Centre for Geo-Information

Thesis Report GIRS-2015-25

Improving Access to Reference Data for Global Land Cover Map Validation

Fahima Amin

August 2015

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Improving Access to Reference Data for Global Land Cover Map Validation

Fahima Amin
(810721014090)

Supervisors:

Dr. Ir. Ron van Lammeren
Ing. Willy ten Haaf

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

August 2015
Wageningen, the Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2015 -25
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Abstract

Global Land Cover (GLC) datasets and available reference datasets for its validation are an important issue now a days for worldwide research related with global changes in relation to agriculture, climate or urban planning. The GOFC-GOLD Land Cover Project Office took a good initiative to support those researches by providing standard quality GLC reference datasets with the detail of validation criteria for free. These reference datasets are reusable; multiple groups of researchers, government and nongovernmental officials use them for different purposes. This thesis presents an exploratory study to web GIS technology to improve the accessibility of reference datasets for GLC maps validation. A vast literature study has been performed to understand the differences between regular reference datasets and the reference datasets GOFC-GOLD offers. An online questionnaire survey has been done with the participation of 12 experts of different research backgrounds to evaluate the existing situation and to identify the lacking of the GOFC-GOLD website. A server side geoportal development has been completed to improve the accessibility of datasets validation with a qualitative approach. However, the evaluation survey of existing and new geo-portal focus on client-side oriented. Evaluation of the server side development has not done yet. The evaluation survey showed clearly that redevelopment of GOFC-GOLD geoportal is essential. Moreover, from the qualitative validation survey, I found, 75% of total respondents said that from the proposed geoportal it's easier to find the reference datasets whereas 67% of total respondents had difficulties to find the datasets from existing geoportal. The new interface is highly appreciated by the 91.66% respondents. Additionally, the overall user satisfaction has improved from 33% to 75%. The participants appreciate the new design. Despite several imperfections of the proposed geoportal, the results allow the conclusion that the methods used to develop a new geoportal, work well. Based on the study outcomes of the proposed server side interface further development of a prototype version is required to reach an implementation stage of the GOFC-GOLD geoportal.

Key Words: GOFC-GOLD, global land cover (GLC) datasets, geoportal, validation, GLC reference data, accessibility.

Acknowledgement:

Foremost, I would like to express my sincere gratitude from the deepest part of my heart to my supervisors Dr. Ron van Lammeren and Ing. Willy ten Haaf for their patience, continuous support and guidance with valuable tips and comments. It was great to have supervisors who gave me a lot of motivation and new ideas for further work. Their guidance helped me to finish my research and writing of the thesis.

I would also like to express gratitude to the government of the People's Republic of Bangladesh for giving permission me to continue the study and to the Netherlands Fellowship Program (NFP) to select me for the funding.

I am also using this opportunity to thank all my teachers and members of the laboratory of Geo-Information Science and Remote Sensing department. I enjoyed every single moment of my MSc at Wageningen University because most of the courses are inspiring, interesting, and student friendly teachers. I learned a lot from them.

Furthermore, I like to thank the participants in my survey, who have enthusiastically shared their precious time during the process of questionnaire survey. In addition, I appreciate that a Ph. D. candidate of Geo-Information Science, Arun Pratihast, helps me for programming. Special thanks to Dr. Brice Mora to provide all the GOFC-GOLD data and documents for my research experiment.

I express thanks to all my friends and fellow students for their encouragement to continue the research and for many great moments. A special thanks goes to Kristin Abraham for proofreading of my report. If you find any mistake in this report, most likely I added it after her revision.

Last but not the least; I would like to thank my parents and my sisters for supporting me spiritually throughout continuing this study.

Fahima Amin

August 2015

Contents

List of Figures

List of Tables

Glossary

AVHRR – Advanced Very High Resolution Radiometer
DBMS – Data Base Management System
EC JRC – European Commission Joint Research Centre
ESA – European Space Agency
ER - Entity Relation
FAO-FRA-RSS – Food And Agriculture Organization – Forest Resource Assessment – Remote Sensing Survey
GDAL - Geospatial Data Abstraction Library
GI – Geographic Information
GIS – Geographic Information System
GLC – Global Land Cover
GLCNMO – Global Land Cover by National Mapping Organizations
GML - Geographic Markup Language
GOFC-GOLD – Global Observation of Forest and Land Cover Dynamics
HTTP - Hypertext Transmission Protocol
IGBP – International Geosphere-Biosphere Programme
ISO - International Organization for Standardization
IT – Information Technology
LC CCI – Land Cover Climate Change Initiative
LCCS - Land Cover Classification System
LC PO – Land Cover Project Office
MERIS - MEdium Resolution Imaging Spectrometer
MODIS – Moderate Resolution Imaging Spectroradiometer
NELDA – Northern Eurasia Land Dynamics Analysis
OGC - Open Geospatial Consortium
OSS – Open Source Software
PPs – Program Points
RS – Remote Sensing
SDI – Spatial Data Infrastructure
SOA - Service Oriented Architecture
SPOT - Satellite for observation of Earth
UN - United Nations
UI – User Interface
UML - Unified Modeling Language
UN-FAO – United Nations Food and Agriculture Organization
VHSR – Very High Spatial Resolution
WMS – Web Map Service
WFS – Web feature Service
WCS – Web Coverage Service
WUR – Wageningen University and Research Centre
WWW – World Wide Web

1. Introduction

1.1 Context

Scientific communities that study forestry and agriculture land use (Fritz et al. 2012), ecology and geographical relations (Belward et al. 1999), climate change (Hese et al. 2005), urban planning, and environmental modelling (Myneni et al. 2002) use Global Land Cover (GLC) datasets for their research work in global, regional and national scale. They use these GLC datasets for natural resource assessments; parameterizing land process models for climate, hydrology, carbon cycle; public health; ecosystem assessment or agricultural activities (Zhao et al. 2014). These datasets are derived by different types of satellite-mounted sensors, which lead to differences in classification methods, and temporal and spatial resolutions (Wu et al. 2008). There is a significant amount of disagreement when these products are compared, because different users have different requirement about the resolution and accuracy level of GLC datasets (Tsendbazar et al. 2014).

GLC datasets have been constructed from the data of Advanced Very High Resolution Radiometer (AVHRR (Loveland et al. 1999)), Satellite for observation of Earth (SPOT-4 VEGETATION (Durpaire et al. 1995)), Moderate-Resolution Imaging Spectroradiometer (MODIS (Friedl et al. 2002)), MEdium Resolution Imaging Spectrometer **(**MERIS (GlobCover 2000-2011)), and Land Cover Climate Change Initiative (ESA 2011). GLC datasets are evolved for information about land cover and land cover dynamics (Herold et al. 2006). These land cover datasets at global scale are the base line datasets which are important assets for the scientists who do their research on global change on different perspective such as climate change, land cover type, urban planning for sustainable development, natural resource management, environmental studies and related fields (Foley et al. 2005, Zell et al. 2012, Sterling et al. 2013).

Most of the GLC datasets are developed with 250m to 1km resolution (Alan H. Strahler et al. 2006). These GLC datasets are increasing with the availability of remotely sensed datasets at global scale. 1 km GLC dataset named GLC2000 was developed from 1km AVHRR (Hansen et al. 2000, Loveland et al. 2000) and 1km SPOT-4 VEGETATION satellite data (Bartholomé and Belward 2005). 500m GLC maps (Friedl et al. 2010) and 250m GLC maps (Zhan et al. 2000) produced from MODIS. 300m GLC maps named GlobeCover developed from MERIS (Arino et al. 2008).

Depending on the input quality, spatial resolution, legend and classification algorithm every GLC data has different limitations such as the resolution of AVHRR is 1km; GlobCover

datasets have been produced using unsupervised classification methods. Noteworthy amount of disagreement arises when these datasets are compared. One of the reasons of this disagreement is the lacking of sufficient real time reference datasets for validation and calibration of GLC datasets (Fritz et al. 2012). Another important reason of disagreement is the lack of interoperability and inter-comparability between the datasets (Herold et al. 2008a). Data users and data producers are frustrated because all GLC datasets are designed as independent dataset and lacking of sufficient information on accuracy of those datasets. Producers need to understand and identify the lacking, which make this disagreement for further development. Furthermore they need to acknowledge that users are confused which dataset is more appropriate for their use purpose (Herold et al. 2008a).

Before using the GLC datasets, a validation is necessary, for studies at regional to global scales. It will help to reduce user's confusion and will help developers to understand the weakness of the datasets for further development. Validation exercises can provide a quantitative accuracy assessment of satellite-derived global land cover datasets (Wu et al. 2008). Validation is the suite of techniques to assess the accuracy of given datasets based on overall accuracy, errors of omission and commission by land cover class, errors analysed by region, and fuzzy accuracy (Alan H. Strahler et al. 2006). Different scientific communities have different requirements for accuracy assessment of GLC datasets (Bontemps et al. 2011, Herold et al. 2011). For the accuracy assessment of GLC datasets high quality reference datasets are needed which will be independent and also their accuracy level will be higher (Tsendbazar et al. 2014). Reference datasets define some standard values to be used by other data sets. These datasets are used only to categorize data or relate data in a database to information beyond the boundaries of the initiative (Inc 2003). "The position or class label of the accuracy assessment site, which is derived from data collected that are assumed to be corrected is known as reference data" (Congalton and Green 2009). These reference datasets are known as independent validation datasets.

Accuracy assessment of GLC datasets should be interpreted for different user needs and use suitable independently validated reference datasets in the assessments. For the efficient use of reference datasets for GLC datasets validation a set of new reference datasets is proposed jointly by Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD) and Boston University (Olofsson et al. 2012). Currently there is no assessment providing information on the use of these datasets beyond their original scope and no implications for specific user applications having different requirements on GLC datasets and their validations (Tsendbazar et al. 2014). Users are demanded frequent and easy accessibility and reusability of GLC reference datasets. Recently GOFC-GOLD has proposed a framework to make a guideline for the user based on their needs, establish the new

database and make it available online to the GLC mapping community and engage partners in the effort of producing reference data for global scale land cover mapping activity (Mora et al. 2014b).

Accessibility of the reference data portal depends on reliability (the server's ability to fulfil the user request), availability (fraction of time a service is available) and performance (how fast the geoserver can fulfil client request) (Boundless 2015). Therefore, accessibility techniques make the reference data portal accessible to all users.

Nowadays Geoportals are used for earth and planetary science, social science, medical science, environmental science and so forth. Web GIS (Geographic Information Science) is the map-centric management system which supports, organize, protect and facilitate access to geographic information products (Esri 2013). The map, product from Web GIS, becomes dynamic, interactive and accessible to a wide selection of users as a visual communication tool in the environment of Web-based GIS functionalities. The highly functional Web-based architecture provides easy access to spatial raster data and facilitates image analysis and its spatial decision support system framework uses for exploratory visual analysis and mapping of the obtained results (Dragićević 2004).

A geoportal is the most open and easy access of the geo-information and remote sensing data for the potential end users which includes World Wide Web (WWW) data, metadata and geo-processing capabilities. Furthermore, it acts as an information providing single access point for discovery, visualization, data retrieval and geo spatial web services (De Longueville 2010). GOFC-GOLD Land Cover Project Office (LC PO) is contributing to provide better information, methods and tools toward the global land cover (GLC) reference datasets and for land cover map accuracy assessment (LCPO 2012-2013). Therefore, there is a need of a Web based Geographic Information System (GIS) solution for these spatial datasets like storage, retrieval and visualization. Figure1 describes the core components of SDI (Spatial Data Infrastructure) which relates to this study and tries to provide a concept about a new reference data portal with an organized data base management system which will facilitates to collect, archive, share, analyse, visualize and simulate GIS and remote sensing data, information and knowledge over the web (Shen et al. 2008).

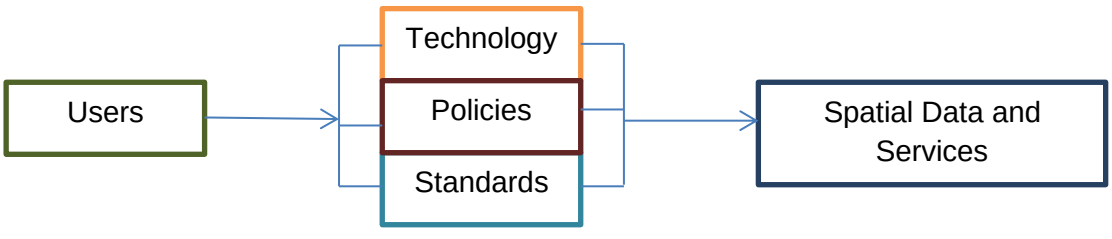


Figure 1: Relation between SDI components (Rajabifard et al. 2002).

A number of GLC datasets have been produced and more are coming in future. In addition, these datasets are very useful for different research field, and can also use for government work purpose or for some social work. To validate these GLC datasets scientific communities need reference datasets. Development of a new set of reference data is expensive. Some reference datasets can be found from different sources that are not authentic enough to use and some of them can be used only once. Reference maps from GOFC-GOLD LC PO are reusable and augmented for different GLC map validation. Furthermore, this study is based on information and datasets of 2014.

1.2 Problem Definition:

There are different geoportals or web based applications to support validation of GLC reference datasets like Geo-Wiki (Fritz et al. 2012) or View-IT (Clark and Aide 2011). They developed web tools to interpret crowd sourced reference datasets, which have limited flexibility like users are allowed to visually estimate the land cover percentage (Clark and Aide 2011, Fritz et al. 2012). On the other hand GOFC-GOLD is providing independent GLC reference datasets with an extensive suit of characteristics (like sample design, response design and level of classifications) for free of cost and easily downloadable (Olofsson et al. 2012, Mora et al. 2014b). The intention of GOFC-GOLD is to provide independent reference datasets. However, currently the GOFC-GOLD portal is not well known to the user community and not well organized. Especially the structural design of this portal is not good enough structured to attract the users. Moreover, I talked with the project personnel about the storage capacity and retrieval system, which seems to cause problems. They do not have any storage facility for the reference datasets nor is their retrieval system user friendly.

The re-usability of these GLC reference maps would be greatly enhanced by making them publicly available in an expert framework that supports users 1) to choose the most suitable reference map based on their needs, and 2) to use the map appropriately for a robust statistically accuracy assessment. An information system for this GOFC-GOLD project requires the storage, retrieval and visualization of the sample plots facilities. This system will allow tracking the progress made on the interpretation of these sites and all reference datasets will be available to the users in a spatial database form from the GOFC-GOLD Project Office data portal.

It is good news for the scientists who use GLC maps for their research work, that reference datasets to validate GLC datasets are available for free in GOFC-GOLD website. The website of GOFC-GOLD PO hosts the first version of a reference data portal to inform on

and support validation of global land cover datasets (Mora et al. 2014b). After visiting the GOFC-GOLD portal, I found that the current version of their reference data portal has been set up with basic features enabling the access to the home page. The home page leads to other pages, which consist of lots of document files and downloadable datasets in zip, pdf or text format. However, these links should be presented in a more prominent way, so that the user can easily spot them.

GOFC-GOLD website is a hosting data portal and does not own the datasets (Townshend et al. 2006). They are contributing to provide better information, methods and tools toward the global land cover (GLC) reference datasets and for land cover dataset accuracy assessment. But the problem is, they didn't explain clearly any information about assessment or consequences of those GLC reference datasets and it limits the value of these datasets to the users requirements in their website (Tsendbazar et al. 2014). Therefore, clear and well-organized metadata of these GLC datasets is the strong requirement of this portal .

Four consolidated reference datasetss (GLC 2000, GlobCover 2005, STEP and VIIRS) are publicly available on the current version of GOFC-GOLD data portal with incomplete metadata information, others are available upon request and access to the others should be based on agreements (Mora et al. 2014b). The problems users are currently facing with this portal are that very few references maps can be found on the GOFC-GOLD website, a lack of user friendly guidance in the portal, the lack of tools for the users and finally there is no option for a user survey to get information about current user needs (Tsendbazar et al. 2014). GOFC-GOLD LC PO did not do any user survey or other experiment to understand the lacking of their data portal.

Along this, the main aim of this research is to determine the present situation and lacking of the GOFC-GOLD geoportal to improve the accessibility of the reference datasets and near future geoportal components according to the technical requirements and user needs as a case study. The outcomes of this study can be the basis of a new geoportal establishment.

1.3 Objectives and Research Questions:

The overall aim of this research is to "improve access of reference datasets for global land cover map validation". To fulfil this objective the following research questions are proposed:

RQ1: What are reference data and how do they validate the global land cover (GLC) datasets?

RQ2: Does the current (2014 version) geoportal of GOFC-GOLD facilitate the access of GLC reference datasets?

RQ3: How to improve the accessibility of reference datasets for validating GLC datasets?

RQ4: Does the geoportal (upgrade) design of GOFC-GOLD geoportal indeed increase accessibility of GLC reference datasets?

1.4 Outline of the Thesis:

Chapter 2 contains a literature review, related to this research, of the validation studies of GLC datasets, improve accessibility of reference data portals, user interface of that reference data portals and research strategies which are important to understand before the research experiment

Chapter 3 describes the methods and techniques that were used to evaluate the current GOFC-GOLD reference data portal, to develop a prototype GOFC-GOLD data portal and to validate this new prototype portal.

Chapter 4 presents the results of the research experiment and provide answers to all research questions.

The last chapter (5) presents the conclusions, discussion and recommendations for future work.

2. Review

2.1 Global Land Cover (GLC) Map:

Global mapping of land cover exists since the last two decades and serves to characterize baseline terrestrial information, which is one of the most important variables for a wide range of societal benefits (GEOSS 2005, Fritz et al. 2011). Nowadays, a number of global land cover datasets are developed to explore the information about land cover and land cover dynamics (Alan H. Strahler et al. 2006, Herold et al. 2008a). As discussed in chapter 1, GLC datasets are independent datasets because they are the product of different national and international initiatives and their standards varies because of using purpose (Herold et al. 2008a, Bontemps et al. 2012). The users are confused about the accuracy and metadata information of different GLC datasets when they want to use these datasets for their research work or other development work. So, both, data producers and users are stressed about different types of land cover datasets, difference in thematic legend and lack of accuracy information (Herold et al. 2008a). To reduce this confusion a well-designed geoportal of reference datasets may have important contribution.

2.2 Validation Strategies of GLC Maps:

“Validation is the cognitive process of establishing a valid proof or the act of validating, finding or testing the truth of something” (Dictionary 2007). A validation process is an ongoing process, which provides a high degree of assurance on the quality of the processed data by establishing scientific evidence.

To use GLC datasets for any research or practical experiences we need to assess the accuracy of that datasets. Accuracy assessment is expensive which force the International Geosphere-Biosphere Programme (IGBP) to develop integrated global validation schemes for multiple GLC datasets (Olofsson et al. 2012). There are several steps of validation processes in GIS (Geo Information Science) and RS (Remote Sensing) research such as i) sampling design, ii) reference data collection, iii) extracting the interpretation results from the reference data, iv) reference data and interpreted data comparison by using appropriate statistical technique, and finally analysing the result (causes and distribution of errors) from different interpretation technique (Chuvieco and Huete 2009).

Quality assessment of maps is known as validation which include assessing the overall accuracy, errors of omission and commission and fuzzy accuracy (Alan H. Strahler et al. 2006, Herold et al. 2009). For the determination of the quality of a map generated from remotely sensed data, accuracy assessment is prerequisite. Positional and thematic accuracy are two types of map accuracy assessment. Positional accuracy uses reference data to estimate the accuracy of the location of map feature (Congalton and Green 2009). Townshend et al. (2006) note that it's sensible for GOFC-GOLD to adopt the MODIS Land Team proposed scheme in Table 1, because accuracy assessment of global land cover datasets is expensive and time consuming. Furthermore, international cooperation helps to get expertise for data collection and the international standard.

Stage 1 Validation: Map accuracy has been estimated using a small number of independent measurements obtained from selected locations and time periods. Validation assessed locally under a limited range of geographic conditions for a limited period of time.
Stage 2 Validation: Map accuracy has been assessed over a widely distributed set of locations and time periods. Validation assessed over a significant range of geographic conditions and for multiple time periods and seasons.
Stage 3 Validation: Map accuracy has been assessed and the uncertainties in the map well established via independent measurements in a systematic and statistically robust way representing global conditions. Validation assessed over the full range of global conditions for all time periods.

Table 1: Levels of Validation adopted from Townshend et al. (2006)

To perform an accuracy assessment suitable reference datasets are needed, which are in good quality. For the scarcity of validation data, reference datasets should be re-useable. An investigation on the efficient use of all available validated and calibrated datasets for GLC datasets is a very new user requirement (Tsendbazar et al. 2014). GOFC-GOLD and Boston University together are generating reference datasets for the validation of GLC dataset, which is possible to use for different and multiple GLC datasets validation (Olofsson et al. 2012, Stehman et al. 2012).

