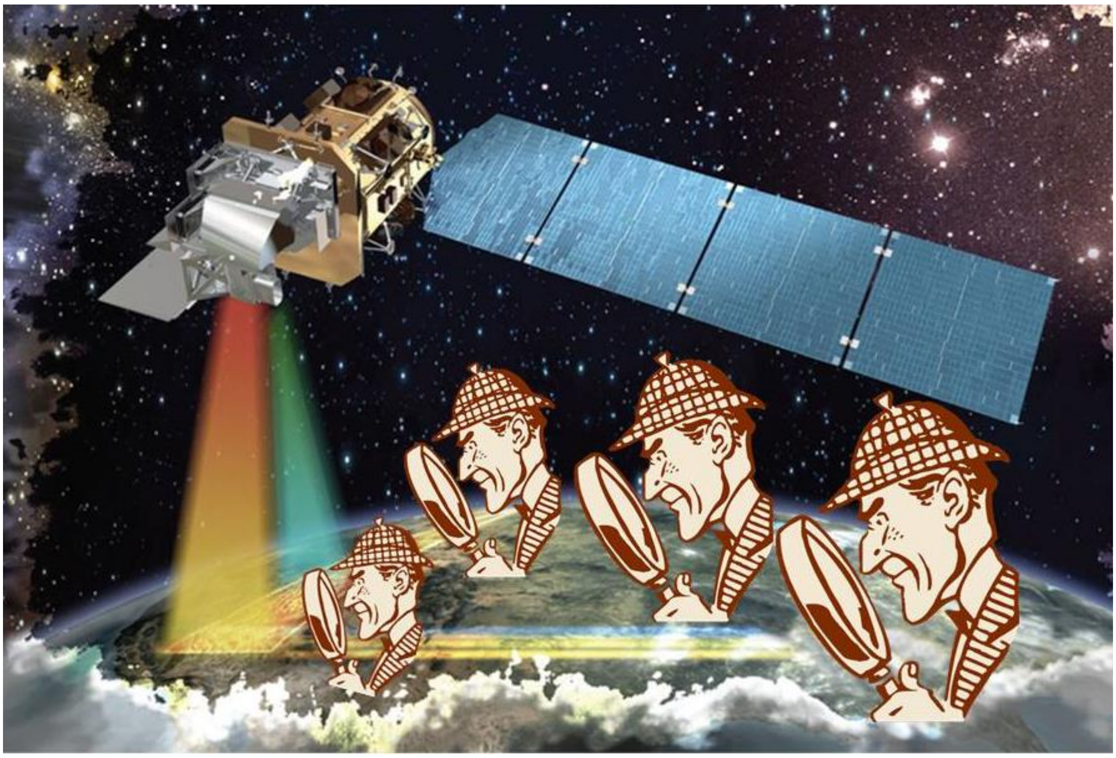
Centre for Geo-Information

Thesis Report GIRS-2014-41

IMPACT OF OPEN LANDSAT DATA ON SCIENCE

Puja Mishra

March 2015



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Impact of Open Landsat Data on Science

Puja Mishra

Registration number 840511574110

Supervisors:

Prof. Dr. Arnold K. Bregt

Dr. Lukasz Grus

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

6th March 2015
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2014 -41
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgement

At the completion of my Master in Geo-information Science at internationally renowned Wageningen University, I would like to take this opportunity to acknowledge the personalities and organizations whose support and contribution endured valuable for me to accomplish my dream.

Firstly, I would like to express my sincere gratitude to Prof. Dr. Arnold Bregt and Dr. Lukasz Grus for their admiring contribution, valuable suggestions, constructive comments and full support during the research.

Secondly, I would like to thank my study advisor Willy ten Haaf and all Geo-information Science and remote sensing instructors of Wageningen University for their respective support and contribution to make the study successful.

Last but not the least I would like to thank all my friends and family members in Nepal; specially my parents and my beloved husband for the inspiration and encouragement at every stage of my life.

Abstract

The tendency of open data has been grown exponentially all over the world in the last decade. Major advantage of open data is that, it provides a major movement towards knowledge sharing, without any cost and permission barriers. The transparency, reusability, availability and reproducibility of open data can increase the scientific process towards the societal benefit. In order to increase transparency, reusability, participation and/or government efficiency, governments around the world start to pay attention to the opening of their data. In line with those trends, U. S. Geological Survey announced in 2008 the opening of Landsat data to make all the Landsat data archive publicly and freely available. After this open data policy, users are allowed to access, freely download and use Landsat imagery products. It is very likely that this initiative have had a significant impact on society, especially on science. However, there is no comprehensive study found that can assess the impact of open Landsat data on science.

In this context, this research aims at assessing the impact of open Landsat data on science. Based on literature review, the possible research methods are investigated to meet the assessment requirements of Landsat data archive. Finally, a fitness for use criteria is used to select the most suitable research method to measure the impact of open Landsat data on science. This assessment uses Systematic Literature Review (SLR) method which focuses on the six indicators to measure the effects of open Landsat: 1) number of publication in which Landsat data used per year; 2) number of publication in different document type; 3) number of publication in different languages; 4) number of publication that use Landsat data per country per year; 5) number of publication in which Landsat data used in different disciplines per year; and 6) number of images used in publication. According to these indicators, search query is constructed to extract the scientific publications in Scopus digital data source from 1972 to 2014. Research tendency was investigated by statically analysing the distribution of publication year, authors, authors with affiliation, title, source title, citation, published country, document type, discipline and language type. After opening Landsat archive from 2008, the result shows that the trend of Landsat use has significantly increased with increasing number of articles (69.43%) and books (0.65%). While production of the conference papers (29.06%) and reviews (0.86%) were relatively reduced almost by half. English language of those journals took the majority of all the publication however after free data distribution policy the publications number of publications in other national languages have slightly increased to languages such as Arabic, Chinese, Korean, Persian, Thai, Ukrainian, Turkish, Spain, Portuguese, Italian, and Finnish, Estonia, Serbian. Similarly, the result shows that after Landsat opening archive the Asian region (mainly China) became the biggest user of Landsat data. Overall, it has also been found that agricultural, ecological/ecosystem science/ management, forest science/management, water resources, land use/land cover, emergency/disaster management/hazard insurance, urban planning and development were dominant application areas of Landsat. Before opening Landsat archive, the scientists were limited to the use of data for large areas. It has been observed that the use of multiple Landsat images using time series and near real time analysis for earth observation and monitoring in large areas have increased significantly. To evaluate the search result of Scopus using SLR method, 3% of Landsat publications (501 in total) from 2002 to 2014 were randomly studied on the basis of abstract, title and keywords. The validation accuracy of five indicators number of publication in which Landsat data used per year; number of publication in different document type; number of publication in different languages; number of publication that use Landsat data per country per year; number of image used in publication were recorded with the highest accuracy (100%), whereas Landsat publication in different discipline was found to have lowest overall accuracy (68.17%). The results suggest that SLR method can help the researchers to assess the overview of global impact of open data on science.

Keywords: Open data, Landsat, science, impact assessment, systematic literature review

Table of Contents

List of Figures

List of Tables

List of abbreviation

EOSAT: Earth Observation Satellite

EROS: Earth Resources Observation and Science

ERTS: Earth Resources Technology Satellite

ETM+: Enhanced Thematic Mapper Plus

EU: European Union

INGOs: International Non-Governmental Organizations

LDCM: Landsat Data Continuity Mission

MSS: Multispectral Scanner

NASA: National Aeronautics and Space Administration

NGOs: Non-Governmental Organizations

NOAA: National Oceanic and Atmospheric Administration

NSPI: Neighbourhood Similar Pixel Interpolator

OLI: Operational Land Imager

REDD+: Reduce emissions from deforestation and forest degradation

RBV: Return Beam Vidicon

SLC: Scan-line corrector

SLR: Systematic literature review

SSR: Solid State Recorders

SWIR: Short wave infrared

TDRSS: Tracking and Data Relay Satellite System

TIRS: Thermal Infrared Sensor

TM: Thematic mapper

TWTA: Traveling wave tube amplifier

USGS: United State of Geological Survey

1. Introduction

1.1. Open Geo-data

The world is heading towards digitalization, where digital data play a pivotal role for information and communication technologies. More than 60% of digital data are related to location on earth surface (spatial data), which are known as geographical data or geo-data (Hahmann and Burghardt 2013). Due to lack of funding and limited opportunity for the general public to reach it, only 10% of geo-data is effectively exploited (Čeh, Smole et al. 2004).

The concept of open data has grown rapidly throughout the world in the last decade. Open data is an emerging term which describes the process of reusing and redistributing data. Major advantage of open data is that, it provides a major movement towards knowledge sharing, without any cost and permission barriers. Information of governmental institutions, business/industry, citizens, science, education, Non-Governmental Organizations (NGOs) and International Non-Governmental Organizations (INGOs) are presented in open data (Gurstein 2011). Thus, regional, national and local governments around the world are eager to adopt an open data strategies to increase the government efficiency, transparency, accountability, citizen participation and economic opportunity (Bauer and Kaltenböck 2011, Huijboom and Van den Broek 2011). In line with the growth of open data, it is also becoming more prominent that if countries want to maximise their gain from open data, their governments need to go beyond simply publishing data on a website. Governments also need to be: a) supplier of the data that user need; b) leader in improving policies, and promoting to the policies through inline agencies for significant use of open data; c) catalysts for fostering a thriving ecosystem of data users, coders and application developers, and incubating new data-driven businesses; and d) and d) users of using open data themselves at national, regional and local administrative levels so as to identify barriers related to the proper usage of the open data and thereby providing guidelines to overcome such barriers while using data within government. The participation of the government on fostering the (re-)use of open data can significantly improve public services. This is supported by the research done by World Bank which argues that innovative approached of reusability of data may improve public services and government efficiency (WorldBank 2014).

Current schemes for geo-data are based on two fold formats: a) vector data and b) raster data (Chang 2010). The vector data format represents the world using points, lines, and polygons (Bregt and Grus 2014). This format is useful for storing data that has discrete boundaries, such as country borders, land parcels and streets (Scarletto 2014). Several initiatives have been launched throughout the world to make these data open. Some of these are: Commons of Geographic Data project led by Harlan Onsrud at the University of Maine, OpenStreetMap provide a growing body of free annotated vector mapping information extracted from contributed GPS traces and the Open Knowledge Foundation supported the Public geo-data campaign to raise awareness of data access issues in the new European Spatial Data Infrastructure Legislation (Malenovský, Rott et al. 2012).

Raster data formats are grid cell based data. Satellite missions are one of the main sources for the collection of raster data throughout the globe. These satellite missions are classified into three spatial resolution categories: low (larger than 30 m), medium (2-30 m) and high resolution imagery (under 2 m). Most of the low and medium categories of satellite images have been acquired (Roy, Wulder et al.

2014) by USGS and NASA, such as Landsat and MODIS satellite data. Recently, a high resolution French satellite called SPOT has made its optical earth observation data archive open and distributes freely to the users (Selding 2014). Furthermore, the European Union (EU) has also announced their future space mission program (Sentinel 1, 2 and 3), and will also follow a free data policy (Malenovský, Rott et al. 2012).

Compared to other satellite missions, Landsat missions offer the following advantages: (1) provide global coverage on a regular basis; (2) available for free; (3) image archive reaches back to 1972; 4) provide responsive delivery of data (Irish 2000). Despite these advantages, limited studies have been conducted to understand the impact of open Landsat imagery to the society particularly on science. Therefore, the major objective of this research is to assess the impact of open Landsat data on science.

1.2. Landsat mission

The Landsat satellite mission is a series of earth-observing satellite missions, jointly achieved by United State of Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) (Wulder, White et al. 2011). The Landsat mission has collected and provided the global space-based earth observation data continuously since 1972. The Figure 1 shows the timeline showing Landsat missions of USGS. Each Landsat satellite has a 5-year life span. Therefore, the Landsat satellite was launched one after another with a gap of 2-3 years. The era of Landsat satellite mission started in 1972, where the first Landsat satellite was called Landsat 1 (or Earth Resources Technology Satellite, ERTS-1). Until now, 8 Landsat satellites have been launched (USGS 2013, Roy, Wulder et al. 2014)., except Landsat 6, remaining 7 have successfully reached the orbit and have efficiently acquire the relevant data. Among all these satellites in the Landsat series, Landsat 8 is the latest one sent to the orbit, and was launched in 11th February, 2013, under Landsat Data Continuity Mission (LDCM). Landsat 8 acquires more than 700 scenes per day. Furthermore, USGS plans to launch Landsat 9 in 2015 (Wulder, Masek et al. 2012, USGS 2013, Roy, Wulder et al. 2014).

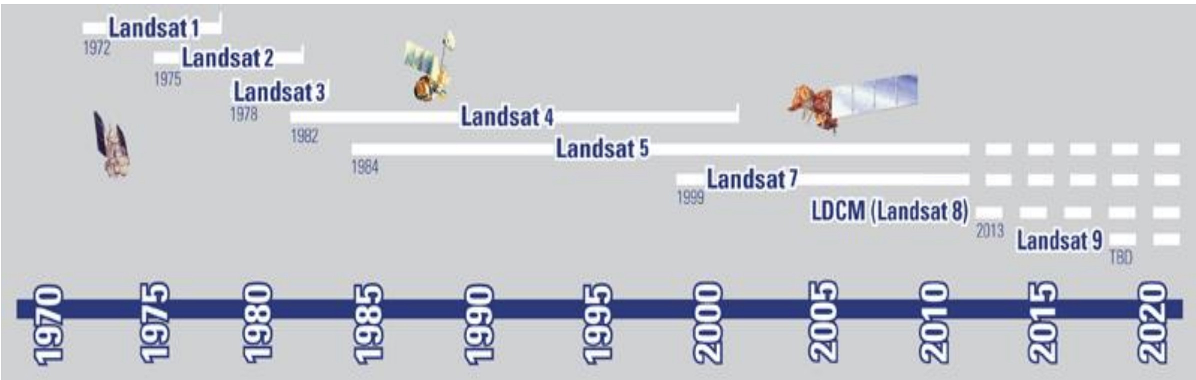


Figure 1. Timeline showing Landsat missions of USGS (Credit:(USGS 2013))

In the past, the data of Landsat mission program was very expensive; the cost per Landsat scene was US$600. Therefore, limited numbers of images were used from the huge Landsat data archives, which raised questions about the value of having such expensive Landsat missions of USGS. In order to resolve this problem and to exploit the Landsat archives effectively, NASA and USGS adopted a new free-data distribution policy in 2008 (USGS 2013). This policy has made all new and archived Landsat imaginary data freely available through USGS web portal. Immediately after the launch of this policy, huge differences were observed in data usage trend. For instance in 2001, before the open data policy, only 25,000 Landsat images were used, whilst in 2010, 2.5 million Landsat images were used (Wulder, Masek et al. 2012).

1.3. Research problem

The Landsat open data policy released in 2008, is considered as the most significant development in the history of Landsat Program (Woodcock, Allen et al. 2008). After the open data policy, users have been allowed to access and freely download Landsat data via internet. Due to the open Landsat archive, it had been anticipated that the use of Landsat data would increase significantly. Even though, the ultimate goal of Landsat mission is for the whole society, which primarily includes government, business companies, citizens, science and many more, limited research has been done to determine the actual impact of Landsat open data policy on the society (USGS 2013). It is assumed that the society has benefited a lot from the open Landsat data. However, a comprehensive study of the impact of open Landsat data has never been undertaken. At the same time, it might be not possible to find single measure that provides a comprehensive indicator of the Landsat data impacts on society.

Considering the existing shortcomings, this research is intended to investigate the impact of open Landsat data on science. Science has been chosen as the research focus due to various reasons which are; a) scientific community can be relatively easy distinguished from the whole society; b) science is big user of Landsat data; and c) the impact of open Landsat data use can be measured by appropriate indicators, for example number of publications.

Michael et al. (2012) provided preliminary impact assessment of Landsat open archive by assessing the number of images downloads from the EROS Data Centre, covering the period from October 2008 to September 2011. They observed that the image download was 3000 scenes prior to October 2008 and it was significantly increased by approximately 6 million scenes till date to September 2011 (Wulder, Masek et al. 2012). Next to the quantitative measures of measuring the use of Landsat data, this research will also focus on qualitative measures to investigate the nature of this use.

With the launch of new Landsat 8 satellite, it had been expected that increasingly more observations are available in the USGS archives for scientific society. However, there is no comprehensive study that can demonstrate the actual impact of these open Landsat data on science. Therefore, this research aims to bridge the current knowledge gap and expand the current state of knowledge of Landsat open data policy and investigate its use by science.

1.4. Research objective and research questions

The main objective of the research is to develop and apply a method to assess the impact of open Landsat data on science.

In order to achieve the main objective of this research, following three research questions have been formulated:
RQ1: What are the possible methods for determining the impact of open Landsat data on science?
RQ2: How can the identified research methods be applied to this research?
RQ3: What insight can be obtained from the result?

1.5. Outline of report

The thesis report of the research is divided into six chapters. Chapter 1 covers the introduction of open geo data, Landsat mission of USGS, research problem, research objective, research questions, and outline of the thesis report. Chapter 2 provides an overview of research background. The approach for answering the three research questions are presented in the methodology Chapter 3. Chapter 4 elaborates three research questions with their results. Chapter 5 provides discussion and Chapter 6 presents the conclusions and recommendations of the research.

2. Research background

The aim of this chapter is to provide theoretical foundation on the current research. The chapter begins with the history of Landsat mission (Section 2.1). The next Section 2.2 reviews the different data policy of Landsat mission. The Section 2.3 provides the different user of Landsat data. Section 2.4 describes the research approach to measure the impact on science. Finally, Section 2.5 presents the conceptual framework about the research to be undertaken based on literature review.

2.1. History of Landsat mission

In 1960, the National Aeronautics and Space Administration (NASA) started research for development and launch of first earth monitoring satellite to meet the needs of resource managers and earth scientists (NASA&USGS 2014). In 1965, William Pecora, director of United State Geological Survey (USGS) proposed the idea of remote sensing satellite program to gather the data for earth observations using a series of Satellite (Pecora 1966). In 1966, USGS and NASA jointly started Earth Recourses observing satellite program called Landsat mission (Lauer, Morain et al. 1997, Goward, Arvidson et al. 2006). The main aim of this Landsat mission was to establish and implement the approaches which ensure the repetitive acquisition of data for earth observation and monitoring global changes. In this case, USGS was the authority for operations, management and maintenance of all ground data processing, reception, product generation, archiving and distribution of data product. Whereas NASA was undertook the responsibility for development, implement and launch of satellite. Considered from 1960s to till date, the Landsat mission has launched seven successful Landsat series of satellites (shown in Table 1) which have provided the longest and continuous record of Earth observation and global monitoring change (Loveland and Dwyer 2012, NASA&USGS 2014). The series of Landsat missions are described below:

a) **Landsat 1**: Landsat 1 satellite was also known as the Earth Resources Technology Satellite-1 (ERTS-1). Landsat 1 was launched as a series of first global earth observation Landsat satellite by NASA with collaboration of USGS on 23rd July 1972 (NASA&USGS 2014). The two sensors Return Beam Vidicon (RBV) and the Multispectral Scanner (MSS) were used to systematically collect images of earth with 80 meter ground resolution. The RBV sensor with three spectral bands was considered as primary instrument. After launch of Landsat 1 satellite, RBV sensor failed with in short period (5th August, 1972). The MSS sensor collected images with four spectral bands (green, red and two near-infrared). The size of each image/scene was 170 km x 185 km with 18 days repeat coverage/cycle of orbit. In addition, data rate was 15 Mbps with direct downlink transmitters from 2.30 minute wide-band video tape recorders. After five year life span, Landsat 1 was expired in 6th January 1978 (Goward, Arvidson et al. 2006, Wulder, Masek et al. 2012, NASA&USGS 2014).

