METHODOLOGY FOR ASSESSING THE USABILITY OF EARTH OBSERVATION BASED DATA FOR DISASTER MANAGEMENT

SWETA LEONARD
February, 2011

SUPERVISORS:
Dr. Ir. W. Bijker
Dr. C.P.J.M. van Elzakker
Mr. A. Kidd, MSc.

METHODOLOGY FOR ASSESSING THE USABILITY OF EARTH OBSERVATION BASED DATA FOR DISASTER MANAGEMENT

SWETA LEONARD
Enschede, The Netherlands, February, 2011

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geo informatics

SUPERVISORS:
Dr. Ir. W. Bijker
Dr., C.P.J.M. van Elzakker
Mr. A. Kidd, MSc.

THESIS ASSESSMENT BOARD:
Prof. Dr. Ir. Alfred Stein (Chair)
Ms. Dr. Ir. S. Zlatanova (External Examiner)
Dr. Ir. W. Bijker
Dr. C.P.J.M. van Elzakker
Mr. A. Kidd, MSc

ABSTRACT

The purpose of this study was to develop a methodology for assessing the usability of Earth Observation data for different phases of disaster management. Another aim was to identify the user groups in disaster management and finally to establish the purpose of the products. Case study application of the methodology on sample datasets produced in response to the Haiti 2010 earthquake was also examined in the study.

Survey and literature review were used to identify the necessary inputs of the methodology. Relevant literature review was carried out to identify user groups and supported by information gathered from survey. Six different data types were explored to establish the purpose of datasets. The survey was done on two user and producer organizations who participated in the response to 2010 Haiti earthquake. The information gathered was used to group related parameters forming major inputs and information containers were created for user requirements, data specifications, comparison platform and ranking. These major inputs were logically aggregated to form the first phase of the methodology to be implemented by the expert and the second part to be carried out by the end-user.

The results were that Earth Observation data are made for direct use when fully processed without possibility of manipulation and indirect use when semi processed with possibility of manipulation. The user groups were identified into major categories of planners, responders and researchers. The application of the methodology resulted into common framework of comparing data specifications against user requirements known as 'quality information template'.

The principal conclusion was that usability of disaster data can be improved by expert evaluation and rating and making the report from experts available to the user through a common interactive platform where they can chose potential datasets and subject them to ranking.

Keywords: Data quality, quality information template, usability, disasters, ranking

ACKNOWLEDGEMENTS

I am grateful to the admission board of the department of Geo-informatics for giving me an opportunity to study at the faculty of geo information and Earth Observation. I appreciate the assistance I received from various lecturers and students during my entire study period at the faculty. I deeply appreciate the support I received from my supervisors Dr. Corn'e van Elzakker, Dr. Wietske Bijker and Mr Richard Kidd during my entire research period.

I am very grateful to my sponsor NUFFIC for the financial support I received throughout my entire study period. I am grateful to the administration of the faculty of Geo information and Earth Observation for providing conducive environment for learning.

Special thanks to members of my family Evelyn, Fischer and Pearson for their unconditional support during my study period.

TABLE OF CONTENTS

1. Quality problems of Earth Observation data in disaster management....1
1.1. Introduction....1
1.2. Motivation and problem statement....2
1.3. Research objectives....2
1.4. Research questions addressed....3
1.5. Innovation of this research....3
1.6. Thesis structure....3
2. Spatial data quality evaluation and production of Earth Observation data for disaster management...5
2.1. Chapter overview....5
2.2. Production of Earth Observation data/products for disaster management....5
2.3. Spatial data quality evaluation....8
2.3.1. Overview of spatial data quality elements....8
2.3.2. Spatial data quality reporting and data transfer....9
2.3.3. Evaluation of external data quality and the concept of ranking....10
2.4. Importance of spatial data quality information to the user of Earth Observation products for disaster management....11
2.4.1. User perspective of data quality and task requirements....12
2.4.2. Spatial data usability....13
2.5. Chapter summary....14
3. Research design and methods....15
3.1. Chater overview....15
3.2. The research design....15
3.2.1. The steps followed in detail....16
3.3. Framework of the proposed methodology....17
3.4. Major inputs of the research design....18
3.5. Data collection....18
3.5.1. Purpose of the Earth Observation datasets....18
3.5.2. Identification of user groups and user requirements....18
3.6. Chapter summary....19
4. result: A methodology for assessing usability of Earth Obsevation data....21
4.1. Chapter overview....21
4.2. User groups and user requirements....21
4.2.1. Planners....21
4.2.2. Responders....21
4.2.3. Researchers....22
4.3. Purpose of EO products....22
4.3.1. Products for direct use....23
4.3.2. Products for indirect use....23
4.4. "The methodology" for assessing usablity of EO data and derived products for disaster management. 24
4.4.1. Application criteria of the methodology....24
4.4.2. Role of users in the different phases of the methodology....28
4.4.3. Purpose of quality information template (QIT) in the developed "the methodology"....29
4.4.4. Importance of QIT to the end-user of EO derived product....29
4.5. Link between data quality description to application of the data/derived products....30
4.6. Chapter summary....30

5. Case Study: The Haiti earthquake 2010 ... 31
5.1. Chapter overview ... 31
5.2. Application of the developed methdology on derived EO data ... 31
5.3. Study area ... 31
5.4. QIT of the EO derived products ... 33
5.4.1. QIT for the individual data quality elements of the derived products ... 33
5.4.2. QIT for the overall quality of the derived products ... 33
5.5. Population of QIT with information of derived products ... 35
5.5.1. Creation of user interface on ZOHO web creator ... 36
5.5.2. Developing interactive QIT and quality information view in ZOHO web creator ... 36
5.6. Description of ranking scenario ... 40
5.6.1. Description of ranking scenario through Analytic Hierarchy Process (AHP) ... 40
5.6.2. Prioritizing end user requirements ... 41
5.6.3. Setting of preferences, calculation of weights and ranking of EO derived products ... 42
5.7. Chapter summary ... 46
6. discussion, Conclusions and recomendations ... 47
6.1. Chapter overview ... 47
6.2. Discussions ... 47
6.3. Conclusion ... 50
6.4. Recommendation ... 50

LIST OF FIGURES

Figure 1: Production of EO data for disaster response modified after (Kerle & Widartono, 2008) 6
Figure 2: Disaster and emergency cycle (URL2.1) 7
Figure 3: ISO 19113 Model of data quality 8
Figure 4: Information communicated to users for assessing quality (Devillers et al., 2005) 9
Figure 5: Quality evaluation model according to ISO 19114 11
Figure 6: User decision contexts and data quality (Byrom, 2003).. 12
Figure 7: Approaches to geographic information quality (Jakobsson, 2006)...................................... 14
Figure 8: Research design flow chart.. 16
Figure 9: Three dimension perception of the right dataset for disaster response.. 17
Figure 10: Thumbnail of some EO based derived products explored (Haiti earthquake of January 12th 2010 response datasets) 22
Figure 11: The methodology for assessing usability of EO based derived products 24
Figure 12: Types of faults within the Haiti region (www.internetgeography.net).. 32
Figure 13: Haiti Earthquake city exposure (USGS) 32
Figure 14: Satellite image damage assessment_ product A.. 34
Figure 15: User interface for accessing rated products and submission of expert evaluation 36
Figure 16: Rating submission form for experts 37
Figure 17: Screenshot of exploring quality of rated product.. 38
Figure 18: Relevance of QIT in reducing information overload 39
Figure 19: possible thumb nail of evaluated information 39
Figure 20: The AHP hierarchy model.. 41
Figure 21: Possible AHP model indicating relationships between data types and spatial data quality elements... 41

LIST OF TABLES

Table 1: List of originations contacted for survey 18
Table 2: Quality information level template for data quality elements 25
Table 3: Quality information level template for data quality elements Cartographic and metadata 26
Table 4: Rating and ranking table 28
Table 5: Sample datasets considered for demonstration....33
Table 6: Application of QIT on usability evaluation of graphic product....34
Table 7: Scale of relative importance (Saaty & Kearns, 1985) 40
Table 8: Pairwise comparison of data type and data quality specification 42
Table 9: Priority vector of data type and data quality....43
Table 10: Relative importance and priority vectors for data type 43
Table 11: Relative importance and priority vectors of data quality specification 44
Table 12: Ranking vectors for data quality specifications and data type 44
Table 13: Illustration of overall rank of all attributes 45

1. QUALITY PROBLEMS OF EARTH OBSERVATION DATA IN DISASTER MANAGEMENT

1.1. Introduction

Disasters[1] happen at certain location on the earth surface; which can be local like a burning factory; extensive like flooding caused by tsunami or regional like damage caused by earthquake. The year 2010 witnessed a number of disasters such as floods in Pakistan (July) and Eastern Europe (May); earthquakes in Chile (February), China (April), and Haiti (January), etc. Management of such disasters require Earth Observation (EO) data, geographic information science (GIS) technologies and related services to provide information before, during and after the disaster has occurred (Pundt, 2008).

The most significant challenge in disaster management is delivering the right information to right group to users at the appropriate time in useful format (Radke et al., 2002). As the development in handling EO data increases, the quality of information delivered to the users is vital (Lorincz et al., 2004) especially with the use of wireless technologies like sensor networks. The quality of EO data fused with geospatial information is affected by lack of reliable and up to date data on reference datasets such as roads, critical infrastructure, place names, topographic information, drainage which are very vital for generating meaningful analysis during disaster management. (Voigt et al., 2007).

Lack of standardized symbology on emergency maps and associated datasets at national and international level (Dymon, 2001) hinders decision making and coordination between parties involved during crucial stages of disaster management (Dymon, 2003). Which means the resulting EO data products exhibit variant quality attributes in their product specification or digital libraries (Mountrakis et al., 2004). The user is then exposed to several products with various thematic representations on the same phenomenon which lacks responsibility and reliability, including those from Disaster Chatter (Ito, 2005).

Analysing quality differences of spatial data is a challenge to both experts and novice users (Boin & Hunter, 2009). The process can be complex when there are many EO data in heterogeneous formats (Abdalla et al., 2007). This is exacerbated by lack of mechanisms in place for making data readily available to all levels of decision making during disaster management (Stevens, 2008). The problem becomes compounded due to complexity in quality information and access regulations including EO data from Disaster Charter (Ito, 2005).

Handling quality problems EO data in disaster management require value adding chain that is compliant with user requirements (Devillers et al., 2007). The user should then be able to visualize quality indicators at different hierarchical levels (Ying & Lei, 2008). Since disaster response has no universal products (Kerle & Widartono, 2008); a rating of EO products on their quality information indicating level of conformance to quality specifications can be useful in setting preferences on potential datasets and establishing their usability.

[1] Disaster is "serious disruption of the functioning of a community or society involving widespread human, material, economic, or environmental losses and impacts, which exceeds the capacity of the affected community or society to cope using its own resources"(UNISDR, 2009).

1.2. Motivation and problem statement

Disaster management is carried out in phases, each with its own demands which are not explicit but may overlap depending on user requirements (URL2.1). The magnitude of a disaster determines the type of organizations who will be involved in the management scenario. Large scale disasters may attract attention from bilateral agencies, the government of the host country, several national and international non-governmental organisations (NGOs) (Radke et al., 2002).

Reporting spatial data quality information to many users in is a challenge because quality is reported as if it was meant for a single usage (Grira et al., 2009) which can be misleading due to different quality perceptions (Brabyn, 1996) by different users especially where many actors are involved like management. The end user of EO data in disaster management have variant characteristics (Fei et al., 2010)related to skills, use context, experience, area of operations, personal opinions, etc. Therefore, addressing usability of EO data requires a flexible approach.

Traditionally, quality of cartographic products were passed unidirectional because maps were made with targeted (specific) user in mind (Hunter et al., 2009). However, currently maps and other EO related products are no longer limited to craftsmen due to technological developments which has widened the user and producer communities through crowdsourcing (Goodchild, 2007; Grira et al., 2009).

Established national and international standards such ISO, OGC, among others; currently dominate data quality and metadata reporting (Joos, 2006). The expected belief is that standards and technology will make data integration very trivial. However, disaster products such maps and other thematic products lack established standard for their production (Fei et al., 2010). Therefore, conflicting interpretations may be drawn from various datasets due to lack of reliable quality information (Devillers & Jeansoulin, 2006) to the end-users.

The global community dedicated to disaster response has made a milestone in making EO data available to most users through mutual understanding and charter agreements (Ito, 2005; Voigt et al., 2007). However, data accessibility is still affected by semantic barriers (Bishr, 1998; Timpf & Raubal, 1996) and bureaucracy even during emergencies (Ito, 2005; Van Oosterom et al., 2005). Incompleteness and lack of inter-operability (Cai & Xie, 2007) creates the need to review the quality of certain products to establish what they could be good for (Nedovic-Budic & Pinto, 1999) to improve the end-user trust in using them.

Making reliable decision during disaster management will depend on the quality of data and representation of that quality information to the users (Su et al., 2009). Organised information showing various categories of quality levels is vital for effective use of EO data for disaster management.

1.3. Research objectives

The main objective of this research is to develop a methodology for ranking quality of different derived image based products for different users in disaster response. The methodology is designed to be used by both experts and end-users in disaster management. The experts rate quality of EO data and make the rating information available to the targeted end-users to set preferences for weighting various datasets. This study therefore, attempts to link product quality description to possible application of the EO data. To achieve the main objective, the following specific objectives are addressed:

i. To identify user groups
ii. To identify the purpose of the products
iii. Link the data quality description to possible application of the data

1.4. Research questions addressed

In order to achieve the above objectives, the following research questions are addressed:

i. What are the categories of user groups?
ii. For what purpose are the derived products made and what level of detail?
iii. What is the link between data quality description to possible application of the data?

1.5. Innovation of this research

This study presents a methodology for ranking quality of different derived image based products for different users in disaster response. The methodology presented here is based on deductive principles where the user is able to select the right product to use based on quality information presented with the product by ranking quality on the basis of theme of interest. This methodology can be applied by end-users in various disasters like flooding, tsunamis, earthquake, among others; whose response generate a lot of products under relief coordinating agencies.

The ratings of conformance are done by the experts or team of experts and then display the rating of the product as thumbnail with original product. The rating harmonizes quality differences, allows the user to set preference and provides a structured report on quality specification of EO data. The information can be displayed on online network where the end-users assessing such information can assign weights and rank products according to their decision preferences.

1.6. Thesis structure

This thesis report is organized into six chapters. Chapter one introduces the study with objectives and research questions being addressed. Information about the study scope and innovations of the study are presented in this chapter as well. The second chapter presents discussion on related work focusing on production of EO data for disaster management, spatial data quality evaluation with overview of spatial data quality elements are discussed as well as data quality reporting. Further discussion is also provided on importance of data quality information to the end user and the chapter end with overview of approaches in usability of spatial data.

Chapter three presents research design and methods used to identify major inputs of the proposed methodology and description of the procedure followed in the development of the methodology. Results are presented in chapter four with chapter five providing information on description of case study application of the methodology developed on sample of EO derived that became available after the Haiti earthquake of January, 2010. Discussion on the outcome of the research objectives are presented on chapter six with conclusion drawn from the study alongside the recommendation.

2. SPATIAL DATA QUALITY EVALUATION AND PRODUCTION OF EARTH OBSERVATION DATA FOR DISASTER MANAGEMENT

2.1. Chapter overview

This chapter presents information on how Earth Observation (EO) products are produced for disaster and crisis management. Descriptions of how spatial data quality is evaluated are also presented to provide information on current practices of quality assessment. The chapter is ends with description that integrates the importance of spatial data quality information to the use and user in EO in disaster management.

2.2. Production of Earth Observation data/products for disaster management

EO and space technologies are key information resources in pre and post disaster management scenarios earthquakes, floods, landslides, tsunamis, forest fires, other natural and technological disasters (Pundt, 2008; Stevens, 2008; Su et al., 2009). In situ measurements take a long time to acquire, therefore, after an occurrence of a major disaster, assistance can be sought from International Charter Space and Major Disaster after eligibility check and charter activation by authorised member (Ito, 2005; Voigt et al., 2007).

Difficulties in data accessibility during emergency for disaster management were the main reason behind the establishment of an international disaster charter during the UNISPACE III conference of 1999. The charter became fully operational by November 2001 with the objective of providing a unified system of space data acquisition and delivery to those affected. The provision of data results from activation by authorised users (Bessisa, 2004). For example, in response to Haiti earthquake of 2010, the charter was activated by UNOOSA[2] on behalf of MINUSTAH[3], public safety of Canada, American earthquake hazards program of USGS[4]. The response program was managed by CNES[5] in collaboration with SAFER project within the framework of the GMES[6] initiative (URL2.2).

Disaster management and response generate plenty of products that are used in various phases of disaster management (Amin & Goldstein, 2008): risk reduction and assessment; early warning and preparedness; immediate response and reconstruction recovery period. The type of disaster influences the kind of space infrastructure to be activated as illustrated in Figure 1.

[2] United Nations Office for Outer Space Affairs
[3] United Nations Stabilization Mission in Haiti
[4] United States Geological Survey
[5] Centre National de'Etudes Spatiales: French government space agency
[6] Global Monitoring for Environment and Security

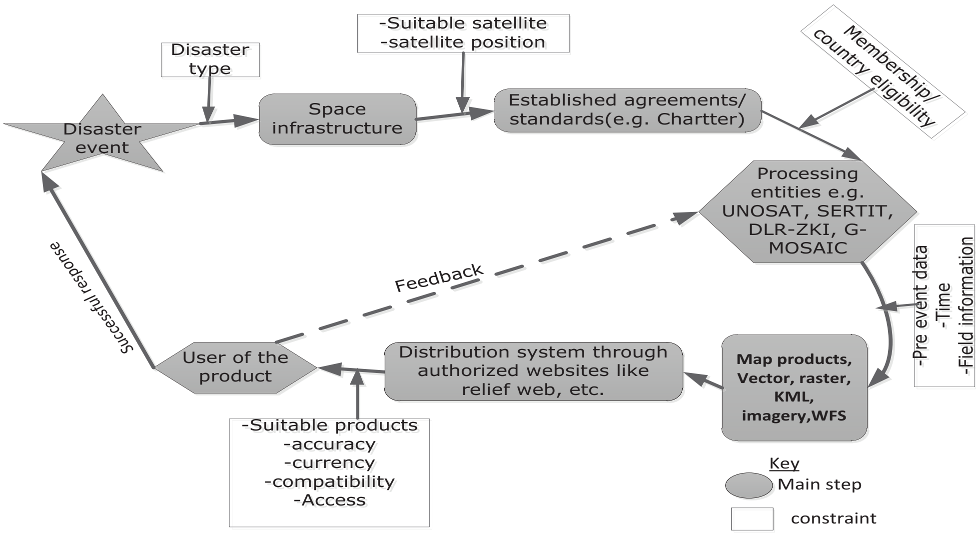


Figure 1: Production of EO data for disaster response modified after (Kerle & Widartono, 2008)

Production of EO data for disaster response involves a number of agencies in the value adding process as illustrated in Figure 1. The main steps in the production of EO data and related products for disaster management are influenced by constraints such as suitability of the satellite position, country eligibility, availability of pre/post disaster /field data and f*eedback* on the use of the product to the processing entities for successful response. Therefore, depending on the supplier, the derived EO product that is available to the end-user will be of variable formats, quality information, accessibility conditions, and at different levels of detail (Dymon, 2001, 2003; Ito, 2005).

Distribution of the disaster response products is currently done through internet on the websites of agencies or organizations dedicated to disaster management. Sometimes in large scale disasters several other volunteer agencies also host disaster management products especially for the phase of response. For instance, Haiti earthquake of 2010 received data from dedicated agencies such as Relief Web, MINUSTAH GIS, CNES, SAFER, DLR, SERTIT as well as volunteer agencies like http://Wiki.Openstreetmap.org, http://www.USHAHIDI.com/, among others. However, these products are distributed without information on the phase of disaster management for which they could be relevant.

The demands of the cycle are not distinct for all phases of disaster management with requirements overlapping before and after the occurrence of the disaster as illustrated in Figure 2.