2.3 Improving Accessibility by Geoportal Design:

A geoportal (i.e. combined product of GIS and internet technologies) is the gateway to access geographic data and information available with directories, search tools and support resources (Maguire and Longley 2005). A geoportal is like a 'one-stop-shop' where users can get access of multiple datasets. It provides online access, visualization and analysis of the

GLC datasets (Han et al. 2015). Fritz et al. (2012) developed a crowdsourcing tool named Geo-Wiki by using web map service (WMS) to increase the availability of GLC training data for validation. The design of a geoportal for spatial datasets has many advantages like easy accessibility, data transparency, independent platform, better visualisation and also cost effectiveness (Kulkarni et al. 2014). Geoportals and SDI encourage and assist people to use Geo-Information (GI) concepts, databases, models and techniques in their work. It also connects GI data provider and user via the medium of internet. Geoportal design depends on the purpose of the use and product of that portal. From my observation on existing GOFC-GOLD portal has been set up with basic features enabling the visualization of metadata information and download of the datasets in a zip or text format, which does not look like user friendly. It is complicated and time consuming for the users to find the reference datasets in that website. Geoportal design should be a user centric web application with applying geo-visualization methods(Figure: 2). However there are always client and server side design involved in a geoportal development.

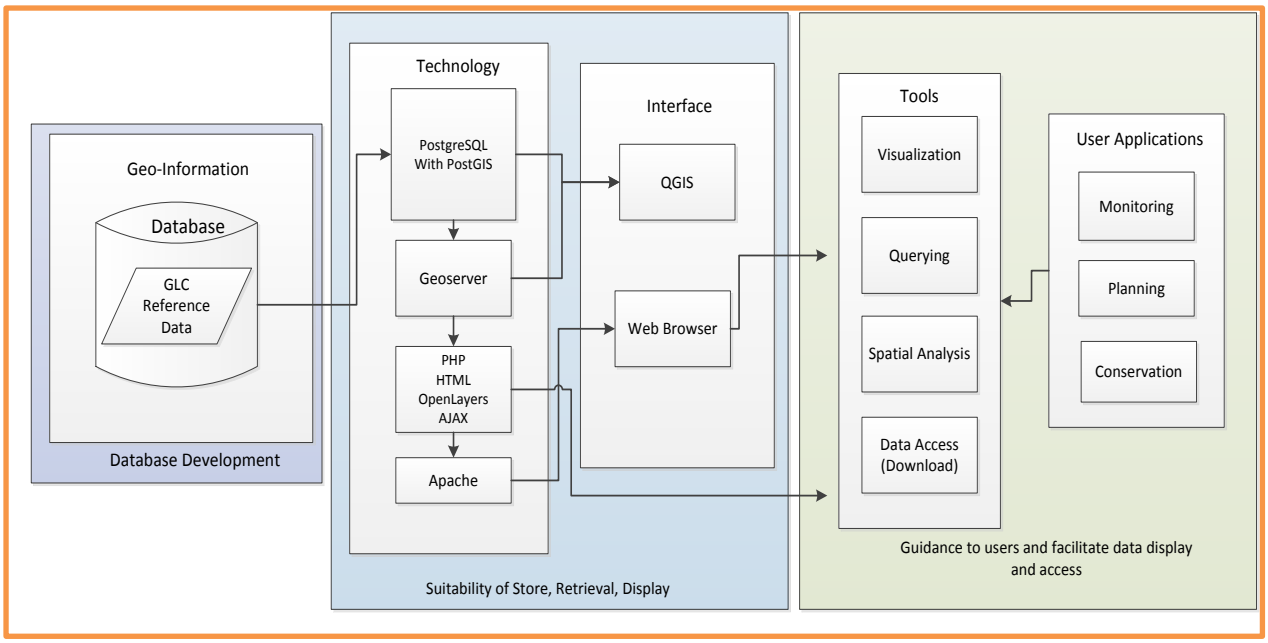


Figure 2: Geoportal design comprised of GIS database, technology generally used tools and user applications. Adopted from Souza et al. (2009)

In the following described concept (from Figure 2) of geoportal can help to improve accessibility of GLC reference maps.

2.3.1 Database Management System

The purpose of a geoportal and SDI is to save time, effort, money and unnecessary duplication in standardization of required reference datasets (Groot 1997). Metadata records are extremely important for SDI to maintain information about the content, source and origin of the data (Crompvoets 2006). SDI aims to create an environment to access and retrieve spatial data sets in an ample, reliable and protected manner. It can also create a setting where participants, users, and spatial data producer can collaborate, and utilize information and communication in a cost-effective way (Manisa and Nkwae 2007). A spatial Database Management System (Spatial DBMS) such as Oracle, PostGIS, SpatialLite and MySQL Spatial extension can store the reference datasets.

2.3.2 Server:

Geo server and map server are free servers to visualize maps and images stored in a database on the web. These servers can be used as a development environment for constructing spatial applications on the internet (Souza et al. 2009). The server component expresses user rights to access various functionalities of the portal.

2.3.3 Interface:

Geoportals are built using underlying World Wide Web (WWW) infrastructure technology and network communication between clients and web servers uses Hypertext Transmission Protocol (HTTP). Thus, a geo-portal is the entrance to collection of information resources which include data sets, facilities, news, tools, tutorials and organized collection of links of other sites (Maguire and Longley 2005).

2.3.4 Tools:

For visualization, to query and to conduct spatial analysis user-friendly tools are essential. Menu bar, different kind of fields, markers, pop-ups, query buttons, layer selections or hyperlinks are the tools users can use to fulfil their requirement. A library is also a searching tool. GDAL (Geospatial Data Abstraction Library) is an open source raster utilities set in a Python wrapper (Zimmerman 2014) which is known as translator library for raster and vector geospatial data formats. As a library it presents a single raster abstract data model and vector data model to the calling application for all supported formats (Doxygen 2005). GDAL will design the raster library which will be easy accessible for the user. There are also some other essential tools for maps such as pan, zoom or get attribute information on the feature.

2.3.5 User Applications:

An application server is the crucial part to visualize the geoweb portal and to prove that the web portal is serving the needs of the user. GeoServer and GeoWebCache are the two most popular open source application servers. The Geographic Markup Language (GML) is a framework to encode features with specific support for geographic information in accordance with the conceptual framework specified in the ISO (International Organization for Standardization) 19100 series of International Standers. GML is designed for the WWW and Web Services. WWW creates possibilities and increased requests for on-demand access to all kinds of information including geographic information (Inc. 2007).

2.3.6 Example of online platform to improve accessibility of GLC reference datasets:

As mentioned before, two good examples for geo-platforms are Geo-Wiki and VIEW-IT. Fritz et al. (2012) developed Geo-Wiki, an online platform, for improving accessibility of GLC datasets for training, calibration and validation to get more accurate land cover information. Geo-Wiki has four classes of services of portal, portrayal, data and catalogue that fulfils the OGC (Open Geospatial Consortium) requirements. Han et al. (2015) designed and developed a web-based system for 30-meter resolution of GLC datasets using a 3-tier SDI (Spatial Data Infrastructure) system model. This web-based system provides online access, visualization of GLC datasets as map, ancillary data; integrate heterogeneous data and services required for GLC data production.

Clark and Aide (2011) introduce Virtual Interpretation of the Earth Web-Interface Tool (VIEW-IT). VIEW-IT is the google Earth plug-in with a web-based application, which has an in-built user interface, basic interpretation criteria, server side storage and automated error checking.

All the platforms discussed above use open source software to develop their geoportal for the GLC datasets.

2.4 Usability Study of Geoportal

Usability study is one of the most important parts of evaluating an information system. This is a system to support people perform their task competently and successfully. Usability depends on several elements like learnability, memorability, effectiveness, efficiency and satisfaction (Henry 2011) and is a very important aspect for map based geoportal

applications. Usability is concern for language, layout and graphics, information architecture, user interface and navigation (Bai et al. 2008). A geoportal has a complex involvement with user interaction like textual search, data retrieval or map interface and geo-search. According to ISO 9241-11 (Al-Kilidar et al. 2005) and ISO 25010 (Bevan 2009) the quality of use of the system including tool, user, tasks and environment is the standard of the usability.

A geoportal is the access point of shared public Geographic Information (GI). Geoportal requires different design than other type of web applications (Peterson 2001, Wachowicz et al. 2008). And it should have at least four basic functions including searching, mapping, publishing and administration of the infrastructure (Tait 2005) in its design list. The simplest use case of a geoportal is to discover the most relevant datasets and a search function is the best option for that. GIS data visualization is also important function because users want to examine the data content before using it (He et al. 2012).

2.5 Conclusion

Nowadays different GLC datasets are available and more are coming in different resolution. Different researchers have different requirement to validate those GLC datasets. Reference datasets are necessary for validation of a datasets. The development of reference datasets is expensive and time consuming. Some international standard GLC reference datasets are freely available on the internet. However, there is a problem of lacking of the easy accessibility of those reference datasets. A clear understanding of GLC datasets and its validation criteria is the prerequisite to do a research on how to improve the accessibility of reference datasets to validate these GLC datasets. To improve the accessibility of GLC reference datasets geoportal development is a solution because a geoportal is the gateway to access all kind of GIS and remote sensing datasets. One hypothesis constructed in this research to solve this problem is to develop/reorganize the portal intotal, which may help to improve the accessibility.

3. Methodology

3.1 Overall Methodology

From the literature study in second chapter, I have found a first impression of accessibility and usability limitations of GLC reference datasets from GOFC-GOLD geoportal. To improve this situation an overall methodology to know more about the existing GOFC-GOLD geoportal is presented along with a questionnaire survey to evaluate the portal. Based on all the collected information, the problems and potentials related to the use of existing GOFC-GOLD portal and accessibility to the reference datasets, a new prototype geoportal will be developed and validated by a small and limited questionnaire survey.

An overall methodology is presented in the following chart-

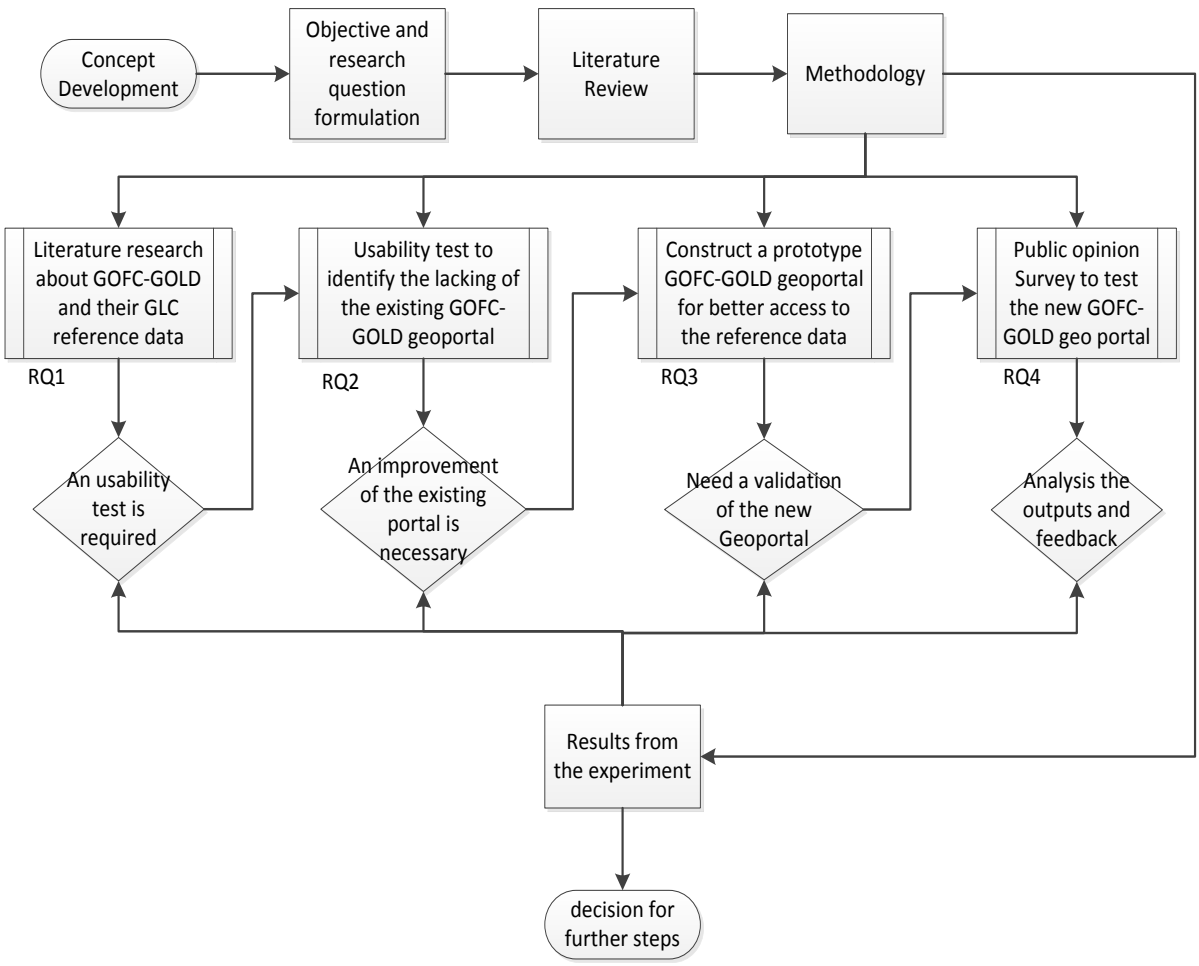


Figure 3: Flow chart of Methodology

RQ1: What are reference data and how do they validate the global land cover (GLC) datasets?

A scientific literature study through internet and university library has been made by using the scientific papers, e-books and websites to examine the reference data and their validation process on global land cover dataset, the users of this data, about GOFC-GOLD and its support to the scientific world by providing GLC reference datasets.

RQ2: Does the current (2014 version) geoportal of GOFC-GOLD facilitate the access of GLC reference datasets?

A usability survey of existing GOFC-GOLD portal has been done to understand the lacking of the portal and identified the weaknesses of the portal needs to be developed.

RQ3: How to improve the accessibility of reference datasets for validating GLC datasets?

A conceptual geoportal with high storage capacity and nice design principle has been made to test if it increases accessibility of the GLC reference datasets. To develop the conceptual portal the following criteria have been considered

- Database management system should have high storage facility, free GIS data loader, free desktop viewers and editor, web-mapping toolkits and free of cost data download facilities.
- Server with known and easy programing language, different output format, easy to connect virtual world, open standard interoperability, open source server.
- Availability of tools to download, scale, pop up information and others.

RQ4: Does the geoportal (upgrade) design of GOFC-GOLD geoportal indeed increase accessibility of GLC reference datasets?

A questionnaire survey on the functionality performance of the new conceptual portal, especially the user interface, has been done to evaluate if it increases the accessibility of GLC reference datasets.

3.2 Case Study GOFC-GOLD

GOFC-GOLD is a coordinated international effort to better understand global changes for forest and land cover observations (LCPO 2012-2013, Mora and Herold 2013). The main goal of GOFC-GOLD is to provide a forum for international information exchange, reflection, data coordination and a framework for establishing the necessary long-term monitoring system of forest, biodiversity and land cover dynamics (GOFC-GOLD and Team 2014). The GOFC-GOLD LCPO is funded by the European Space Agency (ESA) and is located currently on the campus of Wageningen University, to support international assessments of land cover management and environmental treaties and to contribute to natural resources management (Herold et al. 2008b). GOFC-GOLD LCPO provides reference datasets to inform on and support validation of global land cover datasets. Along with understanding the lacking of the GOFC-GOLD portal I would like to explore more about reference datasets, the spatiality and exception of reference datasets GOFC-GOLD offered than others.

3.2.1 Reference Data and validation process of GLC datasets

The reference data acquisition involves collecting measurements and observations about the remotely sensed objects, areas or phenomena and are used to calibrate a sensor or verify information extracted from remote sensing data (Lillesand et al. 2014). Reference data always use for accuracy assessment. Reference data have permissible values, which can be used to validate other data field. This validation is done by comparing reference and interpreted data. Appropriate statistical techniques helps this comparison (Chuvieco and Huete 2009).

The very common procedure to collect the reference data is the principle of statistical sampling design that estimates the possible number of samples, sampling technique, and duration of study. Time critical (for rapid change of ground with time) and time stable (observed materials do not change noticeable with time) are two different ways to measure reference data. There are different methods to collect the reference data like ground base measurement, laboratory spectroscopy, field measurement and spectral reflectance measurement (Lillesand et al. 2014).

GOFC-GOLD LC PO applied UN (United Nations) Land Cover Classification System (LCCS) and IGBP classification schemes (Loveland et al. 2000) to characterize and validate GLC datasets. Legend, sampling design and response design are the baseline validation criteria of GOFC-GOLD (Olofsson et al. 2012). Selecting a large global baseline sample is the first initiative of the global validation database. The sample size will be increased by several iterations and must include a protocol to achieve acceptable precision for any accuracy estimation (Stehman et al. 2012).

For the scarcity of validated data, reference datasets should be re-useable. An investigation on the efficient use of all available validated and calibrated datasets for GLC dataset is a very new user requirement (Tsendbazar et al. 2014). The global validation datasets provide reference data for descriptive and relative accuracy analysis (Olofsson et al. 2012).

GOFC-GOLD adopted area units for reference datasets which are not dependent on a pixel or a block of pixel (Tsendbazar et al. 2014). They can be easily applied for the validation of different datasets because GOFC-GOLD reference datasets are not map dependent. This is one of the most important stratification, which makes this datasets an exception, and which is independent of any GLC dataset and applicable to many datasets. Olofsson et al. (2012) and Stehman et al. (2012) fixed several standards to select the sample blocks for sampling design such as

(1) it satisfies meaning of a probability sampling design to select the baseline sample, augmented sample and combined samples from both of the samples;

(2) it provides adequate sample sizes for rare land-cover classes which is achieved by stratification to assess multiple datasets (these datasets differ from each other in legend and support);

(3) it allows flexibility to change sample size (either 5km x 5km or 1km x 1km) in response to random funding or revised accuracy assessment objectives;

(4) it focuses sample sites in the areas most difficult for land-cover mapping because of two important characteristics one is consistency of the accuracy estimation protocol (sample size increase over time) and the other one is minimize the complexities (related with sample data weight).

Stehman et al. (2012) describe elaborately of all these standards of stratified sampling design for global land cover validation to estimate accuracy by region and land cover classes. Following characteristics (in Table 2) are considered for generating and assessing reference datasets, which are accessible from GOFC-GOLD portal.

1. Legend	➢ Classification Scheme ➢ Number of class ➢ Classifier information provided
2. Sample Design (cost and statistical precision)	➢ Sample unit type and size ➢ Sample size ➢ Sample selection scheme ➢ Sample stratification ➢ Inclusion probability ➢ Minimum Mapping Unit
3. Response Design (reflection on the agreement between map and reference classification)	➢ Source of information ➢ Temporal coverage ➢ Location accuracy ➢ Labelling procedure ➢ Sample verification ➢ Confidence in interpretation ➢ Majority classes and their fraction
4. Current use	➢ Intended application ➢ Other applications ➢ Applied pre-processing ➢ Derived accuracy estimates

Table 2: Characteristics considered generating and assessing reference maps. Based on the literature of Olofsson et al. (2012) and Stehman et al. (2012)

In GOFC-GOLD project, stratified sampling based datasets doesn't use stratifications that are targeted to particular land cover maps (Tsendbazar et al. 2014). In sampling design, stratification (strata represent land-cover classes) incorporates accuracy assessment because it specifies the allocated sample size of stratum (Olofsson et al. 2012). From GlobeCover and LC-CCI (Land Cover Climate Change Initiative) GLC user requirements surveys accuracy, stability, spatial resolution thematic content are most stressed user requirements (Bontemps et al. 2012).

Name	Sampling design	Sample size	Sample unit, size	Source reference data	Legend
GLC 2000	2 stage stratified Cluster sampling	1265 253 PSU 5SSU in each PSU	3 by 3 pixels	Landsat 2000, aerial photographs, thematic maps, NDVI profile	LCCS 22 class
GlobCover 2005	Stratified random sampling	4258 3167 certain	5 by 5 pixel	SPOT VGT-NDVI profile, Google Earth	LCCS 22 class
STEP	Stratified	1780		Landsat, Google Earth	IGBP 17 classes + other classes LCCS in the future
VIIRS	Stratified random sampling	500	5 by 5 km blocks	VHSR (<2-m)	IGBP legend, and LCCS in the future

Table 3: Available Reference datasets of GOFC-GOLD adopted from (Mora et al. 2014c)

Table 3 describes the list of reference datasets with the specifications GOFC-GOLD portal provides for the GLC datasets users.