Table 1. Summary of Landsat mission (adopted from (Goward, Arvidson et al. 2006))

Mission	Launch	Life span (expired)	Sensor	Spatial resolution (m)	No of band
Landsat 1	23rd July 1972	6th January 1978	RBV, MSS	80,80	3 band (RBV) 4 band (MSS)
Landsat 2	22nd January 1975	27th July 1983 (officially)	RBV, MSS	80,80	3 band (RBV) 4 band (MSS)

Mission	Launch	Life span (expired)	Sensor	Spatial resolution (m)	No of band
Landsat 3	5th March, 1978	7th September, 1983	RBV, MSS	40,80	3 band (RBV) 5 band (MSS)
Landsat 4	16th July 1982	14th December 1993	MSS, TM	80,30	4 band (MSS) 7 band (TM)
Landsat 5	1st March 1984	23rd January 2013	MSS (August 1995 turn off), TM (Stop work from November 2011). Then MSS (turn on from November 2011)	80,30	4 band (MSS) 7 band (TM)
Landsat 7	15th April 1999	SLC fail in May 2003	ETM+	15(pan), 30(ms)	8 band
Landsat 8	11th February 2013	--	OLI, TIRS	[15(pan) & 30],100	9 band (OLI), 2 band (TIRS)

b) **Landsat 2:** The second series of satellite Landsat 2 was launched on 22nd January, 1975 by NASA and USGS. Landsat 2 also used same sensor however MSS was considered as primary instrument. The life span of Landsat 2 was seven years (25th February, 1982) and was removed officially from 27th July 1983 (Goward, Arvidson et al. 2006, Wulder, Masek et al. 2012, NASA&USGS 2014).

c) **Landsat 3:** Landsat 3 was launched in 5th March, 1978 to give the continuity of Landsat series mission. Again same sensors (RBV and MSS) and modality was carried by Landsat 3. RBV system was improved by using two cameras with 40 meter ground resolution (panchromatic spectral) in Landsat 3 satellite. On the other hand, MSS used five spectral bands with addition of thermal band. Even though thermal band failed shortly. Finally Landsat 3 was decommissioned on 7th September, 1983 (Goward, Arvidson et al. 2006, Wulder, Masek et al. 2012, NASA&USGS 2014).

d) **Landsat 4:** Landsat 4 satellite was launched on 16th July 1982. It was relatively different than Landsat 1-3. Multispectral Scanner (MSS) and new thematic mapper (TM) sensors were used as primary instruments in Landsat 4. MSS used four band same as previous Landsat satellites however with enhanced spectral and spatial 80m resolution. TM sensor carried seven spectral bands which are blue, green, red, near-infrared, mid-infrared (2 bands) with 30m ground resolution and thermal infrared with 120m ground resolution. The size of per scene was 170 km x 185 km with 16 days repeat coverage/cycle of orbit. Data transmitter Tracking and Data Relay Satellite System (TDRSS) was used with 85 Mbps data rate because direct downlink transmitters was stopped working in Landsat 4. This satellite was decommissioned on 15th June, 2001 and officially closed on 14th December 1993 (Goward, Arvidson et al. 2006, Wulder, Masek et al. 2012, NASA&USGS 2014).

e) **Landsat 5:** NASA launched Landsat 5 with the same sensors and designed as Landsat 4 on 1st March, 1984. This mission continued to operate well beyond its 3-year design life. In 1987, NASA engineers observed failures of components in the attitude control system, propulsion and power modules, the solar array drives, and traveling wave tube amplifier (TWTA) in the wideband communication module of this satellite. This problem was solved by implementing a series of engineering and procedural solutions to sustain collection and transmission of image data to ground stations. Despite the operational degradation of Landsat 5 capabilities, it completed and distributed high quality global earth images more than 28 years of life span. Thus Landsat 5 is also called Longest-operating Earth observation satellite. The USGS declared decommission announcement of Landsat 5 on 21st December 2012 and it was officially stopped on 23rd January 2013 (Goward, Arvidson et al. 2006, Wulder, Masek et al. 2012, NASA&USGS 2014).

f) **Landsat 6:** Landsat 6 was developed and launched on 5th October, 1993 however it did not achieve orbit successfully. And the mission of Landsat 6 failed.

g) **Landsat 7:** After failure of Landsat 6, NASA launched another satellite Landsat 7 on 15th April, 1999. Landsat 7 satellite provides the continuity of Landsat mission and fills the gap of Landsat 6. The aim of this mission was also provide data available for the cost of fulfilling a user request and support Government, international and commercial communities. In Landsat 7, Enhanced Thematic Mapper Plus (ETM+) sensor replicates the competences of the highly successful the TM sensor which was previously used in Landsat 4 and 5. ETM+ instrument have eight spectral band including panchromatic bands with 15m spatial resolution and thermal band with 60m spatial resolution. The transmitter Direct Downlink with Solid State Recorders (SSR) is used with 150 Mbps data rate and 16 day repetitive earth coverage. Compared to previous satellite additional features of ETM+ makes the Landsat 7 more versatile and efficient for global change, earth observation and monitoring assessment or research. Landsat 7 provides 300 to 350 scenes per day. However, the scan-line corrector (SLC) of ETM+ sensor failed since 31st May 2003. As a result approximately 22% of pixel per scene are not scanned and lost the image data (Chen, Zhu et al. 2011, Wulder, White et al. 2011)

To overcome this problem, USGS Centre for Earth Resources Observation and Science (EROS) has developed and tested number of approaches to fill in the data gaps especially for heterogeneous regions. Among the different approaches applied to fill the SLC of problem, one of method is Neighbourhood Similar Pixel Interpolator (NSPI) which was developed to interpolate the values of the pixels within the gaps. It was assumed that the same-class neighbouring pixels around the un-scanned pixels have similar spectral characteristics. So these neighbouring un-scanned pixels show similar patterns of spectral differences between dates. Simulated and actual SLC-off ETM+ images were used to assess the performance of the NSPI. Results indicate that NSPI can restore the value of un-scanned pixels very accurately for large volumes of SLC-off ETM+ data. Furthermore, especially it works well in heterogeneous regions for long time interval or significant spectral changes. The filled images appear reasonably spatially continuous without obvious striping patterns. Supervised classification using the maximum likelihood algorithm was done on both gap-filled simulated SLC-off data and the original “gap free” data set accurately. The gap-filled products generated by NSPI will have relevance to the user community for various land cover applications. (Chen, Zhu et al. 2011, Wulder, White et al. 2011).

h) **Landsat 8:** To give continuity to the Landsat data mission, Landsat 8 was developed to extend the Landsat record and maintain the continuity of global earth observation and monitoring. It was launched on 11th February, 2013 by NASA and USGS. NASA led the design, construction, launch, and on-orbit calibration phases, during which time the satellite was called the Landsat Data Continuity Mission (LDCM). Landsat 8 has two sensors Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The OLI sensor is a push-broom sensor which provides nine spectral bands to collect Landsat imagery data which are visible, near infrared, two short wave infrared (SWIR) and cirrus spectral bands with 30 m resolution each as well as a panchromatic band with 15 m resolution. It has a five-year design life. The two new spectral bands of OLI have capabilities to detect cirrus clouds and coastal zone observations. Another TIRS sensor has two spectral bands with 100m resolution. Direct Downlink with Solid State Recorders (SSR) is used for data transmission with 384 Mbps data rate and 16 days repeat cycle. Landsat 8 provides 450 scenes per day (NASA&USGS 2014, Roy, Wulder et al. 2014).

2.2. Data policy of Landsat mission

The Landsat Program Management Policy for distributing products derived from Landsat data have been intensely affected by the U.S. Government laws and regulations. The U.S. land remote sensing data distribution policy was initiated in the 1967 U.N. Outer Space Treaty, the 1984 U.S. Land Remote Sensing Commercialization Act (Public Law 98-365), the 1986 U.N. Principles on Remote Sensing, and the 1987 U.S. Department of Commerce Private Remote Sensing Licensing Regulations (Draeger, Holm et al. 1997). Landsat Program Management reviewed this policy every five years or sooner, based on new legislation or revisions to the National Space Policy. The specific guidelines for data archiving and distribution were proposed by the U.S. Congress in the Land Remote Sensing Policy Act of 1992, also known as Public Law 102-555 (U.S. Congress, 1992) and by the Executive Branch in the National Space Policy released on 19th September, 1996 (Williams, Goward et al. 2006).

Before launching Landsat-1, NASA and USGS made a cooperative agreement that the Earth Resources Observation and Science (EROS) data centre of USGS should be responsible for processing, achieving and distributing the Landsat data to users (Draeger, Holm et al. 1997, Williams, Goward et al. 2006). In addition, satellite photographic scenes were set price between $8 to $50, and a digital MSS per scene cost $200 (Table 2). For a second time, USGS made cooperative agreement with the National Oceanic and Atmospheric Administration (NOAA). After this agreement, in 1979 U.S. Government transferred all Landsat management and operations from NASA to NOAA. According to Land Remote Sensing Commercialization Act in 1984, NOAA was instructed for privatization of Landsat satellite data (Williams, Goward et al. 2006). NOAA increased the prices for MSS per scene from $650 to $730 and for TM per scene from approx. $2800 to $4400 (Table 2). NOAA started to find a commercial vendor for handling Landsat data and contracted with Earth Observation Satellite (EOSAT) Company (Draeger, Holm et al. 1997). EOSAT had the responsibility for archiving, collecting and distributing Landsat 4 and 5 data as well the responsibility for building, launching and operating the further Landsat satellites (with government subsidies). Again the MSS Price per scene was increased by EOSAT from $660 to $1000 for MSS scene and approx. $3300 to $4000 per TM scene (Table 2). With the increment in Landsat data prices and restricted redistribution the accessibility of the Landsat data to users became troublesome. By 1989, no operational budget was available for Landsat 4 and 5 and NOAA directed to turn off the satellites (Draeger, Holm et al. 1997).

The NOAA decision was strong protested by US Congress, foreign and domestic data users. Finally Vice President of USGS took an intervention to save the Landsat program (Draeger, Holm et al. 1997, Williams, Goward et al. 2006). After this, US Congress facilitated the Land Remote Sensing Policy Act of 1992, which instructed Landsat Program management to build the Landsat 7 and launched in 1999. After two years in 2001, EOSAT returned back all operational responsibility for Landsat 4 and 5 to the USGS then commercial right was again renounced by USGS to sell all Landsat 4 and 5 data, in accordance with the USGS pricing policy to meet the user demand which was made previously (USGS 2013). Again, the failure of SLC issues from May 2003 reduced the usability of Landsat 7 imagery however the Landsat 5 was working (Goward, Arvidson et al. 2006, Wulder, White et al. 2011).

Table 2. US Landsat data price trends (adopted from (Draeger, Holm et al. 1997))

Photographic image					
Year	**Organisation**	**MSS B&W 10-inch Neg.**	**MSS Colour 40-inch Print**	**TM B&W 10-inch Neg.**	**TM Colour 40-inch Print**
1980	USGS	$10	$50		
1982	NOAA	$35	$175	$35	$175
1985	NOAA	$40	$195	$80	$290
1985	EOSAT	90	350	160	500
1989	ESOAT	90	550	300	800
1990	ESOAT	175	1000	500	1500
Digital tapes					
Year	**Organisation**	**MSS/CCT**	**MSS Acquired Fee**	**TM/CCT**	**TM Acquired Fee**
1980	USGS	200			
1982	NOAA	650	790	2800	
1985	NOAA	730	1120	4400	1600
1985	ESOAT	660	1120	3300	1600
1989	ESOAT	660		3600	
1990	ESOAT	1000		3960	

Source: US Geological Survey, National Oceanic and Atmospheric Administration, and EOSAT Company

In 2008, the free and open Landsat data distribution policy was released (http://landsat.usgs.gov/documents/LandsatDataPolicy.pdf). According this policy, all Landsat archives can be accessed freely without any cost through internet (Wulder, Masek et al. 2012, NASA&USGS 2014). Instantly after this policy, all new global acquisitions of Landsat 7 were made freely available over the internet from July 2008 (Table 3). However, it took almost one year to make all historic archives of Landsat data freely available (NASA&USGS 2014).

Table 3. Landsat image archive release schedule after free data distribution policy (NASA&USGS 2014)

Landsat data	**Available over the Internet**
Landsat 7: ETM+ all new global acquisitions	July 2008
Landsat 7: all ETM+ data	September 2008
Landsat 5: all TM (Since 1984) data	December 2008
Landsat 1-5: all TM and MSS (Since1972) data	January 2009
All Landsat (1-8) data free	February 2009

2.3. User of Landsat data

According to Landsat U.S. Geological Survey report, the potential users of Landsat satellite are federal government, state government, local government, private business, non-profit organization, academic institution as scientific group (Miller, Sexton et al. 2011). However, the survey of USGS shows that the academic sectors are very active users of Landsat satellite for their research (Figure 2) (Miller, Sexton et al. 2011). Therefore, the main focus group of this research is the scientific community who uses Landsat data in their research or work.

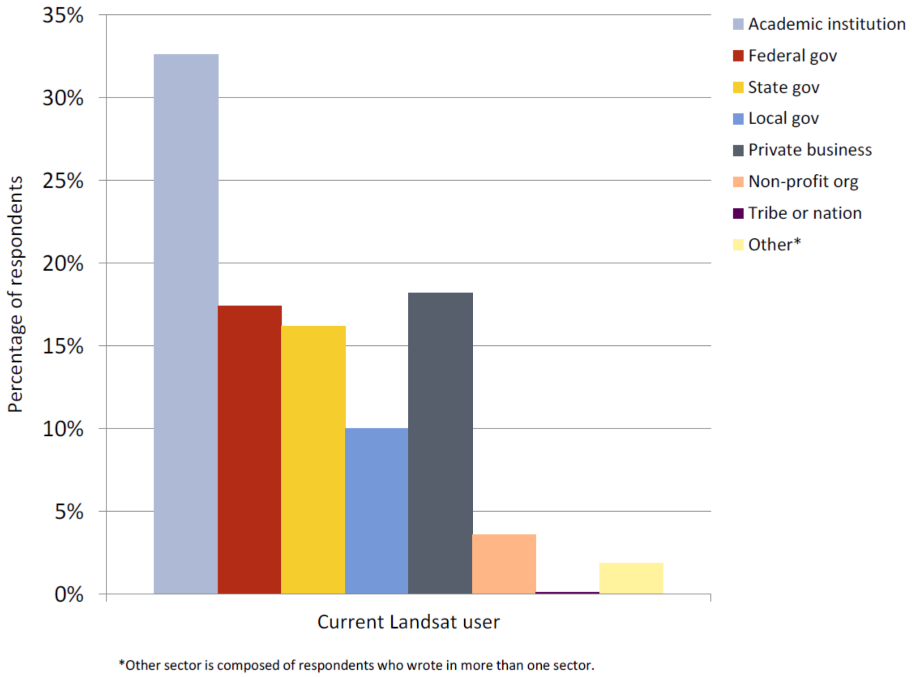


Figure 2. Users of Landsat data [source: (Miller, Sexton et al. 2011)]

2.4. Model of assessing the impact of open data

The first generic model for assessing impact of open topographic data in the Netherlands has been developed by Arnold K. Bregt and Łukasz Grus of Wageningen University. This model has been first used to investigate the effects of open key register topography (BRT). The results have shown that the model is effective to analyse indicators for measuring the effect on the Dutch National Topographic Agency, on society and on the relationship between the National Topographic Agency and the society (Bregt and Grus 2014).

In this research, the modification of this model for assessing the impact of open Landsat data on science has been used which is shown in Figure 3. Even though the model focuses on three types of effects, this research emphases on the external effects on society, and to be precise on science only.

Science is one of the essential part of societies and has immense authority and influence in our society. The main aim of scientist is to understand how the whole world works in very specific way and use this knowledge for predictions of future events (Latour 1987). Scientists are always interested to explore and develop new methods to use data for their research. Most of the time, they publish their research through scientific publication. In Landsat context, Landsat mission was previously aimed for scientific society to explore and develop the new techniques for Earth’s land surface observation. Hence, this research considers science as a society to explore and measure the potential effects on the scientific society before and after releasing open Landsat data.

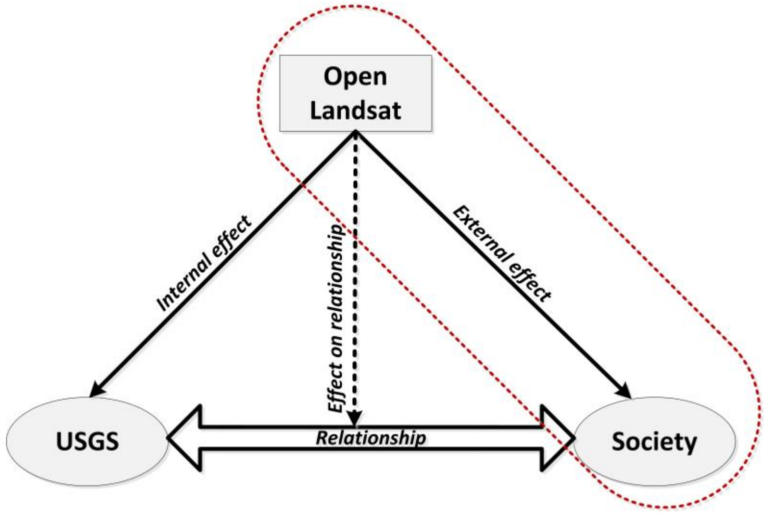


Figure 3. Model for assessing impact of open Landsat data (adopted from (Bregt and Grus 2014))

2.5. Conceptual framework

The previous sections summarised the generic model of assessing the impact of open data. This section presents the conceptual framework that potentially fulfils the requirements of assessing the effects of open Landsat data on science. Figure 4 shows the conceptual framework of this research. The main idea behind the conceptual framework is that it should help to 1) investigate the possible research methods relevant to this study; 2) choose the most suitable research method for this research; and 3) implement the selected research method for impact assessment.