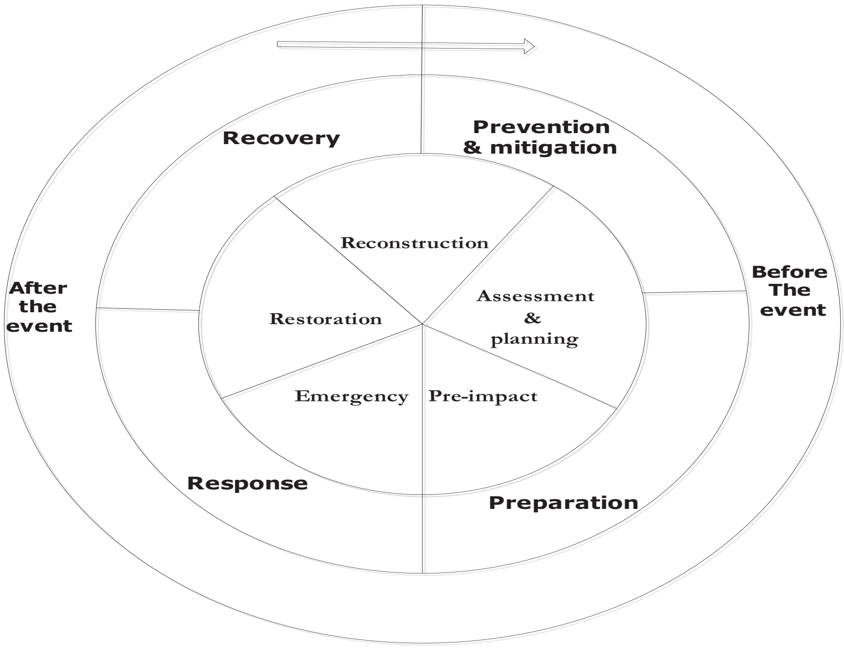


Figure 2: Disaster and emergency cycle (URL2.1)

Production of EO data for disaster management should be organised to meet the demands of the emergency cycle before and after the event as shown in Figure 2. The production of EO data for these phases can be organised to reflect the critical needs of the phase whose product are being produced (Fei et al., 2010; Ying et al., 2009). Since producers cannot predict the needs of all users, the demand of information on all the four phases of disaster and emergency cycle becomes a challenge in the production of response products (Fei et al., 2010; Radke et al., 2002). A summary of the four phases presents unique demands (Amin & Goldstein, 2008; Köhler, 2005)

i. *Prevention and mitigation:* this phase is also known as disaster risk reduction, where the aim is reducing vulnerability within broad context of sustainable development. EO products may provide information about areas that are/may be at risk. Thematic information can be obtained about land cover, drainage, transport, population, physiography and vulnerability.

ii. *Early warning and preparation:* information is provided from well-known institutions in timely manner for effective response preparation. Preparedness is done by activities aimed at giving early warning to those in threatened locations.

iii. *Response:* provision of relief assistance to those affected during/immediately after disaster has occurred for a short or protracted duration. During this phase, critical needs for users are damage evaluation information, location and magnitude of event, possible likelihood of induced events like tsunamis, landslides, flooding; extent of fault rupture and timeframe of aftershock sequence. The EO products of this phase are on high demand and are delivered on rush mode highlighting key affected areas on population, topography, environment and nature of crisis event. The information on these products may include among others potential camping areas, food distribution points, demography, infrastructure, damage extent, damage assessment, etc.

iv. *Recovery* phase is where decisions and actions taken with view of restoring pre-disaster living conditions of the affected communities. The duration of activities in this phase tends to be long and users do have time to plan their actions and projects. EO data can be used to provide detailed damage analysis, reconstruction monitoring, humanitarian situation assessment, trend analysis to generate information that is vital in analyzing the aftermath of the disastrous event.

2.3. Spatial data quality evaluation

Data quality is the ability of the data to meet user requirements (Orr, 1998). Therefore, quality is fitness for use (Chrisman, 1984; Morrison & Veregin, 2010) satisfying both the needs of the customer and the society (Jakobsson, 2006). The purpose of quality evaluation is to determine if the product quality description contains information helping the users to determine if the data is fit for their purpose (Devillers et al., 2007); by examining the relevance of information provided with respect to real or hypothetical world (Joos, 2006) as illustrated in Figure 3. This is because producers' perspective of product specification on a given phenomenon differs from the users' requirements of the same phenomenon.

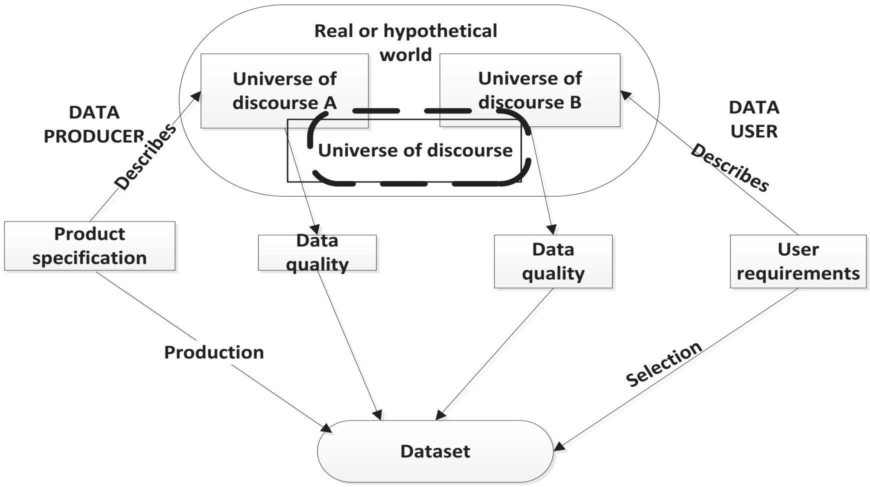


Figure 3: ISO 19113 Model of data quality

From Figure 3, it is demonstrated that producers and users of data use different universe of discourses of the real or hypothetical world. The discourses are defined by real or hypothetical world which makes data quality to be the difference between dataset and universe of discourse. Therefore, product specification developed with prior knowledge of user requirements is a key aspect of the model of data quality in the ISO standard (Joos, 2006). In disaster management, the products' specifications and user requirements should be brought to a common platform, either based on overall quality or on individual quality elements to allow harmonised quality evaluation.

2.3.1. Overview of spatial data quality elements

Completeness: Quality parameter that describes the absence/presence of features, their attributes and their relationships within a given dataset. It is either expressed as commission, depicting presence of excess data, or omission depicting absence of data in a dataset. The user may observe completeness as feature (minimum size of mapped object), attribute or value completeness (Yang, 2007). The parameter can be expressed as Boolean, ratio, integer or percentage depending on the method of evaluation employed by the user.

Logical consistency: Is the quality parameter that allows the user to establish the degree to which the dataset obeys the logical rules of data structures, attributes and relationships. The main forms are:

i. Conceptual consistency: concerned with adherence with rules of the schema
ii. Domain consistency: adherence of values to value domains i.e. domain range.
iii. Format consistency: the degree to which the stored information conforms to physical structure of the dataset.
iv. Topological consistency: correctness of the explicitly encoded topological characteristics.

Positional accuracy: is quality parameter that describes the correct position of features which can be absolute showing closeness to true value or value accepted as being true. Relative accuracy will also show the closeness of relative positions of features to relative positions accepted as being true or gridded data position accuracy showing closeness of gridded positions to true value.

Temporal accuracy: Describes the accuracy of temporal attributes and relationships of features. It can be grouped into accuracy of time measurements showing correctness of temporal references; temporal consistency showing correctness of ordered events which are reported; and temporal validity providing an overview of the dataset with respect to time.

Thematic accuracy: Describes the accuracy of quantitative attributes and the correctness of non-quantitative attributes. It involves classification correctness of classes assigned to features and their attributes.

2.3.2. Spatial data quality reporting and data transfer

In traditional cartography, data was produced with a specific user in mind (Andrienko & Andrienko, 1999) where professional craftsman stated the limits of the quality information transferred with the data to the user. However, currently production of maps and other EO data is no longer limited to professionally trained geo analysts (Morrison & Veregin, 2010) due to current development in crowd sourcing of geo-data or volunteered geographic information (Goodchild, 2007; Grira et al., 2009). It is possible to integrate EO data into various processing platforms and disseminate over the internet. The quality of such products is of course variable depending on the skill of the producer and the tools used.

The data quality in formation reported with the product is expected to allow the end-user to determine fitness for use (URL2.5, 2009). The description of quality information should provide a data quality overview of elements with information to a particular application and a full description indicating how the product specification was evaluated. Quality is therefore linked to the goal to be achieved by the user (Josselin, 2003) whose context changes depending on the nature of the task. Therefore, the level of quality information transferred to the user may range from a dataset of no quality information to those of highly detailed information as illustrated in Figure 4. The quality information of EO data delivered for disaster management will fall within the range described in Figure 4 as well, because quality changes with time.

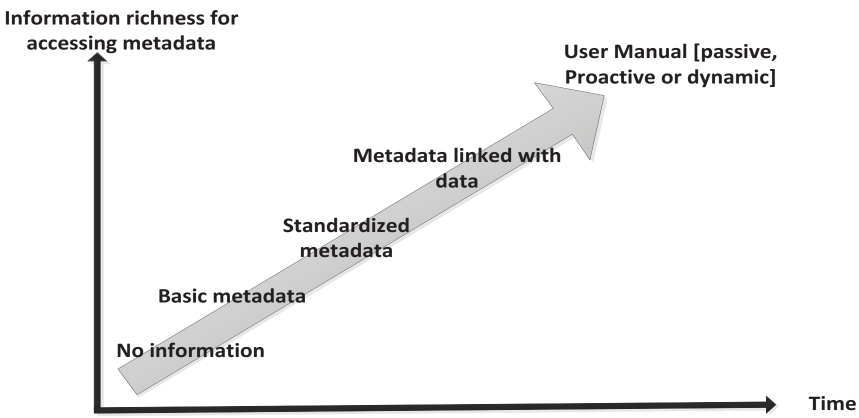


Figure 4: Information communicated to users for assessing quality (Devillers et al., 2005)

From Figure 4, the quality information passed to the user and the level of detail will be influenced by information richness with respect to time. For disaster management, information richness will depend on the time taken to produce the product. Generally, quality may be reported as:

i. Data could be distributed without the quality information at all, which can limit use extent.

ii. Some products contain basic quality information which does not comply to any established standards, with varying degree of information richness from supplier to supplier. Normally, such information is not reported in geospatial terms.

iii. Normalized metadata whose quality report follows the protocols of internal standards which tend to provide homogenous information, normally producer centered.

iv. Quality reported with linked metadata, which allow manipulation and users can manipulate the attributes with tools like Intergraph, ArcGIS, etc.

v. Dataset with quality information in the user manual as passive, dynamic or proactive which may complicate the usability of metadata

Digital spatial data on the web contains quality information reported as metadata, visual cues, legend or at times no quality information (Lowell & Jaton, 1999). The peculiarities of the geospatial and EO data creates difficulty in reporting quality information according to established standards (Huang & Wang, 2007). Implementation of standards in spatial data quality reporting is a challenge (Frank, 1998) especially when handling data from heterogeneous sources (Hutchison, 2010). Disaster management involves data from heterogeneous sources which makes it impractical for the user to assimilate all the quality information on the data (Lowell & Casault, 2004) from all the sources, like those found in context of disaster management. Furthermore, features in the real world are difficult to represent in compatible description due to different ways of representation (Mountrakis et al., 2004). Therefore, spatial data quality is task related (Wadsworth et al., 2009).

The attribute information attached to the features represented by the data are of primary importance while keeping in mind the purpose of the data (Grönlund, 2005). Hence, the user needs to examine the documentation of quality information or reported data quality description to understand the limitations, uncertainty magnitude and direction (Comber et al., 2006) in order to understand the details embedded in the data.

2.3.3. Evaluation of external data quality and the concept of ranking

Quality evaluation should initially be done by producers in line with established standards to provide information on the level of conformance of the product to data quality elements (Comber et al., 2006; Morrison & Veregin, 2010) to avoid other issues like legal consequences (Devillers et al., 2005; Van Oort & Bregt, 2005) and to determine the suitability of the dataset for a given task (Devillers et al., 2007).

Evaluation of external quality is related to determining 'fitness for use' which can be done through trial and error, consulting experts, legality of the dataset, purpose (use) of the data as well as through risk analysis. Quality communicated to the user is still facing a number of barriers (Hunter et al., 2009) due to varied user backgrounds. Datasets with varied aspects of reality in their thematic description should allow the user to establish the effect of themes on the quality of decision making (Frank et al., 2004). Evaluation in terms of themes is relevant to disaster management scenarios where products are also delivered under themes such as transport and logistics, damage assessment, camp management, etc.

The indicators defined for evaluating external quality should provide the user with information that is beyond what is available in the metadata (product) description (Devillers et al., 2002). The indicators could be intrinsic, contextual, representational or based on accessibility and access security (Wang & Strong, 1996). Detailed application of the indicators in external quality evaluation may take the following approaches:

i. Intrinsic nature of the data based on believability, accuracy, objectivity and reputation

ii. Contextual data quality which considers value added relevancy, timeliness, completeness, and appropriate amount of data.

iii. Representation of data quality covering interpretability, ease of understanding, representation consistency, and concise representation.

iv. Accessibility of data quality and access security.

Quality evaluation procedures will depend on the variables represented and the indicators defined by users. Since each group of users has different sets of requirements, conformance level of quality information to task requirements will be different as illustrated by ISO quality evaluation model on Figure 5.

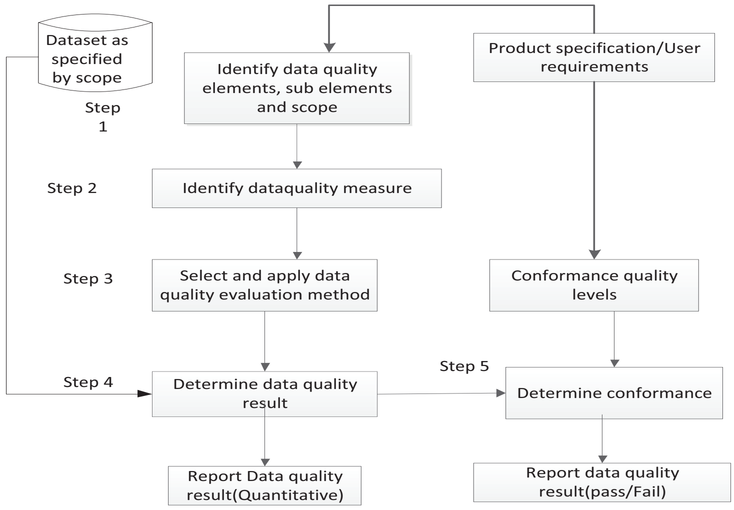


Figure 5: Quality evaluation model according to ISO 19114

Figure 5 indicates the importance of identifying data quality elements and sub elements from product descriptions before the method of quality evaluation is selected. The data quality evaluation method selected will provide either quantitative or qualitative result to be used in establishing the level of conformance of the quality information to user requirements.

Ranking through weighted approach allows the user to identify the most relevant dataset for a given task. Information on spatial rank especially setting of quality preferences is well developed in databases (Yiu et al., 2010); a concept that can be extended to aid decision making in disaster management. Quality information of various products can be visualised at different levels of details such as data formats, semantics or pragmatic aspects of the dataset (Su et al., 2009).

Quality indicators allows the users to visualise information about the dataset in hierarchical level which may assist the user to avoid potential risks associated with the dataset (Devillers et al., 2005; Devillers et al., 2002). Evaluating quality information of EO disaster data in disaster management is challenging due to heterogeneity of datasets, temporal and semantic differences, and lack of ready and accessible quality information.

2.4. Importance of spatial data quality information to the user of Earth Observation products for disaster management

Stages of disaster management require reliable information on quality of datasets that are used for decision making. Since data is collected for a purpose, task centred approach is relevant in disaster management (Erharuyi & Fairbairn, 2005) to allow the user to tailor their tasks based on the product specifications.

Use of acceptable standards in quality descriptions of datasets reduces uncertainty in interpretation of data quality (Lowell & Jaton, 1999). The characteristics EO data can be used to determine the stage of response and task to be performed (Mulí ková et al., 2010). Disaster management comprise of relevant data and information; methods and procedures; tools and technologies; as well as availability of suitable organizational environment (Köhler, 2005) which makes quality information to be very important to the user in handling such diverse pieces of information.

Quality of data depends on the perceived attributes of the product recognizable to the user (Juran et al., 1998) for the task. Therefore, end-user dimension of data quality is dynamic depending on the use at hand (Fisher et al., 2008). Geographical data is unique because the same phenomenon can be modelled differently by different producers (Parker & Stileman, 2005); perception of such quality information by different users is definitely different especially in disaster management where users may not have access to the primary data used in production of the dataset for cross checking.

Objects in real world have infinite classifications and infinite attributes for each classification (Frank, 2001). This influences the way in which spatial data quality is presented to the user and eventually the use of that dataset (Boin & Hunter, 2009). Therefore, a product that is usable becomes acceptable by the users *viz:* dataset information links data quality element to user requirements to determine fitness for use (Jahn & Frank, 2004).

After disaster has occurred there is hardly time to consider many alternatives especially in the immediate aftermath. It therefore, follows that not all users will have adequate time and ability to model the datasets to suit their needs. Instead, most may keep the details as they are in the dataset and build their perspective of use based on information at hand (Josselin, 2003). Most users are not interested in the solution finding process but in the exact solution to their problem (Jahn & Frank, 2004).

2.4.1. User perspective of data quality and task requirements

Disaster management has a myriad of users who have no definite criteria for categorization. Identifying user communities can be based on ontologies derived from description of user tasks (Kuhn, 2001), use of real word concepts indirectly related to data quality (Byrom, 2003), level of skills in handling disaster products (Kuntz et al., 1996) or duration of time over which the disaster task is performed (Erharuyi & Fairbairn, 2005).

The user requirements are centralized on what, when, where and how the information about reality is represented in the product. They determine 'fitness for use' by examining data contents (Boin, 2008) and also rely on any additional information explaining symbols used (Fisher et al., 2008). Identification is followed by assessing user competency in executing the tasks (Parker & Stileman, 2005) with objective of reducing risk and uncertainty (Agumya & Hunter, 1999). Vivid comprehension of user perspective or perception of reality may provide a vivid idea of their requirements as shown in Figure 6.

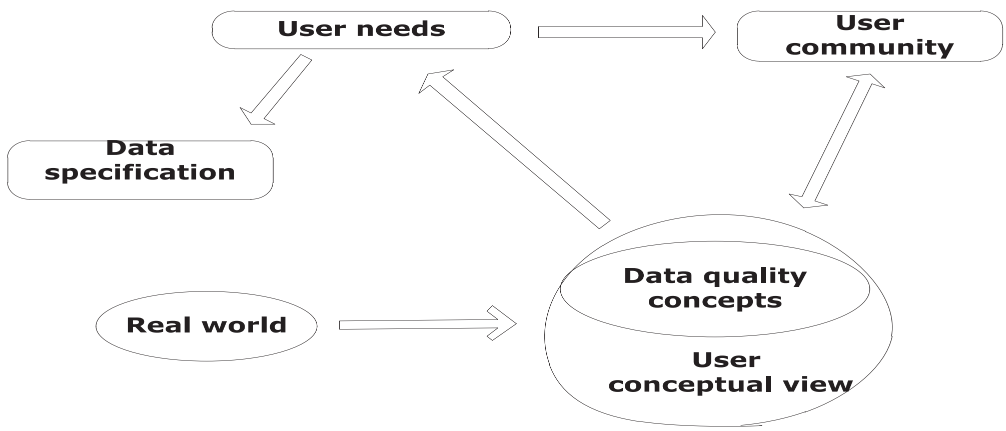


Figure 6: User decision contexts and data quality (Byrom, 2003)

Figure 6 indicates that data quality concepts within the user conceptual view will have effects on the user needs for a given data specification. Which means conceptualization of the real world in dataset influences the kind of quality information that is used in decision context of data quality.

Usability of the dataset centred on the user and the goal to be achieved (Josselin, 2003). Furthermore, data usability can also be centred on the data itself or user and methods available. Acceptable standards can be used by users to determine task requirements. However, the standard should be clearly understood as consistent, complete, flexible and easy to apply to a dataset due to different user responses towards datasets (Morrison & Veregin, 2010) as they determine 'fitness for use'. Some users determine the suitability of the dataset prior to using it; others resort to making choices among the available; while the rest may not consider spatial data quality information at all (Agumya & Hunter, 1999).

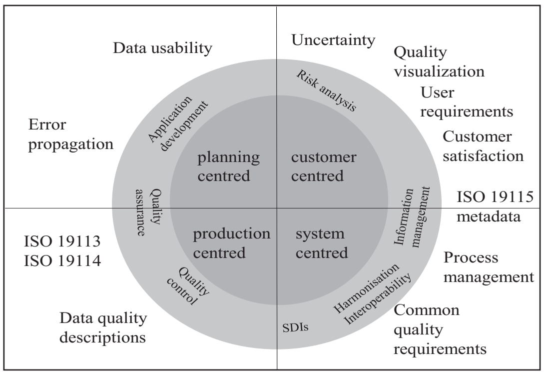
Immediately after the disaster occurrence, users need to know the level of trust they can assign to EO data they want to use. Trust of dataset influences usability and becomes very vital where quick decisions are required within a short time like immediate phases of disaster management such as search and rescue where saving lives and property protection is required with utmost urgency.