An accurate global land cover dataset is very important for different kind of scientific or research work such as forest resource assessment (Zhu and Waller 2003, GOFC-GOLD 2011, Townshend et al. 2012), climate change and modelling (Hese et al. 2005, Hibbard et al. 2010, Verburg et al. 2011, Bontemps et al. 2012), global agricultural (Ramankutty et al. 2008) and land use monitoring (Foley et al. 2005). Based on Tsendbazar et al. (2014) figure 4 shows the suitability of different GLC reference datasets for different user/research purposes and Table 4 shows the specification of the datasets for different users.

Product name	**Resolution**	**Concern Model**	**Reference**
Global Land Cover product	300 m to 1 km	UN-FAO	(Loveland et al. 2000, Bartholomé and Belward 2005, Defourny et al. 2009, Friedl et al. 2010, Bontemps et al. 2011)
Global Land Cover product	.25 to 2.5 degrees	Regional and global climate model	(Bontemps et al. 2012)
Global Forest Cover Loss map	500 m	Global forest change analysis	(Hansen et al. 2003, Hansen et al. 2010)

Table 4: specification of different GLC reference datasets based on different user requirement.

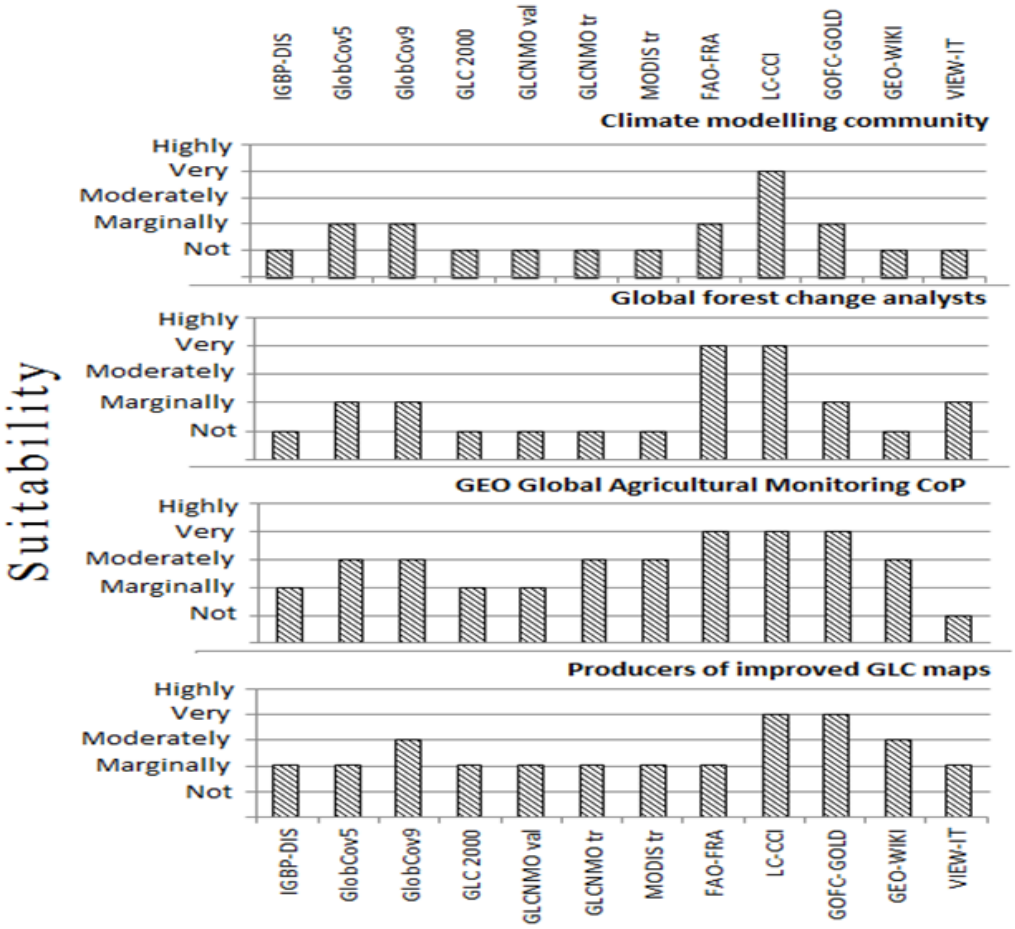


Figure 4: Suitability of status of GOFC-GOLD GLC reference datasets for different users (Tsendbazar et al. 2014)

An example case to understand the use of GLC reference datasets:

Deforestation effects climate by increasing humidity of tropical areas. Satellite imagery from GLC datasets with a statistical sampling strategy were used to calculate the change caused by deforestation (Achard et al. 2002). GLC reference datasets are used for the global forest change assessment and the area estimation (GOFC-GOLD 2011, Olofsson et al. 2013) with the method of sample selection schemes (GOFC-GOLD 2011). However, there are differences between the classes of different GLC datasets. GLC2000 contains tree classes like “Burnt Tree cover”, “Regularly Flooded Tree Cover” and GLOBECOVER use “Open Needle leaved Deciduous or Evergreen Forest” whereas in MODIS tree class is open tree vegetation and covers classes which cannot be unambiguously assigned to either of the two (Pflugmacher et al. 2011). Therefore, reference data with statistical rigorousness, stability in multi-date records, a suitable sample selection scheme and forest definition, and spatial detail suitable for change detection are required for global forest change analysis (Tsendbazar et al. 2014).

GLC datasets are commonly validated by using higher quality independent reference datasets. Before using reference, datasets users have to know the thematic and geometric precision and accuracy of the reference datasets as well as the thematic class distribution and the necessary (Meta) information like time labels and classification procedure. From Olofsson et al. (2012) and Stehman et al. (2012) it's clear that GOFC-GOLD provided reference datasets are independent and also provide all those information.

3.3 Analysis of the accessibility of GOFC-GOLD portal to reference data for GLC validation:

3.3.1 Accessibility:

The very common strategy to give access to the datasets of a geoportal is to offer users three questions: "Where", to specify the latitude and longitude or name of the place; "What", to search term related to attribute; and "When", to limit the search to a specific range (Aditya and Kraak 2009). That means user interface should be operational, well organized and satisfying. Improvement of style sheet and script are the technical parts to increase the accessibility of a geoportal.

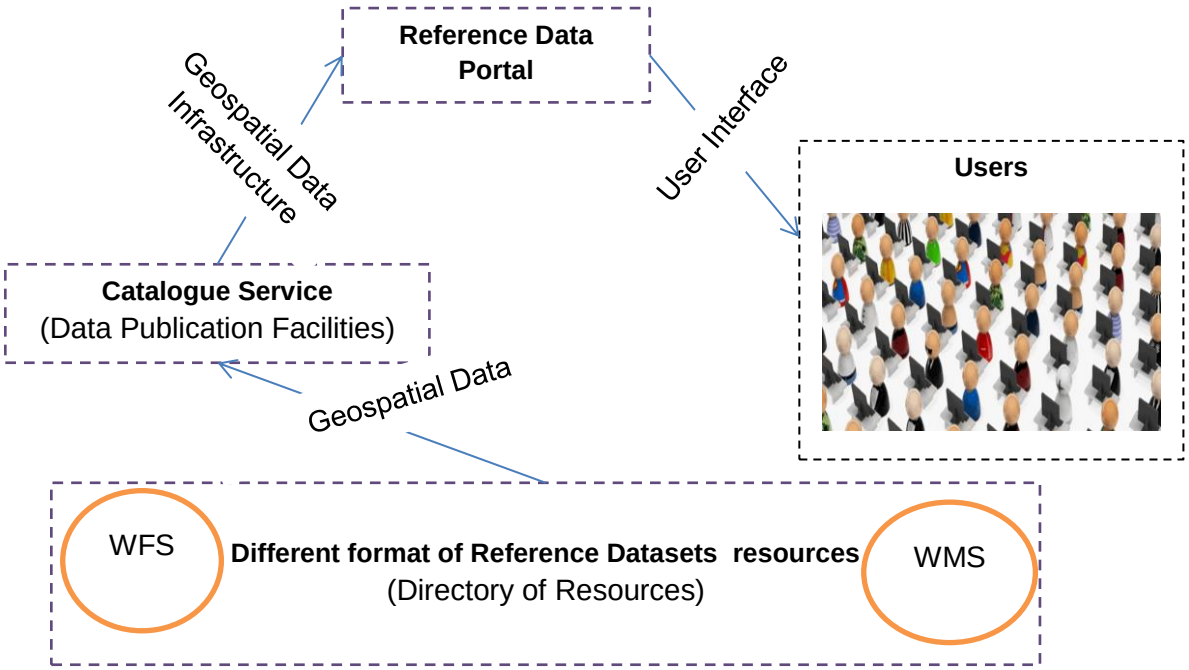


Figure 5: User Accessibility (based on geoportal and accessibility literature)

Accessibility refers to the quality of being available when needed (Vocabulary.com 2015). Web accessibility refers equal access to online content and services for users (Glossary 2005) where user interface should be remarkable, operable and logical (Henry 2007). Geoportals organize directories, tools, support resources, data and application to access Geospatial resources (Maguire and Longley 2005). A successful plan to improve accessibility of the reference data portal addresses reliable source that means a nicely constructed database directory for data storage with a catalogue service to make them available to the user and a user-friendly interface with content development and orientation (Figure 5) of a project. Maximum usability, Design practice and Context & orientation are the main content to increase the accessibility of a geoportal (Figure 6).

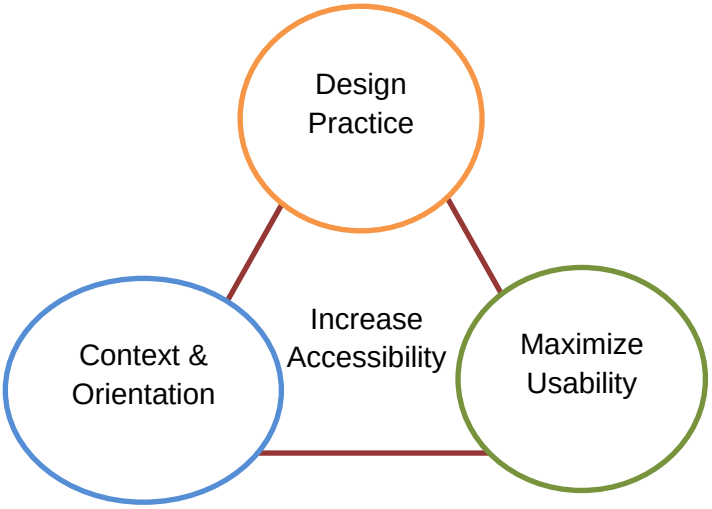


Figure 6: Concept to increase accessibility on the basis of literature from (Wachowicz 2006)

Usability:

Designing a user interface that is effective, usable and simple is known as usability (Henry 2007). Maximizing usability with technical aspects is important for accessibility.

	Accessibility
User Need	Easily accessible and easy download facility to the user
Technology to Address the need	The presentation of the content of the reference datasets, a simple description of the datasets (Crompvoets 2006) which means a well-organized infrastructure easy to access and use

Evaluation	UI should be with catalogue service to publish meta information and tools to search, query and download reference datasets.

Table 5: Main concern to increase accessibility of a reference data portal (from literature study)

Usability depends on several elements like learnability, memorability, effectiveness, efficiency and satisfaction (Henry 2011) and is a very important aspect for map based geoportal applications. Usability focusses especially on language, layout and graphics, information architecture, user interface and navigation (Bai et al. 2008). Because a geoportal has a complex involvement with user interaction like textual search, data retrieval or map interfaces geo-search.

Design Process:

User interface design is focussing on usability areas, user characteristics, background, responsibilities and workflow (Henry 2007). Two main concept of accessibility design are i) ensure smooth transformation by separating content from structure and the presentation style. And ii) content should be logical and navigable which means language should be clear and simple.

Context and Orientation:

W3C provide some guidelines about context and orientation of web accessibility. The title is the most important part of web orientation, which helps to facilitate frame identification and navigation. Moreover, associate labels clearly with their switchers (W3C 1999). Orientation of service/products information, purchase information (free of cost), quality and context information are also important (Bai et al. 2008).

3.3.2 Analysis of the accessibility of existing GOFC-GOLD portal:

An online usability evaluation survey was done to analyse the existing situation of the GOFC-GOLD reference data portal (13th to 17th of April 2015, figure 7) to identify weaknesses and area for further improvement. It was a controlled experiment focused on user needs. The usability test method was a mix of quantitative data and qualitative remarks (Granic et al. 2008) with a less strict heuristic evaluation (Nielsen 1994).

The goal of this usability evaluation survey is to prove our hypothesis that a good interface design helps users discovering, navigating and executing their anticipated data and information. The usability evaluation survey was constructed to discourse, which key features of the portal need to be improved.

The study involved 12 participants with high skills and knowledge about geoportal and geodatabase. According to their working experience and field of work, they are classified in three different groups. Usability test appliance was pilot tested with three participants as first group who are doing MSC thesis (major on GIS and Remote Sensing) to understand the user understanding. Moreover, I received prior feedback to improve and make user understandable survey questions. The other two groups consist of six experts' evaluators from forestry, ecology and agriculture research background who use global land cover maps for their research work and a group of 1st year master students of 'Geo-Information Science' department.

Figure 7: existing GOFC-GOLD reference data portal[1] (2014 Version)

The questionnaire (appendix A) comprised 33 questions in total: four open questions, four task related questions, eleven assertive questions, seven quantitative questions, four general questions and three for personal information. The main purpose of this usability test is to

[1] http://www.gofcgold.wur.nl//

identify user requirements including the use context, tasks to be performed, reasons for using geo-portals, user requirements related with geodatabase and their Meta information, necessary search potentials, and other potential improvements of existing geo-portals. In quantitative questions, the score for every question was calculated as an average mark on a five point Likert scale and every point has a different value. In order to understand the importance of geoportal design in a simple work situation, I created an exercise, a typical task and some user actions. The task was looking for a specific dataset in the existing geoportal of GOFC-GOLD and the participants had to find and download that datasets.

The whole usability test procedure is shown in figure 8.

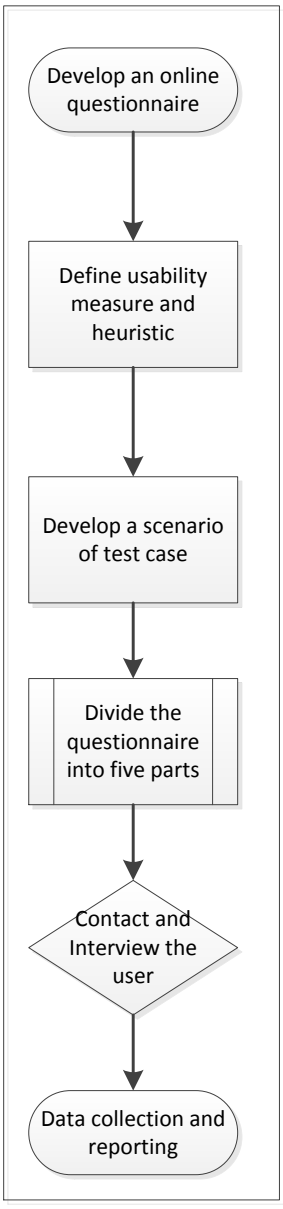


Figure 8: Test procedure and steps

3.4 Design to improve GOFC-GOLD portal:

From the literature research on GOFC-GOLD and its' reference datasets, it is clear that GOFC-GOLD provided reference datasets are potential assets for the research society. However, to access those reference datasets is time consuming and confusing because they do not have structured metadata information or a designed user-friendly geoportal. "Development of a new geoportal" is the third experiment to answer research question 3.

3.4.1 The planning process:

A good plan and a to-do list is the prerequisite of designing a geoportal. I made a priority list of functionalities to design the portal, which is given bellow-

- list the existing GLC reference datasets together in same pop-up button
- view the datasets of the reference plots
- a navigation bar with the facilities of different buttons
- request/search background information of different reference datasets
- report issues along the given reference datasets, e.g. by sending the lack of information, by asking about different datasets that are not in list to upload or update.

In order to be able to develop the necessary functionalities, a good planning is very important. For this reason, the development process was started by developing storage in PostgreSQL to store, retrieve and display the datasets, then build connections between the storage and QGIS for local use purpose and between the storage and GeoServer for online use purpose. After uploading all the reference datasets in GeoServer, for the next step is to write scripts by using JavaScript and HTML for online visualization and related tools to search the information. This is done by aid of a 'planning poker' card deck. This card deck consists of cards with several numbers like 0, ½, 1, 2, 3, 5 and 8. These numbers represent programmer points (PPs). One programmer point was defined as being two-day part of programming for a single person. A card deck was used to assess the required PP's for each user story. Besides this, also priorities were defined, so that more essential data storage is developed first. Priorities were given according to the MoSCoW-principle: M = Must-have, S = Should-have, C = Could-have and W = Would-like, with Must-haves being the most important and Would-likes the least. In addition, I worked on all the functionalities of the prototype Geoportal based on the following priority points.

Title	PP's (amount)	Priority
Create reference database in Postgres	½	M
Connect reference database with geoserver	1	M
Change style of legend	½	S
Add base map as an image of background	8	M
Show a menu on the side of the screen	4	M
Visualize all reference datasets	5	M
Show information on POI's (Points of Interests)	5	S
Show search related tools on the top of the screen	4	S

Table 6: the title describes the list of improvement for the geoportal. The second column represents the number of programmer points (PPs). A priority is given according to the MoSCoW-principle in third column.

3.4.2 Geoportal Design:

Nowadays there are different reference data portals available on internet. Geo-Wiki is one of them, which has a very organized, well informed, and user-friendly geoportal. GOFC-GOLD is also practicing to develop an organized and user-friendly geoportal. there are some specific standards like OGC (Open Geospatial Consortium), INSPIRE, SOA (Service Oriented Architecture) for data sharing procedures and rules (CON 2014). All geoportal architecture and technology need to follow those standards. At this point, it is needed to mention that standards and technological perception should be well suited with the user requirement.

3.4.2.1 System Architecture:

System architecture has been proposed for GOFC-GOLD portal, in order to understand if this design supports improve accessibility of reference datasets from new GOFC-GOLD portal. The following architecture (Figure 9) is based on standardized service interfaces using only open source software. A service-oriented architecture has been developed to test improvement of accessibility of reference datasets. As a standard architecture, this

application framework consists of several components. These components are divided into different layers according to their functionality. Server side components have been more focused on in this research.