There are several research methods used in science for impact assessment such as case study, interview, ethnography, experiment, expert survey, systematic literature review (SLR) etc. (Gable 1994, Navarro Sada and Maldonado 2007, Sloane-Seale 2009, Wen, Li et al. 2012). However all methods cannot be applied directly to assess the impact of Landsat data. Because Landsat data have following unique characteristics: (1) the long and continuous Landsat data archive (1972 to till date); (2) different data distribution policy which affect the Landsat data use; (3) diverse field of scientific applications; (4) data user all over the globe. Hence, the relevant research methods for Landsat data impact assessment must be able to address those characteristics and enable to compare the differences before and after opening the Landsat data. This will be addressed in RQ1 which is described in chapter 3. It will not be possible to apply all the investigated research methods for this research. Therefore, the most suitable research method will be selected based on fitness for use criteria. The fitness for use criteria will be formulated based on the nature of scientific data source and the time limitation of this research. The detail about this process is explained under RQ2 in chapter 3. Finally, the selected method will be applied for assessing the impact of open Landsat data on science. The detail of this research methodology is described under RQ3 in chapter 3.

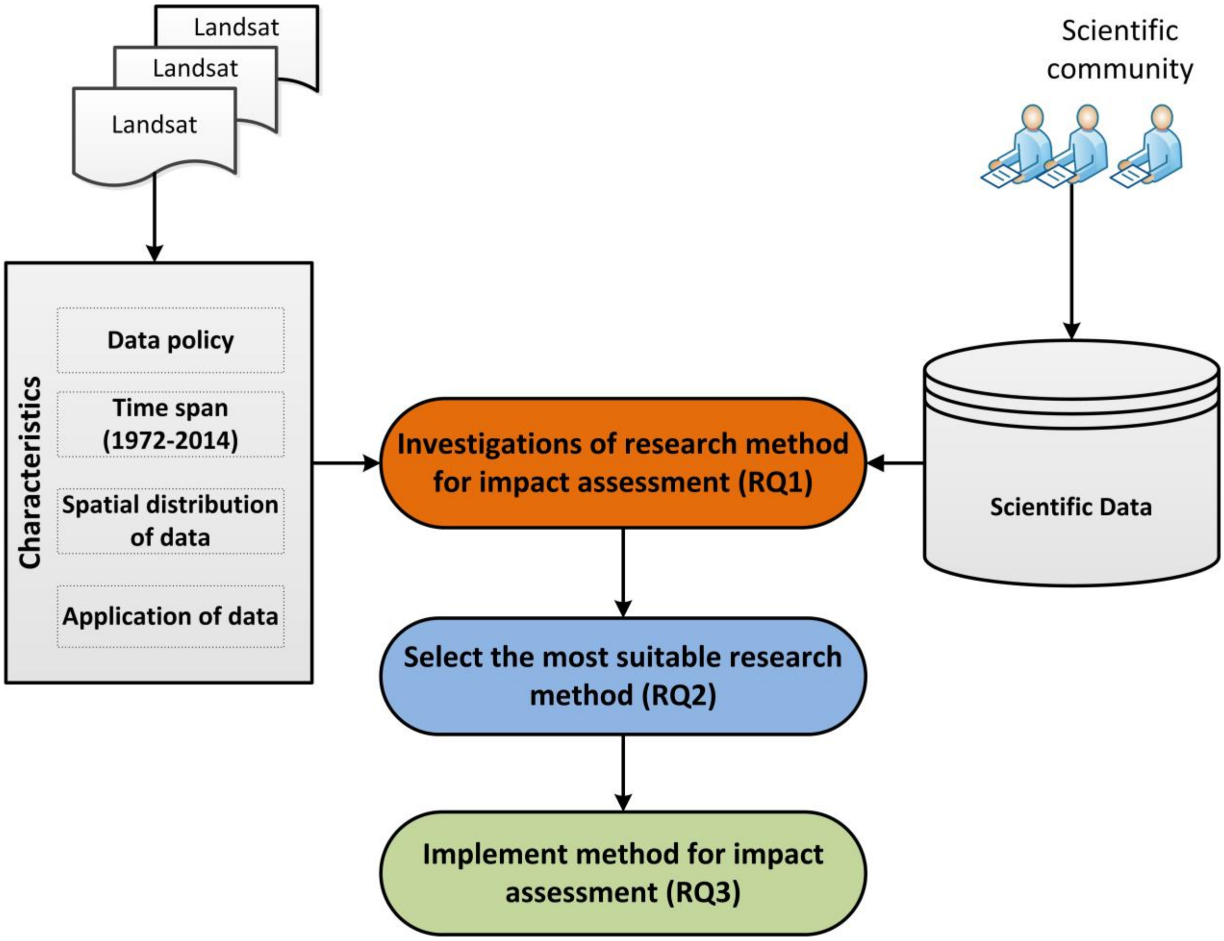


Figure 4. Conceptual framework of the research

3. Research methodology

This chapter contains the methodological approach to address the research questions of the study which is based on theoretical and practical perspective of assessing the impact of open Landsat data on science. This section presents a general plan and approach explaining how the research questions are answered step by step. The research questions have been formulated with the intention of advancing an impact analysis framework for open Landsat data on science. As such, the research questions are related as depicted in Figure 5.

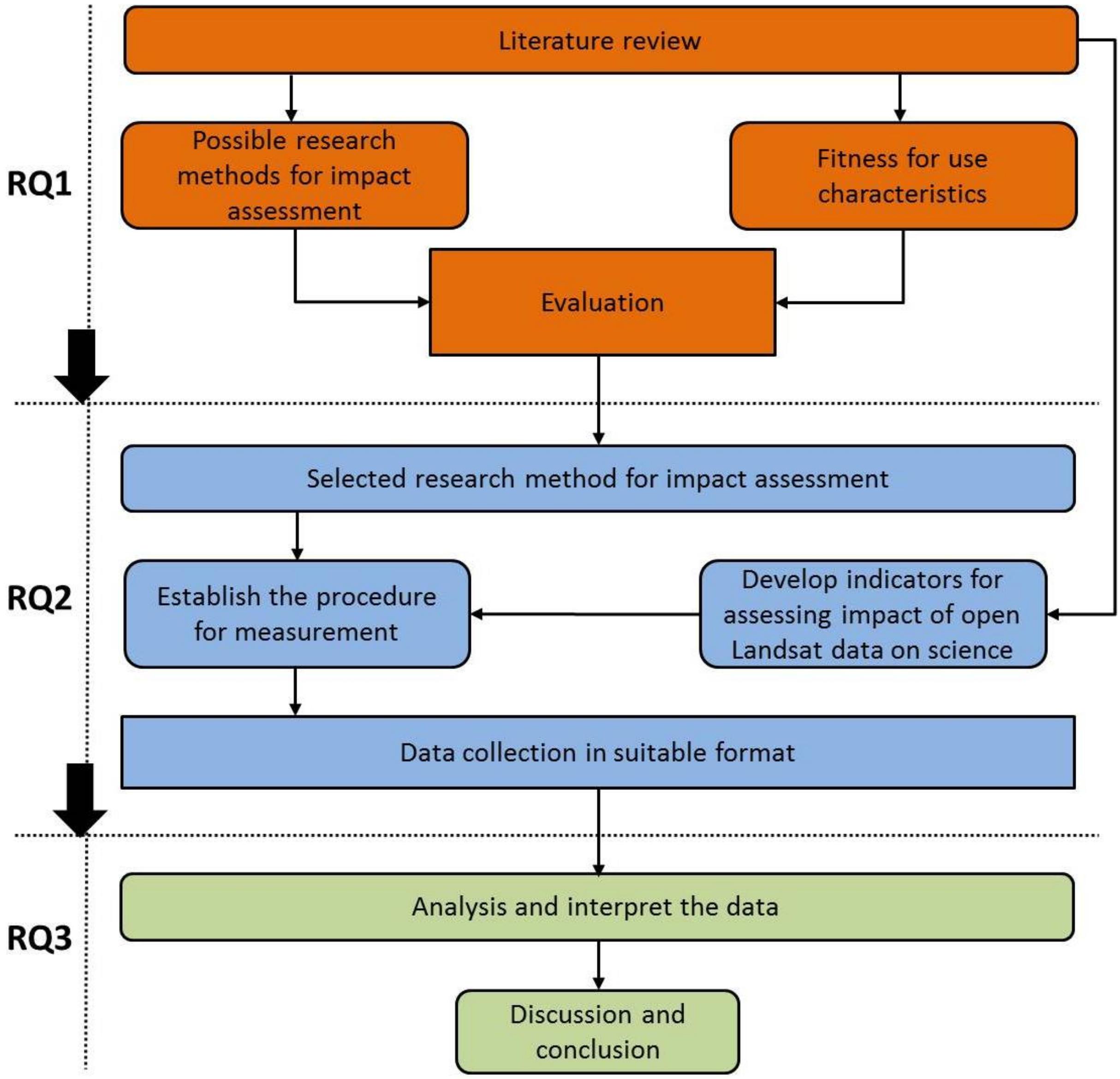


Figure 5. Methodological research framework

RQ1 is intended to explore and select possible research methods for determining the impact of open Landsat data on science. Selected method(s) in RQ1 will contribute to address RQ2 and RQ3. RQ2 will answer the question on how to apply the identified research methods from RQ1. Finally, RQ3 will provide the suitability of the selected method to measure the impact of open Landsat data on science. The results are discussed and hence conclusions are derived accordingly.

RQ1: What are the possible research methods for determining the impact of open Landsat data on science?

The main focus of this research question is to get in-depth understanding about the existing research methods on impact assessment and their applicability in the current research context of the impact of open Landsat data on science. Three steps are followed to answer the RQ1.

- Step 1: In-depth literature review is conducted to find out the relevant research methods that can address the needs of Landsat data characteristics: (1) the long and continuous Landsat data archive (1972 to till date); (2) different data distribution policy which affect the Landsat data use; (3) diverse field of scientific applications; (4) data user all over the globe. Out of these research methods, the most suitable research method is selected.
- Step 2: Due to time limitation, it is not possible to test all the research methods identified in step 1. Therefore, fitness for use criteria has been formulated to assess the effectiveness of the research methods for this research (Agumya and Hunter 1999, Bruin, Bregt et al. 2001, Whitfield 2012). The fitness for use criteria are analysed on the basis of SMART criteria and two science criteria, namely transparency and reproducibility (Downing 2004, Bogue 2005, Laine, Goodman et al. 2007, Goecks, Nekrutenko et al. 2010, Ram 2013). All selected research methods are evaluated based on fitness for use criteria and their specified measurement categories. Fitness for use indicates the method's relevancy for this research and is measured in the scale of high, medium and low which is shown in Table 4. The comparison table is created for all selected research methods (see step 1) on the basis of fitness for use score.

Table 4. Fitness for use criteria for selecting the most suitable research method

SN	Fitness for use criteria	Definition	Characteristics to measure	Score value
1	Achievability of resource	Achievability means availability of resources which defines the regular access of resource and data source.	No access of data source	Low
			If the data source and resource are possible to access but not available when needed	Medium
			If the data source and resource can be assessed any time when required.	High
2	Reliability	Reliability is defined as trustworthiness of resources and data. It can be measured in terms of research bias.	If no bias	Low
			If the research bias introduce due to procedural and subject failure	Medium
			If the research bias introduce due to experimental errors	High
3	Measurability of resources	The measurability of resources means the resource and data source which can be measured.	If the resource and data source provide qualitative data or text format	Low
			If the resource and data source provide both qualitative as well as quantitative data	Medium
			If the resource and data source provide quantitative data	High

SN	Fitness for use criteria	Definition	Characteristics to measure	Score value
4	Time bound	Time bound is defined as the feasibility of time frame for this research.	If large size of data is not possible to be collected within one month	Low
			If large size of data is possible to be collected within one to two months	Medium
			If large size of data is possible to be collected within one to two weeks	High
5	Reproducibility	Reproducibility is defined as the continuously reproducing of the resource or data source.	If there is no possibility of reproducing resources or data sources	Low
			If there is possibility of reproducing resources or data sources but is time consuming	Medium
			If there is possibility of reproducing resources or data sources when needed	High
6	Transparency	Transparency means openness of resources and data sources.	If the resources and data source are not open publicly	Low
			If the resources and data source are partially open	Medium
			If the resources and data source can be assessed publicly	High

- Step 3: Based on the score value the research method(s) with the highest score is selected as suitable research method(s). The selected method serves as an input for the RQ2 and RQ3.

RQ2: How can the identified research method be applied to this research?

The second research question deals with the use of research method selected in RQ1. Four major steps are conducted to answer this research question.

- Step 1: Based on the method description best practice guidelines are formulated to implement the selected research method in a systematic way.
- Step 2: Suitable indicators of Landsat data on science are defined based on the existing literature and Landsat user survey report (Miller, Sexton et al. 2011, Wulder, Masek et al. 2012).
- Step 3: The appropriate measurement methods are defined to acquire the data for each indicator.
- Step 4: The actual measurement of the value of each indicator is conducted. The results are used in RQ3 for further analysis.

RQ3: What insight can be obtained from the result?

The acquired data is analysed systematically using the selected research method from RQ2. The result of the systematic analysis is interpreted to get the impact of open Landsat data on science. This research question aims to synthesize and validate the obtained results; and to draw conclusions by comparing the findings with the existing scientific literature in order to provide insight about the impact of open Landsat data on science. In addition, it presents the general reflections about the advantage and disadvantage of the method which are applied for measuring the impact of open Landsat data on science.

4. Results

4.1. Results for RQ1

RQ1: What are the possible research methods for determining the impact of open Landsat data on science?

The ultimate aim of this research is to develop and apply a method to assess the impact of open Landsat data on science. In this context, it is also crucial to compare the scientific impact before and after the change in Landsat data distribution policy. The open Landsat data encourages scientists to exploit the use of data in diverse scientific innovations. The scientific society is a key platform for research activity, and publishing is a formal mechanism through which scientific community make contributions to the body of knowledge. In order to assess the impact of open Landsat data on science, a suitable method should be implemented. As a first step it is needed to make an overview of potential methods. The literature review is used to search and explore the possible existing research methods for assessing scientific impact. Finally, suitability of the use of identified research methods are evaluated on the basis of fitness for use criteria.

4.1.1. Research methods

Various research methods such as case study, interview, ethnography, experiment, expert survey, systematic literature review (SLR) are used to enable researchers and policy makers to assess the impact on any scientific domain (Gable 1994, Navarro Sada and Maldonado 2007, Sloane-Seale 2009, Wen, Li et al. 2012). Among these various research methods, three research methods case study, survey by experts and systematic literature review (SLR) have capabilities to meet the requirement of Landsat data (section 2.5) and to assess the impact on science for this research (Gable 1994, Cavaye 1996, Sloane-Seale 2009, Wen, Li et al. 2012, Miller, Richardson et al. 2013). In the following sections, these research methods are briefly discussed with respect to the research objective of this research.

a) Case study

Case study provides in depth study of a single unit which is used to narrow down the broad field of study into easily researchable topics. Most of the case studies are flexible in nature and use more than one data sources such as interview, observations, documents or surveys (Schell 1992, Navarro Sada and Maldonado 2007). Sometimes case study are conducted with only one subject and small analysis of data sets, which can be difficult to generalise and rely with a single case (Zainal 2007). In case of impact assessment of open Landsat data on science, multiple case studies will be required for assessing the diverse scientific impact of Landsat. Multiple data sources, multiple investigators and sites may be included for collection of data in the form of interviews, observations, documents or surveys and most of these data sources of case study are in the qualitative format (such as text). The case study method can be time consuming to collect data, and even more time-consuming to analyse the obtained data. Case studies have advantage of giving detailed information of specific subject; however the aggregation of information highly depends on researchers own subjective feeling (high probability of researcher bias). Furthermore it would be challenging to reproduce data for assessing the impact of open Landsat broader science.

b) Expert Survey

Expert Survey is a research method used to investigate and gather the insight information on expert perception, thinking, opinion and their feelings about a specific topic (Navarro Sada and Maldonado 2007). Expert surveys are the preferred research method for developing generalized suggestions based on collecting information using questionnaires or interviews from a certain population. The most common data collection tool for expert survey is a face to face interview, telephone interview, internet based interview (Email, Skype), internet/web based online questionnaire survey etc. (Gable 1994). Web based survey is cost effective and comparatively faster in analysing the data than other research methods (Schleyer and Forrest 2000, Fricker and Schonlau 2002, Andrews, Nonnecke et al. 2007, Navarro Sada and Maldonado 2007, Creswell 2013). However, the result of the survey is dependent on the response rate of experts who provide data for analysis through a questionnaire. Furthermore, it provides the limited resources and sample size for a specific topic. Within the timeframe of this research, it would be difficult to find the large amount of experts or resources of the whole science for assessing the scientific impact of Landsat data.

c) Systematic literature review (SLR)

Systematic literature review (SLR) is one of the most commonly used approaches to gather numerical data of systematic literature search (Kitchenham, Pearl Brereton et al. 2009, Creswell 2013). This approach provides in-depth understanding of various aspects of scientific literature systematically. The SLR is conducted to identify, evaluate, and interpret all the existing scientific evidence relevant to the specific subject. The SLR is a requirement for quantitative meta-analysis which analyse statistical data from a number of scientific studies, performed over a period of time. The SLR was developed initially in medicine as research indicated that expert opinion based medical advice was not as reliable as advice based on the accumulation of results from scientific experiments (Kitchenham, Pearl Brereton et al. 2009, Wen, Li et al. 2012). Since then several researchers have adopted this approach in many domains, e.g. Criminology, Social policy, Economics, Nursing, Software Engineering etc. (InnvÃ¦r, Vist et al. 2002, Welsh and Farrington 2002, Goff, Rose et al. 2007, Keele 2007, Petticrew and Roberts 2008, Beard, Feeley et al. 2009). SLR involves the following steps: data sources, search strategy, selection criteria, data extraction, synthesis and analysis of the extracted data (Keele 2007, Wen, Li et al. 2012). The main data source of SLR is digital electronic database libraries (bibliography) which are free, transparent and assessable at any time such as Google Schlor, PubMed, Web of Science, Scopus etc. These data sources have provided the comprehensive search of diverse scientific databases (publication) over the years. In addition, the data source chosen for SLR has scientifically accepted research results (publication) that are of high scientific value. This is maintained by publishing scientific papers after rigorous reviewing process. Hence, the quality of the data obtained from SLR (source) is significantly high.

4.1.2. Evaluation of research methods

The evaluation of the above mentioned three research methods: case study, expert survey and systematic literature review (SLR) were analysed based on general fitness for use criteria and specific requirements related to the specificity of this research described in section 2.5. The details of evaluation of these research methods on the basis of fitness for use criteria and their specified measured values are described below and are summarized in Table 5.

1. Achievability of resources

Achievability of resources means accessibility or availability of resources regardless of potential difficulties or challenges. In this research context, case study and expert survey have scored *medium* in terms of achievability of resources due to the fact that these research methods are based on human perception, opinion, thinking, participation observation, documents for specific area. Human is one of the main sources for data collection of these two methods which is not available all the time to respond the questions. Moreover, case studies and expert surveys are expensive, and require significant time to gather resources for different domain in case of identifying the scope of more than

42 years of Landsat use. This limits the applicability of these two methods while SLR facilitates free search of domains without being restricted to one. In addition, the data can easily achieve using digital database libraries. Hence, SLR has scored *high*.