2.4.2. Spatial data usability

Spatial data usability and quality improvement is currently studied from several dimensions ranging from tools that improve fitness for use to modification of present models to accommodate flexible structure of quality reporting (Zargar & Devillers, 2009). Data uncertainty engine has been developed to handle error in GI science (Heuvelink et al., 2009). Other methodologies which have been developed include computer based intelligence engines for analysis of uncertainty information (Lowell & Casault, 2004), multidimensional user manual for experts of spatial data (Devillers et al., 2005; Devillers et al., 2007), risk analysis (Agumya & Hunter, 1999).

Standards for spatial data quality reporting have been established by bodies like CEN[7], OGC[8], FGDC[9], OGC, and ISO[10] among others; to guide in publishing information that would promote usability through concept of common understanding. However, despite all these efforts, quality information is still ignored by users of GIS applications (Van Oort & Bregt, 2005) due to difficulty in understanding quality information (Boin & Hunter, 2009) which is considered to be producer oriented (Comber et al., 2006).

The usability concept is well grounded in software engineering as illustrated in Figure 7 where products are made with particular user in mind (Nielsen & Mack, 1994). The usability concept has been extended to aid decision making on the basis of data quality elements (Byrom, 2003; Wachowicz et al., 2003).



[7] European Committee on Standardization
[8]Open Geospatial Consortium
[9] Federal Geospatial Data Committee
[10] International organization on standardization

Figure 7: Approaches to geographic information quality (Jakobsson, 2006)

From Figure 7, it follows therefore, that user satisfaction will revolve around data usability, its associated uncertainty, how the well its quality is described and common quality requirements that users may need as basic blocks for decision making or making judgments (Jahn & Frank, 2004; Jakobsson, 2006). Current usability studies of datasets tend be quite diverse such as an object oriented approach which integrates data quality information with spatial information of the object (Duckham, 2002), usability ranking of spatial attributes (Yiu et al., 2010), and so on. Therefore, disaster management requires a methodology that addresses the quality of products with the user in mind for effective and efficient use EO data.

2.5. Chapter summary

Use of EO data is vital for various phases of disaster management especially large scale disasters. However, quality reporting of the disaster response products is still a challenge due to various actors involved in production and dissemination of such products. The derived products are disseminated over the internet to potential users on websites of various agencies involved in managing the disaster in question. Evaluation of external quality requires indicators that allow the users to assign weight and rank quality information after setting preferences. Therefore, usability of any dataset for disaster management will depend on the level of trust on quality information delivered with the product, the user capability and the task to be performed. Therefore, there is need for a methodology or a new approach that may allow the user to assess usability of various data formats.

3. RESEARCH DESIGN AND METHODS

3.1. Chater overview

This chapter presents information on research design process with major steps illustrating how the research was carried as shown in Figure 8. Description of framework of "the methodology" developed describing the interaction of identifying the right dataset is presented as well as information on data collection methods on purpose of EO datasets, identification of user groups and user requirements and the detailed description of the overall steps taken to achieve the objectives.

3.2. The research design

After identifying the major inputs into methodology, the necessary information was organised under the steps shown in the methodology process. The data on user requirements was collected through primary and secondary methods. The parameters in the information containers of user requirements were identified by examining relevant literature on data quality evaluation. A summary of the methodology process is illustrated by Figure 8.

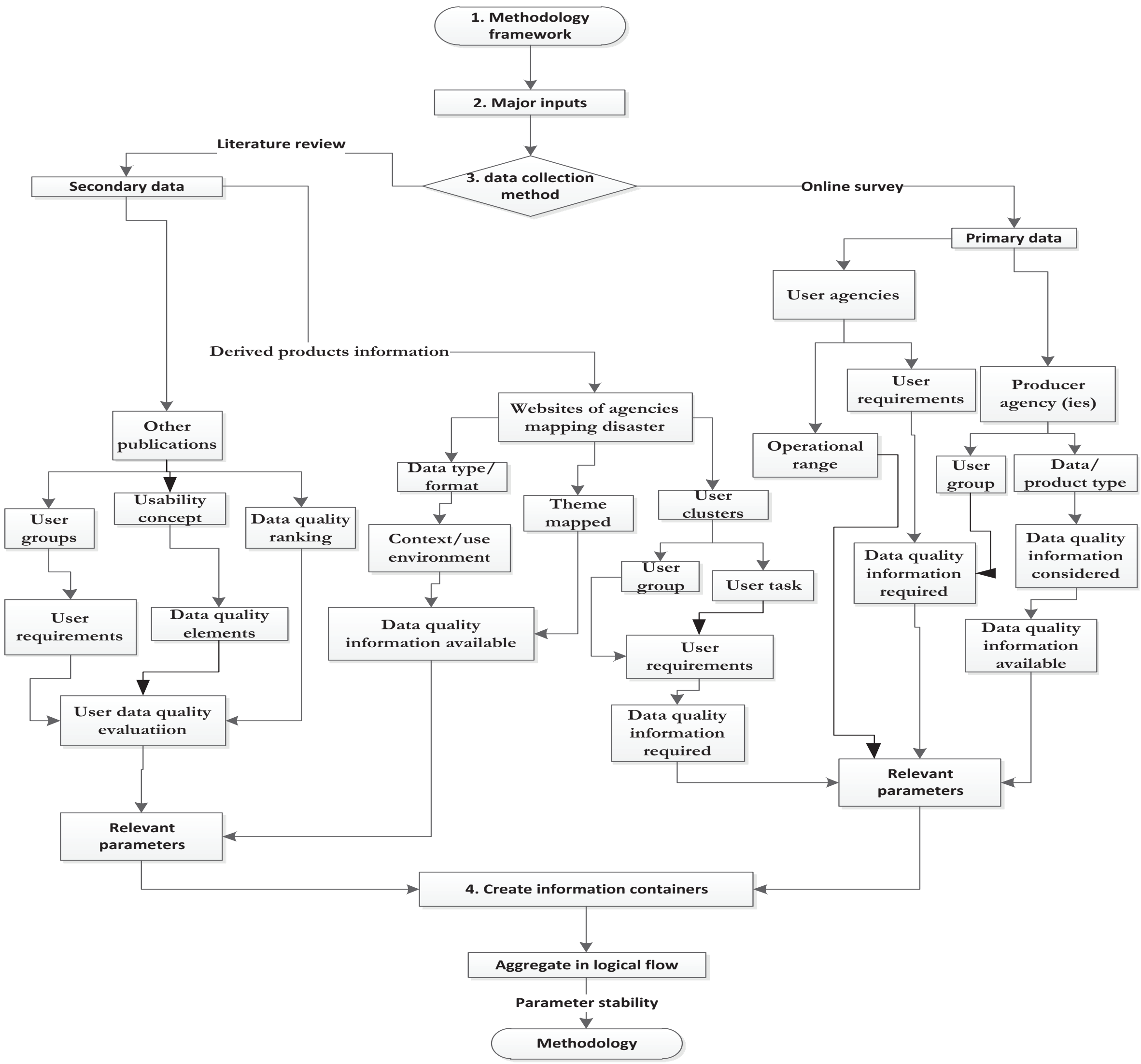


Figure 8: Research design flow chart

3.2.1. The steps followed in detail

1. Begin by establishing the methodology frame work on the task to be achieved by the methodology and identify possible interaction or relations that may exist globally or in detail. The framework was established after review of relevant literature on quality evaluation and production of EO data for disaster management.
2. Identify the necessary inputs that have impact on the outcome of the methodology. A methodology is like system interface, whose major inputs are like actors on a system. Therefore, inputs were identified as actors whose interaction led to a different stage and introduced a new concept in the process of interaction.
3. Decide on the relevant data collection method. Information on inputs of the methodology was collected from primary (online survey) and secondary sources.
4. Create information containers with related parameter values that form the major building blocks of the identified inputs. On the basis of major inputs of the methodology, the parameters for derived products specification and user requirements were identified as those that allow comparison to be made at technical level. Then the possible outputs of the information

containers were designed to have a filter because different combinations are possible as output display.

5. Aggregation of containers in logical sequence to form the methodology with view of the outcome of each container and the user at each stage of integration and comparison filters. Possibilities of parameter stability were used to determine the logical flow where reformulation was considered in containers that would produce a different result from the whole interaction process.

3.3. Framework of the proposed methodology

The methodology is developed on the principle that for a given disaster, timely response depends on the ability of users to identify the right product to use. Well organised organizations of users into groups will make it easier for dedicated data providers to consider general needs of a particular user group and provide information that is suited for that general group. However, current technological development has made EO data and spatial information ubiquitous which creates complexity in identifying quality differences (Mihaila et al., 2001). Therefore, users may require a methodology that allows to evaluate alternatives of quality information offered by various datasets (Jankowski, 1995).

Consider usability or selection of the right dataset for task after a disaster in terms of interacting three orthogonal axes i.e. the user group *i*, the user task *j* and user theme requirements *k* as shown in Figure 9.

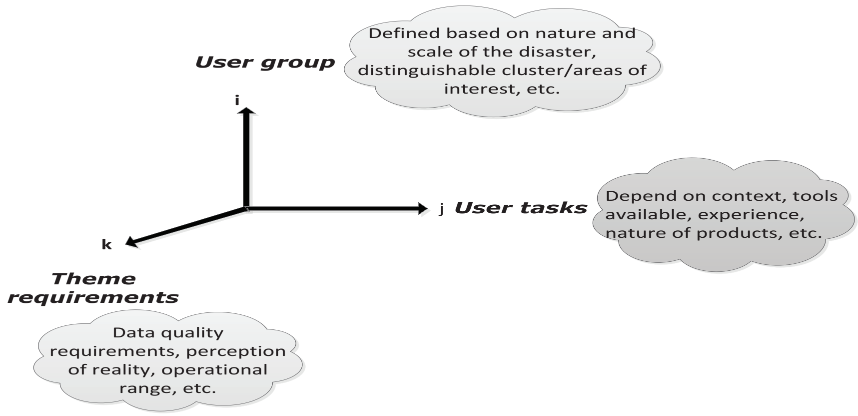


Figure 9: Three dimension perception of the right dataset for disaster response

From Figure 9, every activity in the disaster response will be defined by the user group (planners, responders, researchers, etc.) which are further divided into clusters. Therefore, a right product can be can be defined for a particular user group (i), performing a particular user task (j), based on requirements of a particular user theme (k).

It can be assumed that the right dataset for a task can be identified based on data quality elements reported with it, user requirements in the form of tasks and area of interest of a particular user group. While theoretically orthogonal interaction is expected so that the right product will be found for user group (i) to carry out their user tasks (j) in order to achieve theme requirements of a particular user (k); in practice compromise is normally made because the orthogonal relationship rarely holds in reality. Therefore, the right product will be obtained by choosing the *user group cluster* and *task* to be performed while the *theme requirements* will vary from user to user based on data type selected tools available, level of expertise with GIs data, task environment and duration of operation among others.

The 3D perception tends to establish the interaction and also shows how compromise can be made in reality where users operate. That is it is impractical to have user groups with tasks without theme requirements or operational range. Therefore, to assist users in defining operation range (making choices), the products used for disaster managements will require rating of theme requirements by experts in order to refine their task (use) requirements and where possible subject the potential set of products to preference ranking at personal level.

3.4. Major inputs of the research design

The major inputs that were considered in this research design process are:

i. Publication on user groups, usability of spatial data and data quality evaluation
ii. Information on derived products such as data type, theme and user clusters from websites of disaster mapping agencies.
iii. Primary data from user agencies to establish their requirements and operational range
iv. Primary from producer agencies to establish quality aspects considered in production.

3.5. Data collection

After description of the framework and identification of major inputs, data was collected through primary and secondary methods as summarised under the following sections.

3.5.1. Purpose of the Earth Observation datasets

The application of the EO data will depend on whether it consumed directly (no manipulation by the user) or it can be manipulated by the final user. The purpose and level of processing on various datasets used in disaster management was obtained through literature review and online survey of four organizations that participated in recent disaster management (Haiti).

After the review of relevant literature on purpose of earth observation datasets, sample datasets used in response to Haiti earthquake of 2010 were selected on the basis of data type and explored on processing tools. The purpose was to determine the level of interaction, processing and the kind of quality information that accompanies such EO data for disaster management. The data types explored were key mark-up language (KML) and zipped version of KML (KMZ) were explored in Google earth ; Haiti vector dataset for roads in Arc GIS, graphic products and raster processed pre disaster imagery on explored on ERDAS; and the level of interaction with layers.

3.5.2. Identification of user groups and user requirements

Information on user groups and their requirements were obtained through literature review and survey carried out of the four organizations. The literature consulted for identification of user groups of disaster management highlighted key areas such as: use of quality elements to derive ontology of user tasks (Kuhn, 2001); use of real word concepts indirectly related to data quality (Byrom, 2003); skills of users in handling spatial data (Kuntz et al., 1996); duration over which tasks are performed (Erharuyi & Fairbairn, 2005) among others.

The primary data on user group and their requirements was collected through a short survey on user and producer agencies as shown on Table 1

Table 1: List of originations contacted for survey

Method of data collection	*Type of data collected*	*Collection period*	*Organizations contacted*	*Purpose of collection*

Online survey	Primary data on producers	One month	• DLR • UNOOSA/UNSPIDER	Collect sample information on product specifications and user requirements considered during production producing agencies.
	Primary data on users	One month	• CARE international • SHA(Swiss Humanitarian Aid)	Collect user requirements for generating quality information template (QIT) from users perspective
Literature review	Secondary data	N/A	N/A	Gather information on relevant approaches of usability assessment Download of relevant dataset for quality assessment.

Survey on user organization was designed to collect information on phases of disaster management, types of derived products used and phases of application in disaster management, quality elements considered while selecting dataset for a task as well as how users handle data with no quality information. Ten questions were used for collecting primary data from the *user* and *producer* organizations that were contacted. The questionnaire of multiple choices with radio buttons and comment box were used. It was developed in S*urvey Monkey* and distributed online (*see appendix 1 &2 for questions used in the survey*). The responses from the survey were also collected online. In total twenty questions were administered.

Survey on producer agencies was used to gain insight on the kind of derived products that become available after occurrence of a disastrous event as well as to establish whether quality information is a priority to trusted and dedicated producers of disaster data; determine whether user groups are considered by producers and how the requirements of such users is determined; establish if context of use is considered during production, data sources used in production as well as whether the quality of such products are evaluated after their production. The results of this survey together with information collected from literature review was used to identify relevant parameters of information containers which were later aggregated in the methodology process.

3.6. Chapter summary

The research was carried out in four phases. Description of methodology framework was done in phase one where perception of identifying the right dataset in disaster response was established. This was followed by phase two where major inputs were identified. Data collection was carried out in phase three trough literature review and online survey to identify relevant parameters which were used to create information containers. The information containers were aggregated in logical flow to produce "the methodology".

4. RESULT: A METHODOLOGY FOR ASSESSING USABILITY OF EARTH OBSEVATION DATA

4.1. Chapter overview

This chapter presents "the methodology" developed with detailed description of criteria for its application. The role of users in different phases of the methodology is fully described alongside information requirement for each phase. Results of purpose of EO products and classification of user groups and requirements are presented as well.

4.2. User groups and user requirements

In disaster management, the user group determines the accessibility of response products due to cost, legal or security restrictions. The groups can also be used to establish the parameters of emergency context, related to event, task and phase of disaster management. For example search and rescue map would not be relevant if it is produced several months after occurrence of a disaster unless the disaster is continuous; but a product produced for recovery and reconstruction phase would be relevant several months after the disaster.

Crisis under consideration in disaster management can be earthquake, flooding, volcanic eruption, forest fires, etc. Therefore, this study categorizes user groups and their requirements into three broad groups: *planners*, *responders* and *researchers* on the basis of task requirements and quality information on the theme on which the EO product has been produced that is likely to be communicated to the user. The details of the categories is summarised as follows:

4.2.1. Planners

These categories of users provide the organizational environment for overall management of the disaster in question. Their EO data and related requirements tend to be wide and may be totally different from those interested in specific information. This group may also have authority to control accessibility of certain data types during disaster management. Users under this group may operate in all phases of disaster management and therefore, they can provide technical support (assist in processing EO e.g. MINUSTAH was involved in processing of EO during Haiti earthquake of 2010) and make plans for different phases. Examples of users in this category include the government(s) of the host country (yes) and international disaster coordinating agencies like UNOCHA, UNOOSA, GMES, etc.

4.2.2. Responders

These are frontline users of the derived EO data whose main interest is the activity to be performed with the dataset. They interact with various types and formats of EO derived products such as KML, KMZ, vector, raster, and graphics, to perform tasks related to the phase of disaster on which they operate. In other words, this group puts the information represented in various derived EO products into actual use (dataset usability): i.e. information represented in the EO data and derived product is translated into actual task in the real world.

This category of users may operate in clusters based on their activities, mandates, or nature of the agency (bilateral or NGO, etc.). The EO information for this group should provide information on key areas such as search and rescue, health, sanitation and related services, protection and security, communication infrastructure, transport and logistics, *inter alia.*

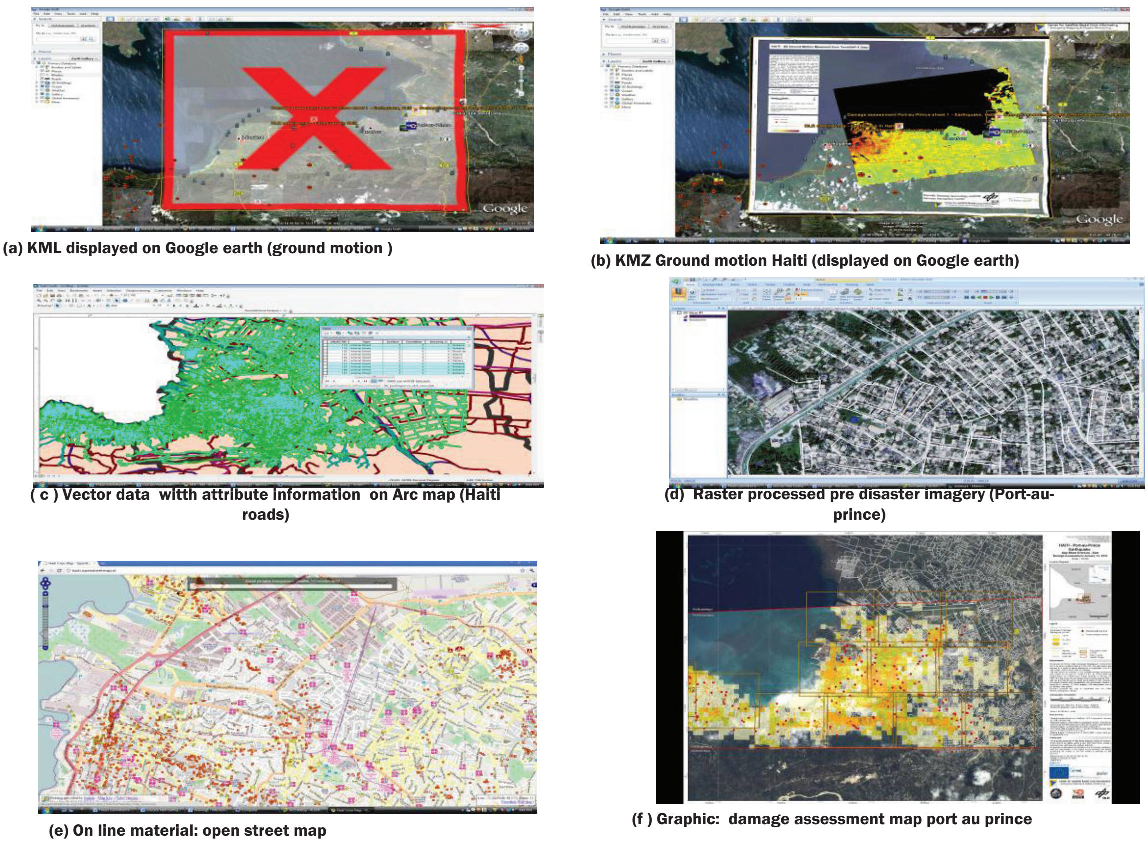
4.2.3. Researchers

The interest of this group is to best understand the event and provide information that best explains the event. Members of this group have the capability to apply technology on EO data and generate information of interest. Therefore, products destined for this group may not necessarily have to be final end products, since they can generate it their own concrete end products. Members for this group could include individual consultants, researchers, institutes, universities, or even research companies.

The user requirements and user groups have to be harmonised during any kind of disaster. The harmonization helps to address the problem of interoperability across user platforms, access regulation, reduce congestion in the network of delivery and help the users to identify the right tools appropriate for each user group.