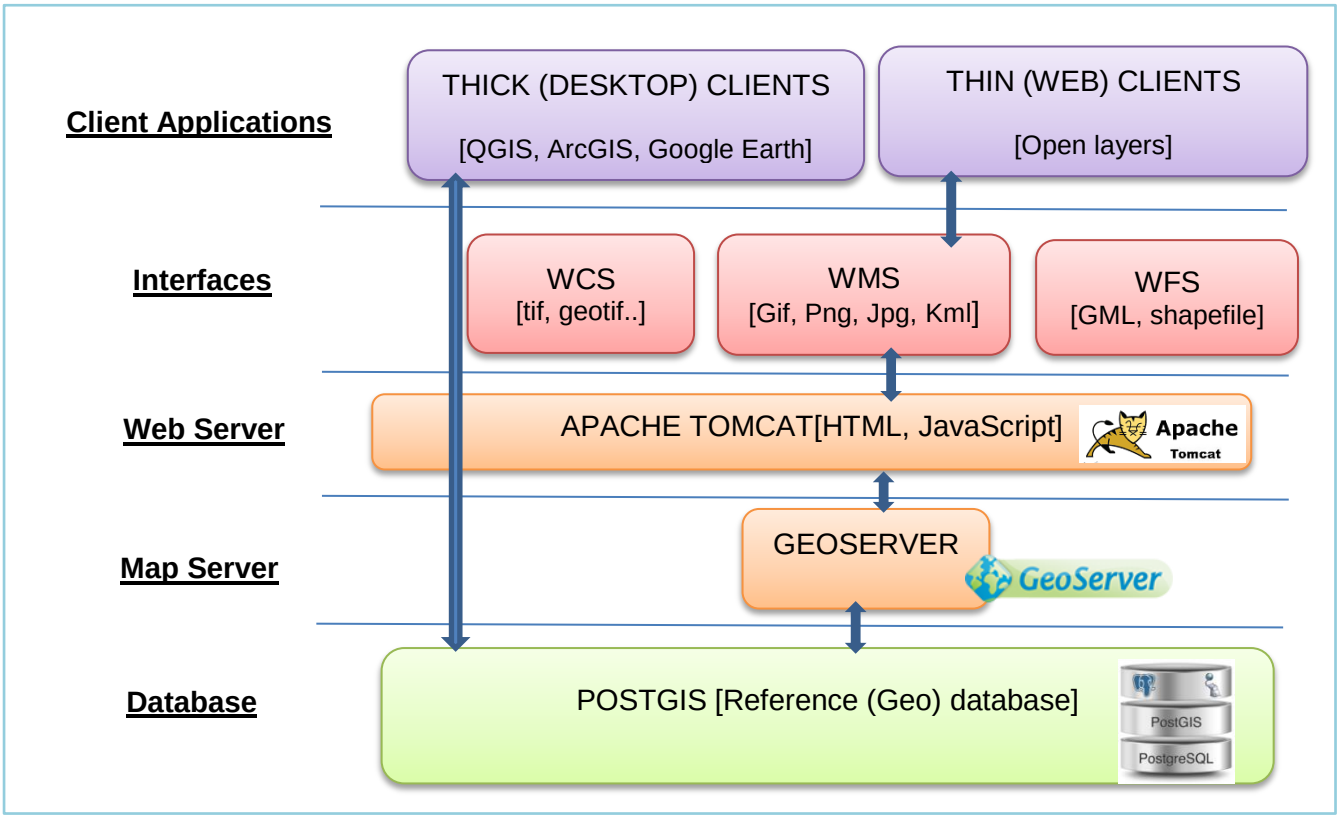


Figure 9: System Architecture of Proposed GOFC-GOLD portal

Database:

The main goal of this research is to provide reference datasets as a useful, articulate and multi-source datasets, which will be easily accessible and well described for the user. This geo-database is a relational database for storing, querying and deploying of spatial data. Moreover, among this architecture it is easy to view and download datasets without any struggle. This database was developed using PostgreSQL with PostGIS. Because PostgreSQL is the object-relational, database management system (ORDBMS) and free open source software. Figure 10 shows the ER (Entity Relation) diagram of the developed reference database which I developed for the server side development of the new geo-portal. Appendix F have a brief description of Database library. As every reference dataset is independent there is no relation between each other in the table. As sample size of the database is too small (only four type of datasets) and the datasets are directly linked with Geoserver there is no query developed.

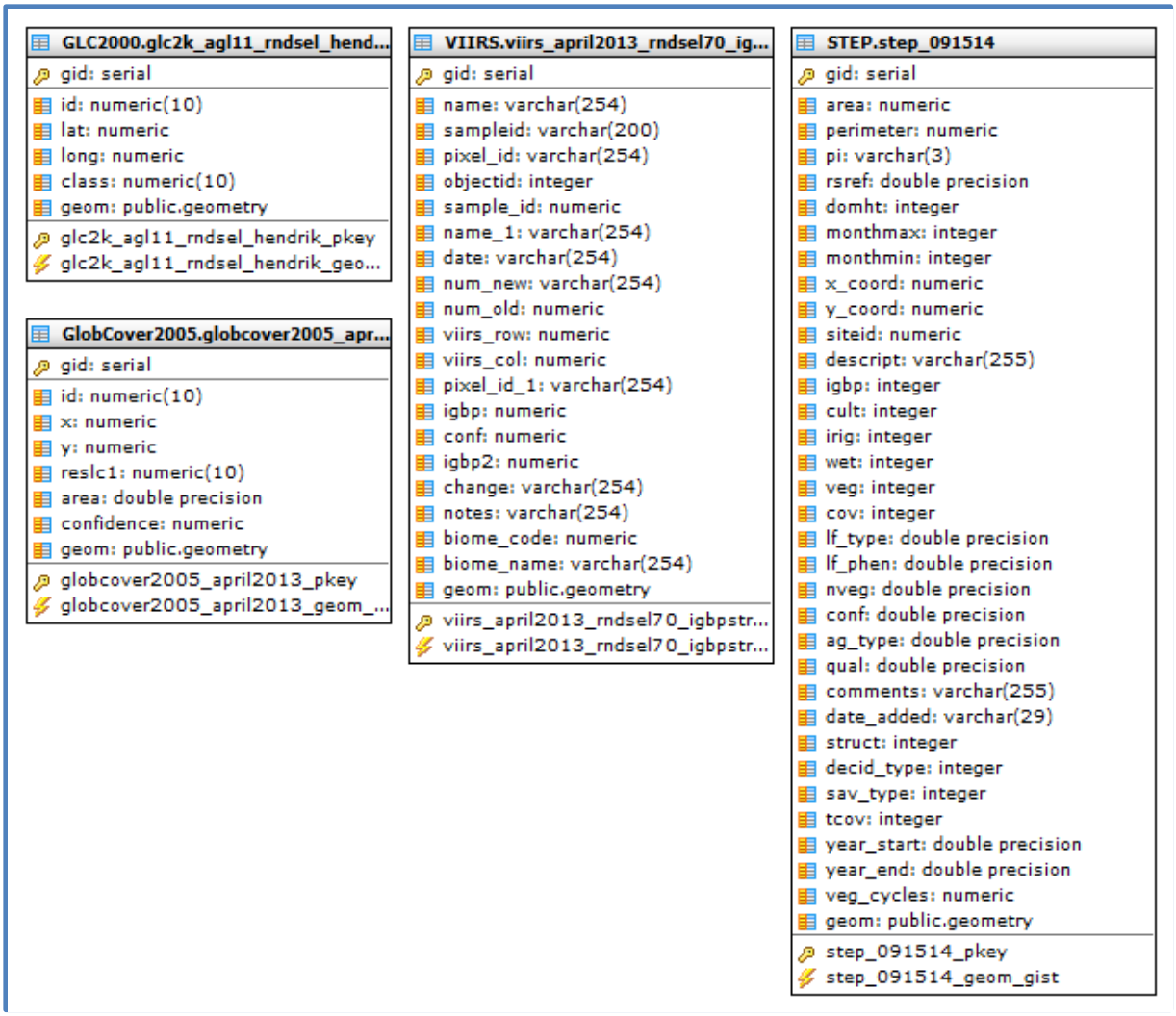


Figure 10: ER Diagram of Reference database

Server:

This portal is developed and organized using apache-tomcat server because this is a standalone server with a cross platform with a servlet container, flexible to run on different operating system, to support HTML and JavaScript, which were used to develop client applications. The functionality of geo-web services (here I used GeoServer) include published map rendering, feature streaming; data projection and geographic and attribute based queries, address geocoding. Apache server visualizes reference datasets with Geoserver support.

Interface:

Interface layer handles WMS (Web Map Service), WFS (Web feature Service) and WCS (Web Coverage Service) functionalities. In this project, WMS has been used to visualize the output on the web page.

Client Applications

Client side applications like full screen view, list of datasets in a pop-up menu, screen zoom level and base layer are developed by using JavaScript, which access the WMS (Web Map Service) data along with background satellite image. Finally, all these functionalities work with the HTML (Hyper Text Markup Language). A client can request GLC reference data from the geo-server using the HTTP-GET method. This method actually sends a request by sending an URL with the server address and a parameter to specify the data a client wants to get access to. The server returns an XML file that contains all requested reference data.

The following diagram (Figure 11) is known as UML (Unified Modeling Language) sequence diagram. This diagram will help to understand the interaction of messages between objects in the prototype geoportal and also will describe the sequence of actions which will be performed to complete a task. There are two types of dimensions of this diagram: a vertical dimension and a horizontal dimension. The vertical dimension represents the passage time and horizontal dimension represents the objects involved in the interaction. This sequence diagram in figure 11 represents four main objects, namely: Bing Maps Panel, WMS Layer Panel, GeoServer and Database.

In the beginning, a user has to access the map presentation by using correct web address. After a user has successfully entered into the system, the Bing map panel displays a satellite image of the world. Then, thematic maps are displayed on top of the satellite image. There is an option for the user to select a specific map for his use purpose.

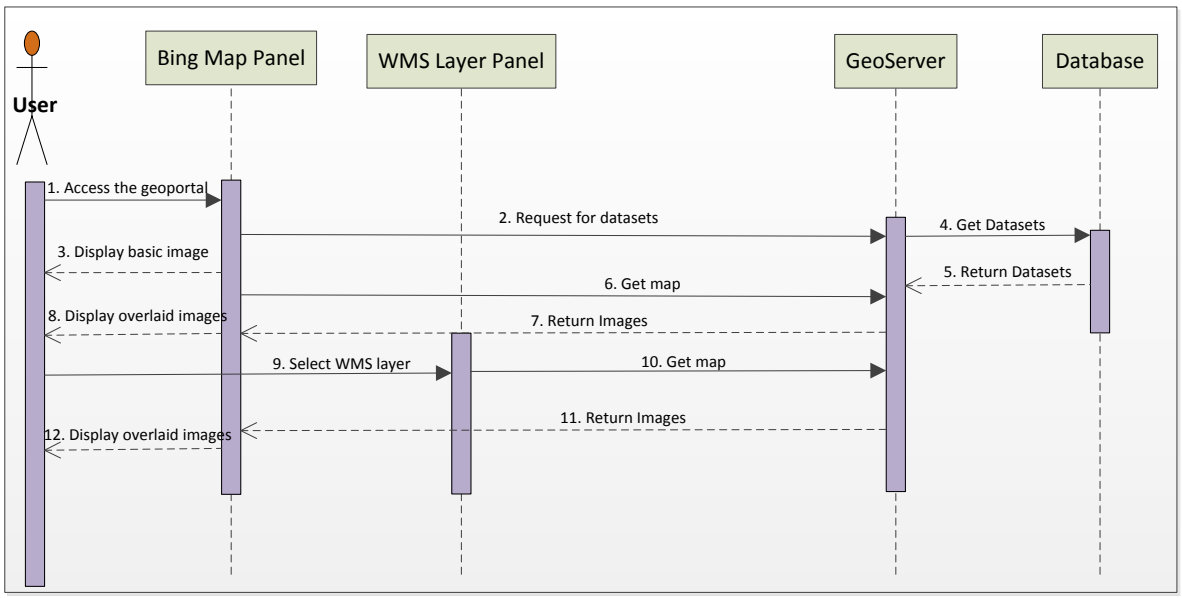


Figure 11: ER sequence diagram of the prototype geoportal

All the software and applications I used to develop the new prototype geoportal are open source software. In appendix E there is a brief discussion about my choice.

3.5 Usability test of conceptual design:

Usability test gives feedback to the system developer and owner about the product, to say whether it is good enough or needs more improvement or to identify the weakest part of the development or the strong part. The following specification in Table 7 shows how to test the design of new geoportal services for GOFC-GOLD.

Content	1. Completeness of datasets 2. Relevance with user needs 3. Status about datasets availability and updates
Satisfaction	1. Awareness of possibilities 2. Time necessary to master a new function
Interaction	1. Access to meta data 2. Tools like Zoom IN/Out, Pan, Identify, Proximity, Layer Selection etc. 3. Use functionality to link with the user and datasets.

Table 7: Specification of the dimension. Adopted from (Abugessaisa and Ostman 2011)

In usability perspective, different demonstration which allow users to search particular dataset and give the result as a map in a specific scale is known as user satisfaction. It improves the efficiency and effectiveness within a spatial context. Time spent is also an important part to determine user satisfaction. To design a usability test questionnaire all the points discussed in table 7 are important.

The focus of this test is to get feedback on content and satisfaction from the users. For the technical design of a geoportal it is important to mention that usability criteria is not enough to fulfil the user needs because usability is only a set of technical guidelines to develop a geoportal (Resch and Zimmer 2013). User experiences and opinions are also very important. In addition, after developing a new geoportal I need to validate my work. So I performed a user survey, to the same persons that have participated in the usability test for accessibility analysis of old GOFC-GOLD portal, to identify if the development is going in a right direction. Survey procedure and steps are shown in figure 11. This questionnaire survey is identical to the procedure of the questionnaire survey I did to find the answer of research question 2, but

with only a very limited number of questions (4 questions) because the development is not fully done. Only the development of the server side SDI has been completed, while the user side development is still remaining. The questions are based on the pictures of the new server side interfaces , because this is a prototype geoportal with limited server access.

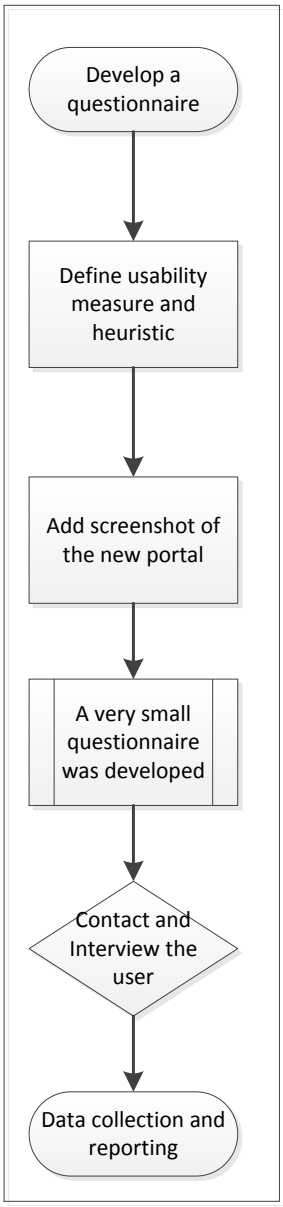


Figure 12: Survey procedure and steps

4. Results

4.1 Findings from Literature Research:

From the study for research question 1 - "**What are reference data and how do they validate the global land cover (GLC) datasets?**", the most important user requirement for this GLC map is to assess the accuracy. Because of the scarcity of validation data, reference datasets should be re-useable. An investigation on the efficient use of all available validated and calibrated datasets for GLC map is a very new user requirement. The sample size of the reference datasets should be large enough to achieve acceptable precision for accuracy estimation of any certain map. Below are some specifications of GOFC-GOLD datasets.

1. GOFC-GOLD adopted area units for their reference datasets which are not dependent on a pixel or a block of pixel. They can be easily applied for different map validation.

2. The sample size of reference datasets are of 500 reference sites with 5x5 km block and can be increased for small secondary sampling unit.

3. Stratified sampling based datasets do not use stratifications that are targeted at particular land cover maps.

4. Reference datasets provided by GOFC-GOLD are not map dependent. This is one of the most important specifications which makes these datasets exception from other reference datasets which is independent of any specific GLC dataset and applicable to different GLC datasets.

5. From GOFC-GOLD LCPO provided reference datasets are suitable for re-use and also examine the characteristics of the datasets. GOFC-GOLD is an ongoing project. They are still improving their reference datasets for validation.

From the summary above about **reference datasets** and their use some important findings are

- Reference maps should be easily accessible in different format like shape files with supporting files, KML, WMS and WFS.
- Users of GLC maps need clear detailed information of the validated reference maps. It will help users to select a suitable reference map and will reduce the uncertainty in their applications.

- Visual interpretation makes reference data more transparent, which will envisage the location accuracy.
- The list of essential information should be included in metadata information of the reference datasets in an organized way for different applications of different analysis.
- To increase the acceptance of the reference datasets you need to describe focus/aim of evaluation, map used for evaluation, corresponding year, spatial resolution, thematic detail, quality flag information and classification scheme in the metadata information.

4.2 Analysis the accessibility of GOFC-GOLD portal:

A usability test for the evaluation of the existing portal was performed with a sample size of 12 participants. Four persons are experts in different research areas of Forestry, Agriculture and Remote Sensing and eight students. There are several issues in practice of the current geoportal implementation, such as missing search options, missing link between datasets and metadata descriptions, and missing filter, sorting and selection functions. In this test, user's satisfaction should not only based on user interface design and guidance but also consider the functional capability and Meta data information.

The respondents could not answer properly to the questions in the section referring to "task performed" (Appendix A) about timing and counting the clicks in the task. Therefore, I skipped those four questions during analysis. From first four open questions I tried to understand how popular the GOFC-GOLD geoportal is and if researchers use this portal frequently or not. Also I examined whether they use the geoportal to get their required datasets, and they were asked to rank the importance of interface or datasets or tools facilities.

From the graph (Figure 13), it is possible to make the priority list of the functionality of a geoportal. The question was "**What do you look for at first step when you open a geoportal?"**

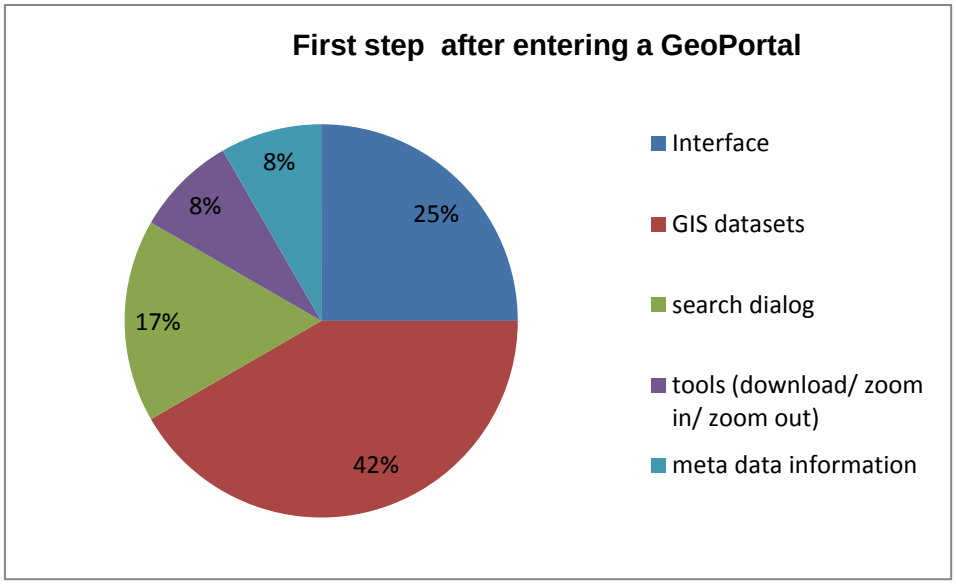


Figure 13: First step after entering a geoportal

It is clear from figure 13, when people enter into a geoportal most of them look for the geo-database. Nevertheless, it is surprising that only a very few of them looks for metadata information which is most important for the datasets. Metadata information proves the authentication of a datasets.

From the question, "**Which function of a geoportal you use frequently?**" (Figure 14) I tried to predict the user demand about the functionality from a geoportal. Half of the respondents looked for the GIS datasets. Second popular function is the search dialog. Users frequently use these functions in their operational procedures.

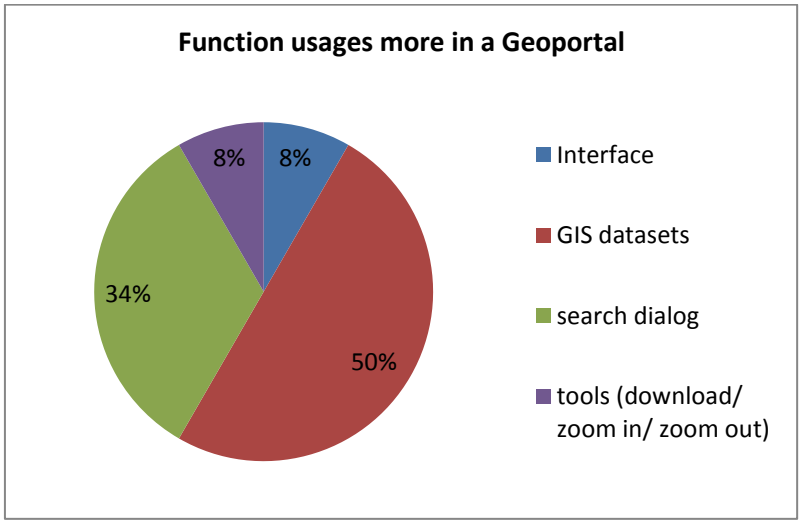


Figure 14: Function usages more in a Geoportal

The participant performed a small task when they participate in this questionnaire survey. They had to look for the reference datasets in existing GOFC-GOLD portal. And following analysis are based on the related questions with that task.

In figure 15 most of the users were not satisfied with the visual appearance, interface design, scattered and incomplete information, lacking of tools functionality and so on. Regarding the content 67% of all respondents have objections about the completeness of the datasets in figure 15. 42% of the respondents didn't find any information about regular updates.

