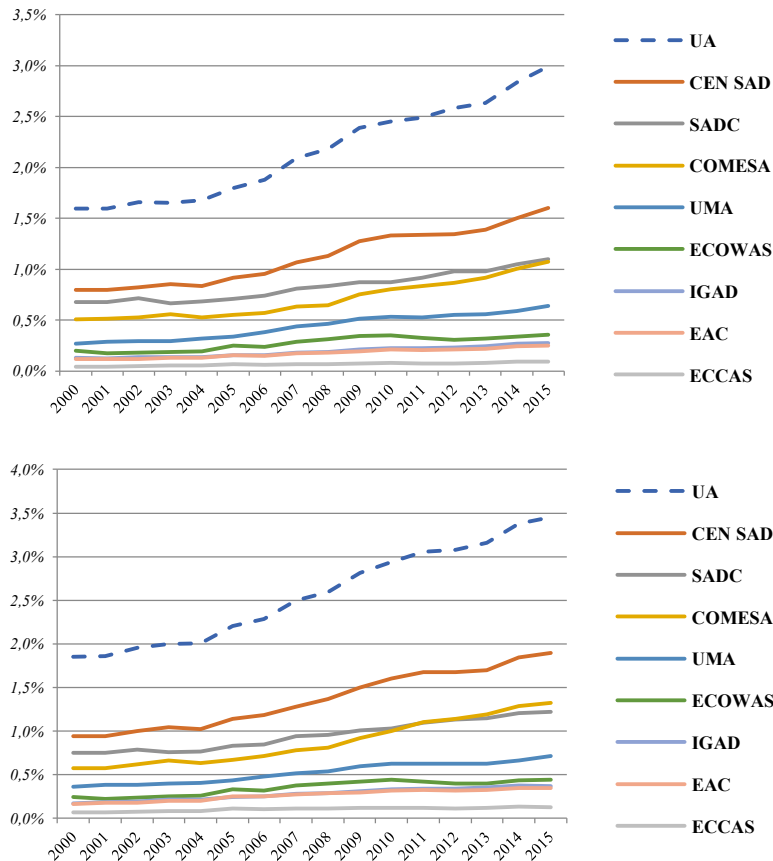
Variable	Nb obs	Mean	Std Dev	Minimum	Maximum
<b>lang_off</b>	1431	0.44	0.50	0.00	1.00
<b>comcol</b>	1431	0.26	0.44	0.00	1.00
<b>dummy_cer</b>	1431	0.51	0.50	0.00	1.00
<b>cer</b>	1431	0.68	0.76	0.00	3.00
<b>CEN_SAD</b>	1431	0.28	0.45	0.00	1.00
<b>COMESA</b>	1431	0.11	0.31	0.00	1.00
<b>EAC</b>	1431	0.01	0.08	0.00	1.00
<b>ECCAS</b>	1431	0.03	0.17	0.00	1.00
<b>ECOWAS</b>	1431	0.07	0.26	0.00	1.00
<b>IGAD</b>	1431	0.01	0.12	0.00	1.00
<b>SADC</b>	1431	0.07	0.26	0.00	1.00
<b>AMU</b>	1431	0.09	0.29	0.00	1.00
<b>dist_tech</b>	1431	0.28	0.30	0.00	1.00
<b>copub</b>	20034	1.99	8.72	0.00	287.0
<b>logdistw</b>	20034	8.04	0.65	5.09	9.19
<b>logpub_i</b>	20034	5.28	2.34	0.00	10.36
<b>logpub_j</b>	20034	4.91	2.26	0.00	10.36
<b>logpastcopub</b>	20034	0.63	1.05	0.00	6.43
<b>logafrique</b>	20034	0.96	1.35	0.00	3.78
<b>logonlyafrique</b>	20034	0.38	0.67	0.00	3.04
<b>logcolony</b>	20034	0.51	0.69	0.00	1.61
<b>lognorth</b>	20034	0.77	1.04	0.00	2.30

### 3.3 Main features of the publications in RECs

Originally proposed to support the economic integration between members and the African Economic Community (AEC), RECs show important heterogeneity in terms of size, roles and abilities to structure the related ‘regional spaces’. The figures 1 confirm that the continent and most RECs have slightly improved their scientific production. As illustrated by the dashed lines, Africa’s relative share has almost doubled in 15 years to about 3% and 3.5% of world scientific publications and co-publications in 2015, respectively.