2. Reliability

Reliability can be measured in terms of research bias. Research bias is defined as the degree of possible experimental errors, failure and selection of research subject that can potentially introduce errors or desired results in the research (Joy 2007). The main focus of this research is to explore the possible methods and implement for assessing the impact of open Landsat data on science. There could be possibility of research bias in procedural, subject error, and also bias due to misunderstanding between researcher and the data source. Therefore, the three research methods case study, expert survey and SLR have given *medium* scored. Despite the fact that SLR scores *medium*, it is worth noting that the data sources (i.e., published papers) are of *high* scientific quality and are rigorously reviewed before publishing.

3. Measurability of resources

To achieve the goal of this research, the research method should be able to measure the whole time span of Landsat data (1972 to till date). However, as discussed earlier in achievability of resources, case study and expert survey commonly focus on human perception, opinion, thinking, participation observation, documents. The available information from the resources of these two methods could be more qualitative (subjective) in the text rather than quantitative (numeric) or both. In addition, it would be difficult to measure the spatial/geographical scope of Landsat data use in scientific field from 1972. For this reason, the case study and expert survey have scored *medium* for this research. Whereas SLR method provides quantitative data sources in a digital format and can measure the available resource of Landsat use in diverse scientific field from 1972 to till date. Therefore, SLR method has scored *high* under this category.

4. Time bound

The case study provides in depth understanding of a specific case. If the case study is chosen for assessing the impact of Landsat data on science, more than one case study should be used to collect the scope of Landsat data used in different scientific domain which is time consuming. Whereas expert survey using a web based questionnaire survey tool provides comparatively faster data entry and analysis than case study. However, finding the experts within the given time frame of this research is challenging. Hence the case study and expert survey have scored *low*. SLR method can provide digital scientific data sources which can able to measure the spatial/geographical scope of Landsat data use in scientific field from 1972 to till date within short period. That is the reason it *has* scored *high*.

5. Reproducibility

If the data is collected once using case study for a specific case, it is difficult to reproduce them again. As compared to case studies, the resources or data sources can be reproduced using expert survey method however the sources are dependent on the respondent. In case of SLR, the data source can be digitally reproduced many times for collecting and analysing the broad period of Landsat scientific database. Therefore the case study, expert survey and SLR have scored *low, medium* and *high* respectively.

6. Transparency

The case study and expert survey have scored *medium* in terms of transparency. Commonly, the data sources of case study and expert survey such as human perception, opinion, thinking, participation observation, documents are not completely opened and accessible publicly. Compared to case study and expert survey, all the available Landsat data in any scientific field can be accessed and are fully transparent for anybody at any time in SLR method. So SLR has scored *high*.

As a summary of the evaluation of the three research methods with respect to fitness for use criteria are marked as *high*, *medium* and *low* in Table 5. It is found that SLR has relatively *high* impact than others in context of this research. So the SLR research method is chosen for further analysis.

Table 5. Summary of the research methods with respect to fitness for use criteria

SN	Fitness for use criteria	Research methods		
		Case study	Expert survey	Systematic literature review (SLR)
1	Achievability of resources	Medium	Medium	High
2	Reliability	Medium	Medium	Medium
3	Measurability of resources	Medium	Medium	High
4	Time bound	Low	Low	High
5	Reproducibility	Low	Medium	High
6	Transparency	Medium	Medium	High

4.2. Results for RQ2

RQ2: How can the identified research method be applied to this research?

Firstly before applying the SLR method which was selected in RQ1 as a relevant research method to achieve the research goal, the assessment categories for assessing the impact of open Landsat data on science and its indicators are formulated on the basis of literature review. Secondly the steps of the SLR methods are discussed to measure the indicators for impact assessment in details.

4.2.1. Indicators for measurement of open Landsat data on science

Indicators are used to assess and measure the status of progress of any activities. The indicators enable decision-makers to assess progress towards the achievement of intended outcomes, goals and objectives. Indicators of science are statistical measurement of science which provides an ideal analytical resource for scientists to understand the condition of science. Scientific indicators of Landsat can determine the influential individuals, institutions, papers, publications, and countries in their field of study as well as emerging research areas that could impact their work. In this research, systematic literature review (SLR) method based on scientific publication from scientific databases was used to measure the impact of open Landsat data on science. For assessing the scientific impact of open Landsat data, it is essential to investigate how the uses of Landsat data and their application has potentially changed before and after they got open. Therefore, the scientific impact of open Landsat data will be assessed by answering the questions: i) has the intensity of Landsat data use changed as a result of making those data open? ii) has the geographical scope of Landsat data applications changed? iii) has the scope of application disciplines of Landsat data changed? To answer those questions six indicators were formulated (two per questions) to measure and assess the impact of open Landsat data on science (Joy 2007, Macauley 2009, Mao, Wang et al. 2010, Miller, Sexton et al. 2011). The definition of each indicator described in Table 6.

Table 6. Indicators of each assessment category

Assessment category	Indicators for measurement
Intensity of Landsat data use	**1) Number of publication in which Landsat data used per year:** This indicator is used to measure the trend of Landsat publication using Landsat data from 1972 to 2014. However the open data distribution policy of Landsat data was announced in 2008. This indicator is used to assess the comparison of scientific publication using Landsat.
	2) Number of publication in different document type: Several peer-review document types (publications) are available in electronic/digital libraries such as: Journal, conference paper, book, book chapter, reviews etc. This indicator provides the comparison of diversity of document types from 1972 to 2014 using Landsat data.
Geographical scope of application	**3) Number of publication in different languages:** Under this indicator, numbers of publication in diverse languages with respect to different Geographical locations are assessed within 1972 to 2014.
	4) Number of publication that use Landsat data per country per year: The indicator assesses the use of Landsat data increase or decrease per country or region over the year. The outcome of this indicator provides the change in geographical spread of using Landsat data from 1972 to 2014.

Assessment category	Indicators for measurement
Scope of application disciplines	**5) Number of publication in which Landsat data used in different disciplines per year:** This indicator is used to assess the third assessment category of the scope of application disciplines using Landsat data. The indicator addresses the use of Landsat data in different discipline from 1972 to 2014.
	6) Number of images used in publication : The main purpose of this indicator is to assess the use of multiple Landsat images in different process from 1972 to 2014. For instance, time series, near-real time applications.

4.2.2. Systematic literature review (SLR)

Systematic literature review (SLR) is one of the well-known methods for reviewing the exiting literature. Several researchers have reported their experiences and lessons learned from applying systematic reviews to different subject area. In this research for assessing the impact of open Landsat data on science, SLR method is chosen as a suitable method. The SLR method includes the following application components: data sources, search strategy, selection criteria, data extraction, synthesis and analysis of the extracted data (Keele 2007, Wen, Li et al. 2012).

a) Data source

Scientific databases are seen as systems of production and dissemination of knowledge (Keresztesi and Oberman 1982). Systematic literature review through scientific digital databases (bibliographic) are the most significant source of information where number of scientific publication are published and available freely (Kitchenham, Pearl Brereton et al. 2009). Currently, there are several electronic citations database. The commonly used scientific digital databases are PubMed (http://www.ncbi.nlm.nih.gov/pmc/), Scopus (http://www.scopus.com/), Web of Science (www. webofknowledge.com), Google Scholar (scholar.google.nl) (Falagas, Pitsouni et al. 2008). These digital databases have been focused on several scientific peer-review publications for example article, conference paper, review, book chapter etc. The comparison of these four scientific databases PubMed, Scopus, web of Science, Google Scholar are summarised in Table 7.

Table 7. Comparison of scientific database PubMed, Scopus, web of Science, Google Scholar (Falagas, Pitsouni et al. 2008)

Characteristics	PubMed	Scopus	Web of Science	Google Scholar
No. of journals	6000 (827 open access)	12850 (500 open access)	8700	No exact data limit
Language	English plus 56 other languages	English plus more than 30 other languages	English plus 45 other languages	English plus any language (No exact data)
Focus field	Medical and biomedical sciences	Physical sciences, health sciences, life sciences, social sciences	Science, technology, social sciences, arts and humanities	Theoretically all field available electronically

Characteristics	PubMed	Scopus	Web of Science	Google Scholar
Period Cover	1950–present	1960-present	1945-present	Theoretically all available electronically
Search by	Title and abstract	Title, abstract, keyword	Title	No specific
No. of keywords allow	No limit	30	15	Theoretically no limit
Updating	Daily	1–2 times weekly	Weekly	Monthly on average

PubMed digital database was developed by the National Library of Medicine (NLM), a division of the National Institutes of Health. It especially focuses on medicine and biomedical science. Whereas Google Scholar databases was developed by Google Inc. It indexes all electronic references on a subject irrespective to any languages. So there is no specific focus field available in Google Scholar. For instance, it cannot support searching journals by abstract, title and keywords separately (Falagas, Pitsouni et al. 2008, Ghafari, Saleh et al. 2012, Miri and Bahmani 2012). Thus PubMed and Google Scholar database are not considered for assessing the scientific impact of open Landsat data in this research. The multidisciplinary databases Scopus and web of science are more relevant for this research. While comparing Scopus and Web of Science databases, it was found that Scopus (with over 12,850 journals) provides substantially more number of journals coverage than web of Science. With comparing the number of journals being indexed Scopus has wider scope and can search by title, abstract and keywords compare to Web of Science compare to Web of Science. Additionally, Scopus has better analysis tools like search filters engine (Falagas, Pitsouni et al. 2008, Ghafari, Saleh et al. 2012, Miri and Bahmani 2012). These tools are crucial for the purpose of measuring different aspects of impacts which is the goal of this research. Because of those reasons, Scopus database is chosen as an appropriate data source for this research.

b) Search Strategy

A search strategy is necessary to perform search at search engine digital source. In SLR, the search string consists of a set of logical expressions that combine keywords and its alternative spellings, synonyms and abbreviations arranged in way that the highest amount of relevant studies is obtained from search engines. Then, the sophisticated search string based on data source is constructed using Boolean "AND" to connect the related keywords and Boolean "OR" to allow synonyms and word class variants of each keyword (Keele 2007, Wen, Li et al. 2012). In this research, search string has been constructed using Landsat as main keywords and has been included synonyms and related terms. The search string was then constructed using Boolean "AND" to connect the related keywords and Boolean "OR" to allow synonyms and word class variants of each keyword. The search string was executed in the digital library Scopus to search Field type as Article title, Abstract, Keywords. Besides the search string, the range of study dates also has to be defined in the search strategy. The study date conducted from 1972 to 2014 because Landsat mission was first launched in 1972 (NASA&USGS 2014). One example of search strategy is given below:

TITLE-ABS-KEY (landsat) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE ,"ch") OR LIMIT-TO (DOCTYPE , "bk"))

c) Inclusion and exclusion criteria

It is likely that some of the results (study data) of a search might contain the keywords but are irrelevant to our research. For instance, a study data are related to editorials, notes, reports and letters which are not systematic peer-review publication are excluded for this research. The study selection not only eliminates irrelevant studies, but also ensures the quality of the study and the scoping of the research. A set of inclusion and exclusion criteria based on the scope of research and the quality of the studies were determined. The inclusion and exclusion criteria are given in Table 8.

Table 8. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
The date of publication years from 1972 to 2014 were selected	Other years exclude
Journal, conference paper, book, book chapter, reviews were taken into account	Editorial, notes, letter, report etc.
First author's country will be used from which address the publication was published	Other exclude
For assessing the third category (scope of application disciplines), Agriculture/ environmental sciences and management/land use and land cover/planning and development/commercial/ human needs/Energy (oil, gas and mineral) application domain were taken into account from the application of Landsat imagery among Landsat users-executive report (Miller, Sexton et al. 2011)	Other domain were excluded

d) Data extraction

The data extraction process is used to collect the information needed to address the research questions. Generally, data extraction provides a set of numerical values (e.g. number of publications, number of document type, confidence intervals). That can be extracted for each search results. These numerical values are important to summarise the results (Keele 2007, Wen, Li et al. 2012).In this research, the scientific publication from Scopus retrieved in comma-separated values (csv) format with following details information: publication year, authors, authors with affiliation, title, source title, citation, published country, document type, discipline, language type etc. Then these data were cross checked for completeness of necessary details because it was found that some of publication had missing title or author or date.

e) Data synthesis

Data synthesis is the process of collecting, tabulating and summarising the results in a descriptive way (Keele 2007). Extracted information about the studies should be synthesised in a consistent manner to answer the review questions. After extraction of data, statistical analysis and mapping were done to analyse and visualize the data to achieve the objective of this research.

To summarize, the actual review process that the SLR used was divided in four main phases. Figure 6, presents an overview of this process. Phase I covers the construction of the search string (query) based on the six indicators which were formulated for assessing the impact of open Landsat data on science. Phase II includes the execution of these string through Scopus database library for the review and retrieving the initial selection of studies. In the Phase III of the secondary selection, the publication database was randomly assessed to test whether the abstract, title and keywords of publication were missing or not. In Phase IV, the inclusion and exclusion criteria were used to determine the quality of the results that were deemed relevant in the third phase. The output of this phase was the final list of primary studies which were retrieved in full text and extracted in CSV file and synthesised. Finally the search results will be validated by randomly study of 3% publications on the basis of title, abstract and keywords from 2002 to 2014 to test the search query results.

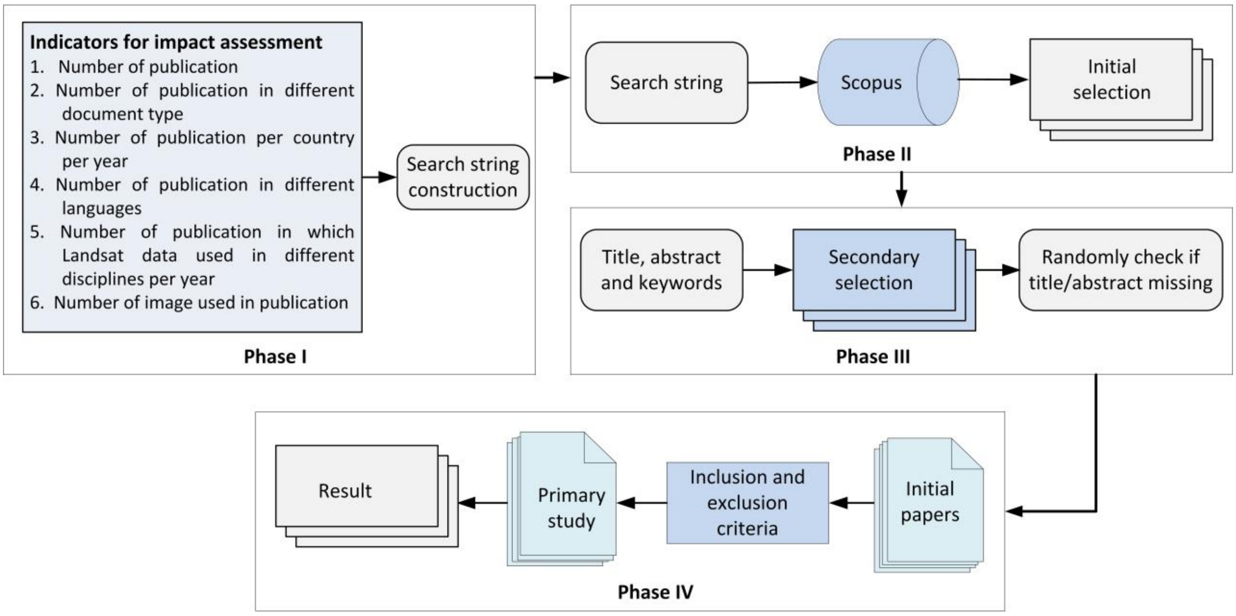


Figure 6. Overview of review process of SLR method

4.3. Results for RQ3

RQ3: What insight can be obtained from the result of applied method?

As mentioned in the section 4.2.2, Scopus data source was chosen for the data collection. On the basis of six measurement indicators developed for assessing scientific impact of Landsat data, search query was constructed and applied in Scopus database. The overview of data collection is presented in following Table 9.

Table 9. Overview of data collection

Data source	**Scopus**
Data collection date	October - November 2014
Search year	1972-2014
Major search filed type	TITLE-ABS-KEY
Main keyword	Landsat
Total number of publication (without inclusion/exclusion criteria)	23,604
Total number of publication (after inclusion/exclusion criteria)	18,116
Randomly validation for search result of Landsat (from 2002 to 2014)	501

In the first phase, the review process was conducted with the construction of search string query for each indicator to assess the impact of Landsat data on science. The constructed list of queries for Scopus using SLR method can be shown in Annex A. Initially, the total number of publications was 23,604 when the search query/string **TITLE-ABS-KEY (Landsat)** ran over the digital library Scopus (in second phase). The list of those publications (23,604) was randomly checked based on title & abstract, if the abstract, title and keywords were missing or not in in third phase. These articles were then evaluated based on inclusion and exclusion criteria (Table 8). After the inclusion and exclusion criteria, there was not found any publication during 1972 to 1974. Thus, the total number of publication from 1975 to 2014 was 18,116 (Figure 6). Finally these publications database were collected in comma-separated values (CSV) format and were analysed using Microsoft Excel. Similarly, the Scopus outcomes were linked with the country list obtained from World Bank for assessing the geographical scope of Landsat (WorldBank 2015). Out of 18,116 (from 1975-2014) publications, 3% of publications from 2002 to 2014 were randomly studied on the basis of title, abstract and keywords to test the search results using search query in SLR method (Table 9). The detailed results and validation of this general data analysis are elaborated and discussed in following section.

4.3.1. Number of publication in which Landsat data used per year

The distribution of Landsat publication per year (from 1975 to 2014) along with the trend-line can be seen in following Figure 7. In the early years from 1975 to 1980 the publication using Landsat was very low. As shown in Figure 7, the use of Landsat data and its publication can be divided in two phases in terms of before and after opening Landsat data in 2008. The trend line shows that before opening Landsat data, the number of scientific publications using Landsat slightly increased during 1981 to 1984, after that it decreased. It could be because of Landsat data price policy and its trends (Table 2), Landsat users were restricted the use of Landsat data for their research in large area. The trend line shows that after opening Landsat data, the number of publications using Landsat also relatively increased. Free data policy of Landsat has drawn more attention for users. However, sometime scientific publication takes too much time to publish and finally to get index in Scopus. Thus the actual impact may not be seen in same year of research using Landsat. For instance, the publication using Landsat decreases in 2014. It is not certain that all studies with a publication date later than 2014 have been covered, since publication/journals may not have been indexed yet in Scopus.