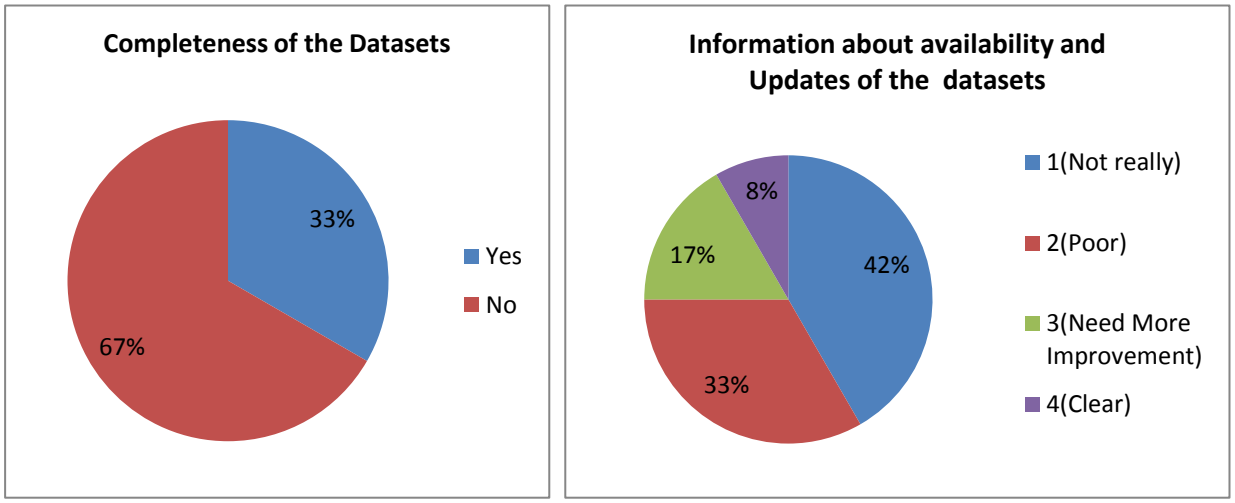


Figure 15: Completeness of the datasets and Information availability and updates of the datasets

To understand the user satisfaction about the portal I asked a question in the questionnaire survey using Likert Scale "**On a scale of 1 to 10 (1 = lowest and 10 = highest), how would you rate this GOFC-GOLD reference data portal?**"

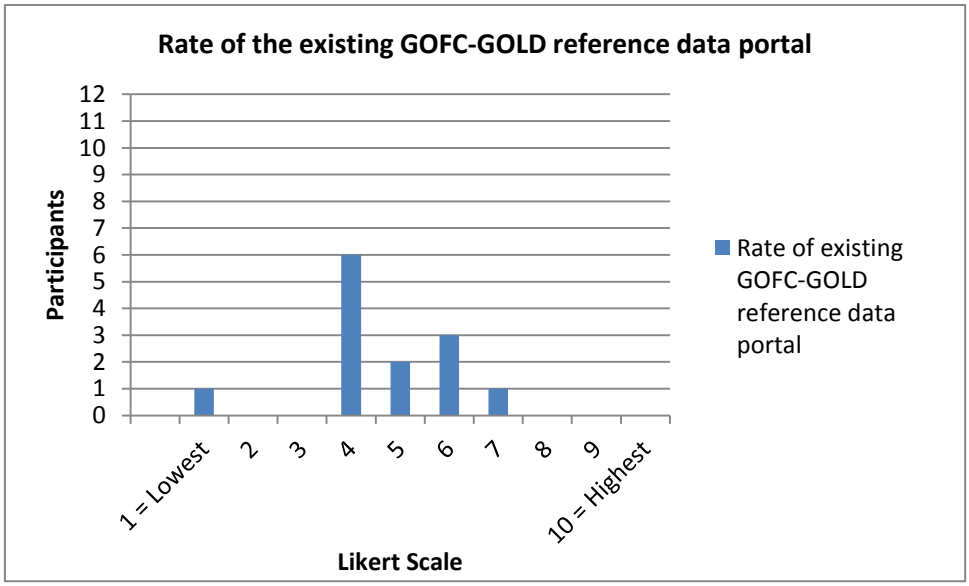


Figure 16: Rate of this existing GOFC-GOLD reference data portal

After performing all tasks 66.6% of the participants (Figure 16) share the opinion that the existing GOFC-GOLD portal is user-friendly enough. The average score of user satisfaction is 40%. The analysis of interaction focuses more on the availability of meta data information of the reference datasets, the link between the user and the datasets and the tools functionalities (like zoom in/out, pan, layer selection, download and so on).

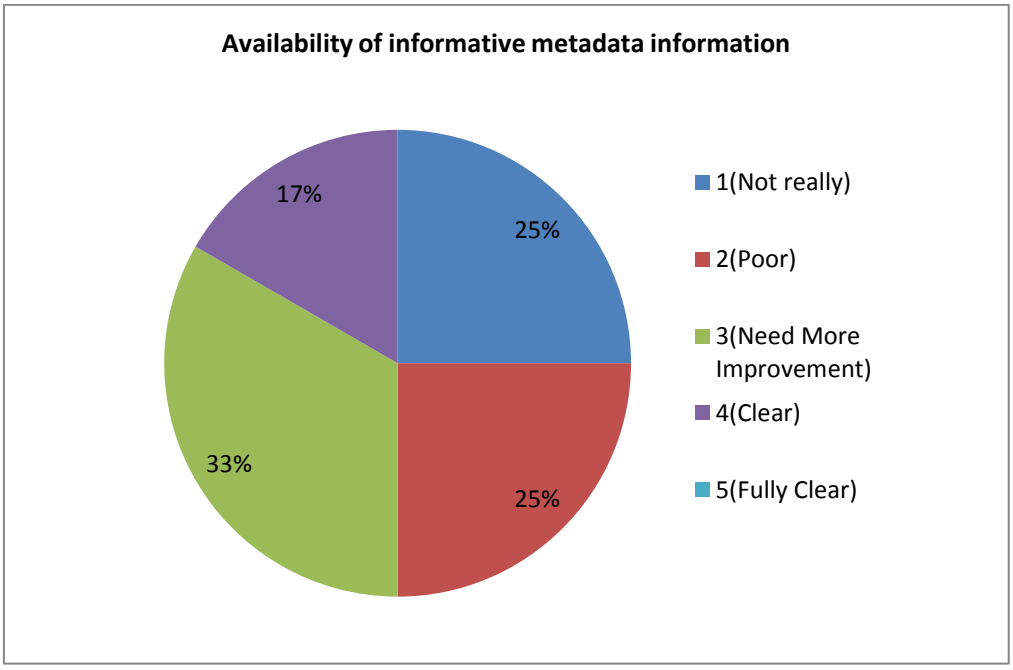


Figure 17: Availability of informative metadata information

The results regarding the evaluation of present metadata information is shown in figure 17. The question was **"Are the meta data information about the reference data clear**

for you?". Figure 17 shows that 25% of the respondent score the meta data information is totally unclear. 25% believe the metadata information is poorly explained. 33% of them think it's need more improvement. However 17% of the respondents judge it is clear. But no one said it's fully clear for them.

To understand the link between user and datasets I asked about the design of the portal **"Is the portal organized enough to find your required datasets?"** and about the user interface **"Is the interface of the portal user friendly?"** And the result is mostly negative, as can be seen below in figure 18.

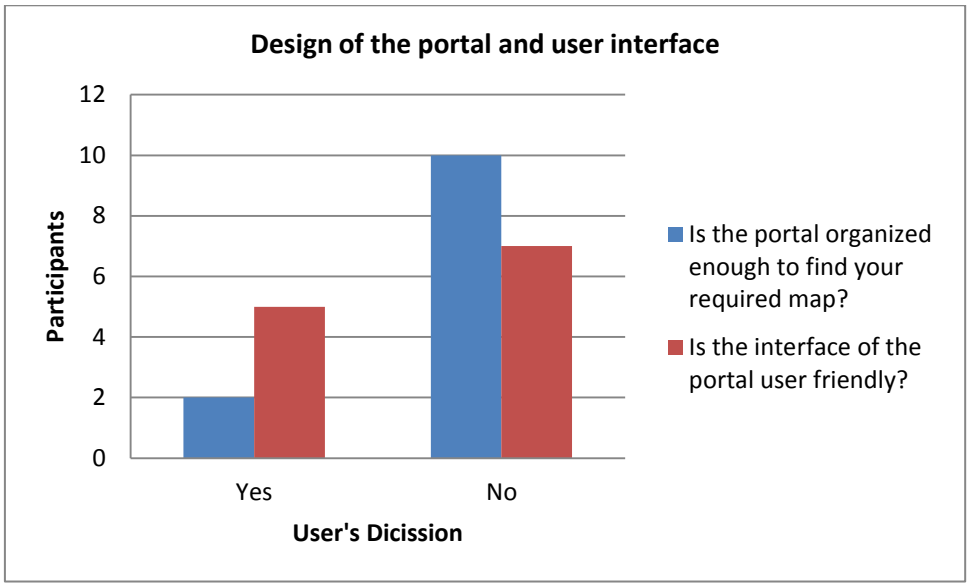


Figure 18: about the design of the existing portal and user interface

Additional results are presented in Appendix "B" which describes more clearly the situation of existing GOFC-GOLD geoportal and the actions needed to improve the accessibility.

4.3 Design of GOFC-GOLD portal:

A new geoportal (Figure 19) has been developed using open source software to check if it helps to improve the accessibility of GLC reference maps provided by GOFC-GOLD. There are two parts of a geoportal design. First part is Server side design and the second part is client side design. In my work, I focus more on server side development that include database construction, user-friendly viewing interface. Because to develop a new geoportal first need to assure the data storage, retrieval and display capacity which

belongs to server side design development. The script to develop new interface are given in Appendix C.

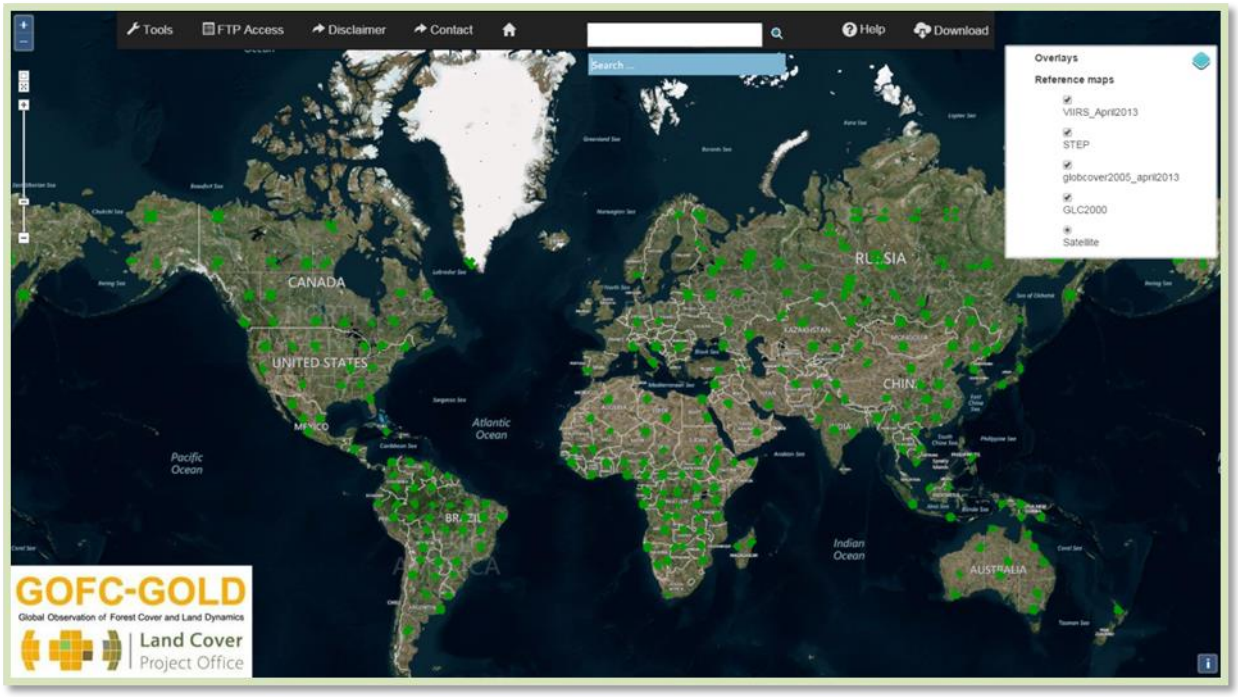


Figure 19: New prototype interface of the portal, the start-up screen with different functionalities combination with pop-up button (the right side white screen with a name list of the reference maps), left side zoom in and zoom out option, on the top a navigation bar.

When starting the new GOFC-GOLD geoportal on different browsers, the first screen that is visualized, displayed in figure 18, shows a satellite image from Bing map. It also shows reference datasets (mainly in this figure the green spots are the GLC2000 reference points), a pop-up button with a list of reference datasets, zoom in zoom out button and a navigation bar with search tool, download button, link with home page, FTP and disclaimer button. To see all datasets you need to use zoom in and out tools. In figure 20 all marked points represents different reference datasets; the arrows indicate the datasets content. With the pop-up button layers of a reference dataset can be turned on or off by clicking on the box underneath the name of the reference dataset.

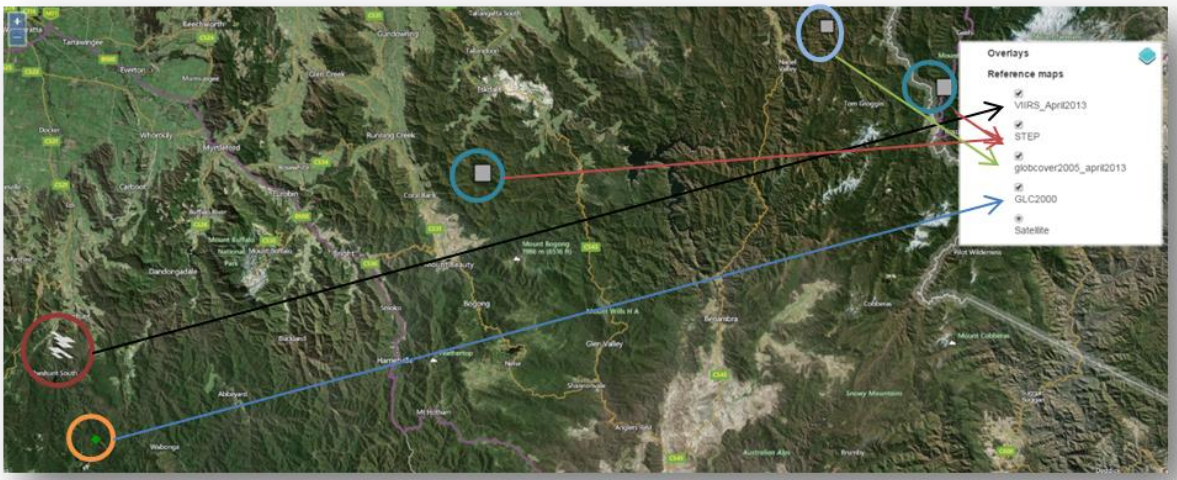


Figure 20: All datasets are in same list and visualized on the same interface together.

A navigation bar in Figure 21 was created which holds different links such as a link to the home page, contact addresses, disclaimer, FTP. A search engine, a dropdown button named Tools, a help button and most important a download button was included in that navigation bar. The Home button links the portal to the GOFC-GOLD website. The FTP link downloads all reports, newsletters and other necessary documents. from the contact button provides access to a list of contact persons of this project. The other three important tools are search engine, which facilitates a data or information search related with GOFC-GOLD Land Cover Project, a download button with the facilities of data download in different format such as KML, GML, WFS, WMS or WPS file. Lastly a help button is created to help the user to navigate through the portal.

Figure 21: Navigation bar with different buttons.

4.4 Usability analysis of new designed GOFC-GOLD portal:

A small questionnaire survey was prepared for the validation of the new prototype GOFC-GOLD geoportal aimed to understand the user acceptance. This survey was the follow up of the evaluation test of the existing GOFC-GOLD geoportal. Only four questions were held in this survey to evaluate the visualization improvement and different kind of tools facilities. All questions are presented in Appendix D.

The first question was about visualization, asking if it is easier for the user to find the reference datasets. The majority of the participants answered this questions with yes, as seen in figure 22.

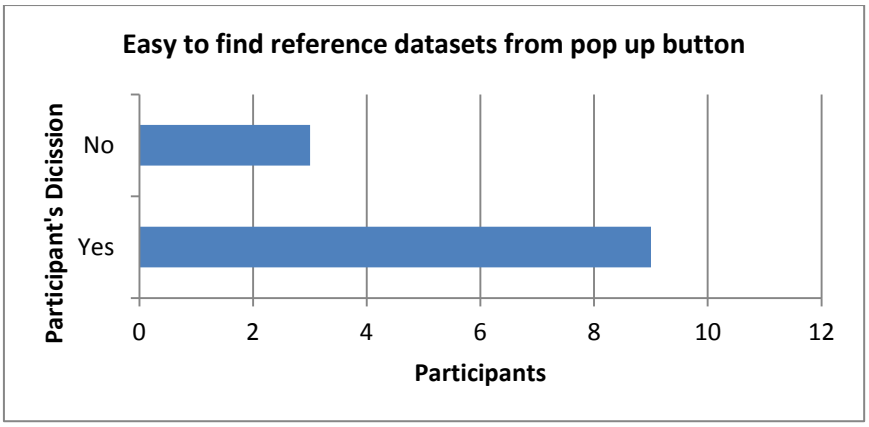


Figure 22: Easy to find reference maps from pop-up button

This graph describes, that out of 12 respondents 9 persons found it was easier to find the reference datasets. For 3 people, 25% of the respondent, it was still difficult, to find their required reference datasets from this new geoportal. However, this is an improvement to the old situation where 67% respondents face difficulties to find their required datasets from the existing GOFC-GOLD geoportal in figure 15 (completeness of the datasets).

The second question was about the comparison between the existing GOFC-GOLD portal and the new designed portal to understand if this change helps to improve the accessibility of reference datasets. And the result is following figure 23.

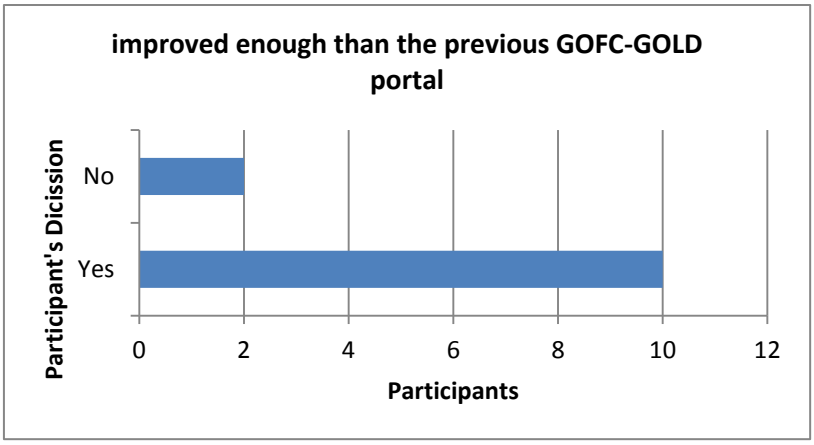


Figure 23: Improved enough than the previous GOFC-GOLD portal

Most of the participants agreed that the prototype portal is improved enough than the existing one. The participants didn't have another option then answering with yes or no. However, to gain insight into more need for improvement, open questions might have been more suitable.

The third question was aiming to evaluate the user-friendly interface. The user interface is one of the most important parts of a geoportal which can visualize a large amount of information interactively. The response from the respondents is presented in figure 24.

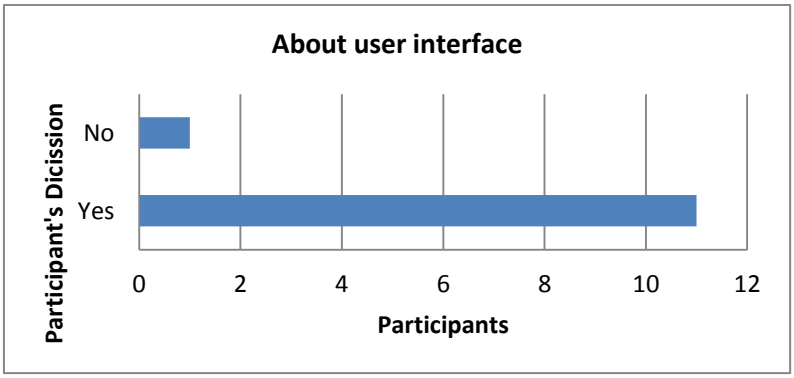


Figure 24: question about new user interface

This graph represents that the new user interface is highly appreciated by the user. If we compare the result from figure 18 and figure 24 it's clear that there is a big change in the results, which shows that the user interface has improved with the new concepts that have been used. In figure 17 41.66% of total respondents like the existing user interface and from figure 24 91.66% of total respondents like the new interface of the prototype geoportal.

The fourth and last question evaluate the user tools (download button as an example) that were created to facilitate the users by providing different tools like pan, zoom in/out, pop up button, download button, search engine and so on. The outcome is shown in figure 25.

