An approach of a data-driven Spatial Decision Support System to manage the effects of Climate Change on agriculture

David Gerardo Navarro Ferrer[1], Tatiana Delgado[2], Gustavo Martin[3], Efren Jaimez[4]

Abstract

Climate change is an important challenge of our era. Enhancing resilience to climate change, particularly to reduce its effects on agriculture is crucial to improve food security and achieve sustainable development.

Persistent drought and extreme weather events affect the agricultural calendar and crops, particularly in developing countries. Crop harvesting is being affected due to higher air temperatures that reduce the daily temperature variability and increase plagues and sicknesses. Sea-level rise and frequent flooding are other factors that threaten food security and efforts to eradicate poverty and achieve sustainable development. Land degradation is one of the indirect effects of Climate Change.

Within the framework of the Project "Environmental Bases for Local Alimentary Sustainability", sponsored by EU and UNDP (2012-2017), this work is aimed to provide decision-making tools to support new Climate Change adaptation policies. Particularly, a decision support solution to analyze land degradation as one of the driving force Climate Change effects is discussed. An index of degradation is computed by a weighted sum according to Saaty prioritization mechanism.

Common spatial analyses from Geographical Information Systems have been used in the study area to combine different factors (layers) and produce the layer of degradation indexes. Such a layer is used as input data to the ETL component in a Business Intelligence solution. As a typical data-driven Spatial Decision Support System, other components as data warehousing, OLAP and spatial reporting are also presented; as well as, the software tools developed to support this BI solution.

Key words: Climate Change, Agriculture, Data-driven Decision Support System, Business Intelligence, Spatial Decision Support System

1. Introduction

Any discussion on technical support for decision-making in different domains should consider a holistic approach to human problem-solving, the concrete environments in which decision support systems (DSSs) will be used, and the acceptance of the system by the user. The ultimate goal is a system under which international organizations, governments, local authorities, and individuals are able to conduct negotiations as well as coordinate and evaluate their own independent decisions.[1]

This paper deals with this problematic situation, providing an approach of a data-driven Spatial Decision Support System to characterize spatially the combination of agriculture vulnerabilities. It could contribute to provide possible adaptation scenarios for the impact of Climate Change on agriculture.

Particularly relevant for this paper is to highlight, from these researches, some key definitions:

[1] Environment Agency, Ministry of Science, Technology and Environment, Cuba ingdavid.gerardo@gmail.com. Environment Modeling Department

[2] Industrial Engineering Faculty, ISPJAE, Cuba, tdelgado@ind.cujae.edu.cu. Information Systems Department

[3] Tropical Geography Institute, IGT, Cuba, gmartin@geotech.cu. Geomatic Department

[4] Geophysical and Astronomy Institute, IGA, Cuba.

Decision Support System: Vercellis defined DSSs as "an interactive computer-based application that combines data and mathematical models to help decision makers solve complex problems faced in managing the public and private enterprises and organizations".[2]

Data-driven Decision Support System: data-driven DSSs emphasize access to and manipulation of a time series of internal company data and sometimes external and real-time data. Data-driven DSSs with online analytical processing provide the highest level of functionality and decision support that is linked to analysis of large collections of historical data.[3-9]

Business Intelligence: The term BI is a popularized, umbrella term coined and promoted by Howard Dresner of the Gartner Group in 1989. It describes a set of concepts and methods to improve business decision making by using fact-based support systems. In general, business intelligence systems are data-driven DSSs.[10]

A first part of this work deals with a multi-criteria evaluation based on an Analytical Hierarchical Processing (AHP) technique. As a result, a land degradation index spatially distributed is obtained. Second part of the paper is dedicated to describe the methodological issues of a Business Intelligence solution to complement the SDSS. According to a data-driven SDSS classification, different components for a Geospatial Business Intelligence Platform are described in integration with business needs, potential reports, spatial visualization tools and analytical results.

1.1. BASAL Project

The project "Environmental Bases for Local Alimentary Sustainability" (BASAL) aims to support Cuba's adaptation to Climate Change and contribute to the country's continued and sustainable economic and social development by improving the national food security. More specifically, the project aims to decrease the vulnerabilities of the national agricultural food production related to climate change and variability; through:

- The identification and reduction of Climate Change impacts which affect major national food production;
- The promotion of good/best practices and provision of useful and accessible information which eventually help farmers produce food in a sustainable manner;
- The provision of scientific- and experienced-based tools and recommendations which enhance the decision making capacity in relation to the Climate Change challenge of key actors at all relevant levels.

This project is targeted to 3 municipalities with different characteristics and important relevance in the economy of Cuba. Güira de Melena is one of them and will be the selected study area for this paper. This mainly rural municipality on the Southern coast of the province Havana produces a large part of the food (fruits, vegetables and meat) reaching the 2M inhabitants of the province, including the capital city of Havana. It is affected by coastal flooding and saline intrusion as far as 6 km from the sea, coastal erosion, and land degradation.[11]

Land degradation in Güira de Melena is the result of the combination of different factors. By determining their priorities regarding soil and vegetation factors, this work builds a solution of Business Intelligence with the view to make diverse analyses supporting decision making.

2. Methods and Tools

Institute of Soil together Institute of Geophysics and Astronomy in Cuba undertakes often different studies to support decision making in this field. However, less frequently they have conducted studies to determine a land degradation composite index. Developing a study to support this problematic becomes in a key factor into the framework of BASAL.

In consultation with experts, the following variables, shown in the Table No. 1, were identified as factors to determine land degradation.