However, Africa still represents a very low share of the world scientific production, reflecting mainly the low number of researchers observed in this world’s region (UNESCO 2015), the limited funding for STI activities as well as other economic and structural features, e.g. economic development stages, priority research fields, institutional and policy support for science (see for instance Irikefe et al 2011, AOSTI 2013).

**Figures 1 – Share of RECs in world total publication and co publications**  
*Publications and Co-publications*



Noteworthy, it seems that this increase has come with a higher quality of overall Africa's publications (Blom et al 2016, Dosso et al 2017). As suggested by the figures 1, the greatest share of the upward trends in scientific publications seems to take place since the mid-2000s. This period also marks the adoption of the African Science and Technology Consolidated Plan of Action (CPA, 2006) as well as the establishment of the 2008 Protocol on Relations between the AU and the RECs and the Joint Africa-EU Strategy (JAES, end 2007)<sup>8</sup>.

Table 3 breaks down the scientific production at RECs levels for the period 2000-2015. It provides the number of distinct publications, the share of co-publications, the top and bottom country and their shares in the REC's total publications and the variability of countries' shares within the REC.

<sup>8</sup> See at [https://ec.europa.eu/europeaid/regions/africa/continental-cooperation/joint-africa-eu-strategy\\_en](https://ec.europa.eu/europeaid/regions/africa/continental-cooperation/joint-africa-eu-strategy_en)

**Table 3 – Main features of publications, Regional Economic Communities (number of countries)**

RECs	Number of publications	Share of inter. Copubl.	Top country	Bottom country	Top share	Bottom share	Coef. Var. (within)
CEN SAD (29)	240,704	49.5%	Egypt	Sao Tome	38.9%	0.01%	221%
COMESA (19)	151,834	55.2%	Egypt	Comoros	61.7%	0.04%	251%
EAC (5)	37,610	80.2%	Kenya	Burundi	49.8%	0.71%	94%
ECCAS (10)	14,297	79.2%	Cameroon	Sao Tome	61.0%	0.28%	175%
ECOWAS (15)	58,047	52.7%	Nigeria	Cape Verde	55.0%	0.26%	188%
IGAD (8)	40,635	75.15%	Kenya	Somalia	46.1%	0.15%	116%
SADC (15)	174,986	49.8%	South Africa	Lesotho	82.7%	0.17%	305%
AMU	94,612	49.85%	Tunisia	Mauritania	42.9%	0.38%	90%
<i>All sample (52 countries)</i>	<i>460,324</i>	<i>48.11%</i>	<i>South Africa</i>	<i>Sao Tome</i>	<i>31.4%</i>	<i>0.009%</i>	<i>258%</i>

Over all, the 52 African countries considered have produced about 460,324 distinct publications including 229,479 international co-publications. The important variability within the RECs mainly relates to the heterogeneous characteristics and endowments of the national science and research systems of the countries. This somehow also raises the issue of the relevance or ability of some of these geographical spaces to favour the regional scientific integration(s). In terms of countries, South Africa leads the scientific production on the continent with about 31% of the total of publications on the period 2000-2015. As also underlined by previous works, other major publishers include Egypt, Nigeria and Cameroon, and Kenya (see also in Tijssen 2007, Pouris and Ho 2014). In terms of fields, the continent's publications are dominated by medical research (35%). Representing about 15% each in African publications, the fields of Fundamental biology and Applied biology-ecology, come as the second most important ones. This pattern seems to hold true when looking at individual RECs, but for the Arab Maghreb Union in which the second top areas are Engineering and Chemistry, accounting for one fifth of the RECs publications each.<sup>9</sup>

#### 4 Empirical results

Over all the continent, South Africa, Egypt, Tunisia, Nigeria and Morocco, respectively, emerge as the most active countries in terms of international co-publications. Among these countries, the role of South Africa and Nigeria as possible enablers of further integration has already been suggested by Toivanen and Ponomariov (2011). Outside Africa, the top five partners include the United States, France, UK, Germany and Saudi Arabia.