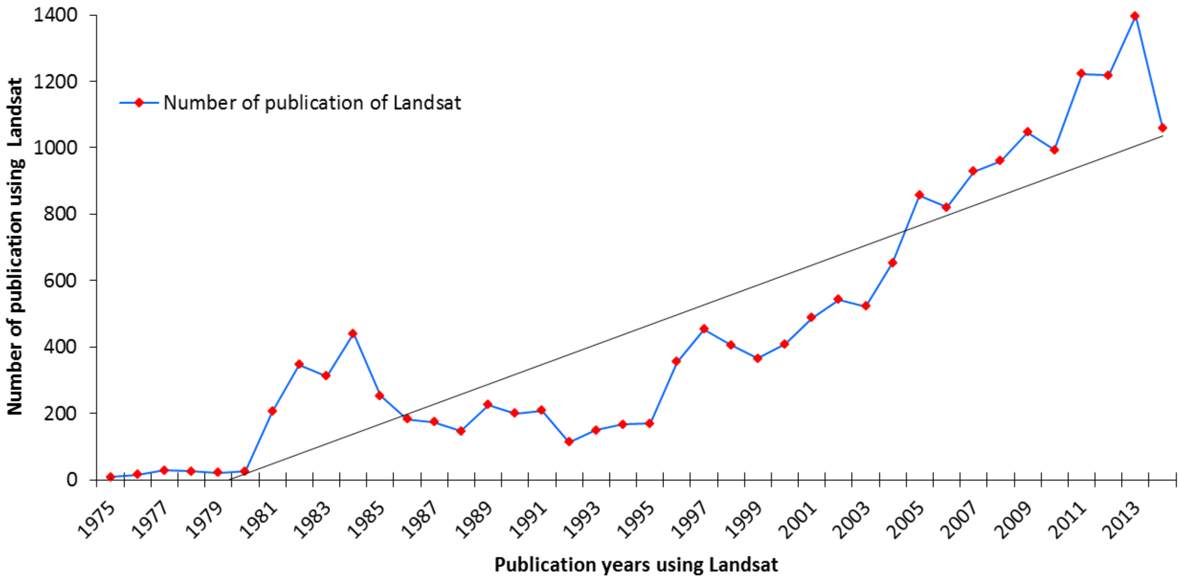


Figure 7. Number of publication in which Landsat data used per year

4.3.2. Number of publication in different document type

Among the total publication of Landsat after selection criteria (18,116), Article (5779) was the most frequently used document type, comparatively conference paper (4179), Book (21), and review (240) from 1975 to 2014. It can be also clear from Figure 8, the article (56.55%) and conference papers (40.89%) were relatively more published than books (0.21%) and review (2.35%) before opening Landsat data (1975 to 2007). However, after opening of Landsat data, user of Landsat data was given more priority for the production of articles (69.43%) whereas conference papers (29.06%) and reviews (0.86%) were reduced almost by half. Though, the productions of books have increased during this phase.

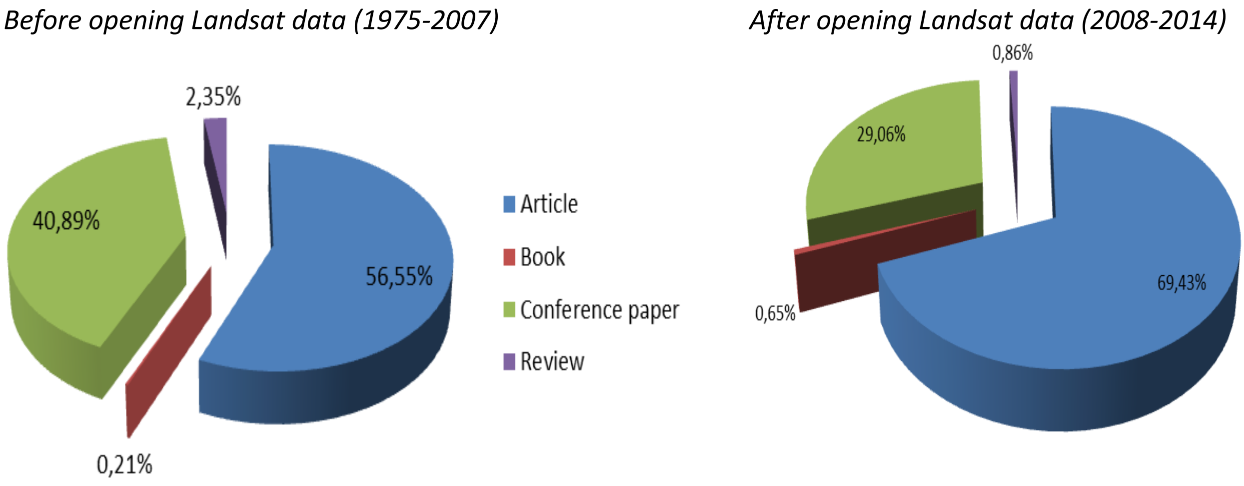


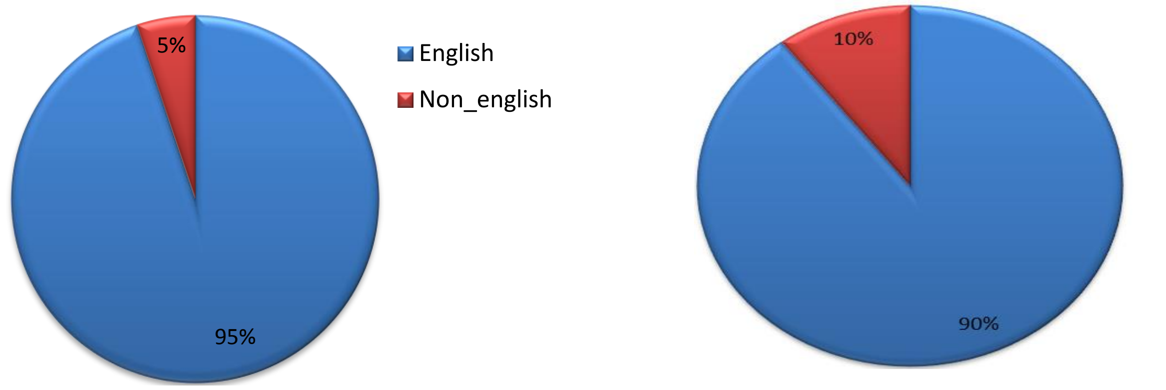
Figure 8. Distribution of document types using Landsat publication

4.3.3. Number of publication in different languages

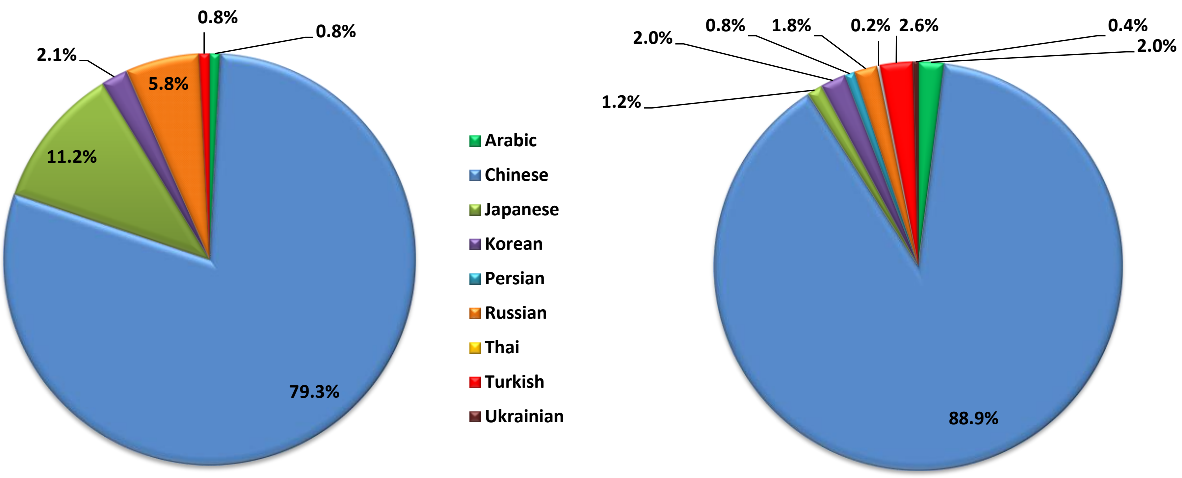
Before opening the Landsat data policy, 95% of all the above documents types were published in English. Whereas after opening Landsat data, about 90% of these document type were published in English and 10% publication used in other languages (Figure 9a). After opening Landsat data policy, Portuguese (51.3%), Spanish (23.9%) and Italian (3.6%) language publication were relatively increased, compared to other European language publications. The other European language publication was also increased slightly which were Catalan (0.6%), Estonian (0.3%) etc. However, the result shows that after opening the Landsat archive, German and French European language were decreased almost by half. Furthermore, Asian languages Chinese (88.9%) and Turkish (2.6%) language publication were highly increased whereas Arabic, Persian, Thai, Ukrainian languages publications were also slightly increased after releasing Landsat data. The detail of the publication in Asian and European languages can be seen in following Figure 9b and 9c.

a. English language verses non-engligh language publications

Before opening Landsat data (1975-2007) *After opening Landsat data(2008-2014)*



b. Comaprision of Asian language publications



c. Comaprision of European language publications

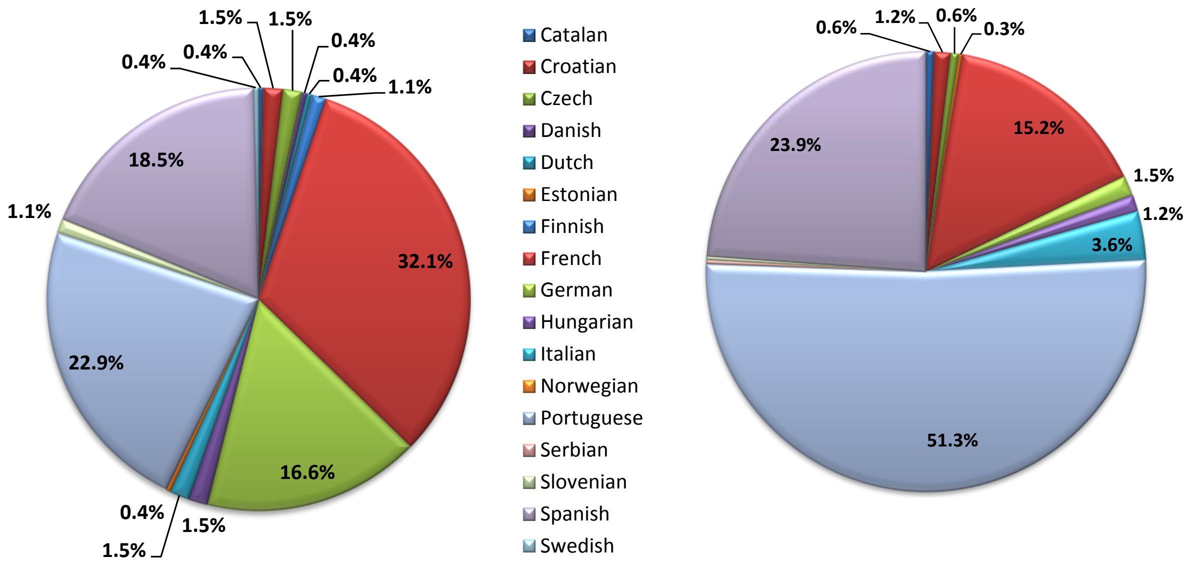


Figure 9. Number of publication in different languages using Landsat

4.3.4. Number of publication that use Landsat data per country/region per year

Looking at a historical trend of Landsat publication by regions, the use of Landsat publication was minimal in the very beginning phase of Landsat mission from 1975-1980 (Figure 10). After this period during 1981 to 1984, North America was the largest users and also publishers compared to others. As mentioned in the data Landsat policy section 2.2, NOAA increased the price of Landsat per imagery and also restricts the redistribution of Landsat imagery for commercialization purpose after 1984. This regulation highly affected the trend of use of Landsat imagery. It can be clearly seen in graph, after opening Landsat data policy Asia shows relatively faster growth of Landsat publication compared to others (Figure 10). This is also because of the biggest user of China which is shown in Appendix C and D. Whereas in North America and Europe, there were also increased of Landsat publication compared to Africa and South Africa. Furthermore, the result shows that before opening Landsat archive in 2008, Africa, South America and Oceania regions were limited users of Landsat however after opening the Landsat; the trend of Landsat use was slightly increased in scientific research.

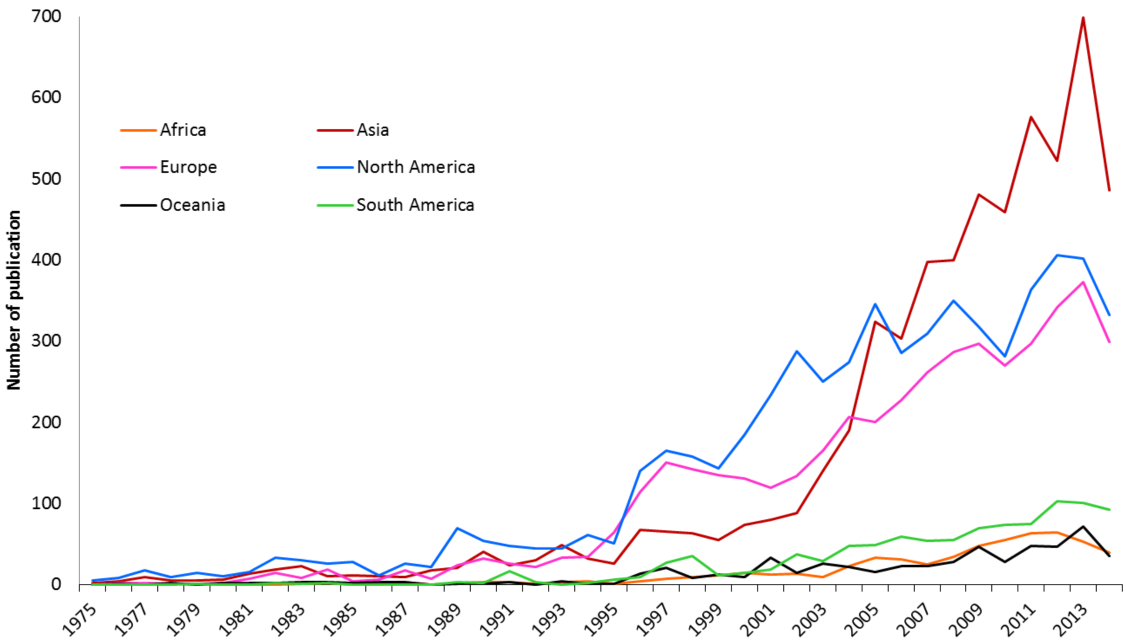


Figure 10. Trend of Landsat publication by region

4.3.5. Number of publication in which Landsat data used in different disciplines per year

Generally, Landsat data have wide variety of applications across several disciplines. In this research, seven applications categories of Landsat, broadly classified from U.S. Geological Survey report was adopted (Miller, Sexton et al. 2011). To assess the number of publications using Landsat data in these seven disciplines, individual applications keywords were used to execute the query with Landsat keyword in Scopus Advanced search. The result of the applications and their categories using Landsat data can be shown in following Table 10. It is also shown that after opening Landsat data the publication of using Landsat data in following individual applications have comparatively increased.

Table 10. Number of publication in which Landsat data used in different disciplines per year

SN	Applications for analysis	Individual applications	Before opening Total publications (1975-2007)	After opening Total publications (2008-2014)
1	**Agriculture**	Agricultural forecasting Agricultural management/ production/conservation	733	996
2	**Environmental sciences and management**	Biodiversity conservation	159	286
		Climate science/change	150	376
		Coastal science/monitoring/management	10	22
		Ecological/ecosystem science/management	908	1421
		Fish and wildlife science/ management	139	145
		Fire science/management	282	320
		Forest science/management	1460	1775
		Geology/glaciology	702	500
		Range/grassland	217	317

SN	Applications for analysis	Individual applications	Before opening Total publications (1975-2007)	After opening Total publications (2008-2014)
		science/management		
		Recreation science/management	25	25
		Water resources (for example, watershed management, water rights, hydrology)	570	792
3	**Land use/land cover**	Land use/land cover	1649	2223
4	**Planning and development**	Assessments and taxation	2132	2682
		Engineering/construction/surveying		
		Rural planning and development	11	17
		Urban planning and development	524	993
		Urbanization		
5	**Commercial**	Cultural resource management (for e.g., archaeology, anthropology)	99	126
		Real estate/property management		
		Software development		
		Telecommunications		
		Transportation		
		Utilities		
6	**Human needs**	Emergency/disaster management	964	1232
		Hazard insurance (for e.g., crop, flood, fire)		
		Humanitarian aid	8	16
		Public health		
7	**Oil/gas/minerals**	Oil and gas/mineral exploration/ extraction	219	186

4.3.6. Number of images used in publication (using Landsat during 1975 to 2014)

Use of multiple Landsat imagery analysis such as time series and near real time was also increased after opening Landsat data which are shown in Figure 11. Before opening Landsat data, Landsat users had to pay a lot for using multiple Landsat imagery in their research. When open Landsat data was released, they were able to download and analyse multiple Landsat imagery from the internet. Thus the scientific publication related to time series and near real time has increased as a result of opening the Landsat archive.

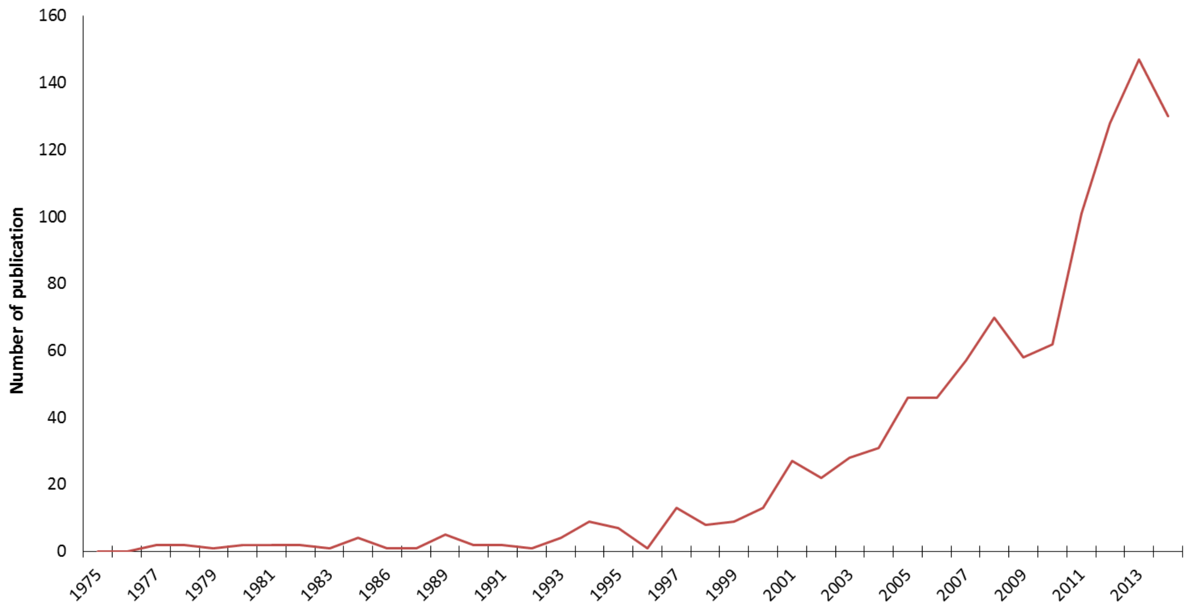


Figure 11. Publication trend of multiple time series Landsat imagery

4.3.7. Validation of search results

As shown in Table 9, out of 18,116 publications (from 1975 to 2014), 3% of publications (501 publications from 2002 to 2014) on the basis of title/abstract/keywords were randomly studied to evaluate the executed search query results of each indicators. The five indicators number of publication in which Landsat data used per year; number of publication in different document type; number of publication in different languages; number of publication that use Landsat data per country per year; number of image used in publication shows 100% correctness of search result (Table 11). The validation result shows that all selected publications, 501 (from 2002 to 2014) were used Landsat data and was published in English. Out of 501 publications, 472 were categorised in articles, 16 were in conference papers, 12 were in review and only 1 used for book. The 69 countries were found the user of Landsat. Similarly, 44 out of 501 publications were found the use of multiple time series and near real time analysis for their research however it was difficult to find the use of Landsat images only on the basis of title, abstract and keywords. Because the authors of publication are not included the number of images use in title, abstract and keywords which also effects the search results.