4.3. Purpose of EO products

Products produced in response to disaster are in different formats with information at various levels of detail. The main category of EO products that were found to be commonly available for disaster management include graphic products especially maps, GIS vector data, image products donated, from archive or bought; products integrated on the web maps, xml data, reports, and so on. Information on sample datasets explored during this study is summarised by Figure 10.



(a) KML displayed on Google earth (ground motion)

(b) KMZ Ground motion Haiti (displayed on Google earth)

(c) Vector data with attribute information on Arc map (Haiti roads)

(d) Raster processed pre disaster imagery (Port-au-prince)

(e) On line material: open street map

(f) Graphic: damage assessment map port au prince

Figure 10: Thumbnail of some EO based derived products explored (Haiti earthquake of January 12th 2010 response datasets)

The exploration of various derived products as shown on figure 10 indicated that, various formats of EO data exist with varying levels of details. The purpose of the derived EO products for disaster management are therefore, grouped in this study under two broad categories namely:

4.3.1. Products for direct use

These comprise of products that are used in their original format as delivered from the supplier. Their nature does not allow manipulation but exploration is possible when displayed on suitable tools. These include graphic products such as maps which can be provided in various resolutions, when availed as soft copy. This study considers products of direct use as fully processed with (good) thematic information in case the product has been classified.

4.3.2. Products for indirect use

These products allow the users in disaster management to customize their requirements because the products can be manipulated with special tools or software before the user can make use of them. These products could be *raw products* collected from field ground measurements or obtained by sensors such as satellites, radar, and other optical devices. The raw products are ideal for professionals of spatial data because they tend to be large in size and may be available in several inconsistent formats which do not allow combination or overlay with any geographical information system in their original form.

Basic processing on EO data such as radiometric correction, geo referencing, and geometric correction result in *semi processed products* which tend to be smaller in size and may be disseminated over the internet. The semi processed products can be manipulated by various users to suit their tasks or to produce new customised final product. For example during the Haiti earthquake, some providers released their semi processed products to assist in disaster management, which were adopted and used by volunteers individuals and volunteer agencies to produce an abundant collection of response products.

4.4. "The methodology" for assessing usablity of EO data and derived products for disaster management

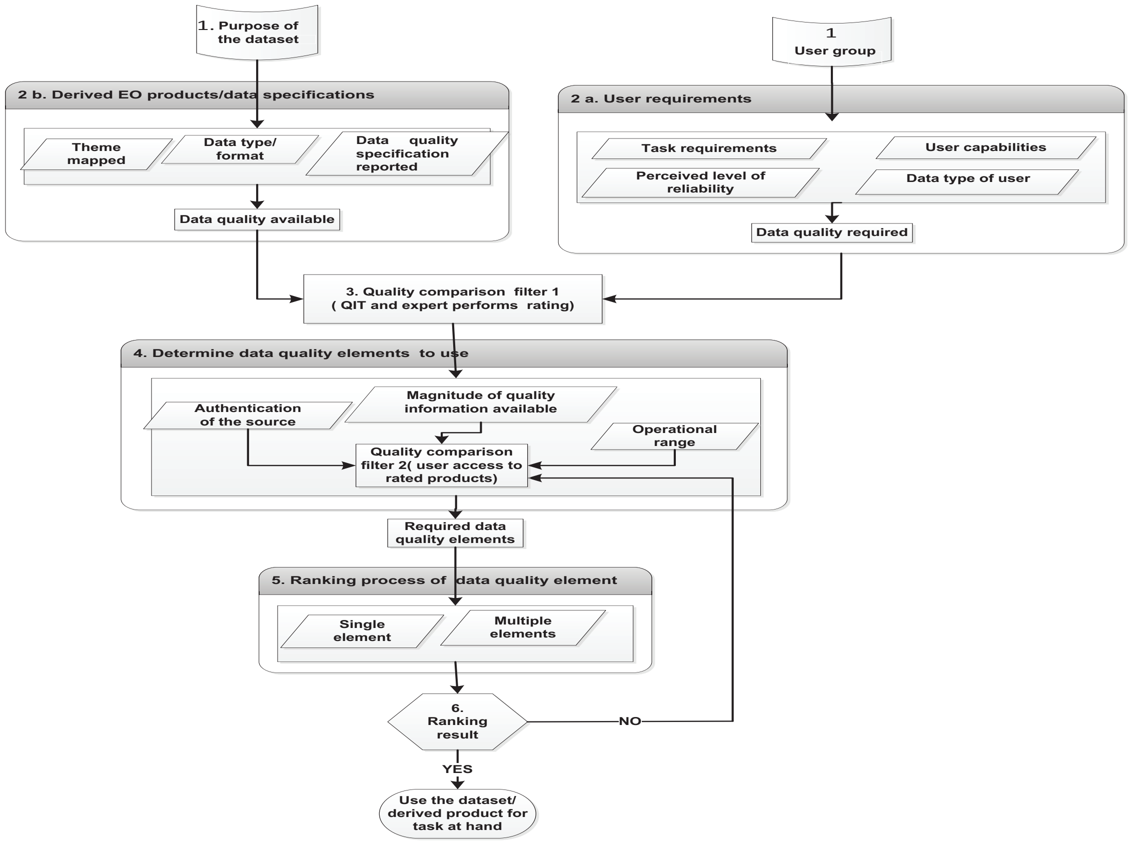


Figure 11: The methodology for assessing usability of EO based derived products

The methodology presented in Figure 11 is designed to be implemented in two phases. The first phase ranges from steps 1 to 3, which involves experts who have sound understanding of data quality information with capability of establishing the level to which a given a product description conforms to the contents in the dataset. The implementation of this phase requires a common platform for comparing quality information of various products for various user groups irrespective of the provider of the product. The platform proposed in this study is the quality information template (QIT) that should be designed by experts in disaster management to provide a common standard to be used for judging the quality of various EO products used for disaster management.

The second phase ranges from step 4 to 6 of the methodology and involves making the rating information available to the end-user in order to interact with the product and rated quality information. The interaction allows determining the necessary parameters of quality elements that can be subjected to ranking.

4.4.1. Application criteria of the methodology

Application of the methodology illustrated on Figure 11 involves two phases of the major six steps described. The steps involved in the *first* phase are:

1. Observe the theme under which the product belongs, the purpose of the product and establish the possible broad category of user group(s) i.e. planners, responders, researchers, etc. Determine

whether the global needs of the possible group of users have been addressed by the nature of quality information available as product specification.

2. Create a common platform for comparing the spatial data quality required and the spatial data quality available i.e. create quality information templates (QIT) with parameters of purely data quality elements as shown in Table 2 or basis of graphic products and general description of metadata as shown in Table 3. In general, this step requires that, the experts evaluating the EO products for a given disaster should develop a heuristic list of parameters that allows harmony in the results of evaluation by various experts who might be in various distant parts of the globe and taking part in the evaluation exercise.

Table 2: Quality information level template for data quality elements

Theme mapped		
Data format being evaluated (graphic, xml, kml, gml)		
Quality information reported as metadata or legend or reports, etc.		
Data quality information	Derived products/datasets	
	Datasets details and operational range	*Defined rating indices*
Scale	Large (much details on small area)	
	Medium (moderate details on moderate area)	
	Small (less details on large area covered)	
Positional accuracy **Level of details in display**	input resolution (scanned maps)	
	Output resolution	
	High resolution	
	Medium resolution	
	Low resolution	
Completeness:	minimum size of mapped objects	
	coverage full and highly detailed	
	full and less detailed	
	partial and highly detailed	
	partial and less detailed	
Attribute accuracy	Attribute present	
	Attribute absent	
	Conformity of attributes high	
	Low conformity of attributes	
Logical consistency	High	
	Medium	
	Low	
	General information on compatibility	
Lineage	Purpose of the dataset	
	Reliability of sources used in production	
	Transformations applied	
	Time of data capture and time of availability	
Time	Consistency of dates of products	
	Currency of datasets and update frequency information	
	Time range covered by the product	
	Time for mastery of the dataset or tools	
	Time taken to access and download the data	

	Validity period	
	Data of specific dates	
No quality information	Possibility of combining with other datasets and supported formats	
	No possibility of combining with other datasets	

Table 3: Quality information level template for data quality elements Cartographic and metadata

Parameter being evaluated	Value	Observation	Preference/operational range
Service provider			Establish whether dedicated provider, volunteer agency/individual
Type of disaster or crisis event mapped			Does the product provide information relevant to disaster in question
Time of production			Determine the relevance of time of production to the time of occurrence of the disaster and availability.
Spatial coverage of area of interest			Identify whether the product explicitly identifies the area that is represented. Bounding box for graphic and spatial coordinates for metadata. Product sheet number for graphic products produced in series.
Presence of overview map			Determine whether spatial orientation of the area represented is provided.
Visibility of the product			Determine coordinate graticule or grid labels for graphic, Observe visibility of features in case of vector data and rate depending on presence or absence of discrepancies.
Map title/theme			Establish the relevance of title to the contents of the product(s) mapped
Type of map background			For topographic products establish if the background promotes better interpretation of the contents.
Specification of sensor used			Establish the relevance of sensor used and specification of that category of sensors e.g. specification of optical sensor should not be reported with radar product and vice versa
Language of production			Determine whether the product is produced in technical language or natural language as well as support for multiple languages.
Information on occlusion			Establish if there are occluded areas in product but not indicated in the quality report or metadata.
Type of quality information reported			Establish if the quality description provides relevant details on the product mapped .i.e. legend, metadata, or report.
Spatial resolution of sensor used in production			Determine if the input resolution is relevant to the level of accuracy reported with the product.
Consistency between map and legend information			Determine whether some items in the legend are not in the map
Contrast between background and thematic information			
Clarity and readability of symbols			Establish relevance of symbols used in representation and whether they allow the user to vividly distinguish information being represented.
Presence of scale information			Determine whether there is scale for analysing minimum size of napped object and relevant printing size.
Information on accessibility and data sharing			Does the product support established standards of data sharing and delivery media?
Responsibility assumption			Establish if reliable information involving trust is provided and whether the product conforms to any standards
Consistency between declared scale and resolution of input			Does the product show consistency between input resolution and scale of production.

images			
description of processing steps			Lineage information of the processing steps and assumptions made during the production. Are the generalization, modification and assumption used in line with the kind of product produced?
Information on quality control procedure			Does the provider indicate measures taken for quality control and are those measures in line with established protocols or standards.
Information on spatial accuracy			Is there information on internal accuracy of the datasets being evaluated and are they within acceptable limits.
Report on known sources of error			Is there information on known sources of errors identified during production and possible consequences on use of the dataset?
thematic accuracy report			Is there information for validating thematic classes
Reference datum/ Reference projection			Determine if the projection used preserves shape and direction.
Information on (logos of the) partner agencies			Is the dataset from single provider or produced collaboratively and what extent are the datasets accurate i.e. is there information on accuracy of all datasets used in production or just overall information final product?
Media used and accessibility			How relevant is the media used in delivering the product
Correct use of symbols			Establish credibility of symbols and any ambiguity that may exist when the product is being interpreted.
Product version number			Determine if the product was registered during production and whether other versions of the product exist. This information is useful for evaluating if the product has been updated.
Authentication			Is there any sign pointing towards the level of quality met by the product? Information regarding the reliability of the sources used in production and any sign of quality assurance or responsibility assumption.
Additional fees (if any)			
Parameter being evaluated	**Value**	**Observation**	**Preference/operational range**
Basic information about the dataset(citations, abstract, purpose, use constraints, contact persons)			Preference on this section will depend on the kind of established metadata standard that is deemed as suitable for evaluation (FGDC, OGC, ISO, etc.).
Data quality information(lineage, processing steps)			
Spatial organization of information(point, vector, raster)			
Spatial reference information (UTM, IDTM, etc.)			
Entity and attribute information (summary, definition, values and units)			
Distribution information(data format, download link, data compression)			
Metadata reference information (time the metadata was created and by who)			

The parameters used in populating QIT should reflect the general characteristics of product type, purpose (use) and the general standard of the quality description that the product type should contain to allow the end user determine quality differences. After identifying the relevant parameters and populating the relevant QIT, examine the nature of quality description of derived EO products to be evaluated.

3. Enumerate gap analysis between the quality information reported on the product specification/quality description and QIT. Determine the level of conformance based on QIT variables and assign conformance rating. The indices used should allow harmonization of parts into whole: i.e. the indices should be normalised between 0 & 1 and populate the template (see *chapter 5 for illustration*). Publish the rating as hard copy or soft copy (depending on the nature of intended use by end-users) alongside the product to allow comparison with other rated products, just like in a shopping catalogue.

The *second* phase is carried out by the (remote) end-user who interacts with rated EO data and products with alongside rating information. The end-user encounters a rating thumbnail and display of the product or links to the rated product. *The context of use of the rated information influences how the user will access the rated EO datasets and determine the relevant quality information to use. Large number of products with wider user community like disaster management products may require the information to be organised on a website with interactive user interface.*

4. The end-user interacts with the product and makes a selection of the potential EO data and derived products with details of interest. The user compares their requirements with authentication of QIT rating, magnitude of quality elements/available parameters, or based on their own personal operational range (minimum quality information that user are willing to accept in order to make decision).
5. Set preferences for all the potential EO datasets selected from stage 4 and establish weighting criteria for decision making. Decide on the number of parameters to rank and establish whether the outcome will result into a single data quality or multiple elements depending on the sub categories under each QIT parameter and calculate the rank as follows: Single parameter rank (R) = value (V) × weight (W) and Product rank (PR) = ∑ R Depending on the number of datasets, the information can be organised as shown in Table 4.

Table 4: Rating and ranking table

Criteria	Products						
Quality information template parameter (QIT)	A			B		…n	
	Weight	Value	Final value	Value	Final value	Value	Rating
Total product rank (∑ R)							

6. Review the result of the rating and make a decision to use the preferred dataset or reformulate the dataset collection and set new preferences. Stop the raking process if the ranking result is conducive for decision making on the task (use) and pick the dataset.

4.4.2. Role of users in the different phases of the methodology

The initial user in the first phase of the methodology is the expert who implements the first three steps of the methodology and handles all the technical details and expertise required to establish the conformance of product specification to the actual contents. The experts do not market or demean any producer of the product rated. They only judge the information content on the basis of QIT parameters and creates harmony by organizing the products either by theme, product type, the kind of quality information available and may set the level of privileges on the template view of quality rating; that allows the user to edit the product rating or view only.

The second phase of the methodology from steps 4 to 6 is designed for the end-user. The end-user explores the rated products and set preferences to perform the process of suitability analysis. The end-user decides on the method of decision making approach to adopt and set the necessary weight on the basis of preferences they set on each potential EO data/derived products.

The conformance of quality specifications of the derived products can be evaluated by established team of experts within coordinating agencies or established quality assessment team in the departments of value adding agencies. However, it worth noting that some producers do not deliver concrete derived products to users; which means, the users have to modify the products before they can use them. Therefore, semi processed derived products should be accompanied with well documented evaluation report which can be used by the prospective end user to set preferences and make relevant choices

4.4.3. Purpose of quality information template (QIT) in the developed “the methodology”

The templates are created as common platform for various EO data as well as a means of harmonizing various user requirements. The template can be designed by a team of EO experts especially from agencies dedicated to disaster management as a means of introducing a standard to quality of various products that may reduce conflicting interpretation of the same dataset from various producers.

The derived products’ specifications may contain different aspects of quality information such as metadata, legend, visual cues or report, etc. Since the quality information may not be explicitly reported, the template sub-categories are created to capture information on product quality that is on lower hierarchy of the data quality elements.

The design of two templates is proposed in the methodology because the user may be interested in rating individual elements or parameters only on specific products or on the overall quality of the product. Furthermore, graphic products may be evaluated on the basis of cartographic principles which examine the overall quality conformance.

4.4.4. Importance of QIT to the end-user of EO derived product

Disaster management involves a large collection of datasets that is made available within a modest period of time. These products contain various characteristics such as abundance and variety of properties, varying levels of quality information, possibility of multiple functionalities, and different aspects of quality reporting from various agencies as well as lack of authentication. Therefore, through expert rating, QIT provides an opportunity to the end-user to access collection of organised products whose quality information has been harmonised and made available. Such information allows the end-user to select a product and perform relative comparison of quality preference. This does not require the user to understand the spatial units of various data types, but to establish the relevance of quality information to task at hand.

The problems associated with characteristics of various EO derived products such as differences in data structure, geographic scale anomalies, duplication of information, and lack of structured quality report are harmonised through QIT. These aspects limit the effective use of such datasets especially when the user cannot quantify the magnitude of quality differences among various preferences that may be available for a task at hand. Through the QIT and ranking, the elucidation of common quality elements of usability can be assessed and quantified through relative preference.

The aim of developing QIT is to reduce information overload on the end-users who are exposed to a lot of unclassified data/derived products produced during disaster response. This is because good information presented poorly may be misunderstood or poor information neatly presented can masquerade as right information. By designing QIT for spatial data, expert teams can review the products to determine the level of conformity of their quality information to the quality specifications reported and assign a rating value on the selected parameters.

4.5. Link between data quality description to application of the data/derived products

EO based products can be useful for disaster management and response if the users in a particular user groups can receive an immediate detailed picture of the environment affected. However, the information being provided to users comes from various sources and is represented in various formats with varying level of details. Users relate the quality information of the derived product or dataset to the representation in reality. Which means quality will deteriorate as divergence is observed between the dataset/derived and aspects it represents in reality (Devillers & Jeansoulin, 2006). Therefore, there is need to set preferences by choosing an approach that assists the user in quantifying preferred elements for decision making.

The methodology described in Figure 11, tries to link the role of the experts and end-user in one model to assess the usability of EO derived products. The challenge associated with derived products is varying levels of information regarding their quality, large volume of products from various suppliers and volunteers. To assess the quality information and usability requires that the EO derived products are brought to a common platform that allows comparison to be carried out. This is the level where expertise knowledge is required to generate quality information template (QIT) for various user groups that address the use needs a particular group.

In the first phase of the developed methodology, the user requirements and the product specifications are subjected to filtering. Since there are large numbers of products used for disaster management, the product specification can be catered for by the information on QIT and the user requirements by making the QIT interactive so that the user is able to select the relevant attributes to their task. To achieve this, it is proposed in this study that the team of evaluators which can be volunteers or appointed by various suppliers can prepare a common template and make it available as online form that can be used by experts to examine whether the product specification conforms to the details in the products made available to the user.

4.6. Chapter summary

This chapter has presented the major steps in the development of the methodology and description of various phases involved. The processes described tries to establish conformance of quality specification to actual contents of the data. The main objective being to assess usability of EO datasets for disaster management by integrating the role of people with expertise in handling spatial data and the preference of actual end-user who is the intended consumer of the produced information. In other words, it is simply 'expert rating with capability of user suitability'.

Actually, it is the product's quality that matters not data quality (Couclelis, 1991). But the product's quality is expressed as data quality element(s) in the product descriptions i.e. metadata, legend, quality reports, etc. The main reason for this is to establish quality elements that convey required information for decision making. Select the dataset of choice based on ranking results. The ranking results will depend on data format/type (nature of derived product) and possible applicability of indices showing differences in levels of detail. This is because the same data may be evaluated differently by different users or the same user may evaluate the same data for different user tasks. This is due to the subjective nature of evaluating spatial data quality.

5. CASE STUDY: THE HAITI EARTHQUAKE 2010

5.1. Chapter overview

This chapter presents a brief description of study area and full application of the methodology on sample datasets. It begins with populating quality information templates (QIT) on sample EO datasets for the two types of templates and the creation of web interface as the platform of interaction between the experts, users and the rated EO data. The chapter ends with a description of a ranking scenario through analytical hierarchy process (AHP).

5.2. Application of the developed methdology on derived EO data

The objective of this case study is to show how the proposed methodology can be used to assess the usability of EO based derived products. The derived products used in illustration were obtained from websites of agencies dedicated to provision of disaster products. The user requirements of various user groups were collected through survey on user and producer organizations involved in the response to Haiti earthquake of January 2010.

5.3. Study area

Haiti is geographically located 18.443^0 N and 72.571^0 W in the Caribbean, western 1/3 of Hispaniola Island, in between the Caribbean Sea and North Atlantic Ocean towards the West of Dominican Republic. The country has a population of approximately 9 Million in an area of 27750 square kilometres where 190 square kilometres is covered with water; as well as a coastline of 1771 km. It lies in the middle of a hurricane belt and is subject to severe storms from June to October, occasional flooding and periodic drought (URL2.3)

The Haiti region lies between Caribbean plate and on the transform plate boundary with the North American plate which is moving west. The movement creates friction which builds pressure later released as earthquake. The earthquake occurred at shallow depth of 10km; West of Port-au-Prince which increased the energy of the seismic waves due to short distance to the earth's surface. The 2010 earthquake also caused uplift and subsidence due to effect of shaking from a magnitude of 7.0 on Richter scale.