<sup>9</sup> See Sooryamoorthy (2018) who exploits Web of Science for a detailed mapping of African publications at the country and field levels on the period 2000-2015.

Tables 4 – Estimates of the baseline model

Empirical Standard Error Estimates						
Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr >  Z
<b>Intercept</b>	-28.7030	5.9153	-40.2967	-17.1093	-4.85	<.0001
<b>logpub_i</b>	0.2511	0.0105	0.2305	0.2716	23.92	<.0001
<b>logpub_j</b>	0.2540	0.0120	0.2305	0.2776	21.12	<.0001
<b>logdistw</b>	-0.3056	0.0237	-0.3520	-0.2591	-12.90	<.0001
<b>logpastcopub</b>	0.7799	0.0195	0.7416	0.8181	39.95	<.0001
<b>comcol</b>	0.1614	0.0365	0.0899	0.2330	4.42	<.0001
<b>dist_tech</b>	-2.1269	0.1724	-2.4647	-1.7891	-12.34	<.0001
<b>lang_off</b>	0.1506	0.0367	0.0788	0.2225	4.11	<.0001
<b>year</b>	0.0135	0.0030	0.0077	0.0193	4.58	<.0001

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr >  Z
<b>Intercept</b>	-28.3506	6.0067	-40.1236	-16.5777	-4.72	<.0001
<b>logpub_i</b>	0.2557	0.0112	0.2337	0.2777	22.77	<.0001
<b>logpub_j</b>	0.2607	0.0130	0.2353	0.2862	20.10	<.0001
<b>logdistw</b>	-0.3683	0.0323	-0.4317	-0.3049	-11.39	<.0001
<b>logpastcopub</b>	0.7708	0.0214	0.7288	0.8128	35.97	<.0001
<b>comcol</b>	0.1847	0.0385	0.1092	0.2602	4.80	<.0001
<b>dist_tech</b>	-2.0698	0.1736	-2.4100	-1.7296	-11.92	<.0001
<b>lang_off</b>	0.1376	0.0370	0.0650	0.2102	3.71	0.0002
<b>dummy cer</b>	-0.1614	0.0376	-0.2352	-0.0877	-4.29	<.0001
<b>year</b>	0.0136	0.0030	0.0077	0.0195	4.52	<.0001

Table 4 and 5 provide the main results of our empirical application. Table 4 gives the estimates for the baseline model and Table 5 includes the effects of belonging to a specific regional community. Our results confirm that the intensity of co-publication between two countries increases with their number of publications or scientific masses, and decreases with their geographical distances.

Similarly, the further apart two countries are in terms of scientific profiles, the lower the intensity of co-publications can be expected. ‘Speaking the same language’ also matters as suggested by earlier descriptive and mapping analyses (Mègnigbeto 2013, Adams et al 2014). The positive effect of past co-publications means that two countries are more likely to collaborate if they had previously collaborated. Indeed, future collaborations may benefit from existing communication channels (and even established routines) and frameworks (institutional, scientific, etc.). At the policy level, this might in turn be leveraged to further support initial collaborations between non-partner African countries with similar research interests (scientific proximity matters) as this might facilitate the establishment of sustained collaborative patterns.

Moreover, the analysis suggests that past colonial ties (France, United Kingdom, the Netherlands and Portugal) still influence significantly the intensity and patterns of co-publications of the African continent. Among them, France and UK rank as the second and third top non-African partners for co-publication. In the next subsection, we explore further

the change overtime of the effect of the colonial ties as well as the importance of the geographical distance in shaping African collaborative science.

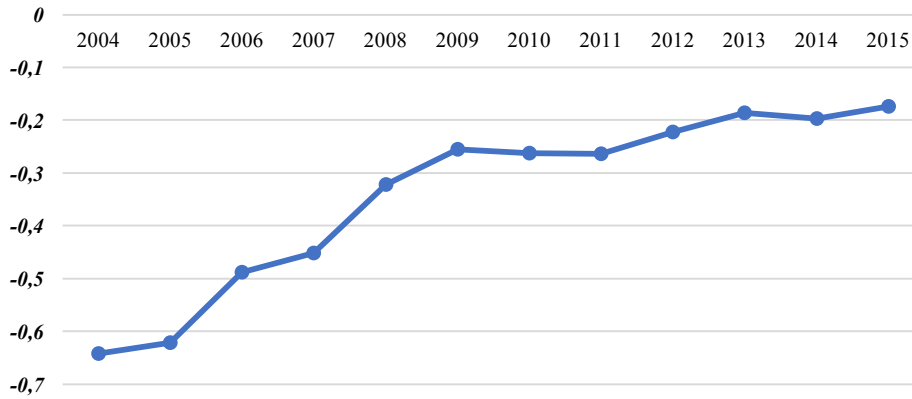
**Table 5 – Estimates of Regional Community (ies) effect**