Table 11. Validation of search results

SN	Indicators	% of correctness	Description of validation results
1	Number of publication using Landsat	100	All publication (501 in total) use Landsat data.
2	Number of publication using Landsat in different document type		Article (472), book (1), conference paper (16), review (12), all classifies under the respective category.
3	Number of publication using Landsat in different languages		501 English publication
4	Number of publication that use Landsat data per country		59 country were found using Landsat data

SN	Indicators	% of correctness	Description of validation results
5	Number of images used in publication (eg time series and near real time)		Out of 501, only 44 papers have used time-series

However number of publication in which Landsat data used in different disciplines indicator have found seven applications disciplines; agriculture, environmental sciences and management, land use/land cover, planning and development, commercial, human needs, and oil/gas/minerals. This makes confusion to distinguish the degree of coverage of Landsat use applications. Common measures of confusion matrix; accuracy user accuracy, producer accuracy and overall accuracy were derived using Foody (2002). The result of validation is presented as confusion matrix in Table 12. It was found that out of 501 publications, 34 publications have fully coverage in agriculture application using Landsat. Whereas land use/land cover application were highly misclassified in agriculture application. Compare to other applications, environmental sciences and management applications have widely coverage of publications (354 publications out of 501). This is because environmental sciences and management application covered 11 individual categories (Table 10). Furthermore, it was also found that the maximum misclassified application was land use/land cover application in environmental sciences and management application. Among 501 publications, 94 publications was found in land use/ land cover application and some were misclassified with agriculture (12), environmental sciences and management application (28), planning and development (7) and human needs (1). However, in case of human need, the maximum publication was also related with land use and land cover. The use of Landsat in commercial and oil/gas/minerals was found very low but 100% accurate. Even though these application were slightly misclassified with other applications. Hence, overall the accuracy of Landsat use in seven disciplines was 68.17%.

Table 12. The confusion matrix of validated search results in different discipline

Sn	Landsat publication in different discipline	Agriculture	Environmental sciences and management	Land use/ land cover	Planning and development	Commercial	Human needs	Oil/gas/minerals	Total	Producer Accuracy (%)
1	Agriculture	34	6	11	1	0	1	0	53	64.15
2	Environmental sciences and management	19	354	50	10	0	10	0	443	79.91
3	Land use/land cover	12	28	94	7	0	1	0	142	66.20
4	Planning and development	1	11	25	42	0	1	0	80	52.50
5	Commercial	1	1	2	3	5	1	0	13	38.46
6	Human needs	2	58	8	0	0	53	0	121	43.80
7	Oil/gas/minerals	1	3	1	0	0	0	7	12	58.33
Total		70	461	191	63	5	67	7	864	
User Accuracy (%)		48.57	76.79	49.21	66.67	100	79.10	100		
over all accuracy (%)		68.17%								

5. Discussion

This chapter discusses the answers obtained from the research questions defined in the earlier chapter. In the first section, the main findings and limitations of the applied research method in this research are discussed. Second, section discusses the results for assessing the impact of open Landsat data on science.

5.1. Critical evaluation of results — impact of open Landsat data on science

To assess the scientific impact of open Landsat data for three assessment categories, SLR method with Scopus was used on the basis of six developed indicators:1) number of publication in which Landsat data used per year; 2) number of publication in different document type; 3)number of publication in different languages; 4)number of publication that use Landsat data per country per year; 5)number of publication in which Landsat data used in different disciplines per year; and 6) number of image used in publication.

The first indicator which was analysed was the number publication using Landsat data per year from 1975 to 2014. The results of number of publication trend in Figure 7 revealed the use of Landsat has relatively increased after opening all the Landsat data. These results are consistent with the research studies conducted by Michael et al. (2012) and (Miller, Sexton et al. 2011) who found that the number of Landsat imagery downloads from December 2008 have slightly increased. In addition, after opening Landsat archive it was also found from the Landsat project statistics, the number of imagery download has significantly increased from 17th August, 2009 (1 million) to 6th January, 2015 (22 million) (USGS 2015) which is shown in Annex A. USGS claim that the increase of Landsat use is due to opening Landsat archive (USGS 2015).

The result of a number of publications using Landsat is compared with the overall trend of Scopus publication (Figure 12). The comparison of both trends suggest that the increasing trend of Landsat publications could be also partially because of the general trend of increasing number of publications rather than just opening of Landsat archives (Wuchty, Jones et al. 2007). However, during 1982 to 1984 period, the use of Landsat shows increasing trend, it could be because of the beginning phase of Landsat and Landsat user may be more fascinated and curious about Landsat use for their research. According to Land Remote Sensing Commercialization Act in 1984, NOAA was instructed for privatization of Landsat satellite data. Hence after this period, the price of Landsat data were increased as well the data were restricted for redistribution. Because of these limitations, the Landsat users were reduced to use Landsat data for their research and the Landsat trend was decline. Again when EOSAT returned back all operational responsibility to the USGS, commercial right was again renounced by USGS to meet the user demand. This also affects the scientific publication and the trend of Landsat use slightly increases. Furthermore, when the open data policy was announced in 2008, the use of Landsat and scientific publication were relatively increased. However the publication need some time to publish and index in Scopus. This effect can be also seen in the trend line of Landsat use in Figure 12.

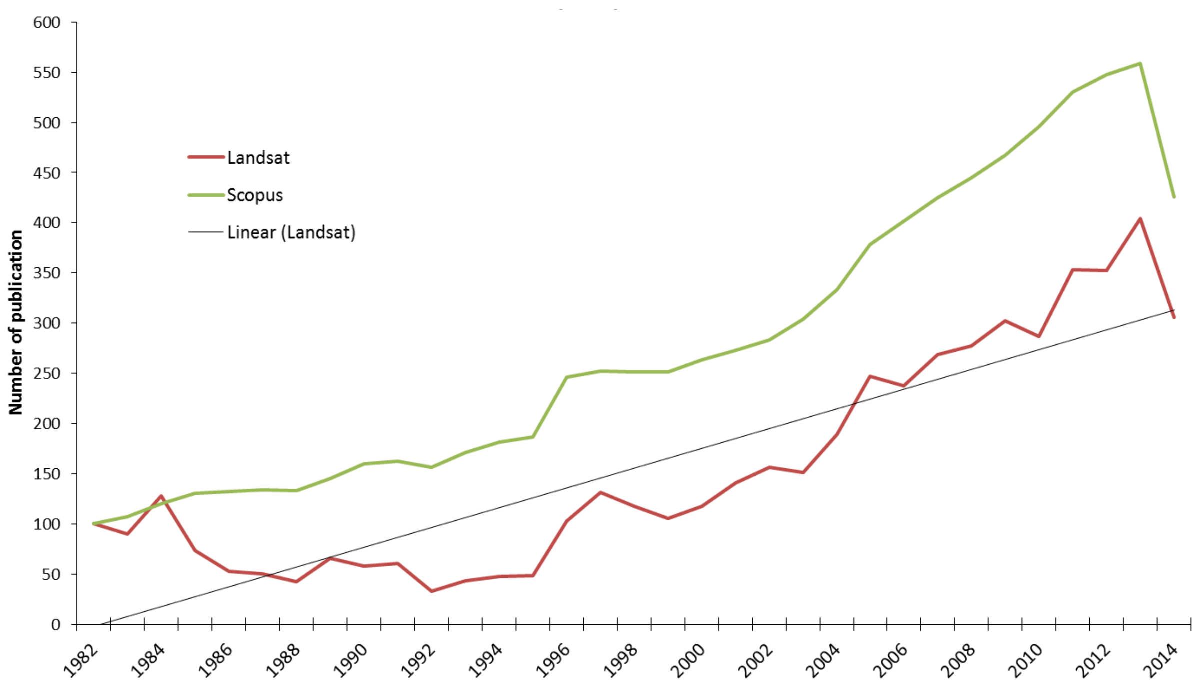


Figure 12. Trend of Landsat and Scopus publication

The second indicator was used to analyse number of publication in different document types which were article, book, conference paper and review. The result of the number of document types revealed that the most of the scientists of Landsat were interested to publish their research on articles compare to conference paper (Figure 8). The rate of conference paper was reduced a lot after opening Landsat archive and slightly increased in books. This could be possible because of scientists were interested to write journal papers rather than the conference papers.

Third indicator that was analysed number of publication in different languages to determine the scientists of Landsat were published their research in which languages. Before opening the Landsat archive, the 95% publications were published in English. However after free data policy 2008, the publications in English were slightly decreased by 90%. Other users from developing country were able to use Landsat data without any price, the publication were slightly increased in other languages too for instance Arabic, Chinese, Korean, Persian, Thai, Ukrainian, Turkish, Spain, Portuguese, Italian, and Finnish, Estonia, Serbian, etc. (Figure 9).

Fourth indicator was used to analyse the geographical scope of Landsat use with number of publication that use Landsat data per country/region. The result shows that after opening archive the biggest user of Landsat was Asian region compare to others which is especially in China. This finding also supports user survey of USGS (2013).

Fifth indicator, number of publication in different discipline using Landsat was analysed to assess the variety of application of Landsat use before and after opening Landsat archive. According to USGS report (2011), the search query on Scopus were executed for only seven application categories which is shown in Table 10 (Miller, Sexton et al. 2011); other two application categories education and Legal/security were excluded. The results revealed that after opening Landsat archive, the Landsat data used in environmental science and management applications were increased drastically. It may be because of agreement done by all parties under the United Nations Framework Convention on

Climate Change (UNFCCC) to mitigate the effects of global climate change (UNFCCC 2015). Whereas, second application category of Landsat was found in planning and development application. Similarly, land use/land cover, human need, agriculture and commercial application slightly increased Landsat use for their research. Overall, it was found that out of all applications, agricultural, ecological/ecosystem science/management, forest science/management, water resources, land use/land cover, emergency/disaster management/hazard insurance, urban planning and development applications widely used Landsat data and the result of using Landsat in various application are relatively consistent with the use of Landsat imagery in various application of USGS report (Miller, Sexton et al. 2011).

The last indicator which was used to assess number of multiple Landsat images used in publication especially time series and near-real time applications which need multiple images for analysis. Time-series studies require systematic acquisition of images in multiple times. In this SLR method with large amount of database, it was difficult to say how many Landsat images have been used by scientist for their research. However, time series and near-real time analysis always require the use of multiple Landsat images. Consequently, these two analysis: time series with Landsat and near-real time with Landsat Keyword were executed on the Scopus advanced search query for assessing the trend. After opening the Landsat archive, there was a significant increase of the use of multiple Landsat images in time series and near real time analysis for earth observation and monitoring (Gutman, Huang et al. 2013).

Free and open data policy of Landsat provides the opportunity to use the analysis at national, regional and global scale. Hansen, Potapov et al. (2013) uses 654,178 time series of Landsat images to perform the global forest cover change analysis from 2002 to 2012. The analysis is viewable via web portal (http://earthenginepartners.appspot.com/science-2013-global-forest) and the screenshot of the portal is shown in Figure 13. Furthermore, free and open data policies of Landsat would enable user for greater use of these data for public good and foster greater transparency of the development, implementation, and reactions to policy initiatives that affect the forests monitoring and reduce emissions from deforestation and forest degradation (REDD+) (De Sy, Herold et al. 2012). Similarly, the result of this research shows that after opening the Landsat archive, trend for using Landsat data for near real time monitoring have increased. More and more scientists are using for near real-time forest and drought monitoring (Verbesselt, Zeileis et al. 2012, Xin, Olofsson et al. 2013). Thus, It can be concluded that opening the Landsat data archive have the biggest positive impact on the near real time and time series analysis applications.

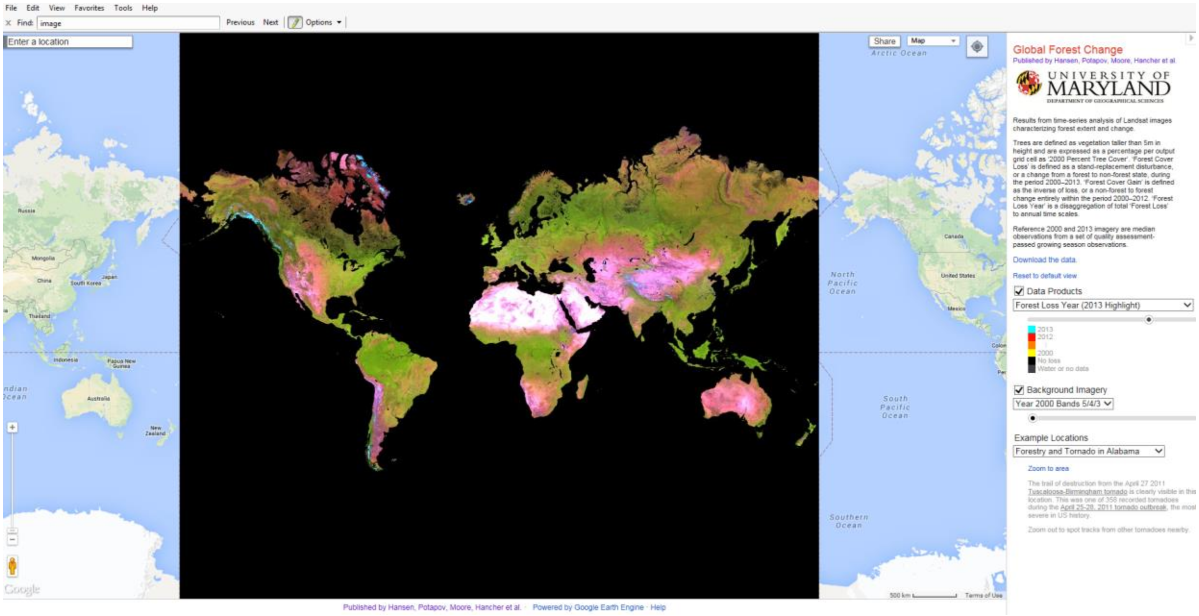


Figure 13. High-Resolution Global Maps of 21st-Century Forest Cover Change (Hansen, Potapov et al. 2013)

5.2. Critical evaluation of applied Research method

The central aim of this research was to explore possible research methods and select out of these possible methods the most suitable method for assessing the impact of Open Landsat data on scientific society. To meet the requirement of Landsat characteristics, all selected research methods are evaluated based on fitness for use criteria and their specified measurement categories (described in Table 4). Although, these measurement categories are defined as "objective" as possible and preferably measurable, still there are various subjective measurements involved. For instance, reliability and transparency itself have subjective measurements. Such subjective measurements can bias the overall result of the findings. Hence, the countermeasures such as measurability of resources, time bound and reproducibility that have taken to minimize the research bias. The SLR method was selected as the most suitable and relevant method for this research (Table 5). Identifying the factors influencing SLR method is important to understanding the role this research method in an impact assessment of open Landsat data in science. The main influencing factors are explained in detail below.

SLR was influenced by two main factors: search query and data source. First, SLR method is query based search approach and the outcome of SLR method is directly dependent on the construction of search query as well as inclusion/exclusion criteria (Keele 2007, Kitchenham, Pearl Brereton et al. 2009, Ghafari, Saleh et al. 2012, Wen, Li et al. 2012). In this research, the multi-level search queries and inclusion/exclusion criteria were used to obtain the results of six Landsat indicators. In the search query construction for the first four indicators, the Landsat as the main keyword with inclusion/exclusion was used to search the results and there was no confusion to obtain the results of the first four indicators. Hence, the search query was generic for the first four indicators which can be shown in Annex B. However, it was difficult to construct the generic query for the fifth indicator, number of publication in which Landsat data used in different disciplines per year hence the results were also influenced by search query. The overall the accuracy of search result in seven discipline was 68.17% (Table 12). This is because multiple disciplines of Landsat has multiple keywords. The Scopus search engine have limitations related to finding and summarizing all the keywords and their synonyms to construct the single search query applicable for Landsat data use. Therefore in this research, the search query was built based on the USGS report of Landsat use where the multiple applications of Landsat were defined on the basis of Landsat user survey (Miller, Sexton et al. 2011). For instance, environmental science and management is multidisciplinary application field and multiple keywords are used to represent the different categories of this field. Hence, it is difficult to construct a generic search query that can extract all relevant data for this field. Similarly, land use/land cover application was also frequently used term in multiple disciplines. Thus, this application was found to be over represented in comparison to other categories (Table 12).

Second, it was observed that the frequency of data collected through SLR largely depends on data source (Ghafari, Saleh et al. 2012). Each data source such as Scopus, Web of Science, Google scholar, PubMed etc. have their own data storage. In this research, Scopus was chosen as a main source of data (Table 7). However, this data source has the following limitations:

1. Indexing process of new publication in this data source is long process and takes time to update. Hence the actual impact can be seen after some time. For instance, this may have consequences that all the publications of Landsat in year 2014 may not be included. This could be the reason of decreasing trend of publications using Landsat in 2014.
2. Search string for Scopus can be repeated for the search result in Scopus itself. While the search string created for Scopus cannot be used in other data sources (web of science, Google scholar, PubMed etc).
3. It is difficult to provide an integrated search across multiple databases.

6. Conclusions and Recommendations

This chapter presents the conclusions of the research and recommendations for future research.

6.1. Conclusions

This research has been carried out with an objective of assessing the impact of open Landsat data on science. The objective is supported by three research questions. The answers of these research questions are summarised as below.

RQ1. What are the possible research methods for determining the impact of open Landsat data on science?

The possible research methods to assess the impact of open Landsat data are: case study, expert survey, and systematic literature review (SLR). An evaluation of these three methods is done based on six aspects of fitness for use criteria, i.e., achievability of resources, reliability, measurability of resources, time bound, reproducibility, and transparency covering the research requirements for measuring the impact of open Landsat data on science. Among these three possible methods, SLR method was selected as the best fit for the purpose of assessing the scientific impact of Landsat data. The research revealed that the SLR is able to collect and analyse the use of Landsat data in science since 1972 to till date. The results of this research also show that SLR method provides 1) transparent of data source 2) consistent meta-analysis information 3) avoid grey literature (not peer-reviewed or unpublished literature) 4) reproducible data source, and 5) no time restriction to collect data.