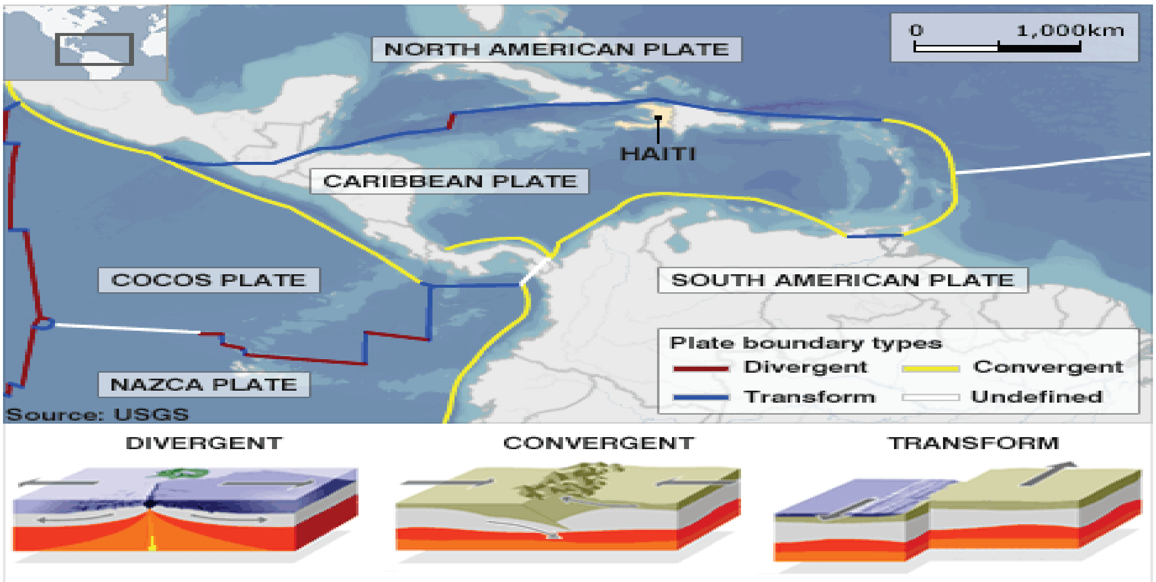
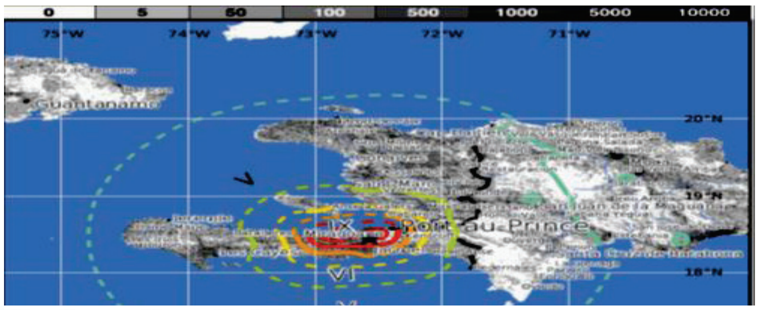


Figure 12: Types of faults within the Haiti region (www.internetgeography.net)

The tectonic summary of Haiti indicates that the plate boundary is dominated by left lateral strike slip motion and compression accommodates approximately 20 mm/year slip with Caribbean eastward movement with respect to the North American plate (USGS) as shown in Figure 12.

The main affected areas were Port-au-Prince, Petit Goave, Leogane, Jacmel and settlements found South West of Haiti as shown in Figure 13. Leogane and Jacmel were the epicentre with 90% of damage. This left over 200000 people dead and approximately 1 million people homeless (USGS) causing estimated damage of US $ 7.8 billion (Orhan et al., 2010).



Selected City Exposure
from GeoNames.org

MMI	City	Population
X	Petit Goave	118k
X	Gressier	26k
X	Grand Goave	49k
IX	Leogane	134k
VIII	Carrefour	442k
VIII	Miragoane	89k
VIII	Port-au-Prince	1,235k
VIII	Delmas 73	383k
V	Verrettes	49k
IV	Santo Domingo	2,202k

bold cities appear on map (k = x1000)

Figure 13: Haiti Earthquake city exposure (USGS)

Disasters like the Haiti 2010 earthquake that affect large segments of population as shown in Figure 13, attract international attention and therefore, receive data from various sources with varying levels of information on their quality require an approach that allows users to set preferences during the process of decision making. Trusting such products or establishing authenticity is necessary in order to determine the level of confidence that can be assigned to them in decision making. These considerations made Haiti a suitable area for study because data is collected for purpose and so in this study, a methodology is developed that can be used as an attempt to establish usability of disaster response products.

5.4. QIT of the EO derived products

5.4.1. QIT for the individual data quality elements of the derived products

For the individual data quality elements, relevant details are entered on sub sections of the data quality elements as shown in Table 5 after exploring the products. The template for individual quality elements is used and populated as shown.

Table 5: Sample datasets considered for demonstration

	Theme mapped	Data type/	QIR		Scale (s)		positional accurcy/level of details			completeness(C)					Attribute accuracy (			
				Larg	medium	smal	input resoluti	output resolu	Mom	CFHI	CFLI	PCLD	PCHD		present	absent	LC	HC
Data s	damage assessm	graphic	legend															
A	damage assessm	graphic	legen	1		1	0.5 m	50 m	2.5 m				1		0			
B	damage assessm	graphic	legen	0		0	0				1				1			
C	damage assessm	graphic	legen	0		0	0				1				1			
D	damage assessm	graphic	legend		1	1	0.5 m	250 m	12.5 m			1				1		
E	damage assessm	graphic	legend		1	1	0.5 m	250 m	12.5 m			1				1		

Logical consistency (LC)					lineage (L)					Time (T)				no quality inf		overall combinat	fit for evalution(FFE	
high	medium	low	purpose	RoS	TA	ToA &Td	CoD	TR	VP	SD	C&UF	TFM		PoC	No_PoC		yes	no
	0		1	0.5	0	1	1	1	1		0.5	1			1	S, PA, C, T		
	1			0	0	1	1	0	0	0	0	1			1	C,AT,LC,L,T		
	1				0	1	0	1		0					1	C,AT,LC,L,T		
	1		1	1		1	1	1							1	S, PA, C,LC,L, T		
	1		1	1		1	1	1							1	S,PA,C,LC,L,T		
																		C

KEY:
- **Mom: minimum size of mapped object**
- **CFDH: completeness full and highly detailed**
- **CFLH: completeness full and less detailed**
- **PCLD: partialy complete and less detailed**
- **PCHD: partialy complete and highly detailed**
- **LC: low consisitency**
- **HC: high consisitency**
- **ToA &ToD: time of acquisition and time of delivery**
- **TA: time of acquisition**
- **CoD:currency of dates**
- **TR: time range covered by the product**
- **VP: validity period**
- **SD: specific dates**
- **C&UF: currency and update frequency**
- **TFM: time for mastery of the product.**
- **PoC: possibility of combining with other dataset**
- **No_PoC: no possibiltyof combining withother datasets**

5.4.2. QIT for the overall quality of the derived products

Graphic products are evaluated on the basis of overall quality and adherence to cartographic principles. Therefore, their heuristic evaluation parameters should be developed with the two aspects in mind. An example of QIT application on cartographic product on Figure 14 is demonstrated by Table 6

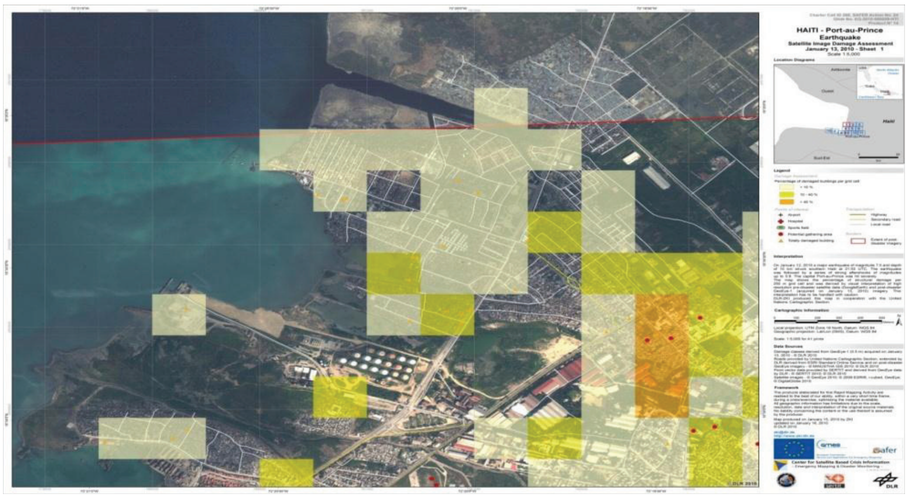


Figure 14: Satellite image damage assessment_ product A

Table 6: Application of QIT on usability evaluation of graphic product

Parameter being evaluated	Value	Observation	Preference/operational range
Service provider	1	Copyright information	Establish whether dedicated provider, volunteer agency/individual
Type of disaster or crisis event mapped	1	Earthquake information provided	Does the product provide information relevant to disaster in question
Time of production	1		Determine the relevance of time of production to the time of occurrence of the disaster and availability.
Spatial coverage of area of interest	1	Bounding box present and sheet number	Identify whether the product explicitly identifies the area that is represented. Bounding box for graphic and spatial coordinates for metadata. Product sheet number for graphic products produced in series.
Presence of overview map	1	Clearly shown	Determine whether spatial orientation of the area represented is provided.
Visibility of the product	0.5	Can be read, but there is mismatch between background and the product grids	Determine coordinate graticule or grid labels for graphic, Observe visibility of features in case of vector data and rate depending on presence or absence of discrepancies.
Map title/theme	0.5	Not properly placed	Establish the relevance of title to the contents of the product(s) mapped
Type of map background	0.5	Satellite background	For topographic products establish if the background promotes better interpretation of the contents.
Specification of sensor used	1	Sensor name and resolution provided	Establish the relevance of sensor used and specification of that category of sensors e.g. specification of optical sensor should not be reported with radar product and vice versa
Language of production			Determine whether the product is produced in technical language or natural language as well as support for multiple languages. Usage of a product in technical language is limited to experts of that field.
Information on occlusion		No details	Establish if there are occluded areas in product but not indicated in the quality report or metadata.
Type of quality information reported	1	Legend and interpretation text.	Establish if the quality description provides relevant details on the product mapped .i.e. legend, metadata, or report.
Spatial resolution of sensor used in production	1	Provided	Determine if the input resolution is relevant to the level of accuracy reported with the product.
Consistency between map and legend information	0.5	Some aspects lack in the legend	Determine whether some items in the legend are not in the map

Contrast between background and thematic information	1		
Clarity and readability of symbols	0.5	Colour repetition of symbols and symbol shape not distinctive	Establish relevance of symbols used in representation and whether they allow the user to vividly distinguish information being represented.
Presence of scale information	0.5	Well stated, wrongly placed	Determine whether there is scale for analysing minimum size of napped object and relevant printing size.
Information on accessibility and data sharing		Not provided	Does the product support established standards of data sharing and delivery media?
Responsibility assumption	0	User is responsible for misuse	Establish if reliable information involving trust is provided and whether the product conforms to any standards
Consistency between declared scale and resolution of input images	0		Does the product show consistency between input resolution and scale of production.
description of processing steps	0	Not provided	Lineage information of the processing steps and assumptions made during the production. Are the generalization, modification and assumption used in line with the kind of product produced?
Information on quality control procedure	0	Not provided	Does the provider indicate measures taken for quality control and are those measures in line with established protocols or standards.
Information on spatial accuracy	0	Limits of spatial accuracy missing	Is there information on internal accuracy of the datasets being evaluated and are they within acceptable limits.
Report on known sources of error	0	Missing.	Is there information on known sources of errors identified during production and possible consequences on use of the dataset?
thematic accuracy report	0	Missing	Is there information for validating thematic classes
Reference datum/ Reference projection	0.5	Mismatch is observed	Determine if the projection used preserves shape and direction.
Information on (logos of the) partner agencies	0.5	Present	Is the dataset from single provider or produced collaboratively and what extent are the datasets accurate i.e. is there information on accuracy of all datasets used in production or just overall information final product?
Media used and accessibility			How relevant is the media used in delivering the product
Correct use of symbols	0.5		Establish credibility of symbols and any ambiguity that may exist when the product is being interpreted.
Product version number	1	Present	Determine if the product was registered during production and whether other versions of the product exist. This information is useful for evaluating if the product has been updated.
Authentication	0	Absent	Is there any sign pointing towards the level of quality met by the product? Information regarding the reliability of the sources used in production and any sign of quality assurance or responsibility assumption.
Cost	0		

The dataset was observed against the heuristics and listed under the column of preferences and rating was applied on the assumption of expert role to illustrate how experts may apply QIT on graphic product.

5.5. Population of QIT with information of derived products

The products were explored on processing software that supported their formats and the details entered against QIT predefined parameters as shown is sections 5.4.1 and 5.4.2. The values entered were in the range of 0 to 1. The procedure was typical to user heuristic evaluation and the ratings were assigned to the datasets.

5.5.1. Creation of user interface on ZOHO web creator

The targeted user of this methodology is a wider group of end-users involved in various phases of disaster management. Therefore, after the expert rating, the information should be published on an interactive platform with user interface that allows the end-user to explore various categories or templates containing the evaluated information.

The user interface on ZOHO web creator was designed to show that it is possible to develop heuristic parameters on a distributed system which can be used by various experts at different locations as baseline for their evaluation. The interface for demonstration of the methodology was designed with templates that allow submission and viewing of the submitted information. During disaster management, depending on the level of privileges set, users can view the rated products and subject them to ranking. They can also provide feedback on a number of datasets they used and such information will appear alongside the rated information. Therefore, it becomes an interactive platform that addresses usability by linking the expert to the user and finally feedback from the user to the system portal which can be accessed by producers as well.

The information can be organized per theme, data type or usability assessment information. In other words, the home page should allow the user flexibility to make a selection of suitable elements to subject to ranking. A demonstration of this approach should result into a home page like the one in Figure 15.

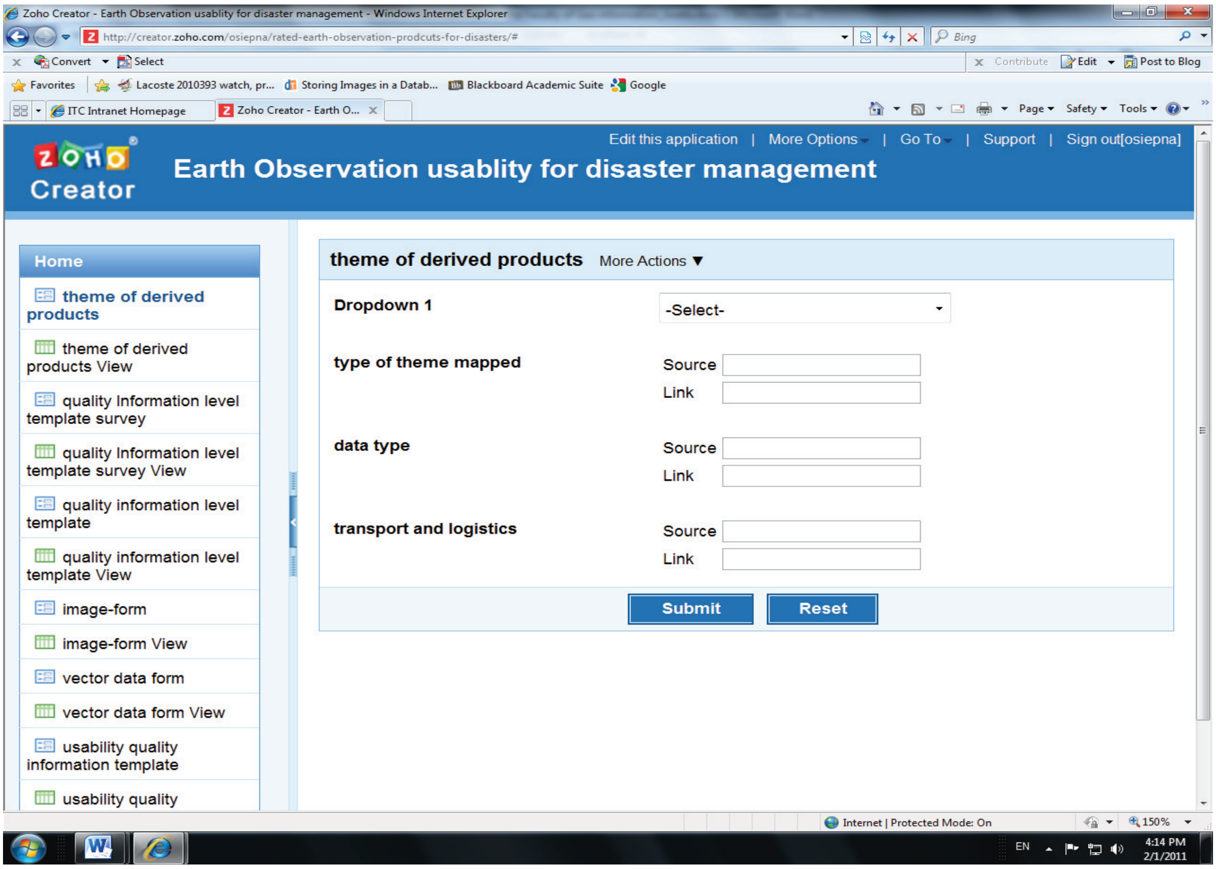


Figure 15: User interface for accessing rated products and submission of expert evaluation

5.5.2. Developing interactive QIT and quality information view in ZOHO web creator

The purpose of evaluations was to determine the conformance of the products to quality specification reported with the product. In this study, the quality information was extracted after examination of various selected datasets and their information recorded in respective fields of QIT. That information was then entered in ZOHO web creator. The web creator was a suitable tool because it was already hosted on the internet and therefore linking datasets from various websites was possible. The purpose of building web demonstration was to determine the relevance of this approach to management of disasters through an established portal.

The ZOHO web creator tools were used to create forms of templates for relevant fields of information input and eventually template views of such forms were generated which allowed viewing the entered fields and respective uploaded. The template forms were designed with drop down menus and radio buttons which allows the end-user to view only the available details of the product and not the whole list used by experts. Then the datasets used for demonstration were located on relevant websites and uploaded into the template forms. The result is an interface that allows the experts to submit several rated products and also allows the end-user to view the rated EO products and subject them to ranking as shown in Figure 16

Once the application is completely built, it can be made free accessible to search engines like Google, Yahoo, Explorer, etc. Depending on the level of privileges allowed, the user accessing the application may have the opportunity to edit the data entry fields which could be ideal in situation where streaming information is necessary or updating the rated information with information obtained from volunteer geographic information (VGI). A typical submission form may resemble the one in Figure 16

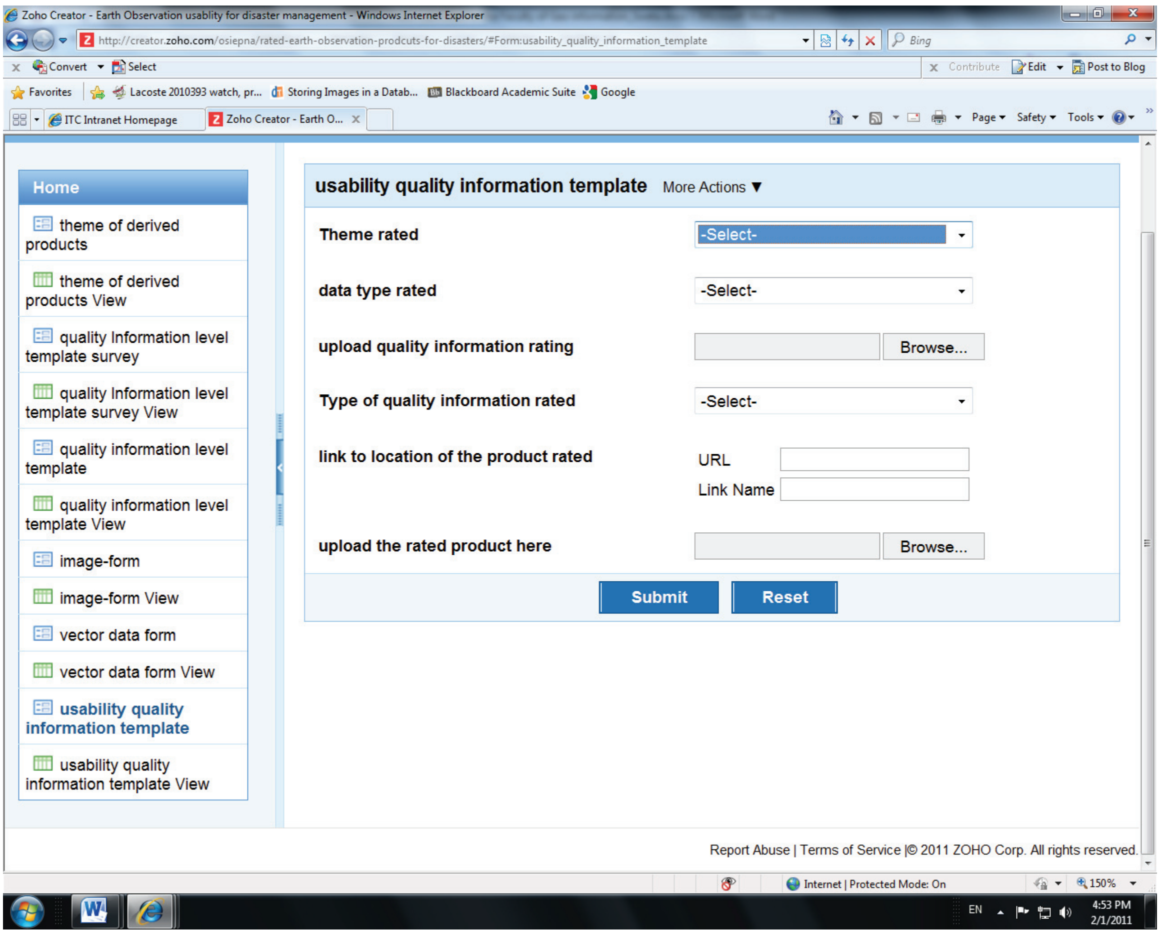


Figure 16: Rating submission form for experts

The objective of creating web template forms was to ensure that once team of evaluators are determined, they can evaluate the products at their various centres of work where they have relevant tools for handling a variety of data types and upload the information of their assessment to the disaster portal to be accessed by the general public.