<b>Analysis Of GEE Parameter Estimates</b>						
<b>Empirical Standard Error Estimates</b>						
<b>Parameter</b>	<b>Estimate</b>	<b>Standard Error</b>	<b>95% Confidence Limits</b>		<b>Z</b>	<b>Pr &gt;  Z </b>
<b>Intercept</b>	-32.0372	5.8866	-43.5747	-20.4998	-5.44	<.0001
<b>logpub_i</b>	0.2805	0.0125	0.2559	0.3051	22.36	<.0001
<b>logpub_j</b>	0.2869	0.0128	0.2618	0.3120	22.42	<.0001
<b>logdistw</b>	-0.2864	0.0364	-0.3578	-0.2150	-7.86	<.0001
<b>logpastcopub</b>	0.7217	0.0215	0.6796	0.7638	33.60	<.0001
<b>comcol</b>	0.2042	0.0376	0.1305	0.2779	5.43	<.0001
<b>dist_tech</b>	-2.0749	0.1614	-2.3913	-1.7585	-12.85	<.0001
<b>lang_off</b>	0.1689	0.0388	0.0929	0.2449	4.36	<.0001
<b>CEN_SAD</b>	-0.1309	0.0513	-0.2316	-0.0303	-2.55	0.0108
<b>COMESA</b>	-0.1783	0.0472	-0.2709	-0.0857	-3.77	0.0002
<b>EAC</b>	0.0780	0.1037	-0.1253	0.2813	0.75	0.4523
<b>ECCAS</b>	0.0938	0.0769	-0.0569	0.2445	1.22	0.2224
<b>ECOWAS</b>	0.2241	0.0779	0.0715	0.3767	2.88	0.0040
<b>IGAD</b>	0.0441	0.0687	-0.0906	0.1788	0.64	0.5214
<b>SADC</b>	0.1041	0.0550	-0.0036	0.2118	1.89	0.0583
<b>UMA</b>	-0.4095	0.0650	-0.5368	-0.2821	-6.30	<.0001
<b>year</b>	0.0150	0.0029	0.0092	0.0207	5.10	<.0001

Focusing on the regional or RECs dimensions, the negative and significant impact of the dummy variable (Table 4) may reflect the high diversity across and within the communities in terms of scientific capability and maturity, and also a limited structuring role of some RECs (Table 5), once other determinants are controlled for (see paragraphs above). This is consistent with the high standard deviation ranges in the scientific production capacity of RECs (Table 3). Rather than concluding that some regions would hamper the development of scientific co-publications, our results suggest that the role of some RECs or regions as integrative scientific areas can be questioned. At the same time, this can stimulate further research on the specific socio-institutional factors, that underpin scientific co-authorship in areas such as ECOWAS or SADC.

Figure 2 reports the coefficients of geographical distance in each of the 12 regressions run for the “dynamic” analysis. The coefficients are reported only if they are significant (this holds true overtime).

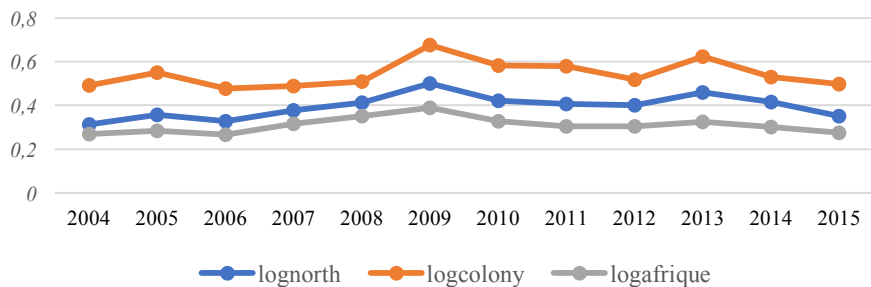
**Figure 2 – Geographical distance coefficients**