RQ2. How can the identified research method be applied to this research?

Systematic literature review (SLR) is applied using Scopus on a specific sequence of steps which are data sources, search strategy, selection criteria, data extraction, synthesis and analysis of the extracted data. In total, 23,604 scientific publications were obtained using Landsat as key word. After applying inclusion and exclusion criteria, 18,116 publications were selected for further analysis since 1975 to 2014. But it is difficult to evaluate all the publication one by one; hence 3% of Landsat publication (501 in total) from 2002 to 2014 was randomly selected to validate the search result of applied SLR method. The validation accuracy of five indicators number of publication in which Landsat data used per year; number of publication in different document type; number of publication in different languages; number of publication that use Landsat data per country per year; number of image used in publication were recorded with the highest accuracy (100%), whereas Landsat publication in different discipline was found to have lowest overall accuracy (68.17%).

Finally, this research has contributed to developed six Landsat indicators. The result shows that out of six indicators, number of publication in which Landsat data used per year; number of publication in different document type; number of publication in different languages; number of publication that use Landsat data per country per year; number of image used in publication easily show the impact of Landsat use. However the numbers of Landsat images are difficult to find only by reading abstract, title and keywords which was not included by author. In case of Number of publication in which Landsat data used in different disciplines per year, several application categories are included in one application domain, it was found over representative application categories during validation. Overall, the result shows that these indicators were capable to measure the overall impact of Landsat data on science.

RQ3. What insight can be obtained from the result?
In this research, some significant points on the worldwide research trends of using Landsat data was obtained throughout the period from 1972 to till date. The effort provided a systematically structural picture as well as the clues to the impacts of open Landsat data in science. The results of these indicators shows that open data policy have revolutionized the utilization of Landsat data throughout the globe. The scientific publications using Landsat data in Asia have increased by 13% where as in Africa and South America have increased by 1%. English was by far the dominant language whereas the scientific publications in native languages (other than English) have increased by 5% after opening Landsat archive. It indicated that the use of opening Landsat becomes more globally exploited. Apparently, the document types of articles and books were increased and the uses of Landsat archive in diverse discipline were significantly increased. As a result of free Landsat data, there is clear evidence that multiple Landsat images for time-series and near-real-time analysis have been used by the scientists significantly more often.

To sum up the conclusion, this research presents the research methods and their potential ways of evaluation to assess the impact of open Landsat data on science. Three methods; case study, expert survey, and systematic literature review (SLR); and six fitness for use criteria; achievability of resources, reliability, measurability of resources, time bound, reproducibility, and transparency are discussed. In addition these criteria, procedures for measurement categories are also defined. Finally, the use of SLR method for assessing the impact of open Landsat data on science is underlined. The findings of SLR result show that Open Landsat data have significant impact on science. Hence, the result analysis by using this new SLR bibliographic method can help relevant researchers realise the panorama of global Landsat use on science and establish the further research direction.

6.2. Recommendations for future research

Following recommendations have been proposed for further research:

1) The fitness for use criteria designed in Chapter 3 are supposed to incorporate entire facets of research methods used for assessing the impact of open Landsat data in science. Currently, only six indicators, four on the basis of SMART criteria and two science criteria, namely transparency and reproducibility, are incorporated in the framework. Further research is required to incorporate all characteristics of open data. At the same time, further research is required to validate the strength of the framework.
2) In this research, only one approach i.e. SLR was used to assess the impact of Landsat data. More than one approach is recommended to the investigation of a research question in order to enhance confidence of the findings.
3) This research used SLR to assess the impact of open Landsat data in science. The following further research steps are identified to improve the contribution of this method.
   - This research used only Scopus data source. Further research is recommended to include and compare the results with other data sources such as web of science.
   - Validation is crucial for SLR method. It is recommended that the experts or train reviewers of Landsat checks the results independently in order to minimize the potential bias.

References

Agumya, A. and G. J. Hunter (1999). "A risk-based approach to assessing the â€˜Fitness for Useâ€™of spatial data." Urisa Journal 11(1): 33-44.

Andrews, D., B. Nonnecke and J. Preece (2007). "Conducting research on the internet: Online survey design, development and implementation guidelines."

Bauer, F. and M. Kaltenböck (2011). "Linked Open Data: The Essentials." Edition mono/monochrom, Vienna.

Beard, J., F. Feeley and S. Rosen (2009). "Economic and quality of life outcomes of antiretroviral therapy for HIV/AIDS in developing countries: a systematic literature review." AIDS care 21(11): 1343-1356.

Bogue, R. L. (2005). " Use S.M.A.R.T. goals to launch management by objectives plan." Retrieved 23 Jan, 2015.

Bregt, A. K. and Ł. Grus (2014). Wat zijn de effecten van een open basisregistratie topografie na twee jaar? Wageningen, The Netherlands, Wageningen University.

Bruin, S. d., A. Bregt and M. v. d. Ven (2001). "Assessing fitness for use: the expected value of spatial data sets." International Journal of Geographical Information Science 15(5): 457-471.

Cavaye, A. L. M. (1996). "Case study research: a multi-faceted research approach for IS." Information Systems Journal 6(3): 227-242.

Čeh, M., D. Smole and T. Podobnikar (2004). "Geodata–are they Accessible and Useful?" AGILE, Heraklion, Crete: 789-794.

Chang, K.-t. (2010). Introduction to geographic information systems, McGraw-Hill New York.

Chen, J., X. Zhu, J. E. Vogelmann, F. Gao and S. Jin (2011). "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images." Remote Sensing of Environment 115(4): 1053-1064.

Creswell, J. W. (2013). Research design: Qualitative, quantitative, and mixed methods approaches, Sage.

De Sy, V., M. Herold, F. Achard, G. P. Asner, A. Held, J. Kellndorfer and J. Verbesselt (2012). "Synergies of multiple remote sensing data sources for REDD+ monitoring." Current Opinion in Environmental Sustainability 4(6): 696-706.

Downing, S. M. (2004). "Reliability: on the reproducibility of assessment data." Medical Education 38(9): 1006-1012.

Draeger, W. C., T. M. Holm, D. T. Lauer and R. J. Thompson (1997). "The availability of Landsat data: Past, present, and future." Photogrammetric engineering and remote sensing 63(7): 869-875.

Falagas, M. E., E. I. Pitsouni, G. A. Malietzis and G. Pappas (2008). "Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses." The FASEB Journal 22(2): 338-342.

Fricker, R. D. and M. Schonlau (2002). "Advantages and disadvantages of Internet research surveys: Evidence from the literature." Field Methods 14(4): 347-367.

Gable, G. G. (1994). "Integrating case study and survey research methods: an example in information systems." European journal of information systems 3(2): 112-126.

Ghafari, M., M. Saleh and T. Ebrahimi (2012). "A federated search approach to facilitate systematic literature review in software engineering." International Journal of Software Engineering and Applications 3(2): 13-24.

Goecks, J., A. Nekrutenko and J. Taylor (2010). "Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences." Genome Biol 11(8): R86.

Goff, A., E. Rose, S. Rose and D. Purves (2007). "Does PTSD occur in sentenced prison populations? A systematic literature review." Criminal Behaviour and Mental Health 17(3): 152-162.

Goward, S., T. Arvidson, D. Williams, J. Faundeen, J. Irons and S. Franks (2006). "Historical Record of Landsat Global Coverage." Photogrammetric Engineering & Remote Sensing 72(10): 1155-1169.

Gurstein, M. B. (2011). "Open data: Empowering the empowered or effective data use for everyone?" First Monday 16(2).

Gutman, G., C. Huang, G. Chander, P. Noojipady and J. G. Masek (2013). "Assessment of the NASA–USGS Global Land Survey (GLS) datasets." Remote Sensing of Environment 134(0): 249-265.

Hahmann, S. and D. Burghardt (2013). "How much information is geospatially referenced? Networks and cognition." International Journal of Geographical Information Science 27(6): 1171-1189.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice and J. R. G. Townshend (2013). "High-Resolution Global Maps of 21st-Century Forest Cover Change." Science 342(6160): 850-853.

Huijboom, N. and T. Van den Broek (2011). "Open data: an international comparison of strategies." European journal of ePractice 12(1): 1-13.

InnvÃ¦r, S., G. Vist, M. Trommald and A. Oxman (2002). "Health policy-makers' perceptions of their use of evidence: a systematic review." Journal of Health Services Research & Policy 7(4): 239-244.

Irish, R. (2000). "Landsat 7 science data users handbook." National Aeronautics and Space Administration (NASA), Report: 430-415.

Joy, M. (2007). Research methods in education, The Higher Education Academy Innovation Way, York Science Park, Heslington, York YO10 5BR.

Keele, S. (2007). Guidelines for performing systematic literature reviews in software engineering, Technical report, EBSE Technical Report EBSE-2007-01.

Keresztesi, M. and C. Oberman (1982). "The science of bibliography: Theoretical implications for bibliographic instruction." Theories of Bibliographic Education, edited by. New York: RR Bowker: 1-26.

Kitchenham, B., O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman (2009). "Systematic literature reviews in software engineering–a systematic literature review." Information and software technology 51(1): 7-15.

Laine, C., S. N. Goodman, M. E. Griswold and H. C. Sox (2007). "Reproducible Research: Moving toward Research the Public Can Really Trust." Annals of Internal Medicine 146(6): 450-453.

Latour, B. (1987). Science in action: How to follow scientists and engineers through society, Harvard university press.

Lauer, D. T., S. A. Morain and V. V. Salomonson (1997). "The Landsat program: Its origins, evolution, and impacts." Photogrammetric Engineering and Remote Sensing 63(7): 831-838.

Loveland, T. R. and J. L. Dwyer (2012). "Landsat: Building a strong future." Remote Sensing of Environment 122(0): 22-29.

Macauley, M. K. (2009). "Earth observations in social science research for management of natural resources and the environment: identifying the Landsat contribution." Journal of Terrestrial Observation 1(2): 6.

Malenovský, Z., H. Rott, J. Cihlar, M. E. Schaepman, G. García-Santos, R. Fernandes and M. Berger (2012). "Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land." Remote Sensing of Environment 120(0): 91-101.

Mao, N., M.-H. Wang and Y.-S. Ho (2010). "A bibliometric study of the trend in articles related to risk assessment published in Science Citation Index." Human and Ecological Risk Assessment 16(4): 801-824.

Miller, H. M., L. Richardson, S. R. Koontz, J. Loomis and L. Koontz (2013). "Users, Uses, and Value of Landsat Satellite Imagery—Results from the 2012 Survey of Users."

Miller, H. M., N. R. Sexton, L. Koontz, J. Loomis, S. R. Koontz and C. Hermans (2011). "The users, uses, and value of Landsat and other moderate-resolution satellite imagery in the United States—Executive report." US Geological Survey Open-File Report 1031: 42.

Miri, S. M. and P. Bahmani (2012). "Indexing in ISI Web of Sciences: The opportunities and threats." Jundishapur J. Microbiol 5: 381-383.

NASA&USGS. (2014, December 5, 2014). "History of Landsat (From the Beginning)." Retrieved 09 December, 2014, from http://landsat.gsfc.nasa.gov/?page_id=2281.

Navarro Sada, A. and A. Maldonado (2007). "Research Methods in Education. Sixth Edition - by Louis Cohen, Lawrence Manion and Keith Morrison." British Journal of Educational Studies 55(4): 469-470.

Pecora, W. T. (1966). Earth Resources Observation Satellite (EROS): A Department of Interior Program to Utilize Space-acquired Data for Natural and Human Resource Management. Washington D.C., U.S. Department of Interior: 20.

Petticrew, M. and H. Roberts (2008). Systematic reviews in the social sciences: A practical guide, John Wiley & Sons.

Ram, K. (2013). "Git can facilitate greater reproducibility and increased transparency in science." Source code for biology and medicine 8(1): 7.

Roy, D., M. Wulder, T. Loveland, W. CE, R. Allen, M. Anderson, D. Helder, J. Irons, D. Johnson and R. Kennedy (2014). "Landsat-8: Science and product vision for terrestrial global change research." Remote Sensing of Environment 145: 154-172.

Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.E, R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y. Shuai, R. Trezza, J. Vogelmann, R. H. Wynne and Z. Zhu (2014). "Landsat-8: Science and product vision for terrestrial global change research." Remote Sensing of Environment 145(0): 154-172.

Scarletto, E. A. (2014). "Mapping the Literature of GIS." College & Research Libraries 75(2): 179-201.

Schell, C. (1992). "The value of the case study as a research strategy." Manchester, UK: University of Manchester, Manchester Business School: 1-15.

Schleyer, T. K. and J. L. Forrest (2000). "Methods for the design and administration of web-based surveys." J Am Med Inform Assoc 7(4): 416-425.

Selding, P. B. d. (2014, 23 Jan, 2014). "France To Make Older Spot Images Available to Researchers for Free." Retrieved 01 September, 2014, from http://www.spacenews.com/article/civil-space/39234france-to-make-older-spot-images-available-to-researchers-for-free.

Sloane-Seale, A. (2009). "Research Design: Qualitative, Quantitative, and Mixed Methods Approaches." Canadian Journal of University Continuing Education 35(2).

UNFCCC. (2015). "United Nations Framework Convention on Climate Change." Retrieved 28 February 2015, from http://newsroom.unfccc.int/.

USGS. (2013, 09/19/13). "Landsat Project Description." Retrieved 26 Septembet, 2014, from http://landsat.usgs.gov/about_project_descriptions.php.

USGS. (2015, 01/23/2015). "Landsat Project Statistics." The Landsat Archive Retrieved 18 Feb, 2015, from http://landsat.usgs.gov/Landsat_Project_Statistics.php.

Verbesselt, J., A. Zeileis and M. Herold (2012). "Near real-time disturbance detection using satellite image time series." Remote Sensing of Environment 123: 98-108.

Welsh, B. C. and D. P. Farrington (2002). Crime prevention effects of closed circuit television: a systematic review, Citeseer.

Wen, J., S. Li, Z. Lin, Y. Hu and C. Huang (2012). "Systematic literature review of machine learning based software development effort estimation models." Information and Software Technology 54(1): 41-59.

Whitfield, P. H. (2012). "Why the provenance of data matters: assessing “fitness for purpose” for environmental data." Canadian Water Resources Journal 37(1): 23-36.

Williams, D. L., S. Goward and T. Arvidson (2006). "Landsat: Yesterday, Today, and Tomorrow." Photogrammetric Engineering & Remote Sensing 72(10): 1171-1178.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder and R. Wynne (2008). "Free access to landsat imagery." Science 320(5879): 1011.

WorldBank. (2015). "World Bank Data Catalog." Retrieved 10 February, 2015, from http://data.worldbank.org/country.

WorldBank, T. (2014). Transport & ICT Global Practice, Transport & ICT Global Practice: 20.

Wuchty, S., B. F. Jones and B. Uzzi (2007). "The increasing dominance of teams in production of knowledge." Science 316(5827): 1036-1039.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland and C. E. Woodcock (2012). "Opening the archive: How free data has enabled the science and monitoring promise of Landsat." Remote Sensing of Environment 122: 2-10.

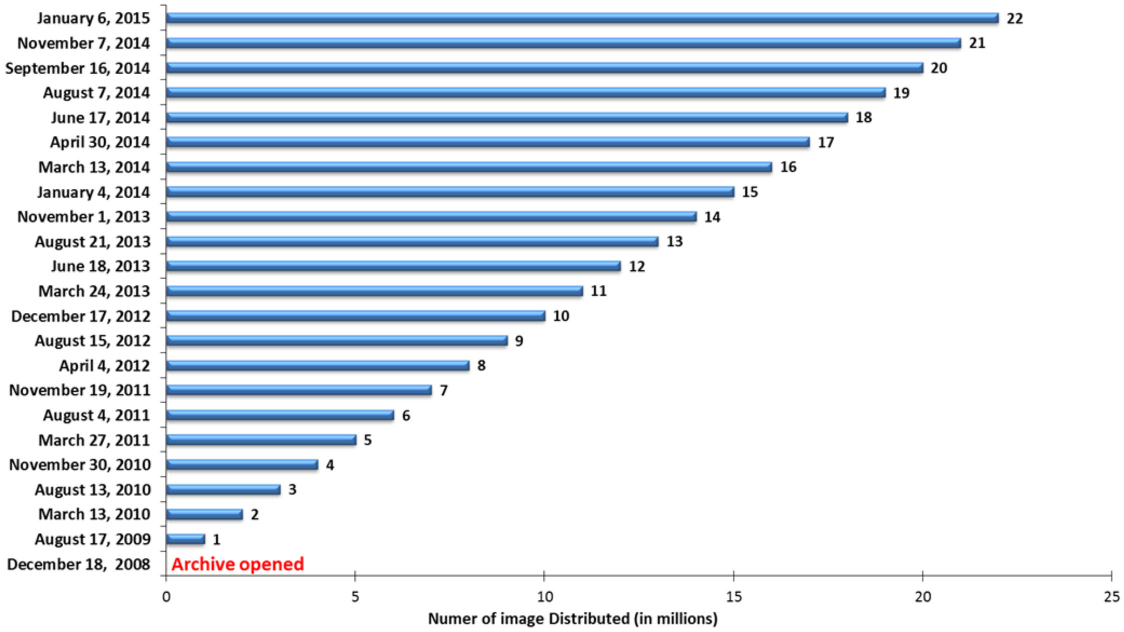
Wulder, M. A., J. C. White, J. G. Masek, J. Dwyer and D. P. Roy (2011). "Continuity of Landsat observations: Short term considerations." Remote Sensing of Environment 115(2): 747-751.

Xin, Q., P. Olofsson, Z. Zhu, B. Tan and C. E. Woodcock (2013). "Toward near real-time monitoring of forest disturbance by fusion of MODIS and Landsat data." Remote Sensing of Environment 135: 234-247.

Zainal, Z. (2007). "Case study as a research method." Jurnal Kemanusiaan(9): 1-6.