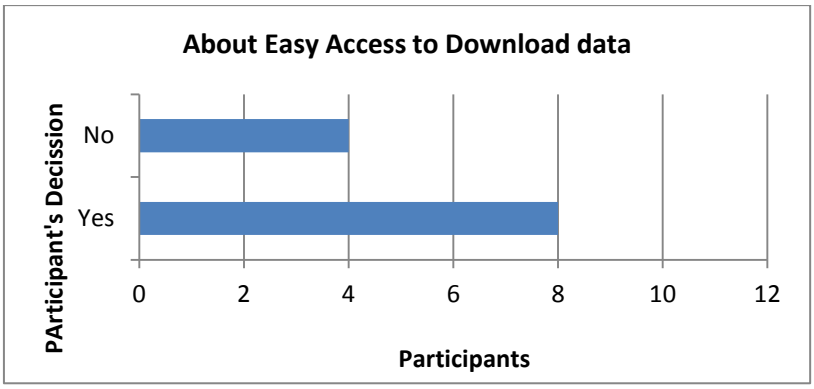


Figure 25: Easy access to download data

From the result of this graph it can be seen that, a large improvement on the client side of the application is required. One third of the total sample group is not satisfied with these tools.

5. Conclusion, Discussion and Recommendation

5.1 Conclusion:

For research work and for the accuracy assessment of GLC datasets we need suitable reference datasets which offer in good quality. To check the quality of GLC datasets the user should have knowledge about the error structure of land cover data to improve their product and prediction. To meet the user requirement well-informed, appropriate GLC reference datasets and a well described validation procedure should be included in the geoportal. This information will be important for the users to select their required reference datasets and is useful to increase the re-usability of the reference datasets. Multi-data records of reference datasets can provide accurate error estimation and uncertainty assessment.

The aim of this research was to improve access of reference datasets for global land cover map validation. To reach the aim four research questions were formulated to examine and evaluate the existing geoportal and a new prototype version of an improved geoportal.

A **literature study** has shown that GOFC-GOLD provided reference datasets are valuable assets for the researchers of the whole world. However, currently these datasets are not easily accessible or fully authentic.

To understand the situation of the existing GOFC-GOLD geoportal an **evaluation survey** was performed. The test consisted of a small task with several related questions. All the participants could finish the task well and the results indicated that there are serious issues that need to be change in the existing GOFC-GOLD geoportal. On the subject of data retrieval and data access, results of the questionnaire expressed that most of the users were in agreement about weak map interface, a lack of search dialogue and unorganized metadata information. The results also identified several problems related to spatial data access and worldwide sharing. In order to provide GLC reference datasets with the given facilities from GOFC-GOLD reference data portal, the existing portal has to be improved, including the creation of an organized database management system with sufficient data storage capacity and guidelines with structured Meta information for the user. A user-friendly interface with flexible tools developed which enable users to download reference datasets for their scientific research purpose.

A **prototype geoportal** has been developed as an example case there are two parts of geoportal development, server side development and client-application side

development. I focused more on server side development which consists of database development, connection with a geoserver and visualisation of datasets as web maps with satellite background image. However, it has been difficult to validate the server side work. So only the server side interface get more focus for the validation. A design for client applications still needs to be developed. A sequence diagram was delivered to visualize the detail of the reference datasets I used for my research purpose. The sequence diagram was delivered to visualize series of actions that can be performed by the user in the prototype geoportal. SDI, geoportal and geospatial web service concepts and technologies were used.

A **usability test** has been performed on the prototype geoportal interface. The result shows that more than 75% participants believed that the new geoportal (prototype) was capable to improve the accessibility of the reference datasets and the overall user satisfaction increase remarkable.

5.2 Discussion

Nowadays, Global Land Cover datasets takes a very special place in the research fields related with different global changes, agriculture, urban planning and so forth. GOFC-GOLD Land Cover Project Office took a good initiative to support those researches by providing standard quality GLC reference datasets with the detail of validation criteria.

An in-depth study which is reported in chapter 2 has been done to understand clearly about reference data and the validation criteria. Normally most of the reference datasets are developed from field surveys and on a specific research purpose, which is expensive and make it unsuitable to use for validation in other research work GOFC-GOLD provided reference datasets are free for use and are applicable for different GLC datasets. From the GOFC-GOLD website, I found only the shape files of reference datasets with a brief description of their validation process. However, reference datasets provided by GOFC-GOLD are re-useable but the metadata framework is not organized. An organized metadata framework help user as a guideline to use the proper reference datasets for specific applications.

In the 3rd GOFC-GOLD land monitoring symposium the GOFC-GOLD geoportal issue was discussed and they mention the existing portal as a prototype stage (Mora and Herold 2013). The aim of my research was to understand how a web based GIS solution

related with spatial data storage, retrieval and visualization of the datasets can improve access to these GOFC-GOLD reference datasets for global land cover data validation.

Han et al. (2015) developed a web-based system to support the GLC data production. For the development he used Open Source Software (OSS) and off-the-shelf (commercial). Fritz et al. (2012) also use OSS. Like them, I decided to use open source software because OSS are for free. More about my decision is described in Appendix E. But both of them developed the web-GIS architecture for commercial purposes and to collect and provide crowed sourcing GLC reference data. Thus, they focus more on the client-side application part of geoportal development whereas I focussed more on the server side architecture development. Their work is in implementation level and my one is in formal level. At this state, search tool, download tool (need more development), catalogue service, metadata design (over all client side design development) is still missing, and can be achieved in implementation level with a nice client-application side design development.

GOFC-GOLD LCPO start reorganizing their geoportal with a new concept from July 2015. The GOFC-GOLD website which leads access to the geoportal is still available during this time. They will added a download option of reference datasets in different formats. Within the map they divided the sample area in different classifications but there is no clear explanation about this classification, which might cause confusion. In my prototype geoportal design, maps are underplayed with a background satellite image, which is the start page of the portal. Also there is a link with the home page of the website a download button. The search option and different user tools on the navigation bar which have been proposed in my prototype are still missing in their new redevelopment. Another important difference between the redeveloped and the prototype version is that the redevelopment is based on CartoDB software service which is known as cloud computing platform to get GIS and Web mapping tools. Unlike the proposed OSS solution. GOFC-GOLD have to pay for that service.

5.3 Recommendations:

Recommendations for further improvement of Geoportal:

1. A user comments/needs a survey form, blog page; problem and solution page should be in the portal to know about the updates, regular contact with the users,

about new research who are interested to use GOFC-GOLD reference datasets and what are their requirement.

2. By adding a dropdown options for the entire list of datasets to give an easy access to the users.
3. A download button with several download options like pdf, jpg, img, shp, kml is required.
4. Each and every dataset should have individual pages with clear metadata information which will be linked with the dropdown button to collect the metadata and validated reference datasets.

All these recommendations should be included as part of a client-side application design.

Recommendations for Assessment improvement of GLC reference datasets:

5. To meet the user requirement a well-informed appropriate GLC reference datasets and map validation procedure should be well described. This information will be important for user selection and also useful for increasing the re-usability of the reference datasets. Multi-data record of reference datasets can provide accurate error estimation and uncertainty assessment. Standard baseline validation criteria should be described in detail in metadata information. Like

- Legend
  - Classification Scheme
  - Number of class
  - Classifier information provided
- Sampling Design (cost and statistical precision)
  - Sample unit type and size
  - Sample size
  - Sample selection scheme
  - Sample stratification
  - Inclusion probability
  - Minimum Mapping Unit

- Response design (reflection on the agreement between map and reference classification)

    - Source of information
    - Temporal coverage
    - Location accuracy
    - Labelling procedure
    - Sample verification
    - Confidence in interpretation
    - Majority classes and their fraction

- Current use

    - Intended application
    - Other applications
    - Applied pre-processing
    - Derived accuracy estimates

The given proposed structure is for the further development of the metadata information of GOFC-GOLD provided reference datasets. From my research point of view it would be user friendly and easy accessible if the future development will follow these recommendations.

References:

Abugessaisa, I. and A. Ostman (2011). "Testing-sdi: e-government prospective, requirements, and challenges." International Journal of Public Information Systems **7**.

Achard, F., et al. (2002). "Determination of Deforestation Rates of the World's Humid Tropical Forests." Science **297**: 999-1002.

Aditya, T. and M. J. Kraak (2009). Geoportals and the GDI Accessibility. Handbook of Research on Geoinformatics. H. A. Karimi**:** 42-50.

Al-Kilidar, H., et al. (2005). The use and usefulness of the ISO/IEC 9126 quality standard. Empirical Software Engineering, 2005. 2005 International Symposium on, IEEE.

Alan H. Strahler, et al. (2006). "GLOBAL LAND COVER VALIDATION: RECOMMENDATIONS FOR EVALUATION AND ACCURACY ASSESSMENT OF GLOBAL LAND COVER MAPS."

Arino, O., et al. (2008). "GlobCover: The most detailed portrait of Earth." European Space Agency Bulletin **2008**: 24-31.

Bai, B., et al. (2008). "The impact of website quality on customer satisfaction and purchase intentions: Evidence from Chinese online visitors." International Journal of Hospitality Management **27**: 391-402.

Bartholomé, E. and A. S. Belward (2005). "GLC2000: a new approach to global land cover mapping from Earth observation data." International Journal of Remote Sensing **26**: 1959-1977.

Belward, A. S., et al. (1999). "The IGBP-DIS global 1-km land-cover data set DISCover: A project overview." Photogrammetric Engineering and Remote Sensing **65**: 1013-1020.

Bevan, N. (2009). "International standards for usability should be more widely used." Journal of usability Studies **4**: 106-113.

Bontemps, S., et al. (2011). GlobCover 2009: Products description and validation report. ESA Glovecover project. O. Arino and K. V.

Bontemps, S., et al. (2011). "Revisiting land cover observations to address the needs of the climate modelling community." Biogeosciences Discuss. **8**: 7713-7740.

Bontemps, S., et al. (2012). "Revisiting land cover observation to address the needs of the climate modeling community." Biogeosciences **9**: 2145-2157.

Boundless (2015). "GeoServer in Production ". Retrieved 18-02-2015, from http://boundlessgeo.com/whitepaper/geoserver-production-2/.

Chuvieco, E. and A. Huete (2009). Accuracy Assessment. Fundamentals of Satellite Remote Sensing CRC press.

Clark, M. L. and T. M. Aide (2011). "Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) for Collecting Land-Use/Land-Cover Reference Data." Remote Sensing **3**: 601.

CON, E. (2014). EVALUATING GEOPORTALS. Geodetic and Geographic Information Technologies Department. THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES, MIDDLE EAST TECHNICAL UNIVERSITY**:** 108.

Congalton, R. G. and K. Green (2009). Assessing the accuracy of remotely sensed data: principles and practices, CRC press.

Crompvoets, J. W. H. C. (2006) Global Developments of National Spatial Data Clearinghouses.

De Longueville, B. (2010). "Community-based geoportals: The next generation? Concepts and methods for the geospatial Web 2.0." Computers, Environment and Urban Systems **34**: 299-308.

Defourny, P., et al. (2009). "Accuracy assessment of a 300 m global land cover map: the GlobCover experience." Proceedings of the 33rd International Symposium on Remote Sensing of Environment.

Dictionary (2007). Validation. The Free Dictionary by Farlex.

Doxygen (2005). "Geospatial Data Abstraction Library ". Retrieved 25-09-2014, 2014, from http://www.gdal.org/index.html.

Dragićević, S. (2004). "The potential of Web-based GIS." **6**: 79-81.

Durpaire, J. P., et al. (1995). "Spot-4 vegetation instrument: Vegetation monitoring on a global scale." Acta Astronautica **35**: 453-459.

ESA (2011). "First official release of a new generation of satellite-derived global land cover products." Retrieved 08-10-2014, 2014, from http://www.esa-landcover-cci.org/?q=node/156.

Esri (2013). "Esri International User Conference." from http://events.esri.com/uc/QandA/index.cfm?fuseaction=answer&conferenceId=CCAEEE69-1422-2418-7F1D0EB8490B776D&questionId=4881.

Foley, J. A., et al. (2005). "Global Consequences of Land Use." Science **309**: 570-574.

Friedl, M. A., et al. (2002). "Global land cover mapping from MODIS: algorithms and early results." Remote Sensing of Environment **83**: 287-302.

Friedl, M. A., et al. (2010). "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets." Remote Sensing of Environment **114**: 168-182.

Fritz, S., et al. (2012). "Geo-Wiki: An online platform for improving global land cover." Environmental Modelling & Software **31**: 110-123.

Fritz, S., et al. (2011). "Highlighting continued uncertainty in global land cover maps for the user community." Environmental Research Letters **6**: 044005.

GEOSS (2005). The Global Earth Observation System of Systems GEOSS 10-Year Implementation Plan.

GlobCover (2000-2011). "Welcome to the European Space Agency GlobCover Portal." Retrieved 8-10-2014, 2014, from http://due.esrin.esa.int/globcover/.

Glossary, M. (2005). "The Motive Web Design Glossary ". Retrieved 23-02-2015, 2015, from http://www.motive.co.nz/glossary/web-access.php#afterMenu.

GOFC-GOLD (2011). A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forests, and deforestation. F. Achard, S. Brown, R. D. Fries et al. GOFC-GOLD Project Office, hosted by Natural Resources Canada, Alberta, Canada.

GOFC-GOLD and F. M. a. M. Team (2014). "Fire Monitoring & Mapping Implementation Team." Retrieved 27-08-2014, 2014, from http://gofc-fire.umd.edu/.

Granic, A., et al. (2008). Experience with Usability Testing of Web Portals. WEBIST 2008, Proceedings of the Fourth International Conference on Web Information Systems and Technologies, Volume 2, Funchal, Madeira, Portugal, May 4-7, 2008.

Groot, R. (1997). "Spatial data infrastructure (SDI) for sustainable land management." ITC Journal: 287-294.

Han, G., et al. (2015). "A web-based system for supporting global land cover data production." ISPRS Journal of Photogrammetry and Remote Sensing **103**: 66-80.

Hansen, M. C., et al. (2003). "Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm." Earth Interactions **7**: 1-15.

Hansen, M. C., et al. (2000). "Global land cover classification at 1 km spatial resolution using a classification tree approach." International Journal of Remote Sensing **21**: 1331-1364.

Hansen, M. C., et al. (2010). "Quantification of global gross forest cover loss." Proceedings of the National Academy of Sciences **107**: 8650-8655.

He, X., et al. (2012). "Geoportal usability evaluation." Int. J. Spat. Data Infrastruct. Res **7**: 88-106.

Henry, S. L. (2007). "Just Ask: Integrating Accessibility Throughout Design." from http://www.uiaccess.com/accessucd/background.html.

Henry, S. L. (2011). "Understanding Web Accessibility ". Retrieved 20-02-2015, 2015, from http://www.adobe.com/macromedia/accessibility/pub/acc_sites_chap01.pdf

Herold, M., et al. (2009). Translating and evaluating land cover legends using the UN Land Cover Classification System (LCCS).

Herold, M., et al. (2008a). "Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets." Remote Sensing of Environment **112**: 2538-2556.

Herold, M., et al. (2008b). Report of the GOFC-GOLD/CEOS Working on Land Cover Change Accuracy Assessment as part of the Third GOFC-GOLD Land Cover Symposium Friedrich-Schiller University.

Herold, M., et al. (2011). "User Requirements Documents: Land Cover CCI." Université Catholique de Louvain (UCL)-Geomatics, Louvain-la-Neuve, Belgium.

Herold, M., et al. (2006). "A joint initiative for harmonization and validation of land cover datasets." Geoscience and Remote Sensing, IEEE Transactions on **44**: 1719-1727.

Hese, S., et al. (2005). "Global biomass mapping for an improved understanding of the CO2 balance - the Earth observation mission Carbon-3D." Remote Sensing of Environment **94**: 94-104.

Hibbard, K., et al. (2010). "Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling." International Journal of Climatology **30**: 2118-2128.

Inc, A. (2003). "What is Reference data?" Reference Data Portal. Retrieved 04-03-2015, 2015, from http://www.refdataportal.com/index.cfm?fuseaction=refdata&mode=whatisit.

Inc., O. G. C. (2007). OpenGIS® Geography Markup Language (GML) Encoding Standard. C. Portele. **3.2.1**.

Kulkarni, A. T., et al. (2014). "A web GIS based integrated flood assessment modeling tool for coastal urban watersheds." Computers & Geosciences **64**: 7-14.

LCPO, G.-G. (2012-2013). "Welcome to the GOFC-GOLD Reference Data Portal Beta version opened on April 2013 ". Retrieved 12-09-2014, 2014, from http://www.gofcgold.wur.nl/sites/gofcgold_refdataportal.php.

Lillesand, T., et al. (2014). Remote Sensing and Image Interpretation, Wiley.

Loveland, T. R., et al. (2000). "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data." International Journal of Remote Sensing **21**: 1303-1330.

Loveland, T. R., et al. (1999). "An analysis of the IGBP global land-cover characterization process." **65**: 1021-1032.

Maguire, D. J. and P. A. Longley (2005). "The emergence of geoportals and their role in spatial data infrastructures." Computers, Environment and Urban Systems **29**: 3-14.

Manisa, M. and D. B. Nkwae (2007). Developing Botswana Spatial Data Infrastructure: From Concept to Reality. Hong Kong SAR.

Mora, B. and M. Herold (2013). Report of the 3rd GOFC-GOLD Land Monitoring Symposium. Annual Report 2012, Wageningen University.

Mora, B., et al. (2014b). A Framework for the Validation of Global Land Cover Reference Datasets, Wageningen University and Research Center.

Mora, B., et al. (2014c). "GOFC-GOLD EFFORTS IN SUPPORT OF GLOBAL LAND COVER MAPPING ACTIVITIES." Retrieved 07-04-2015, 2015, from http://lcluc.umd.edu/Documents/ScienceTeamMtg/2014APR/Poster/Brice_Mora_Poster.pdf.

Myneni, R. B., et al. (2002). "Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data." Remote Sensing of Environment **83**: 214-231.

Nielsen, J. (1994). Usability inspection methods. Conference companion on Human factors in computing systems, ACM.

Olofsson, P., et al. (2013). "Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation." Remote Sensing of Environment **129**: 122-131.

Olofsson, P., et al. (2012). "A global land-cover validation data set, part I: fundamental design principles." International Journal of Remote Sensing **33**: 5768-5788.

Peterson, M. P. (2001). "the development of map distribution through the internet." Department of Geography University of Nebraska at Omaha.

Pflugmacher, D., et al. (2011). "Comparison and assessment of coarse resolution land cover maps for Northern Eurasia." Remote Sensing of Environment **115**: 3539-3553.

Rajabifard, A., et al. (2002). "Future directions for SDI development." International Journal of Applied Earth Observation and Geoinformation **4**: 11-22.

Ramankutty, N., et al. (2008). "Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000." Global Biogeochemical Cycles **22**.

Resch, B. and B. Zimmer (2013). "User Experience Design in Professional Map-Based Geo-Portals." ISPRS International Journal of Geo-Information **2**: 1015-1037.

Shen, S., et al. (2008). Sensor web oriented web-based GIS. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). **5373 LNCS:** 86-95.

Souza, C., Jr., et al. (2009). "Web-oriented GIS system for monitoring, conservation and law enforcement of the Brazilian Amazon." Earth Science Informatics **2**: 205-215.

Stehman, S. V., et al. (2012). "A global land-cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land-cover class." International Journal of Remote Sensing **33**: 6975-6993.

Sterling, S. M., et al. (2013). "The impact of global land-cover change on the terrestrial water cycle." Nature Climate Change **3**: 385-390.

Tait, M. G. (2005). "Implementing geoportals: applications of distributed GIS." Computers, Environment and Urban Systems **29**: 33-47.

Townshend, J. R., et al. (2006). "A revised strategy for GOFC-GOLD." GOFC-GOLD Report **24**.

Townshend, J. R., et al. (2012). "Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges." International Journal of Digital Earth **5**: 373-397.

Tsendbazar, N. E., et al. (2014). "Assessing global land cover reference datasets for different user communities." ISPRS Journal of Photogrammetry and Remote Sensing.

Verburg, P. H., et al. (2011). "Challenges in using land use and land cover data for global change studies." Global Change Biology **17**: 974-989.

Vocabulary.com (2015). accessibility The dictionary

W3C (1999). "Web Content Accessibility Guidllines." Retrieved 23-02-2015, 2015, from http://www.w3.org/WAI/GL/WD-WAI-PAGEAUTH/

Wachowicz, M. (2006). On a conceptual framework for multimodal interaction on the Web: a usability perspective. Proceedings First International Workshop on Mobile Geospatial Augmented Reality, Alberta, Canada.