After submission the user accessing the website or portal with the rated products should be able to select rated products from the homepage and find information about the product. The usability portal can be designed in such a way that depending on level of privileges allowed those accessing it can have the opportunity to view the rated information and associated product just like shopping catalogue of amazon or product reviews of mobile phones in GSM arena. Typical information display to the user should resemble illustration of Figures 17, 18 and 19.

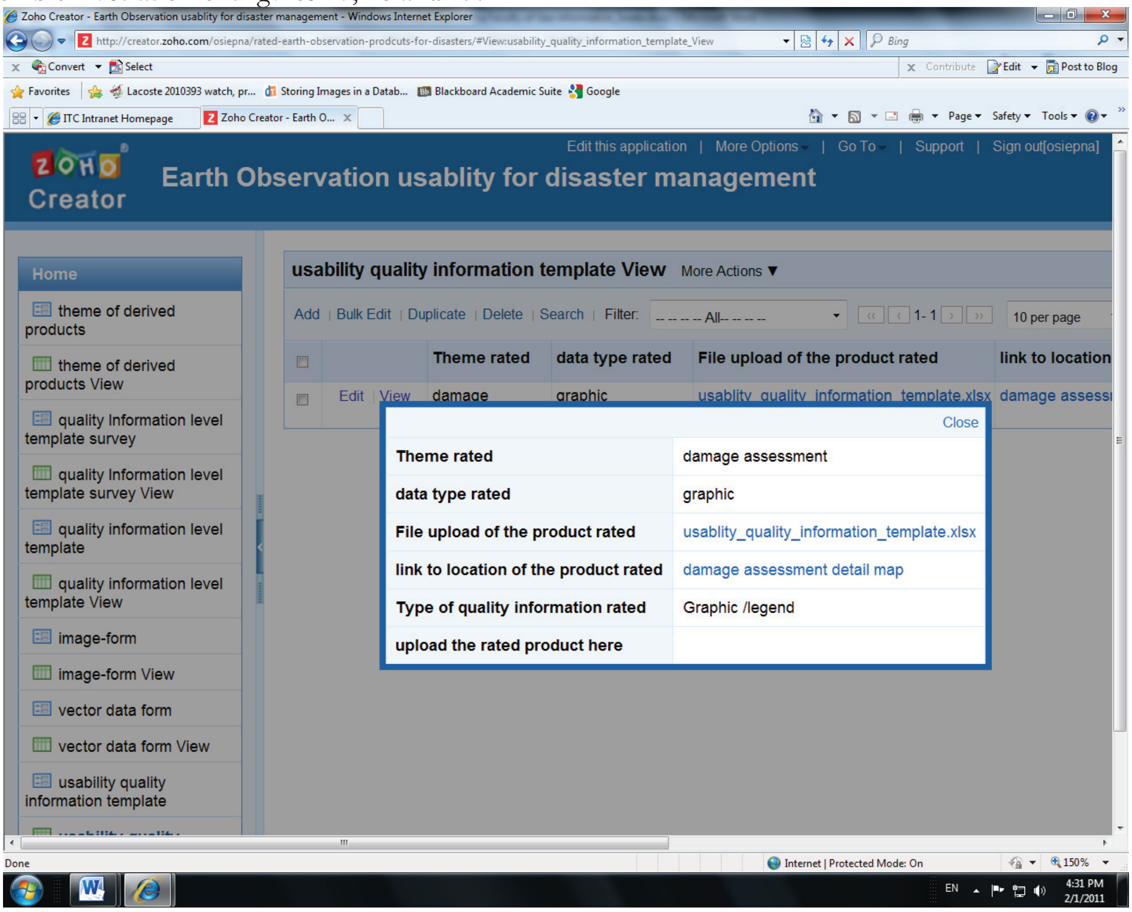


Figure 17: Screenshot of exploring quality of rated product

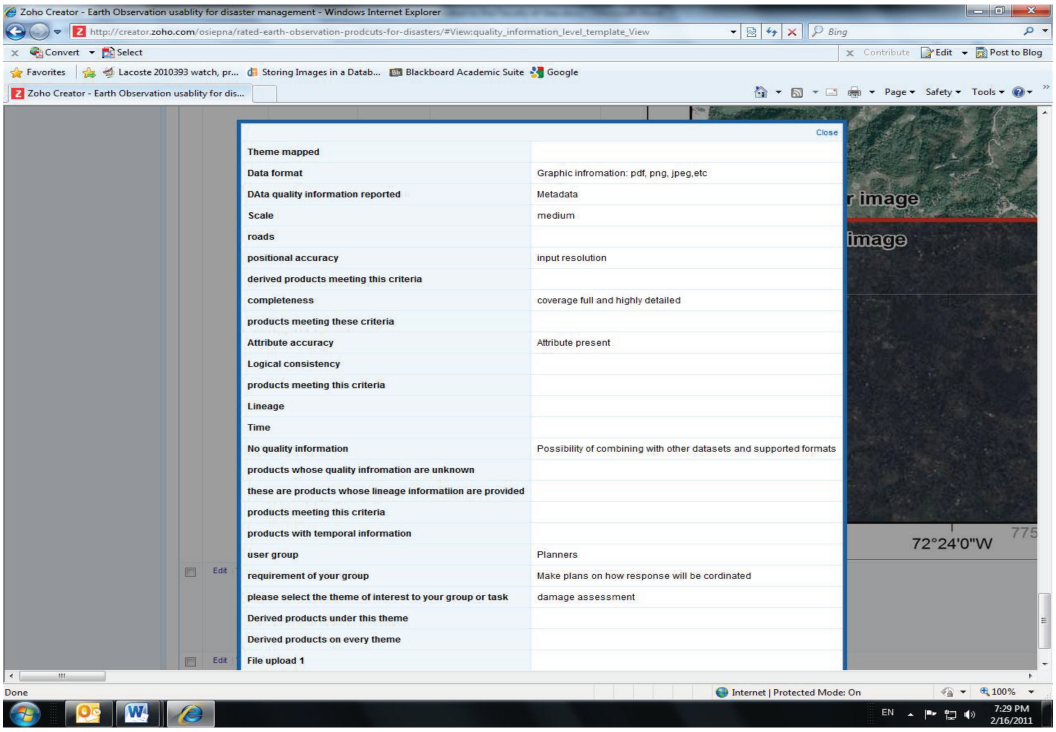


Figure 18: Relevance of QIT in reducing information overload

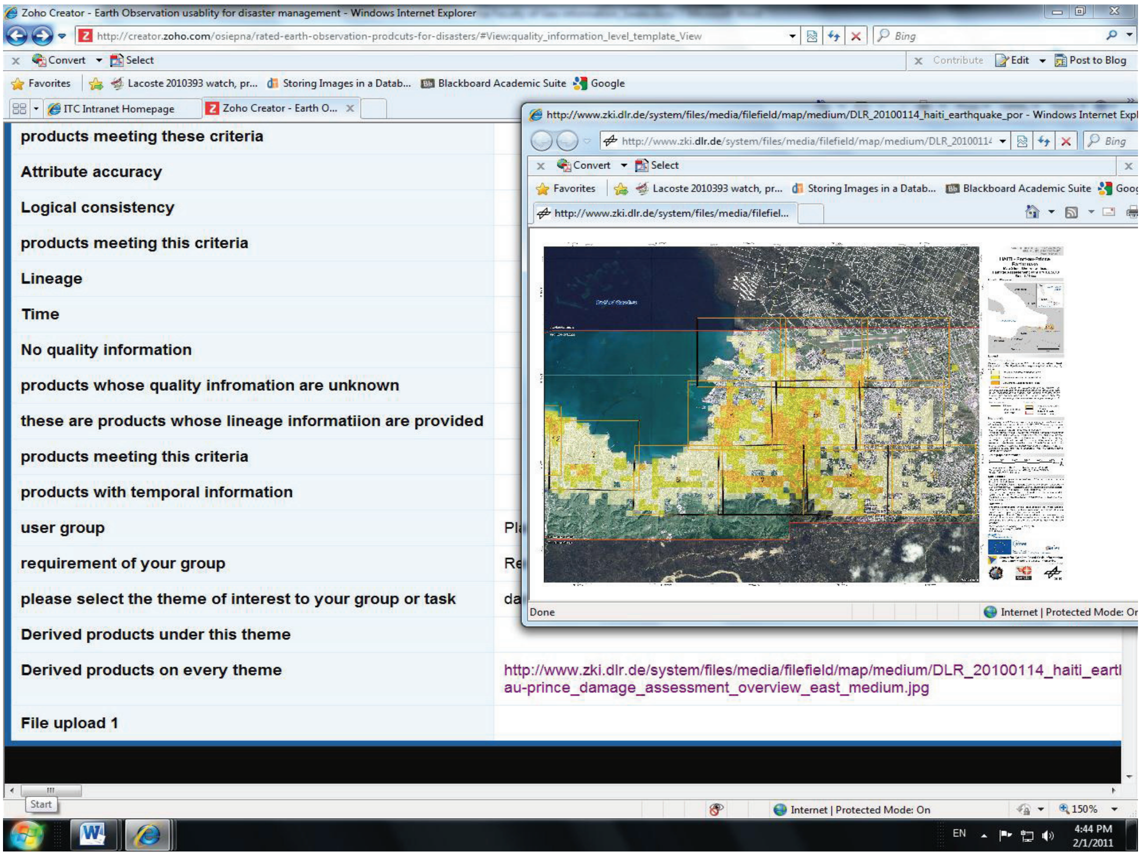


Figure 19: possible thumb nail of evaluated information

The primary purpose of evaluating the products is to improve the use and reduce information overload on the final user interested in making decisions towards a particular task. This is the primary purpose of *filter 2* in the methodology, where the end user access only those details that are present in the product and not the whole heuristic list used by experts during formal quality evaluation. This evaluation of quality to level of conformance can be a step further towards authenticating EO data for disaster management, which at present do not have any established formal standard for producing crisis data.

The derived EO products do not have any water mark or formal stamp of quality conformance. This study therefore, introduces the concept of authentication at the level of experts as a way of improving the use of EO data for disaster management. The evaluation does not brand a given product as bad or good, offers organised platform for comparison and selection of those attributes which the final user feels are the most relevant to subject to ranking among the possible datasets.

5.6. Description of ranking scenario

The selection of potential datasets is complete at this level and the next step is to select the relevant method for ranking. This methodology does not tie the user to a specific method of ranking; instead it is recommended that the method selected should allow determining quality differences between products. After the disaster has occurred, there are several users who are involved in disaster response and management. Every user's task has a set of data requirements which are critical for decision making and at times there may be different levels for rejecting or accepting the dataset, hence, the freedom to the user of method to carry out weighing and ranking.

5.6.1. Description of ranking scenario through Analytic Hierarchy Process (AHP)

Comparing overall quality of each product in relation to other products is not a straight forward process because a number of quality elements are considered before final selection of a given product is done. AHP has been used since 1980 in making complex decisions with a number of attributes (Saaty, 1980). The necessary steps are as follows:

i. The initial process is to define the problem or the task to be performed on a particular theme. The objective of ranking is to produce a value that represents the overall quality of a derived product based on input of selected data quality elements or parameters being evaluated. It is also possible to generate values that show relative importance of individual data quality elements in relation to each other. The AHP theme of disaster management (assess usability of a given number of number of datasets by evaluating their quality) and decide on task requirements (based on data quality elements defined by QIT parameters judgement entries for every product being considered). The problem definition outlines what the user would want to know and relevant attributes of the problem.

ii. The problem definition is followed by determining relative weights for each data quality element being compared (comparison attributes) using the values in AHP Table 7.

Table 7: Scale of relative importance (Saaty & Kearns, 1985)

Relative importance	***Value***
Assigned for equal contribution of two elements or quality attributes i.e. equal importance	1
Judgement slightly favours one element over another i.e. moderate importance	3
Experience and judgement strongly favours one element over another i.e. strong importance	5
Activity is strongly favoured and its dominance demonstrated i.e. Very strong importance	7
Evidence favouring one activity is of the highest possible order i.e. Extremely strong importance	9
Intermediate values when compromise is necessary.	2,4,6,8

The information in Table 6 is then used as guide in comparing two quality attributes at a time in the pairwise matrix comparison. The values are useful in gauging the level of importance when an evaluator assigns value to an activity or elements being evaluated.

iii. Create pairwise comparison matrices for every lower level governed by matrix on higher level because the effects of higher level matrix contribute to the effect of lower level matrix. The result is a square level matrix judgement indicating dominant element among the others; where $[n^2-n]/2$ judgements are necessary inputs in developing every matrix.
iv. Consistency of pairwise comparison is determined using the *Eigen value* consistency index evaluated using the departure of λ *max* from the number compared (n) with relative random entries which results in consistency ratio.
v. Compute priority vector and hierarchy synthesis by summing all the values in every column and dividing by resulting total weight.
vi. Determine the consistency of the whole hierarchy by multiplying every element by consistency index.

5.6.2. Prioritizing end user requirements

The use of AHP allows the user to prioritize their requirements as illustrated on Figure 20. Therefore, the decision making process can be structured by arranging conceptually related components into finer details under the overall goal in the hierarchy depicted like a tree diagram.

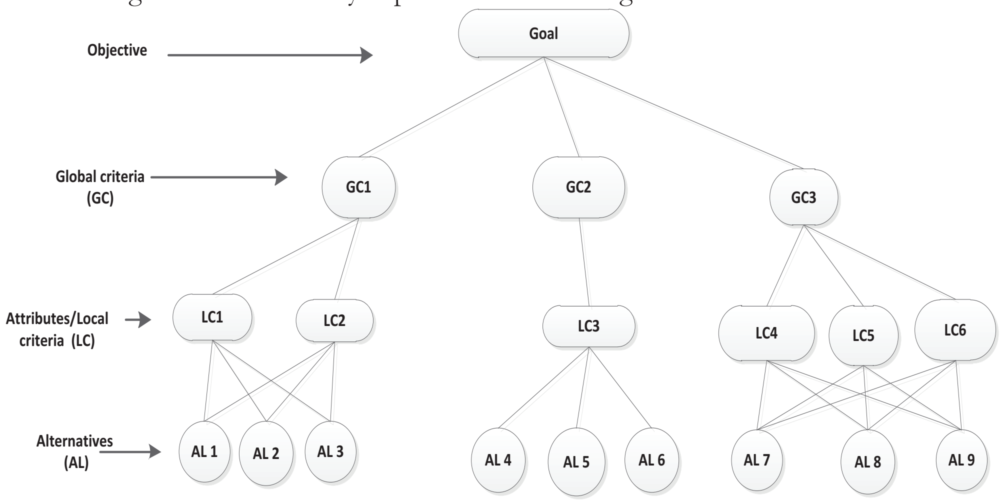


Figure 20: The AHP hierarchy model

Priority vectors are values that are most consistent with the pairwise values entered and refine user requirements into more precise statements of expectation (ReVelle et al., 1998). Therefore, the data quality information to be ranked in this case study can be modelled as shown in Figure 21.

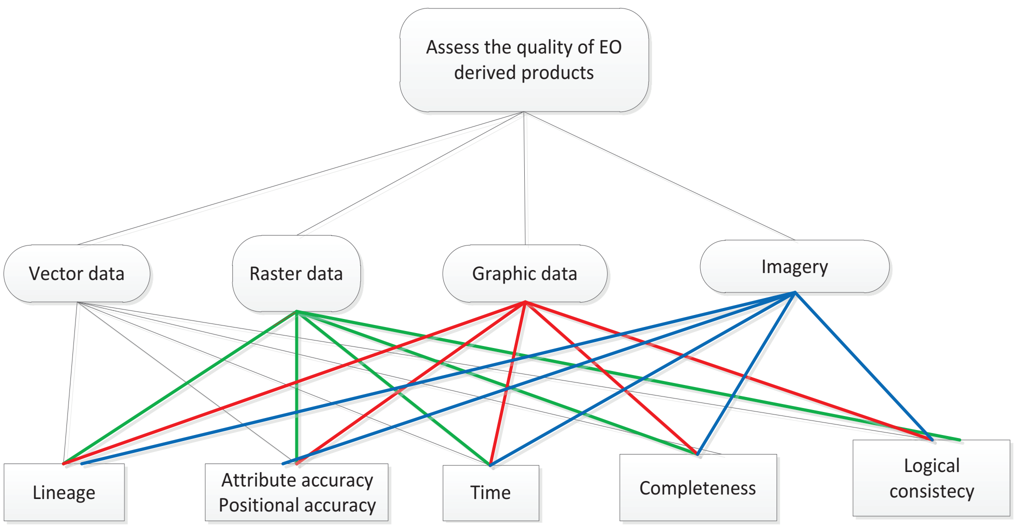


Figure 21: Possible AHP model indicating relationships between data types and spatial data quality elements

The illustration on Figures 20 and 21 shows how AH P allows structuring decision making process to be structured by arranging conceptually related components under higher level dimension. The objective is defined (task definition in disaster management) followed by global criteria (task requirements) or local criteria (attribute information data quality elements relevant for the task). The number of levels will depend on the significance of information that can be derived from the new level created. The alternatives can be perceived as sub categories of data quality elements which contribute to the overall quality of the datasets being assessed for their usability.

The decision making process can be structured by arranging conceptually related components into finer details under the overall goal in the hierarchy depicted in Figure 20. The preferences are set based on such relationships to allow calculation of weights as discussed in Section 5.6.3.

5.6.3. Setting of preferences, calculation of weights and ranking of EO derived products

This process involves steps 4 and 5 in the proposed methodology which follows after the user has identified the relevant attributes and the products to compare. The attributes compared here data quality elements lineage, logical consistency, attribute accuracy, time, lineage and positional accuracy. These are selected for purpose demonstration although it is possible to model any number of quality attributes. Without well-structured techniques, evaluation of different attributes becomes difficult even for a small number of attributes.

After selection of relevant attributes to evaluate on selected datasets, the weights of each comparison attributes are calculated on basis of relative importance. The relative importance Table 7 is used as guide in setting priority when analytical hierarchy process (AHP) is used for multi criteria decision making. When relative importance is between the defined scales, then the intermediate value is used, for example between 1 and 3, the value 2 is used and so on. The size of the matrix Table will depend on the number of attributes that are being evaluated. Suppose out of five datasets A, B, C, D and E in Table 5, a given user decides to examine A, B, and C after examining the values represented in QIT. Then the next step is to compute relative weights for assessment of quality to determine usability based on user preferences, then the initial variables could be *data type* and *data quality specifications*.

Selection of suitable relative importance values for the global variables being analysed and calculation of respective weights forms the initial process. In this example suppose data quality specifications falls in between "moderate importance (3)" and "strong importance (5)" compared to data type, then intermediate value 4 is used from the table. This literally means that data quality specifications will have 4 times influence in determining the weights of data quality information available among the selected datasets. This is summarised by table 8 where comparison of parameters to themselves yield value one because of equal importance. The relative importance of data type to quality specification was already stated as 4 in this example.