Annexes

A. Opening the Landsat archive from December 2008 (USGS 2015)



B. The list of search queries for SLR method using Scopus data source

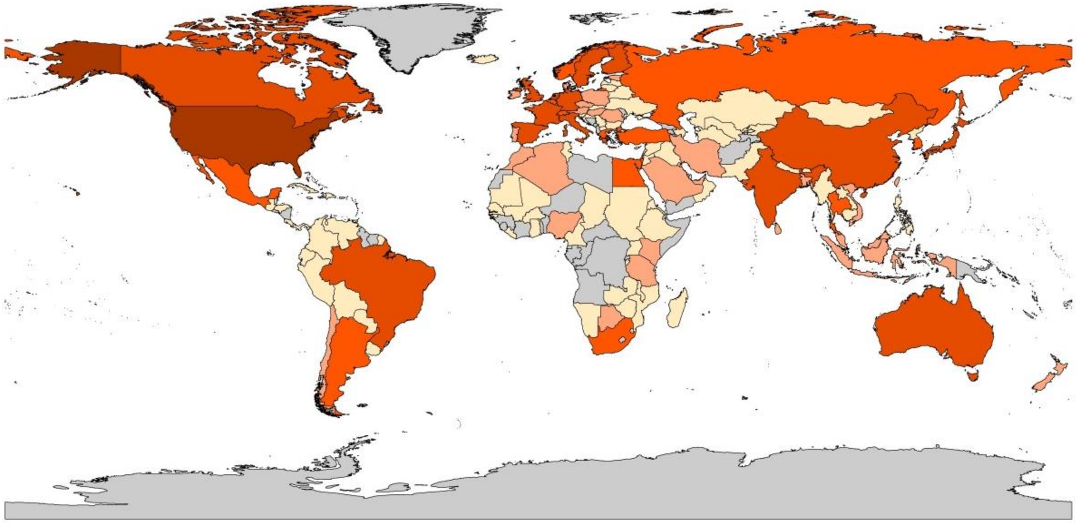
SN	Type in Landsat		Applied Search queries in this research	Number of Publication
1	**Landsat**		TITLE-ABS-KEY (landsat)	23604
2	**Landsat in agriculture Discipline**		TITLE-ABS-KEY (landsat) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "p") OR LIMIT-TO (SRCTYPE , "k") OR LIMIT-TO (SRCTYPE , "b")) AND (EXCLUDE (PUBYEAR , 2015))	18116
3	**Landsat in agriculture**	Agricultural forecasting/ Agricultural management/ production/conservation	TITLE-ABS-KEY (landsat*) AND TITLE-ABS-KEY (agricult* OR "Agriculture forecast*" OR "Agriculture manage*" OR "Agriculture produce*" OR "Agriculture conserv*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	1729
4	**Environmental sciences and management**	Biodiversity conservation	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("biodivers*" OR "biodiversity conserv*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	445
5		Climate science/change	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("climate change*" OR "climat* change*" OR "climat* science") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p"))	526
6		Coastal science/ monitoring/management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("coast* monitor*" OR "coast* science" OR "coast* manage*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (SRCTYPE , "p")) AND (EXCLUDE (PUBYEAR , 2015))	32
7		Ecological/ecosystem science /management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("ecolog*" OR "ecosys*" OR "ecolog* science" OR "ecolog* manage" OR "ecosys* science" OR "ecosys* manage*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	2329
8		Fish and wildlife science/ management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("fish*" OR "wildlife" OR "fish* and wildlife*" OR "fish* and wildlife science" OR "fish* and wildlife manage*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	284

SN	Type in Landsat		Applied Search queries in this research	Number of Publication
9		Fire science/management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("fire manage*" OR "fire*" OR "fire science") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	602
10		Forest science /management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("forest manage*" OR "forest*" OR "forest science") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	3235
11		Geology/glaciology	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("geolog*" OR "glaciolog*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	1202
12		Range/grassland science /management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("grass* manage*" OR "grassland*" OR "grass-land" OR "grass* science") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p"))	534
13		Recreation science/ management	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("recreat* manage*" OR "recreat*" OR "recreat* science") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p"))	50
14		Water resources (for example, watershed management, water rights, hydrology)	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("water* manage*" OR "watershed*" OR "water resourc*" OR "hydrolog*" OR "water right") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	1362
15	**Land use/ land cover**	Land use/land cover	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("Land use*" OR "Land cover*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	3872
16	**Planning and development**	Assessments and taxation/ Engineering/construction /surveying	TITLE-ABS-KEY ("landsat") AND TITLE-ABS-KEY ("engineer*" OR "construct*" OR "survey*" OR "Assess*" OR "tax*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	4814

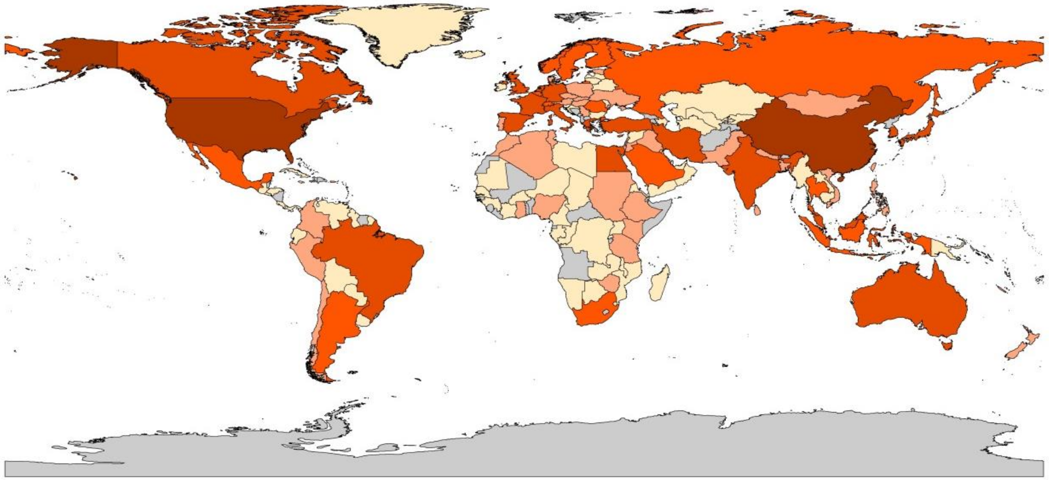
S N	Type in Landsat		Applied Search queries in this research	Number of Publication
17		Rural planning and development	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("rural plan*" OR "rural develop*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p"))	28
18		Urban planning and development/Urbanization	TITLE-ABS-KEY ("landsat") AND TITLE-ABS-KEY ("urban*" OR "urban* plan*" OR "urban* develop*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	1517
19	**Commercial**	Cultural resource management (for example, archaeology, anthropology) /Real estate/property management/Software development/Telecommunications/Transportation/Utilities	TITLE-ABS-KEY ("landsat") AND TITLE-ABS-KEY ("Cultur* resource* manage*" OR "archaelog*" OR "anthropolog*" OR "real estate" OR "property manage*" OR "software develop*" OR "telecommunicat*" OR "transportat*" OR "utiliti*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip")) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d")) AND (EXCLUDE (PUBYEAR , 2015))	225
20	**Human needs**	Emergency/disaster management/ Hazard insurance (for example, crop, flood, fire)	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("emergenc*" OR "disaster manage*" OR "hazard* insurance" OR "crop*" OR "flood*" OR "fire*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (PUBYEAR , 2015)) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d"))	2196
21		Humanitarian aid/ Public health	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("human* aid" OR "public health*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (SRCTYPE , "p")) AND (EXCLUDE (PUBYEAR , 2015))	24
22	**Oil/gas/ minerals**	Oil and gas/mineral exploration /extraction	TITLE-ABS-KEY (landsat) AND TITLE-ABS-KEY ("oil*" OR "gas*" OR "metal develop*" OR "mineral develop*" OR "mineral explor*" OR "mineral extract*") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk")) AND (EXCLUDE (SRCTYPE , "p") OR EXCLUDE (SRCTYPE , "d")) AND (EXCLUDE (PUBYEAR , 2015))	405
23	**Time series analysis and near-real time**	(TITLE-ABS-KEY (landsat)) AND (TITLE-ABS-KEY (mapping OR monitor* OR time-series* OR "time series analysis" OR time AND series AND analysis OR "change detection" OR ndvi OR "near real time" OR "near-real time" OR "real time")) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ip") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ch")) AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "p") OR LIMIT-TO (SRCTYPE , "b") OR LIMIT-TO (SRCTYPE , "k")) AND (EXCLUDE (PUBYEAR , 2015))		1036

C. Number of Landsat publication in world

a) Use of Landsat data before open data policy



b) Use of Landsat data after open data policy

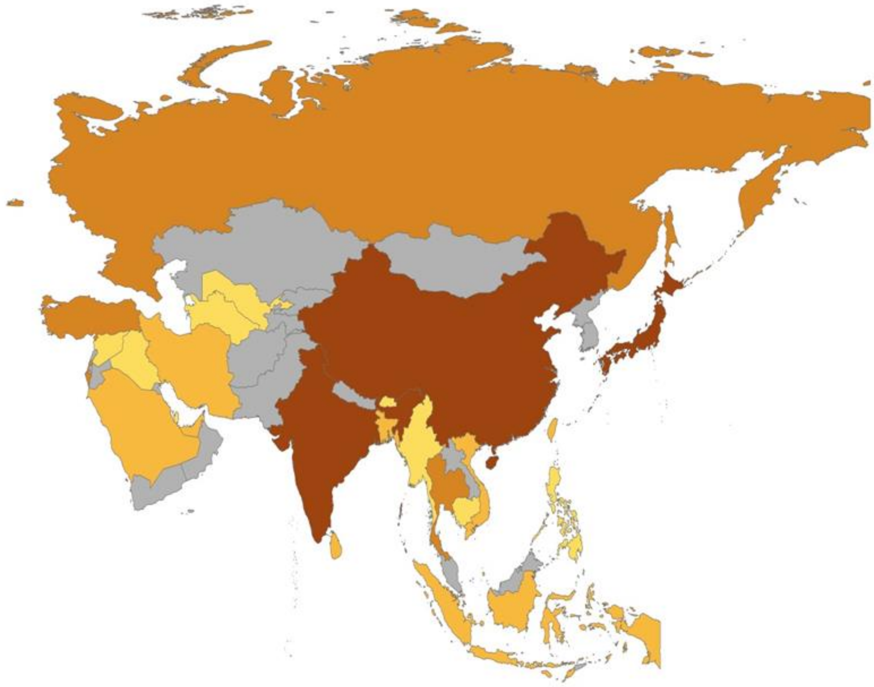


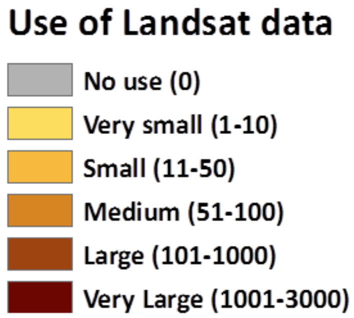
Use of Landsat data

- No use (0)
- Very small (1-10)
- Small (11 - 50)
- Medium (51 - 100)
- Large (101 - 1000)
- Very Large (1001- 3000)

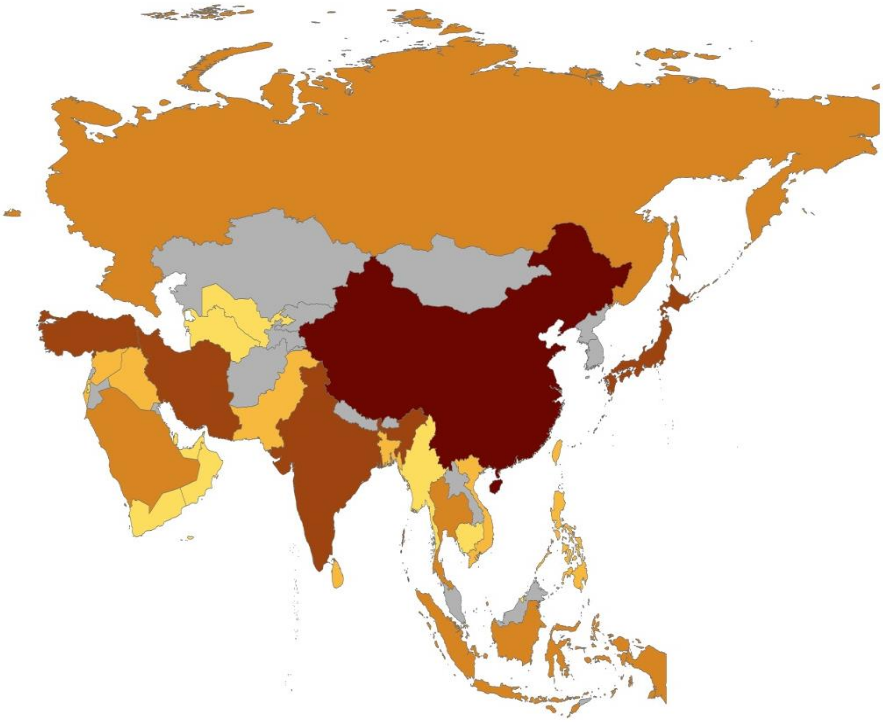
D. Landsat publication in Asia

a) Use of Landsat data before open data policy



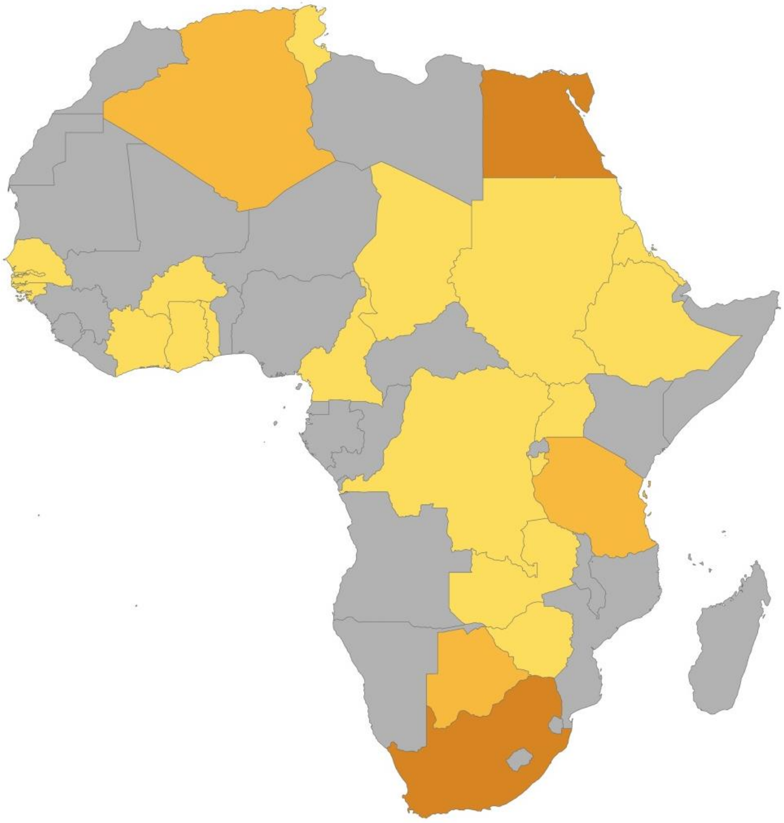


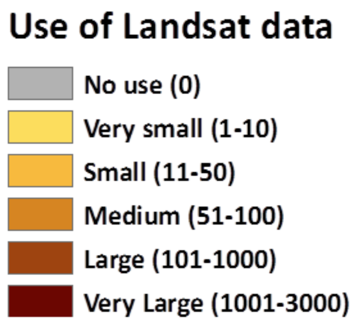
b) Use of Landsat data after open data policy



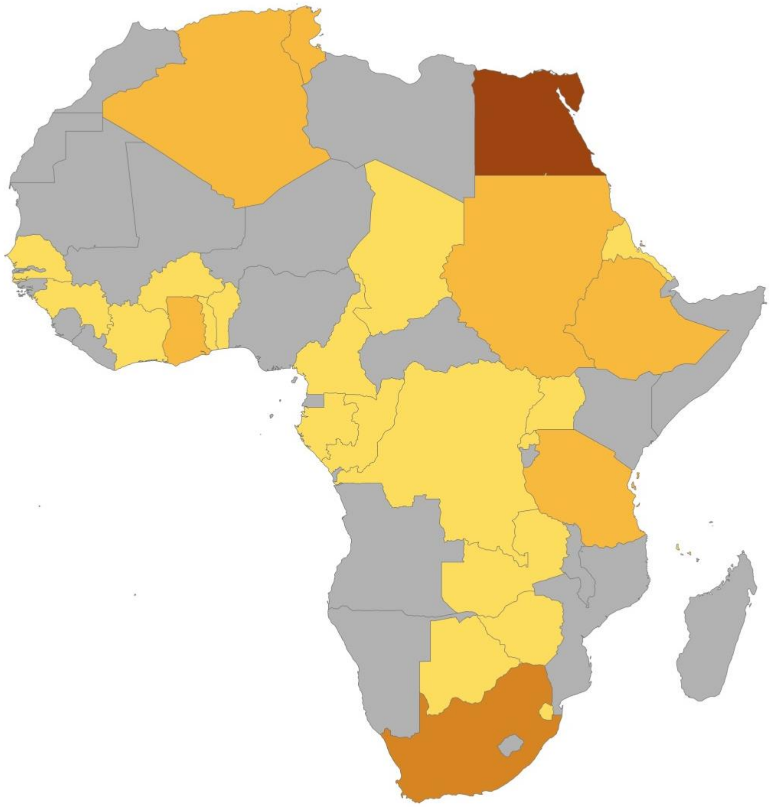
E. Landsat publication in Africa

a) Use of Landsat data before open data policy



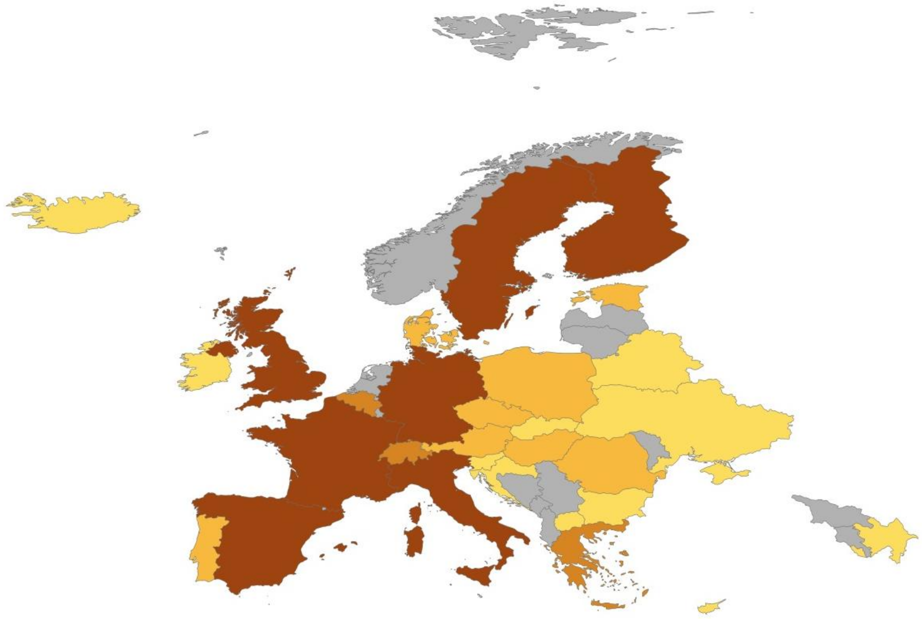


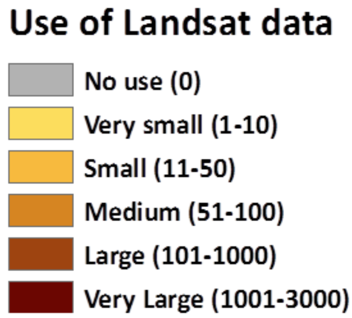
b) Use of Landsat data after open data policy



F. Landsat publication in Europe

a) Use of Landsat data before open data policy





b) Use of Landsat data after open data policy