Expectedly, the geographical distance shows a decreasing effect. As previously mentioned, the reduced importance of the geographical distance effect can signal further integration or greater international co-authorship networks. Besides, the possible impacts of greater institutional and policy support and availability of funding (or grants) possibilities for collaborative projects, this trend can also be related to the few improvements in the last decades in the air traffic in Africa as well as the higher penetration of internet and related software and applications. All together they facilitate for instance regular (face-to-face) exchanges, collective meetings (conferences, workshops) and easier communications for international scientific collaborations and co-authorship.

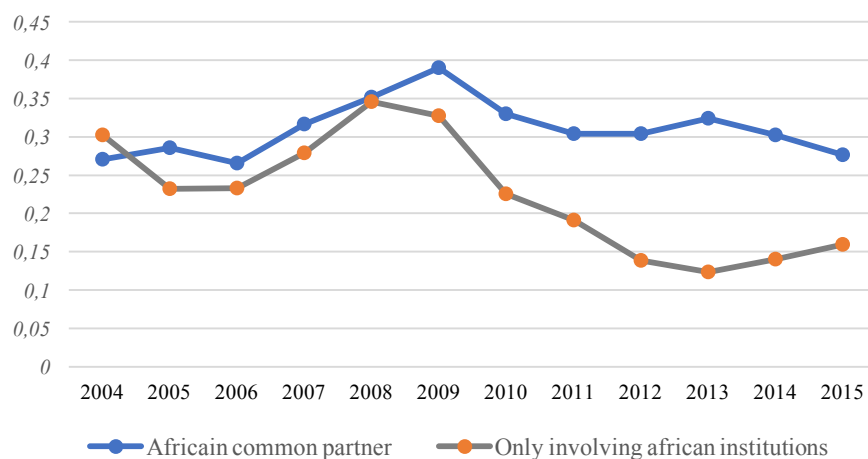
Figure 3a and 3b report the effect over time of having specific common partners (log of number of countries in common). Over all they confirm that having a third common partner always increase the co-publication intensity. Figure 3a shows that this holds true (i) when the third partner is an African country (grey line), (ii) even more when the third partner is from the political North (e.g. developed countries, blue line), and (iii) most importantly when this third partner is a former colony (i.e. France, United Kingdom, the Netherlands and Portugal, orange line).

**Figure 3.a – Common partner effect**



In Figure 3b, the grey line reports the common African partners when only the co-publications produced by African countries are considered.

**Figures 3.b – Common partner effect**



The difference between the two lines tells how much involving out-of-the-continent partners is important for establishing collaboration within the continent. This effect seems to increase over time or ‘having an external non-African partner’ is increasingly important. This confirms the important influence of the global science on the publication patterns of the continent or that increasing convergence is taking place. In particular, it seems that this trend has been more accentuated since the end of the 2000s.

## 5 Conclusion

The rise of an ‘African’ collaborative science is still underway. While the continent shows a few signals of increased local dynamics and continental upward trends in the production of scientific publications, national and regional knowledge and research systems still need fundamental institutional, social and financing frameworks. Our results tend to confirm that African scientific collaborations are increasingly influenced by the presence of external non-African partners, especially through the legacy of common past colonial ties, which is not decreasing overtime. This is also consistent with the presence of important research centres or networks originating, for instance, from France or the UK on the continent. Yet, our study cannot confirm whether the current shares are too “high” or denote a clear situation of dependence (AOSTI 2014); country specific features should play a key role.

Moreover, the study also supports a multidimensional approach to strengthen the regional scientific integration. Indeed, further actions to improve the physical (air, road traffic and related infrastructure) and digital connectivity (access to fast internet, to international and African knowledge sources, networks etc.) and social links across the continent (e.g. regular languages training and educational and cultural exchanges programs) may be explicitly

combined with existing and future policy plans or strategies for a collaborative science within Africa. Together, they contribute to attenuate the negative effects of geographical distance(s).