Wachowicz, M., et al. (2008). The effects of web mapping applications on user satisfaction: an empirical study. International perspectives on maps and the Internet, Springer**:** 397-415.

Wu, W., et al. (2008). "Validation and comparison of 1 km global land cover products in China." International Journal of Remote Sensing **29**: 3769-3785.

Zell, E., et al. (2012). "A user-driven approach to determining critical earth observation priorities for societal benefit." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5**: 1594-1602.

Zhan, X., et al. (2000). "The 250 m global land cover change product from the Moderate Resolution Imaging Spectroradiometer of NASA's Earth Observing System." International Journal of Remote Sensing **21**: 1433-1460.

Zhao, Y., et al. (2014). "Towards a common validation sample set for global land-cover mapping." International Journal of Remote Sensing **35**: 4795-4814.

Zhu, Z. and E. Waller (2003). "Global Forest Cover Mapping for the United Nations Food and Agriculture Organization Forest Resources Assessment 2000 Program " Forest Science **49**.

Zimmerman, C. (2014). "Software Options for Seeding and Tiling." Retrieved 16-09-2014, 2014, from http://opengeoportal.org/software/resources/software-options-for-seeding-and-tiling/.

Appendix A

Questionnaire: Current GOFC-GOLD portal

Usability Evaluation Questionnaire

Welcome and thank you for agreeing to participate in this evaluation session. It will take 30 minutes or a little bit more to finish this questionnaire survey.

The purpose of this user study is to evaluate the GOFC-GOLD geoportal, its design, data accessibility, navigation, interfaces and so on. Than it will be easier for me to figure out which part of the existing portal needs to be improved. I prepared this questionnaire to get feedback from the user of these reference datasets and the geoportal experts. As a volunteer in this study, your participation will be appreciated. You will fill out some questionnaires while using the portal to work through specific tasks.

There are two part of this questionnaire. In first part, you have to do a small task to answer the questions. Please count time when you will do the test. And in the second part, you will only answer the questions.

PART 1

Open Questions regarding the awareness of GOFC-GOLD portal

GOFC-GOLD is land cover project of European Space Agency and they are providing free reference data sets for Global Land Cover map validation in their GOFC-GOLD geoportal. How well do you know about GOFC-GOLD geo-portal?1 = not well, 5 = very well

	1	2	3	4	5	
not well	○	○	○	○	○	very well

How often do you use this portal?1 = never use, 5 = frequently use

	1	2	3	4	5	
never use	○	○	○	○	○	frequently use

What do you look for at first step when you open a geoportal?

- search dialog
- GIS datasets
- tools (download/zoom in/ zoom out)
- interface
- meta data information
- Other:

Which function of a geoportal you use frequently?

- search dialog
- GIS datasets
- tools (download/ zoom in/ zoom out)
- interface
- meta data information
- Other:

Task Performed

During the task you have to count the clicks and time to answer following five questions

Please download the VIIRS reference map from the GOFC-GOLD geoportalTo visit the portal please go for this link http://www.gofcgold.wur.nl/ . and note the time duration of visiting the portal.

:

How many times do you click on different heading to find the VIIRS reference map?

How many click do you need to find your required reference map?

How long time do you need to find your required reference map?

:

Task Related Questions (Design, Interface and tools)

Do you find the required reference map easily when you open the geoportal?

- Yes
- No

Is the portal organized enough to find your required map?(because a geoportal should have one interface to visualize all information but you are using several interfaces to find the required map.)

- Yes
- No

Is the interface of the portal user friendly?

- Yes
- No

Do you think GOFC-GOLD need to improve the design interface to make it more accessible for the users?

- Yes
- No

Do you find any button as a tool (not inside of a text paragraph) on the map to download reference maps?

- [ ] Yes
- [ ] No

Can you download reference maps in different format like shape/image/kml file

- [ ] Yes
- [ ] No

Do you find any tools to select your research area?

- [ ] Yes
- [ ] No

Do you find any information about validation criteria or validation process of these reference map?(the required reference data you are going to download is validated.)

- [ ] Yes
- [ ] No

Is the information in an organized way?

- [ ] Yes
- [ ] No

Do you think you can use this reference maps directly to validate your maps?

- [ ] Yes
- [ ] No

Do you found all criteria of a geoportal in this GOFC-GOLD portal?criteria means data services, catalog services, design of the portal etc.

- [ ] Yes
- [ ] No

Questions related to use of the Portal

Are the meta data information about the reference data clear for you?1 = not really, 2 = poor, 3 = need more improvement, 4 = clear, 5 = fully clear

	1	2	3	4	5	
Not Really	○	○	○	○	○	Fully Clear

does the meta data content about the reference map fit your requests?1 = useless, 2 = useable anyhow, 3 = useable, 4= informative, 5 = very informative

	1	2	3	4	5	
Useless	○	○	○	○	○	Very informative

Quality of map visualization (customized legend, adaptability of map's size, attribute queries, information about data quality, support for all reference system)1= very bad, 2 = bad, 3 = good, 4 = good enough, 5 = perfect

	1	2	3	4	5	
Very bad	○	○	○	○	○	Prefect

Could you locate the terms and conditions of downloading and using this reference maps?1= not found, 2 = hard to find, 3 = complicated, 4= described, 5 = well decribed

1	2	3	4	5

not found	○	○	○	○	○	well described

Could you find information on regular updates of reference maps?1= no information found, 2 = very few, 3 = ok, 4= good, 5 = good enough

	1	2	3	4	5	
No information found	○	○	○	○	○	good enough

Do the resolutions of the portal (given bellow in the images) match your expectations?1= poor, 2 = ok, 3= medium, 4 = high 5 = very high

	1	2	3	4	5	
Poor	○	○	○	○	○	Very High

How good is the categorization of information about reference maps on the portal?1= useless, 2 = ok 3 = less informative 4=informative, 5 = very informative

	1	2	3	4	5	
Useless	○	○	○	○	○	Very informative

General Questions

On a scale of 1 to 10 (1 = lowest and 10 = highest), how would you rate this GOFC-GOLD reference data portal

1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

Describe how the use of symbols and tools could be improved

Describe your experience with GOFC-GOLD reference data portal

If you have any additional comments about GOFC-GOLD portal

Personal Question

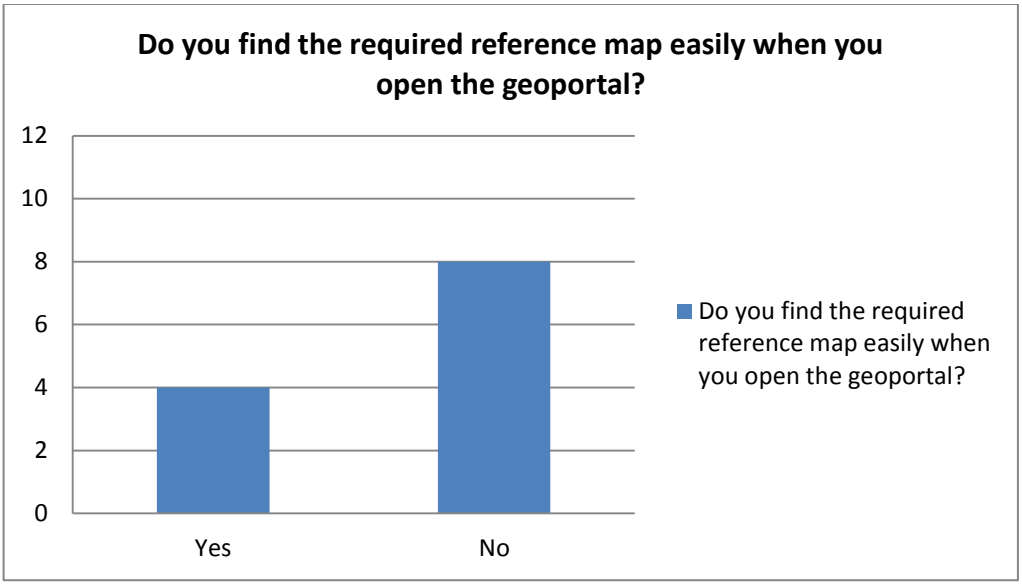
Your name:

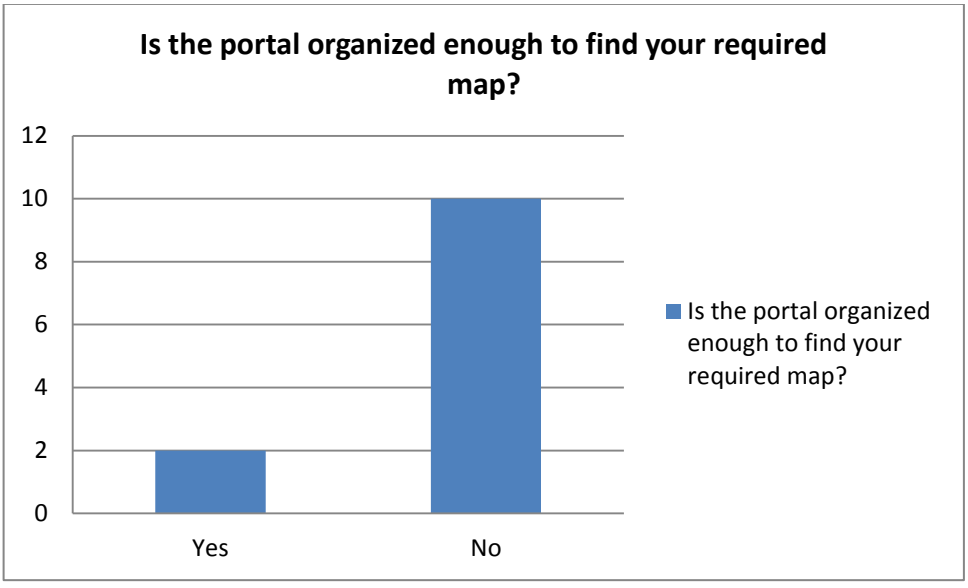
Email address:

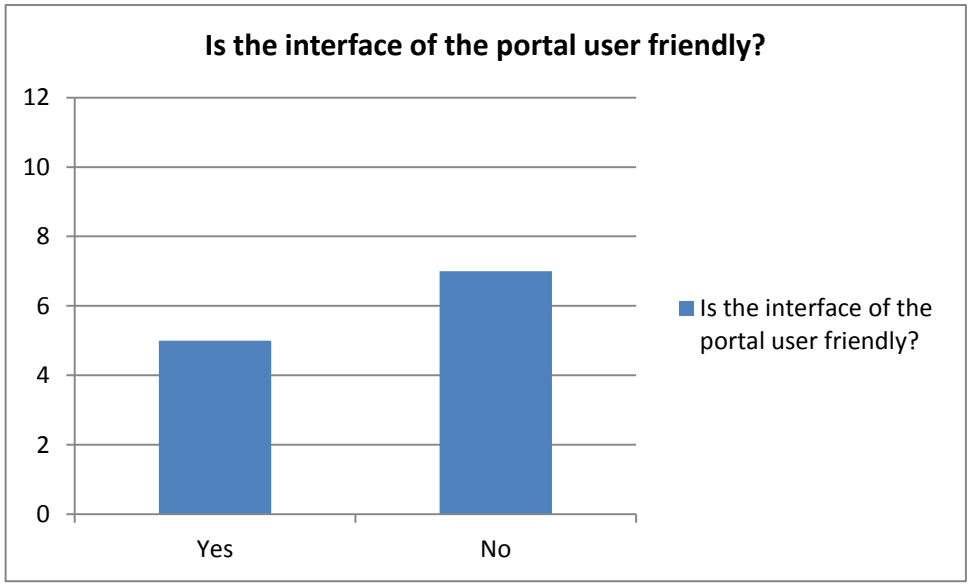
Organisation / Research domain:

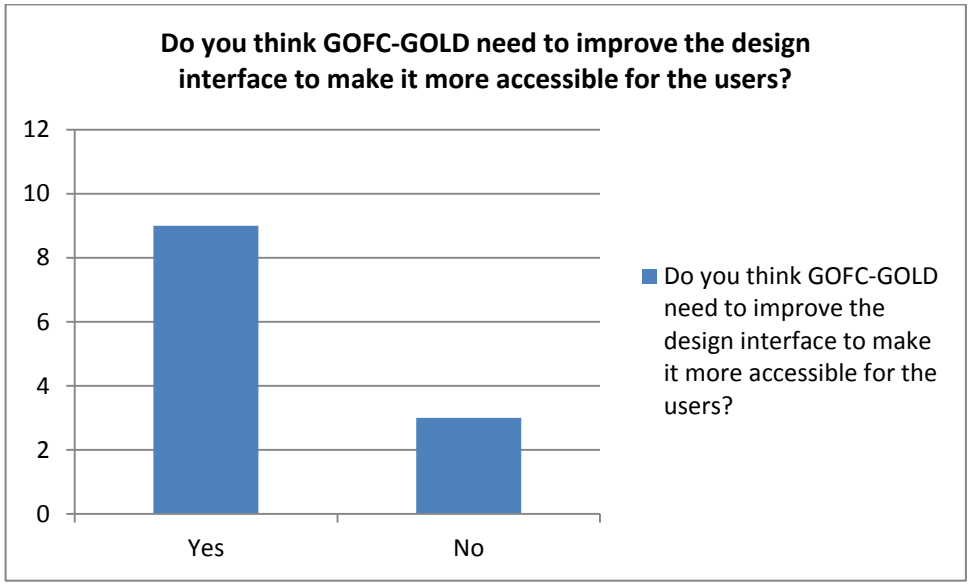
Appendix B
Questionnaire Results

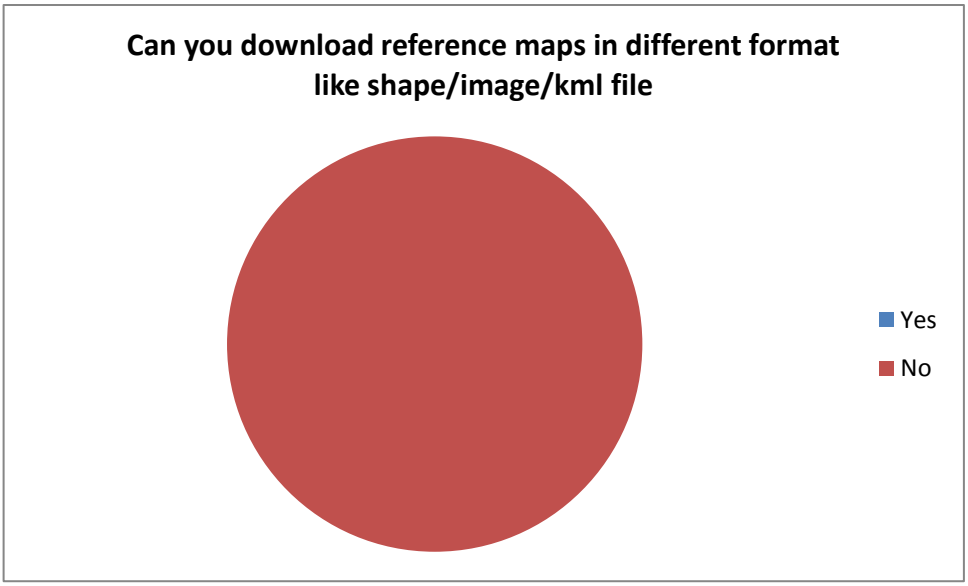
Results current GOFC-GOLD portal:



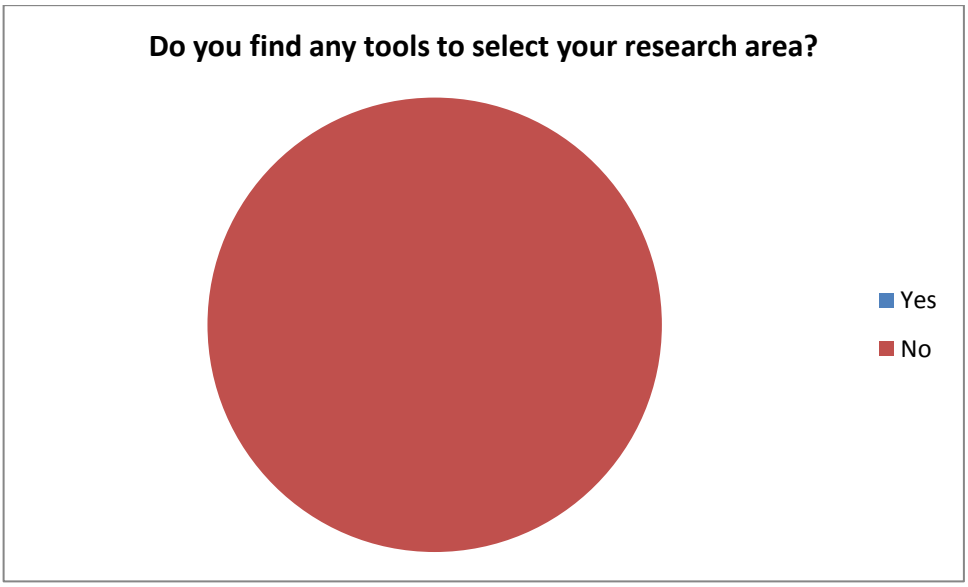


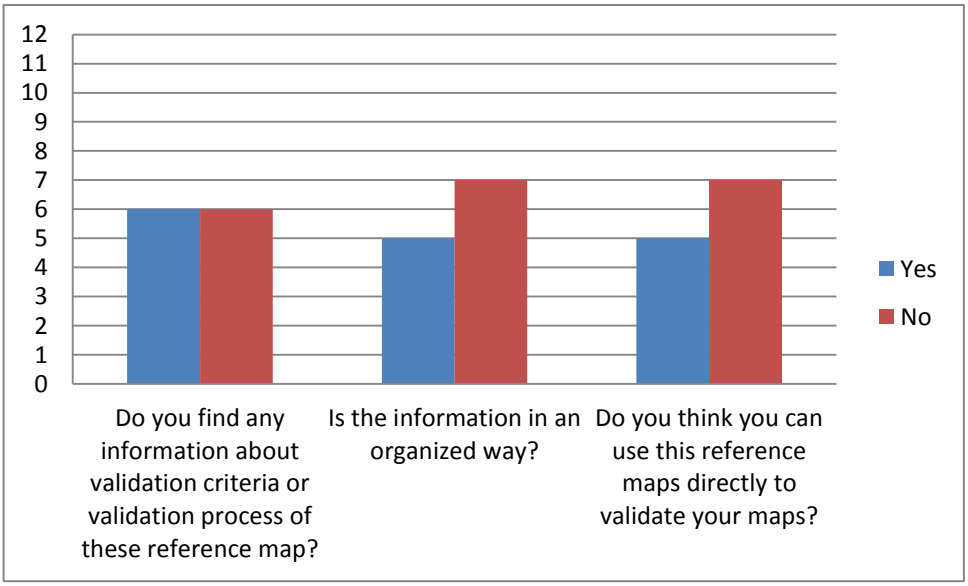

Is the interface of the portal user friendly?
12
10
8
6
4
2
0
Yes
No
Is the interface of the portal user friendly?


Do you think GOFC-GOLD need to improve the design interface to make it more accessible for the users?
12
10
8
6
4
2
0
Yes
No
Do you think GOFC-GOLD need to improve the design interface to make it more accessible for the users?



These two results looks a bit different. Because both of the result is 100% no. The actual problem is in existing (year of 2014) GOFC-GOLD geoportal. There was only one option to download the datasets as shape file. And there is no tools to select the interested area. You have to download full datasets and then using local software you can select your interested area.





Describe how the use of symbols and tools could be improved

- I am confused by the headings and arrangement of different product names, functions, metadata, and download links. Maybe it is helpful to make
- the names and metadata more obvious. The download links also need to be more obvious and state which format of the products I can select;

- add search tool

- use the different color/ bigger letter size for the symbol which indicates a link

- use a common web-symbol

- Small screen showing location of world where you are looking at. This would be especially helpful in situations when you had zoomed in more.
- I do not see a button to remove selected
- The option to layer the different maps together (using different symbols for each layer) would be helpful.
- I couldn't find the portal as shown above....
- I am completely lost. I could only find a download function. I cannot find any visualization tool. Is there one?
- It's okay I think
- A search dialog would help.

Describe your experience with GOFC-GOLD reference data portal

- Confusing; even don't know which is really reference data for which products, where I can see the metadata for more than half hour. Finally find some separated information in somewhere description of data.
- So far I never really use GOFC-GOLD as my reference data
- Which one? The one they have now?

That one makes more the impression of being a info-portal where you can get information about data, but I did not know that it was available for downloading geo-data.

The text is too much in focus, the map looks like a standard google maps product and reference data is presented in red dots. I was not even sure what they mean, and why they only represent a sample. Also optically appeared in the same layout on the reference map. I wasn't able to distinguish between datasets or show more see more than one. If data is represented by points it would be nice to have different colours for each dataset.