Table 8: Pairwise comparison of data type and data quality specification

	Data type	Data quality specification
Data type	1	0.25
Data quality specification	4	1

Values that are diagonal to one another in comparison matrix are inverses of each other by definition, thus we have ¼ as the value of relative importance of data type to data quality specification. The priority vector of these two variables will be column sum to the ratio of each entry, thus we have priorities as shown in Table 9

Table 9: Priority vector of data type and data quality

	Data type	Data quality specification
Data type	1/5	0.25/1.25
Data quality specification	4/5	1/1.25

The final values of priority vector are then calculated by average of each row which yields [1/5 + 0.25/1.25]/2 = 1/5 for data type and [4/5 + 1/1.25]/2 = 4/5 for data quality speciation. Therefore, the global criteria of weighting will be data type **0.2** for all datasets whose attributes are selected on the basis of data type and **0.8** for those datasets selected on basis of data quality specifications. An end user interested in selecting datasets on these two attributes will use a similar approach to find a value to multiply with the rating values in QIT.

The computation of sub level attributes follows where attributes of the two categories are considered, that is: for data type the attributes are vector, raster, graphic, or imagery while for data quality specifications, the attributes are lineage, attribute /positional accuracy, time, completeness, time and logical consistency. The next step is to compute priority vector for the data type and data quality specification. The calculations will follow similar approach, except the size of pairwise matrix becomes 4 by 4 instead of 2 by 2 as illustrated in Tables 9 and 10, since there are four variables. The relative priorities are determined based on information from QIT and resulting information will be similar to the one shown in Table 10.

Table 10: Relative importance and priority vectors for data type

Relative importance					
	Vector data	Raster data	Graphic	Imagery	
Vector data	1	2	5	7	
Raster data	1/2	1	3	4	
Graphic	1/5	1/3	1	2	
Imagery	3	1/4	1/2	1	
Total	**4.70**	**3.58**	**9.50**	**14**	
Priority vectors					
	Vector data	Raster data	Graphic	Imagery	
Vector data	0.213	0.559	0.526	0.5	
Raster data	0.106	0.279	0.316	0.286	
Graphic	0.043	0.092	0.105	0.143	
Imagery	0.638	0.070	0.053	0.071	

Relative pairwise comparison of the four attributes will depend on the values selected by the user on relative importance, thus vector is compared to raster, raster to graphic and then graphic to imagery. Suppose for a given user, vector data is of "between moderate importance and equal importance (value 2)", vector data is "of strong importance (value 5)" than graphic, and vector data is of "very strong importance (value 7)" than imagery; and that raster data is of" moderate importance value (3)" to graphic, raster is also "slightly above moderate importance (value 4)" and finally graphic to imagery as "value 2".

The computation yields priority vectors for vector data (**0.45**), raster (**0.247**), graphic (**0.096**) and imagery (**0.208**). These values are then used as weights to select a product based on data type and values are then used to multiply with selected fields of QIT. Similar approach is used to generate weights for data quality specifications, thus suppose the user is interested in three quality elements lineage, time and positional accuracy, then the pairwise matrix will look like the one shown in Table 11.

Table 11: Relative importance and priority vectors of data quality specification

Relative importance			
	Time	Lineage	Positional accuracy
Time	1	3	7
Lineage	1/3	1	2
Positional accuracy	1/7	1/2	1
Total	**1.476**	**4.5**	**10**
Priority vectors			
	Time	Lineage	Positional accuracy
Time	0.678	0.667	0.700
Lineage	0.226	0.222	0.200
Positional accuracy	0.097	0.111	0.100

The final values of priority vectors are time (**0.681**), lineage (**0.216**) and positional accuracy (**0.103).** From Tables 11 and 12, it can be observed that entity values of priority vectors will range between 0 and 1. Therefore, the normalised values of each priority vector become the *Eigen values* which tend to be consistent with the pairwise comparison values entered. At this stage, the relative weights have been determined and the next is to assess the quality of the datasets whose rating values were used to determine the relative weights for calculation of weights.

Assuming that the user wants to select a product on the basis of both data type and data quality description, then if there are 3 derived products A, B and C, then there will be seven elements to consider i.e. those of data type and data quality specification. The ranking vector is then calculated based on user preference per product type and the attribute information of either data type or data quality specification. The ranking vectors are computed in similar manner as the priority vectors. However, the choice is based on a single attribute for all the data types, for example if I was to choose on the basis of vector data, then my preference will be a, b, c, d; in other words, the user makes preferences on basis of attributes being considered for every single dataset being evaluated. Ranking vectors are then computed for all the variables being considered which results in the information displayed in Table 12.

Table 12: Ranking vectors for data quality specifications and data type

	Vector data				
Datasets	A	B	C	Ranking vector	….n
A	1	2	1	0.435	
B	0.5	1	3	0.384	
C	0.5	0.3	1	0.18	
Datasets	A	B	C	Ranking vector	
A	1	2	3	0.512	
B	0.5	1	4	0.360	
C	0.33	0.25	1	0.127	
Graphic					
Datasets	A	B	C	Ranking	

				vector	
A	1	3	5	0.648	
B	0.333	1	2	0.230	
C	0.2	0.5	1	0.122	
Imagery					
Datasets	A	B	C	Ranking vector	
A	1	2	4	0.532	
B	0.5	1	5	0.366	
C	0.25	0.2	1	0.102	
Time					
Datasets	A	B	C	Ranking vector	
A	1	2	3	0.539	
B	0.5	1	2	0.297	
C	0.33	0.5	1	0.163	
Lineage					
Datasets	A	B	C	Ranking vector	
A	1	2	4	0.571	
B	0.5	1	2	0.286	
C	0.25	0.5	1	0.143	
Positional accuracy					
Datasets	A	B	C	Ranking vector	
A	1	4	5	0.665	
B	0.25	1	3	0.231	
C	0.2	0.33	1	0.104	

At this stage all the relevant priority vectors have been calculated for the three levels as shown in Tables 10-13. The overall aggregate ranking for every dataset being evaluated is obtained by the sum of the product of weights of the three levels as shown in Table 13 whose priority vectors have been computed. Therefore, for the various datasets the computations will take the following form.

Table 13: Illustration of overall rank of all attributes

	Global criteria vector	Attribute priority vector	Datasets being evaluated						
			A	Final value	B	Final value	C	Final value	
Vector data	0.2	0.45	0.435	0.03915	0.384	0.03456	0.18	0.0162	
Raster	0.2	0.247	0.512	0.0252928	0.360	0.017784	0.127	0.0062738	
Graphic	0.2	0.096	0.648	0.0124416	0.230	0.004416	0.122	0.0023424	
Imagery	0.2	0.208	0.532	0.0221312	0.366	0.0152256	0.102	0.0042432	
lineage	0.8	0.216	0.571	0.0986688	0.286	0.0494208	0.143	0.0247104	
Time	0.8	0.681	0.539	0.2936472	0.297	0.1618056	0.163	0.0888024	

Positional accuracy	0.8	0.103	0.665	0.054796	0.231	0.0190344	0.104	0.0085696	
Total usability value				0.546128		0.302246		0.151142	
Percentage				55%		30%		15%	

The results of Table 13 are valid for decision making for the user who has done pairwise comparison on the three datasets. However, if the number of attributes or datasets is changed, the same argument does not hold and a new set of computations is done on the new set of preferences. This makes the ranking approach user centred because they set their own preferences and generate their own weights for the datasets being evaluated for a given task and nothing more. However, due to subjective nature of preference setting, results of one user should not be taken as an absolute value for all the users in a given user group, because every use case is different with its own set of task requirements.

5.7. Chapter summary

The application of the methodology on sample datasets has shown that it is possible to generate quality information on quality specifications of products and rate them through use QIT parameters. It was also shown that end-users, experts and products can be brought together on common platform of web interface (ZOHO) as quick and flexible means of uploading the rating and immediate access by the user. The web platform was seen as suitable for allowing users to view rating of various datasets just like shopping catalogue in order to select potential datasets. The ranking of potential datasets showed, it is possible to establish quality differences based on user preference.

6. DISCUSSION, CONCLUSIONS AND RECOMENDATIONS

6.1. Chapter overview

This chapter presents discussions on user groups, purpose of derived products and application of the methodology on case study. Conclusions drawn from research objectives and recommendations are also highlighted in this chapter as well.

6.2. Discussions

The ultimate goal of producing EO data and other related products is to benefit the end-user during disaster and crisis management. Since there are several end-users involved with various user task requirements, user centred approach for assessing usability of EO derived products is necessary. The challenge in disaster management is that, most end-users are not the producers of such datasets, although there could be a few cases where producers are also consumers of their own products. Therefore, since data is collected for a purpose which is 'not known' to the user accessing the product at disaster portal, assessing usability of such products is necessary.

The methodology presented in this study has addressed assessment of EO usability by identifying user groups, establishing purpose for which derived products are made, creating a common platform (QIT) for comparing product specifications to user requirements and link of quality description to possible application of the datasets achieved through the whole methodological flow.

This study has identified three user categories: *planners, responders* and *researchers* on the basis of *quality characteristics of EO datasets* and *task requirements* in different phases of disaster management. This is because due to current development in EO technologies and advancement caused by internet revolution, the problem is not the amount of data for disaster management; but the quality and utility. The classification of user categories adopted in this study contradicts the approaches taken by Fairbairn (2005) on the basis of duration of tasks; Byrom (2003) user needs and Kuntz et al., (1996) who classified users of disaster data in terms of their skills(*scientific user, commercial user* and *naïve user*) in handling EO data.

The reason for categorising user groups on the basis of quality characteristics of EO datasets and task requirements is to emphasize the fact that it the data is collected to meet the needs of certain group of users. Through this approach, the possible level of details that each product should contain can be determined and possible use contexts alongside the level of processing that may be required for each group. It is also possible to predict data accessibility mechanisms and purpose of the products. This classification integrates the quality of a product to its potential user.

This basis of categorization used in this study to identify user groups contrast with previous classifications of user groups, because knowing the duration of over which task is performed (Erharuyi & Fairbairn, 2005) does not provide information on the level of conformance of the dataset to the user task or how good the dataset is, but provide information on how long the user group will operate. Categorising user groups on the basis of user skills (Kuntz et al., 1996) limits the target user group technological know-how and addresses the user characteristics solely without considering the quality the products they want to use in disaster management. Therefore, the classification adopted in this study may assist providers of EO data to consider targeted users of their products in terms of quality they need to make decision. Which means user groups should be identified by considering important parameters of emergency context such *event*, *task* to be performed by user groups and *phase*. Similar parameters of emergency context were also identified by (Mulí ková et al., 2010).

The findings of this study indicates that derived EO product are produced are produced for direct use with no possibility of user modification and indirect use with possibility of user manipulation and modification before using the product. When providers or suppliers of various disaster products make explicit description of the purpose of the dataset, then the end-users accessing such products would probably find it easy to orient themselves to possible level of compromise associated with datasets meant for a given user group. This may include caveats on the contents of the products, the possible details that providers can deliver with a given product and the level of trust that producers and users can have on the products.

The level of details that shall be required by a particular group of users will be different and when that is taken into consideration, then use of such products would be enhanced and disaster management would be more efficiently managed when conformance is established and the level of detail indicated. By knowing the purpose, providers can determine the level of processing that is relevant for particular use context. Since purpose influences context of use and level of details to be considered during processing, it is a vital aspect in usability of EO data for disaster management.

Feedback from users can be obtained through use of interactive approach of user reviews, where depending on privileges from the web administrator may allow the user to post opinions or also provide their rating. Similar to how products are rated on consumer portals such as amazon, general subscriber module arena (GSM arena) where user can compare the properties of the products and select the potential set of products for subject to ranking.

The temporal component of EO data can be easily monitored because the system can always indicate various versions of the products that have been released from the expert evaluators. The temporal component can be useful in monitoring highly demanded data types, knowing the relevant parameters and determining those that need review, determining the update frequency of EO datasets, identifying critical areas that require immediate attention as well as knowing the magnitude the magnitude of use and the volume of products delivered to assist disaster management scenario.

The approach presented in this methodology contrasts sharply with current methods of spatial data usability like multidimensional user manual (Devillers et al., 2005) which is based on spatial distance and ranking by quality preferences (Yiu et al., 2010) which is based on minimum bounding box. These approaches require the end user to work with certain data type and have knowledge of spatial units, because they involve use of coordinates. The methodology presented in this study allows the experts to handle the technical and expertise aspects of data and end-user to evaluate relevance of content to task by setting preferences and assigning weight to generate a rank.

The ranking results from the end-user like those in section of case study favours the decision making of the particular use context and the same set can be evaluated differently by same user when preferences are changed. The methodology allows aggregation of various decision rules for setting preferences and establishing utility.

Derived products used for disaster management exhibit a variety of characteristics such as differences in data structure, geographic scale anomalies, duplication of information, and lack of structured quality report. These aspects limit the effective use of such datasets especially when the user cannot quantify the magnitude of quality differences among various preferences that may be available for a task at hand. Through the QIT and ranking, the elucidation of common quality elements of usability can be assessed and quantified through relative preference.

Quality awareness has been the field of field of craftsmen and technical experts in various fields. However, during disaster response, not all products are produced by experts and technically well trained craftsmen. Therefore, if the EO data from various disasters dedicated agencies can be produced according to a given standard and rating done by established team of experts, then the level of authenticity introduced will assist the end-users on deciding whether to integrate to integrate their potential choices with crowd sourced data or not.

Linking quality description to possible application of the data requires integrated approach that does not overload the user with unnecessary information to consider for decision making. The case study application demonstrate the importance of using distributive web portal to rate products because once standards, guidelines and rules of engagement have been are established, independent rating can be done by the experts trough distributed network and the rating automatically updated as they soon as the evaluator submits them.

The strengths of this approach are grounded in the fact that:

i. The methodology addresses the targeted user group from the perspective of both data quality characteristics and task requirements for decision making
ii. Targeted end-user do not analyse spatial units but sets preferences on the quality of the product after expert rating which takes care of evaluating the technical and expertise aspect of EO data.
iii. The user has the freedom to decide on the method of ranking and how they assign weight based on their own characteristics such as skills, experience, etc.
iv. Users only rank objective information and do not need to download information they will not use, but can interact freely in distributed system after conformance rating has been carried out by expert reviewers. T
v. The end-user accessing the EO product does not have to analyse complex spatial data structures where only coordinates are ranked and the user must know the bounding box as well as specification of spatial coordinates.

The QIT used in this approach provides harmonised way of organizing reported data quality information and reducing complexities of datasets due differences in data structures, duplication of information and also helps to understand data anomalies when various fields in the evaluated datasets are analysed.

The limitations of this methodology however, may revolve around lack of ready evaluators/reviewers immediately after the occurrence of a disaster. The methodology also relies solely on formal assessment because reviewers may not access the primary information from producers of such datasets. Moreover, the primary reason behind data commissioning remains unknown during this evaluation process. Despite these drawbacks, it remains a viable approach which creates a potential platform for monitoring the quality of EO data produced for disaster management.

6.3. Conclusion

The main objective of this study was achieved by the developed methodology with detailed criteria for its application. Example application of the methodology on sample datasets provided results that indicated that both the expert and end-user phase of the methodology were practical and quality differences among the sample datasets could be assessed. This objective was achieved through literature review, data collected from primary sources and case study application on EO data where user interface for demonstration on web platform (ZOHO) was created.

The results of this study have shown that knowing the user groups is useful in identifying the global requirements of the user category involved in disaster management. The results of the study also showed that the products are made for direct and indirect use which also has impact on the level of processing done on the product. The results of case study application on methodology developed has shown that a common platform such QIT can be useful in assessing usability of various EO data because it can allow distributive evaluation of data quality parameters by experts. Application of the methodology on sample datasets indicated that the method could be useful in harmonizing quality differences between various products and reduces information overload on the end-user. The end-user ranks the decision preferences on various products after expert rating on the product. Therefore, this methodology can be applied by various experts to rate products provided the necessary heuristics or parameters for a particular user group have been identified.

6.4. Recommendation

This study has presented a methodology for assessing usability of EO data for disaster management. However, there are still a number of challenges of use of EO data for disaster management that future research should address:

i. Developing heuristic parameter for each user group could be time consuming for a group of experts every time a disaster has occurred. Therefore, there is need for creating standard in producing EO data both for mapping (Dymon, 2003) and digital libraries (Parker & Stileman, 2005).

ii. Developing fully functional web portal for assessing usability with ranking tool and possibility of accommodating multi criteria decision making capability so that users only manipulates the weights after selecting the potential set of datasets.

LIST OF REFERENCES

Abdalla, R., Tao, C. V., & Li, J. (2007). Challenges for the Application of GIS Interoperability in Emergency Management. In J. Li, S. Zlatanova & A. G. Fabbri (Eds.), *Geomatics Solutions for Disaster Management* (pp. 389-405): Springer Berlin Heidelberg.

Agumya, A., & Hunter, G. (1999). A risk-based approach to assessing the 'Fitness for Use'of spatial data. *URISA Journal, 11*(1), 33-44.

Amin, S., & Goldstein, M. (2008). *Data against natural disasters: establishing effective systems for relief, recovery, and reconstruction*: World Bank Publications.

Andrienko, G., & Andrienko, N. (1999). Interactive maps for visual data exploration. *International Journal of Geographical Information Science, 13*(4), 355-374.

Bessisa, J. (2004). Use of the International Charter Space and Major Disasters For Damage Assessment. *International Archives of Photogrammetry and Remote Sensing (IAPRS), 35*.

Bishr, Y. (1998). Overcoming the semantic and other barriers to GIS interoperability. *International Journal of Geographical Information Science, 12*(4), 299-314.

Boin, A. (2008). *Exposing Uncertainty: Communicating Spatial Data Quality via the Internet*. Saarbrücken, Germany: VDM Verlag.

Boin, A., & Hunter, G. (2009). Do spatial data consumers really understand data quality information. In A. Stein, W. Shi & W. Bijker (Eds.), *Quality aspects in spatial data mining* (pp. 215-224). Boca Raton:CRC

Brabyn, L. (1996). Landscape classification using GIS and national digital databases. *Landscape Research, 21*(3), 277-300.

Byrom, G. (2003). *Data Quality and Spatial Cognition: the perspective of a National Mapping Agency*. Paper presented at the International Symposium on Spatial Data Quality, Maybush, Southampton UK.

Cai, C., & Xie, K. (2007). Measuring Data Quality of Geoscience Datasets Using Data Mining Techniques. *Data Science Journal, 6*, S738-S742.

Chrisman, N. (1984). The role of quality information in the long-term functioning of a geographic information system. *Cartographica, 21*(2/3), 79-87.

Comber, A., Fisher, P. F., Harvey, F., Gahegan, M., & Wadsworth, R. (2006). Using Metadata to Link Uncertainty and Data Quality Assessments. In A. Riedl, W. Kainz & G. A. Elmes (Eds.), *Progress in Spatial Data Handling* (pp. 279-292): Springer Berlin Heidelberg.

Couclelis, H. (1991). Requirements for planning-rellevant GIS: a spatial perspective. *Papers in Regional Science, 70*(1), 9-19.

Devillers, R., Bédard, Y., & Jeansoulin, R. (2005). Multidimensional management of geospatial data quality information for its dynamic use within GIS. *Photogrammetric Engineering & remote sensing, 71*(2), 205-215.

Devillers, R., Bedard, Y., Jeansoulin, R., & Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science, 21*(3), 261-282.

Devillers, R., Gervais, M., Bédard, Y., & Jeansoulin, Y. (2002, 21-22 March). Spatial data quality: from metadata to quality indicators and contextual end-user manual. *Joint Workshop on Spatial Data Quality Management,* 21–22.

Devillers, R., & Jeansoulin, R. (2006). *Fundamentals of spatial data quality*. London ;: ISTE.

Duckham, M. (2002). A user-oriented perspective of error-sensitive GIS development. *Transactions in GIS, 6*(2), 179-193.

Dymon, U. (2001). *'A standardized framework for hazards management mapping'*. Paper presented at the Proceedings 20th International Cartographic Conference, Beijing, China,.

Dymon, U. (2003). An Analysis of Emergency Map Symbology. Int. *Journal of Emergency Management, 1*(3), 227-237.

Erharuyi, N., & Fairbairn, D. (2005). Task-Centred Adaptation of Geographic Information to Support Disaster Management. In P. Oosterom, S. Zlatanova & E. Fendel (Eds.), *Geo-information for Disaster Management* (pp. 997-1008): Springer Berlin Heidelberg.