Furthermore, the findings question the structuring and integrative role of some RECs with respect to scientific collaborations across the member countries. Still, a few communities such as ECOWAS and SADC, or West and South geographic regions, seem to yield a positive influence on the intensity of co-publications between their members, once the geographical distance and the language and the scientific proximity are accounted for (among other). Further studies should be undertaken in order to identify the specific drivers (e.g. role of a specific African country-hub) or potential ‘good-practices’ in these regional spaces.

Finally, it should be mentioned that our study suffers of some limitations. Even if bibliometric studies can provide useful results and policy recommendations, it is worth pointing out that our analysis is only an indirect assessment of the policy initiatives that have been promoted both nationally and internationally in the region over a long time period. Bibliometric-based measures concern only production and dissemination of knowledge, grasping in this way only the output of process. Finally, we rely on a somehow biased data set (Thompson Reuters - WoS), because it includes mainly international journals published in English. Therefore, we cannot capture fully the geography of regional collaborations, in particular if domestic co-publications are issued in African journals that are not listed in Web of Science.

## References

- Adams, J., Gurney, K., Hook, D. and Leydesdorff, L. (2014), International collaboration clusters in Africa. *Scientometrics*, 98(1), 547–556.
- Adams, J., King, C. and Hook, D. (2010), *Global research report: Africa*. Leeds: Thomson Reuters.
- African Observatory of Science, Technology and Innovation, AOSTI (2014), *Assessment of Scientific Production in the African Union, 2005–2010*, Malabo: The AOSTI.
- , AOSTI (2013), Science, technology and innovation policy-making in Africa: An assessment of capacity needs and priorities. Working Paper No. 2. Malabo, Equatorial Guinea. Malabo: The AOSTI.
- African Union Commission, AUC (2014), *AU Strategy for Science, Technology and Innovation for Africa 2024, STISA-2024*, Addis Ababa: African Union.
- Arocena, R. and Sutz, J. (2001), Changing knowledge production and Latin American universities, *Research Policy*, 30(8), 1221–1234.
- Arvanitis, R., Waast, R. and Gaillard, J. (2000). Science in Africa: A bibliometric panorama using PASCAL database. *Scientometrics*, 47, 457–473.
- Bergstrand J.H. (1985), The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence, *The Review of Economics and Statistics*, 67 (3), 474–481.
- Blom, A., Lan, G. and Adil, M. (2016). *Sub-Saharan African science, technology, engineering, and mathematics research: A decade of development*. Washington, D.C.: International Bank for Reconstruction and Development, The World Bank.
- Boschma R. A. (2005), Proximity and innovation: a critical assessment, *Regional Studies*, 39 (1), 1–14.