- In terms of downloading the data, this was not difficult. Since this is the most important thing, I would say it is therefore a good portal. Other functions should be optional.
- Poor
- Too less information on the data portal.
- I would have to find something particular that I need for a real case to say something about this

If you have any additional comments about GOFC-GOLD portal

- Improve it to be more user friendly, especially the user interface
- I am wondering if I'm looking at the right portal... I don't find any data except for a few shape files....
- It is not a type of portal i am used to.

Appendix C

Script: new GOFC-GOLD geoportal demonstrator

JavaScript:

To add a satellite image the code is

```
(function() {
    var map = new ol.Map({
        target: 'map',
        layers: [
            new ol.layer.Group({
                'title': 'Reference maps',
                layers: [
                    new ol.layer.Tile({
                        title: 'Satellite',
                        type: 'base',
  //source: new ol.source.MapQuest({layer: 'sat'})
                        source: new ol.source.BingMaps({
                        key: 'Ak-dzM4wZjSqTlzveKz5u0d4IQ4bRzVI309GxmkgSVr1ewS6iPSrOvOKhA-CJlm3',
                                                imagerySet: 'AerialWithLabels'
                                                })
                    }),
```

Add WMS layer

```
                    new ol.layer.Tile({
                     title: 'GLC2000',
                     source: new ol.source.TileWMS({
                    url: 'http://localhost:8080/geoserver/GOFC-GOLD/wms',
 //outputFormat: 'GML2/ GML3/ KML/ shape-zip/ json/ javascript/ csv'
                    params: {'LAYERS': 'GOFC-GOLD:GLC2000'},
```

```
                                        serverType: 'geoserver'
                                    })
```

Fix viewing resolution

```
        view: new ol.View({
                            center: ol.proj.transform([35.9, 7.46], 'EPSG:4326', 'EPSG:4326'),
            zoom: 2.5
        })
    });
});
```

HTML script: (this script is to visualize the datasets on the web)

HTML script is to visualize the datasets on the web.

```
<html>
  <head>
    <meta charset="utf-8" />
    <title>GOFC-GOLD Land Cover Project Office</title>
    <meta name="viewport" content="initial-scale=1.0, user-scalable=no, width=device-width">


<style type="text/css">
      .popover {
        z-index: auto;
      }
      .popover-content {
        min-width: 180px;
      }
    </style>
```

```
  </head>

<body>
        <div id="map"></div>
        <div style="display: none;">
     <div id="popup" title="Welcome to GOFC-GOLD"></div>

    <script src="lib/ol.js"></script>

    <script src="src/ol3-layerswitcher.js"></script>
    <script src="addlayer.js"></script>

        <script src="src/jquery.min.js"></script>
        <script src="src/bootstrap.min.js"></script>
  </body>
</html>
```

Appendix D
Questionnaire new GOFC-GOLD geoportal

Home Page of the existing and new GOFC-GOLD geoportal

Figure 1: Existing GOFC-GOLD geoportal (version 2014)

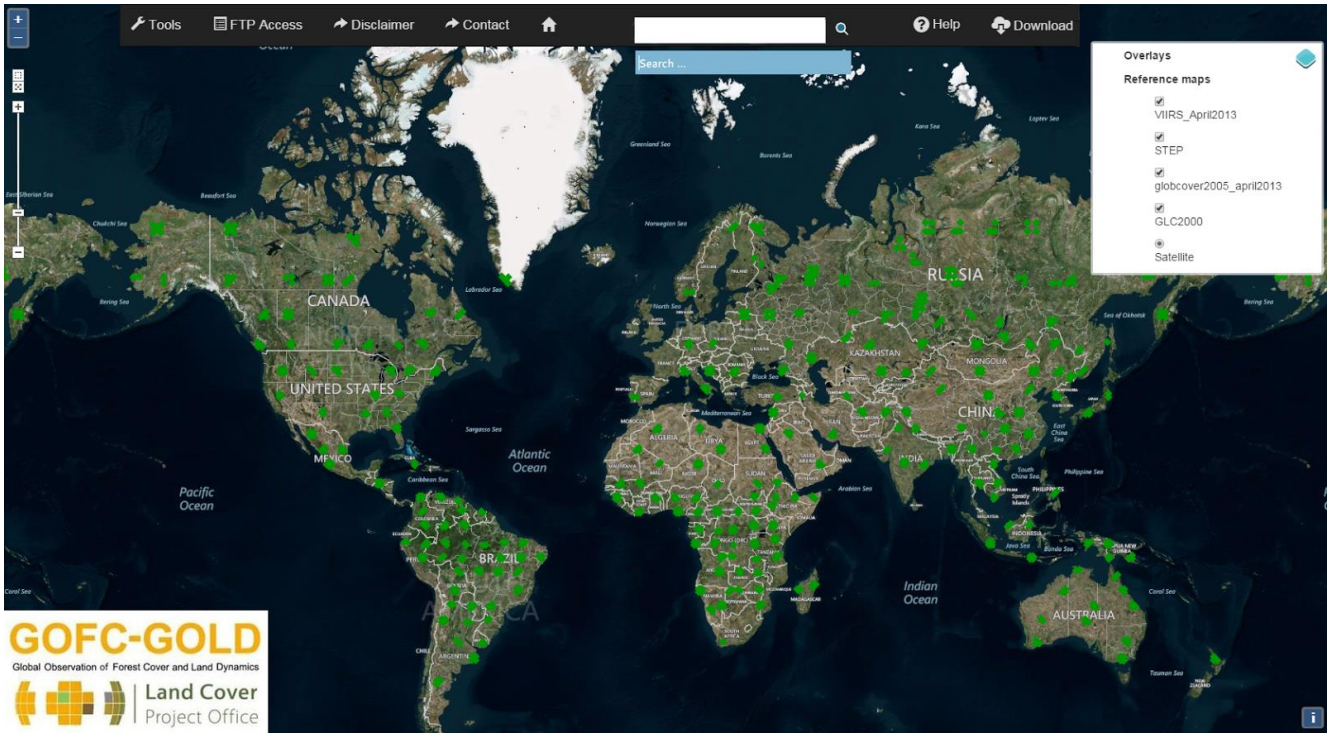


Figure 2: Proposed interface GOFC-GOLD geoportal

Questions for new geoportal

please look at the given image "outlook of the new GOFC-GOLD geoportal" and answer the following questions.

Pop-up Button

Navigation Bar

In this new portal is it easy for you to find reference maps?

- [ ] Yes
- [ ] No

Do you think it is improved enough than the previous GOFC-GOLD portal?

- [ ] Yes
- [ ] No

Does the user interface looks better now?

- [ ] Yes
- [ ] No

Do you find download button now?

- [ ] Yes
- [ ] No

Appendix E

Justification of my choice:

All the information I gathered in the following tables are from different websites of those software. For the rating system, I used my personal preferences. The one I found have most and best options like the server's ability to fulfil the user request or how fast the geoserver can fulfil client request and are of use for my geoportal design. I gave the maximum score of '+++' signs. In case I did expect limited and lower quality options I scored less '+' sign. My scores haven't been reviewed.

Best choice for database management system:

	PostgreSQL	**Rates**	**MySQL**	**Rates**
Feature	support subqueries, stored procedures, subqueries, cursors or views	++	Doesn't support subqueries, stored procedures, subqueries, cursors or views	-
Free GIS Data Loaders	included shp2pgsql, OGR2OGR, QuantumGIS SPIT, SHP loader for PostGIS also developed by Morten using SharpMap.NET various others	+++	OGR2OGR, shp2mysql.pl script	+
Flexibility	Flexible because lots of spatial functions and fairly efficient indexing and lots of open source and commercial support and upcoming ESRI ArcGIS 9.3 supports it too	+++	Limited spatial functions. Some commercial (MapDotNet, Manifold.net), Open source GIS tools gaining steam but still more behind PostGIS.	+
Web Mapping ToolKits	Manifold, MapDotNet, ArcGIS 9.3, UMN Mapserver, GeoServer, Feature Server, MapGuide Open Source (using beta FDO driver)	++	UMN Mapserver, GeoServer, MapGuide Open Source	+
Cost	Free (Open Source)	++	Free (Commercial)	+
Schemas	Yes	++	No	-
Free Desktop Viewers and Editors	OpenJump, QuantumGIS, GvSig, uDig	+++	GvSig	+

Table: Arguments to choose the database management system

Suitable Server Application to visualize maps online:

	GeoServer	**Rates**	**MapServer**	**Rates**
Language	Java based	+	C/C++ with a lot of PHP	++
Output	produces KML, GML, Shapefile, GeoRSS, PDF, GeoJSON, JPEG, GIF, SVG, PNG and more	+++	Produces WMS, WFS and WCS	+
Flexibility	Easy to connect virtual world	++	Hard to connect virtual world	-
Interoperability	publishes data from any major spatial data source using open standards	+++	Develop environment for building spatially-enabled web mapping applications and services	+
Cost	Free (Open Source)	++	Free (Open Source)	++
Server	Open source Server	+	Open Source application	-

Table: Arguments to choose the server for sharing geospatial data

Suitable Server to add user tools and visualize the maps:

	Apache tomcat	**Rates**	**IIS Windows Server**	**Rates**
Connection capacity	Standalone Server with cross platform	++	Single platform	+
Flexibility	Flexible. Can run on different operating system.	+++	Fixed. Can run only on windows.	-
Cost	Free	+	Free	+
Status	Web Server and a servlet container	++	Web server	+
Program	Plain JavaScript with HTML	+	--	

Table: Arguments to choose the server to visualize the geospatial data with retrievals option.

From the above discussion my choices for different software and programs are below:

Software/Programme/Server	Logic for my choice
PostgreSQL	Support several open source desktop viewers and editors. High storage capacity. Support several web mapping toolkits. Huge built-in query.
GeoServer	Open source server and easy connection with virtual world. A lot options for output format. Interoperability system is very high.
Apache tomcat	It is free of cost; connection capacity is very high and flexible for all operating system. Bonus it has servlet container

Table: Logic of my choice

Appendix F

The Database Library (Description of attributes):

GLC2000:

Table Data dictionary report - glc2k_agl11_rndsel_hendrik

Generated: 21-8-2015 15:19:32
Server: PostgreSQL 9.3 (x86) (localhost:5432)
Database: reference_db
Schema: GLC2000

Columns

Name	Data type	Not Null?	Primary key?	Default	Comment
gid	integer	Yes	Yes	nextval('"GLC2000".glc2k_agl11_rndsel_hendrik_gid_seq'::regclass)	
id	numeric(10)	No	No		
lat	numeric	No	No		
long	numeric	No	No		
class	numeric(10)	No	No		
geom	geometry(1107456)	No	No		

Constraints

Name	Type	Definition	Comment
glc2k_agl11_rndsel_hendrik_pkey	Primary key	(gid)	

Report generated by pgAdmin III

Table properties

Property	Value
Name	glc2k_agl11_rndsel_hendrik
OID	28145
Owner	postgres
Tablespace	pg_default
ACL	
Of type	
Primary key	gid
Rows (estimated)	877
Fill factor	
Rows (counted)	877
Inherits tables	No
Inherited tables count	0
Unlogged?	No
Has OIDs?	No
System table?	No
Comment	

```
-- Table: "GLC2000".glc2k_agl11_rndsel_hendrik

-- DROP TABLE "GLC2000".glc2k_agl11_rndsel_hendrik;

CREATE TABLE "GLC2000".glc2k_agl11_rndsel_hendrik
(
  gid serial NOT NULL,
  id numeric(10,0),
  lat numeric,
  "long" numeric,
  class numeric(10,0),
  geom geometry(Point,4326),
  CONSTRAINT glc2k_agl11_rndsel_hendrik_pkey PRIMARY KEY (gid)
)
WITH (
  OIDS=FALSE
);
ALTER TABLE "GLC2000".glc2k_agl11_rndsel_hendrik
  OWNER TO postgres;

-- Index: "GLC2000".glc2k_agl11_rndsel_hendrik_geom_gist

-- DROP INDEX "GLC2000".glc2k_agl11_rndsel_hendrik_geom_gist;

CREATE INDEX glc2k_agl11_rndsel_hendrik_geom_gist
  ON "GLC2000".glc2k_agl11_rndsel_hendrik
  USING gist
```

GlobCover2005:

Table Data dictionary report - globcover2005_april2013

Generated: 21-8-2015 15:26:33
Server: PostgreSQL 9.3 (x86) (localhost:5432)
Database: reference_db
Schema: GlobCover2005

Columns

Name	Data type	Not Null?	Primary key?	Default	Comment
gid	integer	Yes	Yes	nextval('"GlobCover2005".globcover2005_april2013_gid_seq'::regclass)	
id	numeric(10)	No	No		
x	numeric	No	No		
y	numeric	No	No		
reslc1	numeric(10)	No	No		
area	double precision	No	No		
confidence	numeric	No	No		
geom	geometry(1107476)	No	No		

Constraints

Name	Type	Definition	Comment
globcover2005_april2013_pkey	Primary key	(gid)	

Table properties

Property	Value
Name	globcover2005_april2013
OID	28248
Owner	postgres
Tablespace	pg_default
ACL	
Of type	
Primary key	gid
Rows (estimated)	186
Fill factor	
Rows (counted)	186
Inherits tables	No
Inherited tables count	0
Unlogged?	No
Has OIDs?	No
System table?	No
Comment	

```
-- Table: "GlobCover2005".globcover2005_april2013

-- DROP TABLE "GlobCover2005".globcover2005_april2013;

CREATE TABLE "GlobCover2005".globcover2005_april2013
(
  gid serial NOT NULL,
  id numeric(10,0),
  x numeric,
  y numeric,
  reslcl numeric(10,0),
  area double precision,
  confidence numeric,
  geom geometry(MultiPolygon,4326),
  CONSTRAINT globcover2005_april2013_pkey PRIMARY KEY (gid)
)
WITH (
  OIDS=FALSE
);
ALTER TABLE "GlobCover2005".globcover2005_april2013
  OWNER TO postgres;

-- Index: "GlobCover2005".globcover2005_april2013_geom_gist

-- DROP INDEX "GlobCover2005".globcover2005_april2013_geom_gist;

CREATE INDEX globcover2005_april2013_geom_gist
  ON "GlobCover2005".globcover2005_april2013
  USING gist
  (geom);
```

STEP:

Table Data dictionary report - step_091514

Generated: 21-8-2015 15:35:23
Server: PostgreSQL 9.3 (x86) (localhost:5432)
Database: reference_db
Schema: STEP

Columns

Name	Data type	Not Null?	Primary key?	Default	Comment
gid	integer	Yes	Yes	nextval('"STEP".step_091514_gid_seq'::regclass)	
area	numeric	No	No		
perimeter	numeric	No	No		
pi	character varying(3)	No	No		
rsref	double precision	No	No		
domht	integer	No	No		
monthmax	integer	No	No		
monthmin	integer	No	No		
x_coord	numeric	No	No		
y_coord	numeric	No	No		
siteid	numeric	No	No		
descript	character varying(255)	No	No		
igbp	integer	No	No		
cult	integer	No	No		
irig	integer	No	No		
wet	integer	No	No		
veg	integer	No	No		
cov	integer	No	No		
lf_type	double precision	No	No		
lf_phen	double precision	No	No		
nveg	double precision	No	No		
conf	double precision	No	No		
ag_type	double precision	No	No		
qual	double precision	No	No		
comments	character varying(255)	No	No		
date_added	character varying(29)	No	No		
struct	integer	No	No		
decid_type	integer	No	No		
sav_type	integer	No	No		
tcov	integer	No	No		
year_start	double precision	No	No		
year_end	double precision	No	No		
veg_cycles	numeric	No	No		
geom	geometry(1107476)	No	No		

Constraints

Name	Type	Definition	Comment
step_091514_pkey	Primary key	(gid)	

Table properties

Property	Value
Name	step_091514
OID	28652
Owner	postgres
Tablespace	pg_default
ACL	
Of type	
Primary key	gid
Rows (estimated)	2833
Fill factor	
Rows (counted)	not counted
Inherits tables	No
Inherited tables count	0
Unlogged?	No
Has OIDs?	No
System table?	No
Comment	

```
-- Table: "STEP".step_091514

-- DROP TABLE "STEP".step_091514;

CREATE TABLE "STEP".step_091514
(
  gid serial NOT NULL,
  area numeric,
  perimeter numeric,
  pi character varying(3),
  rsref double precision,
  domht integer,
  monthmax integer,
  monthmin integer,
  x_coord numeric,
  y_coord numeric,
  siteid numeric,
  descript character varying(255),
  igbp integer,
  cult integer,
  irig integer,
  wet integer,
  veg integer,
  cov integer,
  lf_type double precision,
  lf_phen double precision,
  nveg double precision,
  conf double precision,
  ag_type double precision,
  qual double precision,
  comments character varying(255),
  date_added character varying(29),
  struct integer,
  decid_type integer,
  sav_type integer,
  tcov integer,
  year_start double precision,
  year_end double precision,
  veg_cycles numeric,
  geom geometry(MultiPolygon,4326),
  CONSTRAINT step_091514_pkey PRIMARY KEY (gid)
)
WITH (
  OIDS=FALSE
);
ALTER TABLE "STEP".step_091514
  OWNER TO postgres;

-- Index: "STEP".step_091514_geom_gist

-- DROP INDEX "STEP".step_091514_geom_gist;

CREATE INDEX step_091514_geom_gist
  ON "STEP".step_091514
  USING gist
  (geom);
```

VIIRS:

Table Data dictionary report - viirs_april2013_rndsel70_igbpstrat

Generated: 21-8-2015 15:40:11
Server: PostgreSQL 9.3 (x86) (localhost:5432)
Database: reference_db
Schema: VIIRS

Columns

Name	Data type	Not Null?	Primary key?	Default	Comment
gid	integer	Yes	Yes	nextval('"VIIRS".viirs_april2013_rndsel70_igbpstrat_gid_seq'::regclass)	
name	character varying(254)	No	No		
sampleid	character varying(200)	No	No		
pixel_id	character varying(254)	No	No		
objectid	integer	No	No		
sample_id	numeric	No	No		
name_1	character varying(254)	No	No		
date	character varying(254)	No	No		
num_new	character varying(254)	No	No		
num_old	numeric	No	No		
viirs_row	numeric	No	No		
viirs_col	numeric	No	No		
pixel_id_1	character varying(254)	No	No		
igbp	numeric	No	No		
conf	numeric	No	No		
igbp2	numeric	No	No		
change	character varying(254)	No	No		
notes	character varying(254)	No	No		
biome_code	numeric	No	No		
biome_name	character varying(254)	No	No		
geom	geometry(1107476)	No	No		

Constraints

Name	Type	Definition	Comment
viirs_april2013_rndsel70_igbpstrat_pkey	Primary key	(gid)	

Table properties

Property	Value
Name	viirs_april2013_rndsel70_igbpstrat
OID	28545
Owner	postgres
Tablespace	pg_default
ACL	
Of type	
Primary key	gid
Rows (estimated)	3667
Fill factor	
Rows (counted)	not counted
Inherits tables	No
Inherited tables count	0
Unlogged?	No
Has OIDs?	No
System table?	No
Comment	

```
-- Table: "VIIRS".viirs_april2013_rndsel70_igbpstrat

-- DROP TABLE "VIIRS".viirs_april2013_rndsel70_igbpstrat;

CREATE TABLE "VIIRS".viirs_april2013_rndsel70_igbpstrat
(
  gid serial NOT NULL,
  name character varying(254),
  sampleid character varying(200),
  pixel_id character varying(254),
  objectid integer,
  sample_id numeric,
  name_1 character varying(254),
  date character varying(254),
  num_new character varying(254),
  num_old numeric,
  viirs_row numeric,
  viirs_col numeric,
  pixel_id_1 character varying(254),
  igbp numeric,
  conf numeric,
  igbp2 numeric,
  change character varying(254),
  notes character varying(254),
  biome_code numeric,
  biome_name character varying(254),
  geom geometry(MultiPolygon,4326),
  CONSTRAINT viirs_april2013_rndsel70_igbpstrat_pkey PRIMARY KEY (gid)
)
WITH (
  OIDS=FALSE
);
ALTER TABLE "VIIRS".viirs_april2013_rndsel70_igbpstrat
  OWNER TO postgres;

-- Index: "VIIRS".viirs_april2013_rndsel70_igbpstrat_geom_gist

-- DROP INDEX "VIIRS".viirs_april2013_rndsel70_igbpstrat_geom_gist;

CREATE INDEX viirs_april2013_rndsel70_igbpstrat_geom_gist
  ON "VIIRS".viirs_april2013_rndsel70_igbpstrat
  USING gist
  (geom);
```