Fei, W., Renqiang, W., & Shaobo, Z. (2010, 25-26 Dec. 2010). *Key Issues in Mapping Technologies for Disaster Management.* Paper presented at the Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference

Fisher, P., Comber, A., & Wadsworth, R. (2008). Semantics, Metadata, Geographical Information and Users. *Transactions in GIS, 12*(3), 287-291.
Frank, A. (1998). Metamodels for data quality description. *Data quality in Geographic Information: From error to uncertainty*, 15-29.
Frank, A. (2001). Tiers of ontology and consistency constraints in geographical information systems. *International Journal of Geographical Information Science, 15*(7), 667-678.
Frank, A., Grum, E., & Vasseur, B. (2004). Procedure to Select the Best Dataset for a Task. In M. J. Egenhofer, C. Freksa & H. J. Miller (Eds.), *Geographic Information Science* (Vol. 3234, pp. 81-93). Berlin: Springer Berlin / Heidelberg.
Goodchild, M. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal, 69*(4), 211-221.
Grira, J., Roche, S., & Bédard, Y. (2009). Spatial data uncertainity in the VGI word: going from cosumer to producer. *Geomatics Solutions for Disaster Management, 64*(1), 61-71.
Grönlund, A. (2005). Methodology for Making Geographic Information Relevant to Crisis Management. In P. Oosterom, S. Zlatanova & E. Fendel (Eds.), *Geo-information for Disaster Management* (pp. 121-128): Springer Berlin Heidelberg.
Heuvelink, G. B. M., Temme, A. J. A. M., Schoorl, J. M., & Claessens, L. (2009). Chapter 5 Geostatistical Simulation and Error Propagation in Geomorphometry. In H. Tomislav & I. R. Hannes (Eds.), *Developments in Soil Science* (Vol. Volume 33, pp. 121-140): Elsevier.
Huang, Q., & Wang, F. (2007). *A methodology for definition and usage of spatial data quality rules.* Paper presented at the Geoinformatics 2007: Geospatial Information Science, Nanjing, China.
Hunter, G., Bregt, A., Heuvelink, G., Bruin, S., & Virrantaus, K. (2009). Spatial Data Quality: Problems and Prospects. [10.1007/978-3-540-88244-2_8]. 101-121.
Hutchison, V. (2010). Spatial Data Management Through Metadata: Global Concepts, Formats, Tools and Requirements (pp. 223-231).
Ito, A. (2005). Issues in the implementation of the International Charter on Space and Major Disasters. *Space Policy, 21*(2), 141-149.
Jahn, M., & Frank, A. (2004). *How to increase usability of spatial data by finding a link between user and data.* Paper presented at the "7th AGILE Conference on Geographic Information Science" Heraklion, Greece.
Jakobsson, A. (2006). *On the future of topographic base information management in Finland and Europe.* Helsinki University of Technology, Helsink.
Jankowski, P. (1995). Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Systems, 9*(3), 251 - 273.
Joos, G. (2006). *Data Quality Standards.* Paper presented at the Workshop 2 - International Standards Seminar I Shaping the Change XXIII FIG Congress, Munich, Germany.
Josselin, D. (2003). Spatial data exploratory analysis and usability. *Data Science Journal, 2*, 100-116.
Juran, J., Godfrey, A., Hoogstoel, R., & Schilling, E. (1998). *Juran's quality handbook*. New York: McGraw-Hill.
Kerle, N., & Widartono, B. (2008). Geoinformation-Based Response to the 27 May Indonesia Earthquake–an Initial Assessment. *Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters*, 11-23.
Köhler, P. (2005). User-Oriented Provision of Geo-Information in Disaster Management: Potentials of Spatial Data Infrastructures Considering Brandenburg/Germany as an Example. *Geo-information for Disaster Management*, 171-179.
Kuhn, W. (2001). Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science, 15*(7), 613-631.
Kuntz, S., Streck, C., Kessler, C., GmbH, V., & Lavalle, C. (1996). Earth Observation for Identification of Natural Disasters-EOFIND. *INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING, 31*, 207-212.
Lorincz, K., Malan, D. J., Fulford-Jones, T. R. F., Nawoj, A., Clavel, A., Shnayder, V., et al. (2004). Sensor networks for emergency response: challenges and opportunities. *Pervasive Computing, IEEE, 3*(4), 16-23.
Lowell, K., & Casault, P. (2004). *Why aren't we making better use of uncertainty information in decisionmaking?* Paper presented at the Proceedings of 6th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences
Lowell, K., & Jaton, A. (1999). *Spatial accuracy assessment: land information uncertainty in natural resources*: CRC.
Mihaila, G., Raschid, L., & Vidal, M. (2001). Source selection and ranking in the websemantics architecture using quality of data metadata. *Advances in Computers, 55*, 89-119.

Morrison, J., & Veregin, H. (2010). Spatial Data Quality. In J. Bossler (Ed.), *Manual of Geospatial Science and Technology, Second Edition* (Vol. null, pp. 593-610). New york: CRC Press.
Mountrakis, G., Stefanidis, A., Schlaisich, I., & Agouris, P. (2004). Supporting Quality-Based Image Retrieval Through User Preference Learning. *Photogrammetric Engineering and remote sensing, 70*(8), 973-981.
Mulí ková, E., Šafr, G., & Stan k, K. (2010). *Context map: A tool for cartography support in crisis management*. Paper presented at the 3rd International conference on Cartography and GIS, Nessebar, Bulgaria.
Nedovic-Budic, Z., & Pinto, J. (1999). *Understanding interorganizational GIS activities: A conceptual framework*.
Nielsen, J. e., & Mack, R. L. e. (1994). *Usability inspection methods*. New York: Wiley & Sons.
Orhan, A. e., Backhaus, R. e., Boccardo, P. e., & Zlatanova, S. e. (2010). *Geoinformation for disaster and risk management : examples and best practices*. Copenhagen: International Federation of Surveyors (FIG).
Orr, K. (1998). Data quality and systems theory. *Communications of the ACM, 41*(2), 66-71.
Parker, C., & Stileman, M. (2005). Disaster Management: The Challenges for a National Geographic Information Provider. In P. Oosterom, S. Zlatanova & E. Fendel (Eds.), *Geo-information for Disaster Management* (pp. 191-214): Springer Berlin Heidelberg.
Pundt, H. (2008). The semantic mismatch as limiting factor for the use of geospatial information in disaster management and emergency response *Geospatial Information Technology for Emergency Response* (pp. 243).
Radke, J., Cova, T., Sheridan, M. F., Troy, A., Mu, L., & Johnson, R. (2002). Challenges for GIS in Emergency preparedness and response. *Greene, RW, Confronting Catastrophe: A GIS Handbook, ESRI*.
ReVelle, J., Moran, J., & Cox, C. (1998). *The QFD handbook*: John Wiley & Sons Inc.
Saaty, T. (1980). The analytical hierarchical process: New York: McGraw Hill.
Saaty, T., & Kearns, K. (1985). *Analytical planning*: Pergamon press Oxford.
Stevens, D. (2008). Increasing the Use of Geospatial Technologies for Emergency Response and Disaster Rehabilitation in Developing Countries. In S. Nayak & S. Zlatanova (Eds.), *Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters* (pp. 57-71): Springer Berlin Heidelberg.
Su, Y., Peng, J., & Jin, Z. (2009). Assuring Information Quality in Sharing Platform for Disaster Management. *Journal of Software, 4*(9), 915.
Timpf, S., & Raubal, M. (1996). Experiences with metadata. *7th Int. Symposium on Spatial Data Handling, SDH'96, Delft,The Netherlands*, IGU,12B31-12B43.
UNISDR. (2009). Terminology on Disaster Risk Reduction. Retrieved 18/11/2010, from http://www.unisdr.org/eng/terminology/UNISDR-terminology-2009-eng.pdf
URL2.1. FP6–Priority 2.3. 2.9 Improving Risk Management Integrated Project. Retrieved 13/10/2010, from http://www.oasis-fp6.org/
URL2.2. Charter Activation: Earthquake Haiti. Retrieved 5/8/ 2010, from http://www.disasterscharter.org/web/charter/activation_details?p_r_p_1415474252_assetId=ACT-287
URL2.3. CIA The World Factbook. Retrieved 31/8/2010, from https://www.cia.gov/library/publications/the-world-factbook/geos/ha.html
URL2.5. (2009). VictorianspatialCouncil
Retrieved 22/09/2010, from http://victorianspatialcouncil.org/cms/library/attachments/SIMF%20Data%20Quality%20Guidelines%20Edition%202_September%202009.pdf
Van Oort, P., & Bregt, A. (2005). Do users ignore spatial data quality? A decision-theoretic perspective. *Risk Analysis, 25*(6), 1599-1610.
Van Oosterom, P., Zlatanova, S., & Fendel, E. (2005). *Geo-information for disaster management*: Springer Verlag.
Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., & Mehl, H. (2007). Satellite Image Analysis for Disaster and Crisis-Management Support. *Geoscience and Remote Sensing, IEEE Transactions on, 45*(6), 1520-1528.
Wachowicz, M., Hunter, G. J., & Bregt, A. K. (2003). Understanding Spatial Data Usability. *Data Science Journal, 2*, 79-89.
Wadsworth, R., Comber, A. J., & Fisher, P. F. (2009). Latent Analysis as a Potential Method for Integrating Spatial Data Concepts *Research Trends in Geographic Information Science* (pp. 123-133): Springer Berlin Heidelberg.
Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of management information systems, 12*(4), 33.

Yang, T. (2007). Visualisation of Spatial Data Quality for Distributed GIS. *The University of New South Wales, Sydney*.
Ying, S., Jie, P., & Zhanming, J. (2009). *Information Quality Assurance Models for Experts Assessing in Disaster Management* (Vol. 4).
Ying, S., & Lei, Y. (2008, 21-22 Dec. 2008). *Assuring Image Quality in Spatial Data Sharing Platform for Disaster Management*. Paper presented at the Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing. ETT and GRS 2008. .
Yiu, M., Lu, H., Mamoulis, N., & Vaitis, M. (2010). Ranking Spatial Data by Quality Preferences. *IEEE Transactions on Knowledge and Data Engineering, 99*(PrePrints).
Zargar, A., & Devillers, R. (2009, 1-7 Feb. 2009). *An Operation-Based Communication of Spatial Data Quality*. Paper presented at the Advanced Geographic Information Systems & Web Services, 2009. GEOWS '09. International Conference on.

Appendices

Appendix 1: Survey questions to producing agencies

"Disaster management: Quality of products derived from Earth observation" Exit this survey

1. Introduction to the survey[Producing agencies]

This survey is designed to collect information about quality aspects considered by producing agencies during the production of disaster response products. The main aim is to develop a methodology that will allows users of products from various suppliers to compare quality as they evaluate various potential products for a given task. This information will be compared with user requirements to establish a link between what users may expect from producers immediately in the aftermath of a disaster in order to make informed choices as they select the product to use. This research considers products explicitly and not the organization that produces them.

The results of the study WILL NOT be pointing at a particular organization as the one providing suitable products than the other; but as a point of infromation of the quality aspects considered by producers of disaster response products.I hope the analysis derived from information that you will provide here will be beneficial in ensuring the products you produce are useful to the potential users. This is part of data collection for my theses project and i will appreciate your efforts of filling it. Thank you.

For comments or questions concerning this survey, contact me through this email [sweta23946@itc.nl], Thank you.

General instructions:
1.left mouse click as you select/check your choices
2. Once you are through, remember to click DONE at the bottom of the page.

1. Kindly fill in the following details[type in the boxes provided]

Name

Organization/company

Job title

Years of experience

Email

Major types of products produced for disaster management

2. After activation of disaster charter following the Haiti earthquake of January 2010, what kinds of Earth observation products did you to make available? [Please check, more than one answer is allowed]

- ☐ Printable maps in PDF, JPEG, etc.
- ☐ GIS vector layers
- ☐ on line resources data like KML.KMZ, shape files
- ☐ Unprocessed images
- ☐ GIS data is made available as report

Other (please specify)

3. Did these products contain information on their quality?

Yes

No

If YES, what information did you provide?

If No, how did you portray the quality of the product to enable its usability?

4. After the Haiti earthquake of 12 January 2010,who were the SPECIFIC users of your products [please check, more than one is allowed]

- [ ] Planners:organization(s) in charge of coordinating the disaster i.e. Government of Haiti, MINUSTAH GIS,etc
- [ ] Police and fire organizations for performing search and rescue
- [ ] Military for security and associated services
- [ ] Health and emergent services users
- [ ] Camp management users
- [ ] Water and sanitation organizations
- [ ] Emergency shelter services organizations
- [ ] Food and logistics organizations
- [ ] Users in the communication infrastructure and emergency operation center
- [ ] Users in eventual reconstruction and restoration of pre-disaster condition
- [ ] Researchers:Users who apply technology to the products derive more more meaning by generating models, etc

Other (please specify any additional information in the text box below)

5. How does your organization define user requirements or the kinds of products you produce?

[Please check, more than one answer is allowed]

- ☐ Based on user requests
- ☐ Determined by coordinating agencies
- ☐ We have already established standards on what to produce

How do you create your user inventory? (please, type in the text box below)

6. Which of the following quality component/element(s) are most rellevant in your oganization during the creation of METADATA and for which derived product?

Lineage: showing history of product i.e. purpose, and sources used in production

Attribute accuracy: emphasis placed on consistency of attributes and feature representation

Positional accuracy: the concern is ensuring that features are placed on the correct location

Logical consistency: the data obeys the rules of data structure, attributes and relationships

Completeness: you show which features are present/absent and extent of area covered

Currency: the current up to datedness is well documented

[please check as appropriate]

	Lineage	Attribute accuracy	Positional accuracy	Logical consistency	Completeness	Currency(Time)
Printable maps in PDF, JPEG, etc.	☐	☐	☐	☐	☐	☐
GIS vector layers	☐	☐	☐	☐	☐	☐
on line resources data like KML.KMZ, shape files	☐	☐	☐	☐	☐	☐
Unprocessed images	☐	☐	☐	☐	☐	☐

GIS data is made available as report	☐	☐	☐	☐	☐	☐

Other (please specify any additional information in the text box provided below)

7. Which contexts of use for the end-user do you consider during your production of disaster response derived products? [please check, multiple answers allowed]

- ☐ Office desktop
- ☐ Portable digital devices like PDA, laptops, mobile phones, etc.
- ☐ Field operation especially immediate rescue activities

Other (please specify any additional information in the text box provided below)

8. What are the sources of data used by your organization to produce the disaster response products? [Please check, more than answer is allowed]

- ☐ Existing topographic maps,
- ☐ Satellite images
- ☐ Aerial photography
- ☐ Sketch maps
- ☐ Field information collected by personnel deployed

Other (please specify any additional information in the text box provided below)

9. What is the update strategy of the products from your organization during disaster response

phases? [Please check, single answer only]

	Daily	weekly	Every fortnight	Monthly	6-12 months	Never updated
Recovery and reconstruction	○	○	○	○	○	○
Early recovery	○	○	○	○	○	○
Immediate resposne	○	○	○	○	○	○
mitigation and preparation: risk assessment	○	○	○	○	○	○

Other (please specify any additional information in the text box provided below)

10. How is auditing of the quality of the products you produce done? [Please check, only one answer is allowed]

○ Established bodies like OGC,ISO,ICA,etc

○ coordinating agencies in charge of the disaster in question

○ We have established quality assessment team

○ We don't consider quality evaluation

Other (please specify any additional information in the text box provided below)

Done

Appendix 2: Survey questions to User Organizations

Appendix 2: : Survey questions to user organizations

[SURVEY PREVIEW MODE] Disaster management: User of earth observation based products" Survey

Disaster management: User of earth observation based products" Exit this survey

1. Introduction and reason for the study[User Organizations]

This survey is designed to collect information about quality aspects considered by users as they make choices where there are several potential datasets which are fit for use during disaster response and related activities. The main aim is to develop a methodology that would allow users to rank and compare various products in terms their quality from a variety of several potential products.

Through your participation in this survey, I expect to understand the kinds of products you use and the quality elements you consider while evaluating kinds of products to use in disaster response.

I hope you will take time to fill this short survey, without people like you, research on usability of derived products used for disaster response and management would not be conducted. Which means problems experienced by user of the disaster products will only be discovered by participation of respondents like you.

Thank you.

For comments or questions concerning this survey, you can contact me on this email [sweta23946@itc.nl].

General instructions:
1. left mouse click as you select/check your choices
2. Once you are through,remember to click DONE at the bottom of the page.
3. questions with star (*) requires at least one answer

1. Kindly fill in your personal details details below [please type the details in the text boxes provided]

Name

Name of your organization

Position/job title

Years of experience in disaster management

Email address

2. Which phase(s) of disaster management do you operate?[you can check/tick more than one]

- ☐ Immediate response: relief phase for search and rescue,evacuation,provision of basic human needs, etc
- ☐ Early recovery: initial phases where basic facilities are provided like expanded health care, sanitation and education
- ☐ Recovery and reconstruction:actions and decisions are taken to restore pre-disaster living conditions
- ☐ Mitigation and prevention: risk assessment phase for reducing vulnerability and determining type of relief needed.

Other (please specify by typing text in the box below)

http://www.surveymonkey.com/...K_FOR_COLLECTION&sm=Z2QUeRG6aQmEH2voahteSvV1umHZ0XRRTJYkUq%2fpYQs%3d[2/19/2011 12:50:40 PM]

***3. Which of the following Earth Observation derived product(s) did you use/still using after the Haiti Earth quake of January 2010?[Please check, more than one is allowed]**

	Recovery and reconstruction	Early recovery:immediate basic needs	Mitigation and prevention: risk assessment	Immediate response
Raster data as classified/non classified image(s)	☐	☐	☐	☐
Map products in the form of pdf, png,	☐	☐	☐	☐
Vector data in the from shape files	☐	☐	☐	☐
on line resources in the form of shape files, KML, KMZ,etc	☐	☐	☐	☐
Spatial information provided in form of reports	☐	☐	☐	☐

Other (please specify by typing text in the box below)

***4. Which of the following data quality elements did you consider while selecting Earth observation derived product(s)?[please check, multiple answers allowed]**

	Raster data as classified/non classified image(s)	Map products in the form of pdf, png,	Vector data in the from shape files	on line resources in the form of shape files, KML, KMZ,etc	Spatial information provided in form of reports
completeness: the extent of area covered by products that are available	☐	☐	☐	☐	☐
positional accuracy showing correctness of features in the ground	☐	☐	☐	☐	☐
attribute accuracy: data fields,presence or absence of attribute information	☐	☐	☐	☐	☐
Lineage: history of dataset showing purpose,sources used in production,associated transformation,and assumptions	☐	☐	☐	☐	☐
logical consistency:correct positioning of landscape objects in relation to each other	☐	☐	☐	☐	☐
time: when the					

product became available	☐	☐	☐	☐	☐

Other (please specify by typing text in the box below)

5. How did you handle products when no quality information is provided? [check one only]

- ○ Combine with other datasets
- ○ Evaluate quality by comparing with similar products
- ○ I don't use the product with no quality information
- ○ don't know
- ○ Did not consider

Other (please specify by typing text in the box below)

6. Which combination of products do you prefer immediately before starting any activity in disaster response/management?[Please check, more than one answer is allowed]

- ☐ Vector data in shape files and kml
- ☐ Map products in printable formats
- ☐ on line resources in kml,kmz,or shape files
- ☐ Spatial information provided as report
- ☐ raster information provided as classified/unclassified images
- ☐ Unprocessed image

Other (please specify by typing text in the box below)

*** 7. Which of the following spatial data quality elements did you consider while evaluating the datasets? [Please check to rank with 1 being the most most prefered and 5 the least]**

	1 most important	2 important	3 necessary	4 slightly necessary	5 least important
Attribute accuracy: it provided opportunity to evaluate the theme of interest for decision making	○	○	○	○	○
Positional accuracy: the features represented in the dataset were perceived to be close to their position in reality	○	○	○	○	○
Up to datedness: the accuracy of temporal attributes and relationships showed current information.	○	○	○	○	○
Data interoperability: the dataset allowed more room for easy download, was in the right format	○	○	○	○	○
Logical consistency: the representations of the objects allowed me to establish the relationship of objects to each other like differences between valley and ridge.	○	○	○	○	○
Lineage: the dataset provided clear purpose, sources used in production and accuracy level which could be achieved	○	○	○	○	○
Completeness: from the dataset, I was able to know extent and coverage that I expected.	○	○	○	○	○

8. Are there cases in which decisions have been made based on unsatisfactory quality information? [please check one]

○ Yes

○ NO

9. Are there cases in which decisions have been made based on unsatisfactory quality information?[please, type]

If yes, what was the problem faced? For example incorrect positioning of features, etc

If No, what do you do with the products whose quality does not satisfy your requirements?

Yes

NO

10. Would you be interested in the outcome of this research? [check only one]

- Yes
- No
- does not matter

Done