- Boshoff, N. (2009), Neo-colonialism and research collaboration in Central Africa. *Scientometrics*, 81(2), 413–434.
- Boshoff, N. (2010), South-South research collaboration of countries in the Southern African Development Community (SADC). *Scientometrics*, 84(2), 481–503.
- Cameron, A.C. and Trivedi, P. K. (2013), *Regression Analysis of Count Data*, 2nd edition, Econometric Society Monograph No.53, Cambridge University Press, 1998 (566 pages).
- Cassi, L., Morrison, A. and Rabellotti, R. (2015), Proximity and scientific collaboration: Evidence from the global wine industry, *Tijdschrift voor economische en sociale geografie*, 106(2), 205–219.
- Confraria, H., and Godinho, M. M. (2015). The impact of African science: A bibliometric analysis. *Scientometrics*, 102, 1241–1268.
- Cusmano, L., Morrison, A. and Rabellotti, R. (2010), Catching-up trajectories in the wine sector: a comparative study of Chile, Italy and South Africa, *World Development*, 38(11), 1588-1602.
- Deutsches Institut für Entwicklungspolitik / German Development Institute, DIE (2010), *African Developments: Higher Education and Research Capacities: Africa in the Globalising Knowledge Society*. Briefing paper 13/2010.
- Dooso, M. et al. (2017), Research and Innovation, pp. 217–226, in European Commission, Joint Research Centre (2017), *Science for the AU-EU Partnership: building knowledge for sustainable development*, Publications Office of the European Union, Luxembourg, 320 pages.
- Frenken, K., Hardeman, S. and Hoekman, J (2009), Spatial scientometrics: towards a cumulative research program, *Journal of Informetrics*, 3(3), 222–232.
- Guns, R. and Wang, L. (2017), Detecting the emergence of new scientific collaboration links in Africa: A comparison of expected and realized collaboration intensities, *Journal of Informetrics*, 11(3), 892–903.
- Hardeman, S., Frenken, K., Nomaler, Ö. and Ter Wal, A. J. (2015), Characterizing and comparing innovation systems by different ‘modes’ of knowledge production: A proximity approach. *Science and Public Policy*, 42(4), 530–548.
- Hilbe, J. (2011). *Negative Binomial Regression*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511973420
- Hoekman, J., Frenken, K. and Tijssen, R.J.W. (2010), Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe”, *Research Policy*, 39(5), 662-673.
- Irikefe, V., Vaidyanathan, G., Nordling, L., Twahirwa, A., Nakkazi, E. and Monastersky, R. (2011). Science in Africa: View from the front line. *Nature*, 474, 556–559.
- Katz, J.S. and Martin B.R. (1997), What is research collaboration? *Research Policy*, 26 (1), 1-18.
- Mazzoleni, R. and Nelson, R. R. (2007). Public research institutions and economic catch-up. *Research policy*, 36(10), 1512-1528.
- Mêgnigbêto, E. (2013), International collaboration in scientific publishing: the case of West Africa (2001–2010). *Scientometrics*, 96(3), 761–783.
- Mongeon, P. and Paul-Hus, A. (2016), The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228.
- Montobbio, F. and Sterzi, V. (2013), The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations. *World Development*, 44, 281–299.

- Mowery, D. C. (1992), The US national innovation system: Origins and prospects for change. *Research Policy*, 21(2), 125–144.
- New Partnership for Africa's Development, NEPAD (2006), *Africa's Science and Technology Consolidated Plan of Action*. Office of S&T of the New Partnership for Africa's Development (NEPAD), Pretoria, South Africa.
- Picci, L. (2010), The internationalization of inventive activity: A gravity model using patent data, *Research Policy*, 39, 1070–1081
- Ponds, R.H.F., Oort, F.G. van and Frenken, K. (2007), The geographical and institutional proximity of research collaboration, *Papers in Regional Science*, 86, 423–443.
- Pouris, A. and Ho, Y.-S. (2014). Research emphasis and collaboration in Africa. *Scientometrics*, 98, 2169–2184.
- Pouris, A. and Pouris, A. (2009). The state of science and technology in Africa (2000–2004): A scientometric assessment. *Scientometrics*, 79, 297–309.
- Rafols I., Ciarli T., and Chavarro, D. (2016), Underreporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice, Presentation <https://contestedagronomy.files.wordpress.com/2016/03/10-rafols-database-biases-in-rice-research.pdf>
- Scheringell, T., Hu, Y. (2011), Collaborative knowledge production in China. Regional evidence from a gravity model approach, *Regional Studies*, 45 (6), 755–772.
- Sooryamoorthy, R. (2018), The production of science in Africa: an analysis of publications in the science disciplines, 2000–2015, *Scientometrics* (2018). <https://doi.org/10.1007/s11192-018-2675-0>
- Tijssen, R. J. W. (2007). Africa's contribution to the worldwide research literature: New analytical perspectives, trends, and performance indicators. *Scientometrics*, 71, 303–327.
- Toivanen, H. and Ponomariov, B. (2011). African regional innovation systems: Bibliometric analysis of research collaboration patterns 2005–2009. *Scientometrics*, 88, 471–493.
- United Nations Educational, Scientific and Cultural Organization, UNESCO (2015), *UNESCO Science Report: towards 2030*, Paris, France, retrieved from [http://en.unesco.org/unesco\\_science\\_report](http://en.unesco.org/unesco_science_report)
- World Bank (2014), *A decade of development in Sub-Saharan African science, technology, engineering and mathematics research*. Washington, DC: World Bank Group. <http://documents.worldbank.org/curated/en/237371468204551128/A-decade-of-development-in-sub-Saharan-African-science-technology-engineering-and-mathematics-research>