Soil	Vegetation
Effective depth	Vegetation Formation
Rockiness	Intervention Degree
Stoniness	Vegetation density
Slope	
Soil Erosion	
Salinity	
Saturation	
Soil sub-type	

Table 1: Factors identified in the model.

2.1. Mathematical Model. Analytic Hierarchy Process Method

The eigenvalue method for assessing attribute weights and single-attribute value functions is part of a general methodology called "Analytic Hierarchy Process"; it consists in structuring the decision problem in a hierarchical manner, constructing numerical evaluations associated with all levels of the hierarchy and aggregating them in a specific fashion, formally a weighted sum of single-attribute value functions.[12]

To make a decision in an organized way to generate priorities we need to decompose the decision into the following steps:

1. Define the problem and determine the kind of knowledge sought.
2. Structure the decision hierarchy from the top with the goal of the decision, then the objectives from a broad perspective, through the intermediate levels (criteria on which subsequent elements depend) to the lowest level (which usually is a set of the alternatives).
3. Construct a set of pairwise comparison matrices. Each element in an upper level is used to compare the elements in the level immediately below with respect to it.
4. Use the priorities obtained from the comparisons to weigh the priorities in the level immediately below. Do this for every element. Then for each element in the level below add its weighed values and obtain its overall or global priority. Continue this process of weighing and adding until the final priorities of the alternatives in the bottom most level are obtained.[13, 14]

Following the AHP proceeding, the main decision objective was identified as follows: "*determining spatially a degradation index in the municipality Güira de Melena*". To determine such an index, a set of factors grouped in two classes (soil and vegetation) were identified.

Saaty´s scale is usually divided into 9 intervals, as depicted below in the Table No. 2.

Value	Definition	Explanation
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favor one activity over

		another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation

Table 2: Saaty´s scale to prioritize criteria (Source: [13])

If one wants to apply AHP in a multiple criteria decision problem, pair-wise comparisons of the alternatives must be performed for each criterion; criteria must also be compared in a pair-wise manner to model their importance.

By executing the Saaty mechanism, weights were determined in each level of the AHP hierarchy. The tables No. 3, No. 4 and No. 5 show the values given by experts considering factors identified above.

Degradation	Soil	Vegetation	Eigenvalue
Soil	1	5	0,636986
Vegetation	1/5	1	0,258285

Table 3: Matrix of pairwise comparisons for the top level of Land Degradation Hierarchy (Consistency Index = 0,047725)

Soil	**Effective depth**	**Rockiness**	**Stoniness**	**Slope**	**Erosion**	**Salinity**	**Saturation**	**Soil Subtype**	**Eigenvalue**
Effective depth	1	2	3	5	5	5	7	9	0,341260
Rockiness	½	1	2	3	3	5	5	7	0,219196
Stoniness	1/3	1/2	1	2	3	4	5	5	0,156623
Slope	1/5	1/3	1/3	1	1	3	4	5	0,096792
Erosion	1/5	1/3	1/3	1	1	2	3	3	0,077594
Salinity	1/5	1/5	1/4	1/3	1/2	1	2	3	0,049920
Saturation	1/7	1/5	1/5	1/4	1/3	½	1	2	0,034017
Soil Subtype	1/9	1/7	1/5	1/5	1/3	1/3	½	1	0,024598

Table 4: Matrix of pairwise comparisons Land Sub-criteria (Consistency Index = 0,041883)

Vegetation	**Vegetal formation**	**Degree of Intervention**	**Density of Vegetation**	**Eigenvalue**
Vegetal formation	1	3	5	0,636986
Degree of Intervention	1/3	1	3	0,258285
Density of Vegetation	1/5	1/3	1	0,104729

Table 5: Matrix of pairwise comparisons for vegetation criteria (Consistency Index = 0,047725)

After this process, it is necessary to evaluate functions u_i that evaluate the alternatives on each criterion i and in coefficients of importance ki. Each alternative a is then assigned an overall value v(a) computed as:

$$v(a) = \sum_{i=1}^{n} k_i \, u_i(a)$$

and the alternatives can be ranked according to the values of v.[13]

To develop AHP for land degradation, some tools of Geographical Information Systems (GIS) were used. Raster calculation tools from ArcGIS were particularly useful. As a result, a layer of cadastre parcels associated with each land degradation index and other attributes was obtained.

Portrayal of land degradation index layer in a map was itself a contribution to decision makers. Nevertheless, they needed any kind of visual tool where they could manage dynamically different views (maps, tables, graphs) at the same time, filtering or aggregating information for time, land use, territories, etc.

For this reason, the research includes also elements of Business Intelligence, by building a multidimensional structure (data warehouse) capable to provide different analyses (OLAP) with their corresponding visualization tools. Considering effects of Climate Change occur in long period of time, these BI tools could become in a crucial support to decision making.

3. Results

A data warehouse (or smaller scale data mart) is a specially prepared repository of data created to support decision making. Data are extracted from source systems, cleaned and scrubbed, transformed, and placed in data stores[15].

The design of data warehouses and data marts is based on a multidimensional paradigm for data representation that provides at least two major advantages: on the functional side, it can guarantee fast response times even to complex queries, while on the logical side the dimensions naturally match the criteria followed by knowledge workers to perform their analyses [2]. The multidimensional representation is based on a star schema which contains two types of data tables: dimension tables and fact tables.

According to the application scenario, land degradation can be interpreted as the “fact” in the data warehouse design. In the same way, time, location and land use would represent “dimensions”. Figure 1 shows a star schema for the data warehouse design.

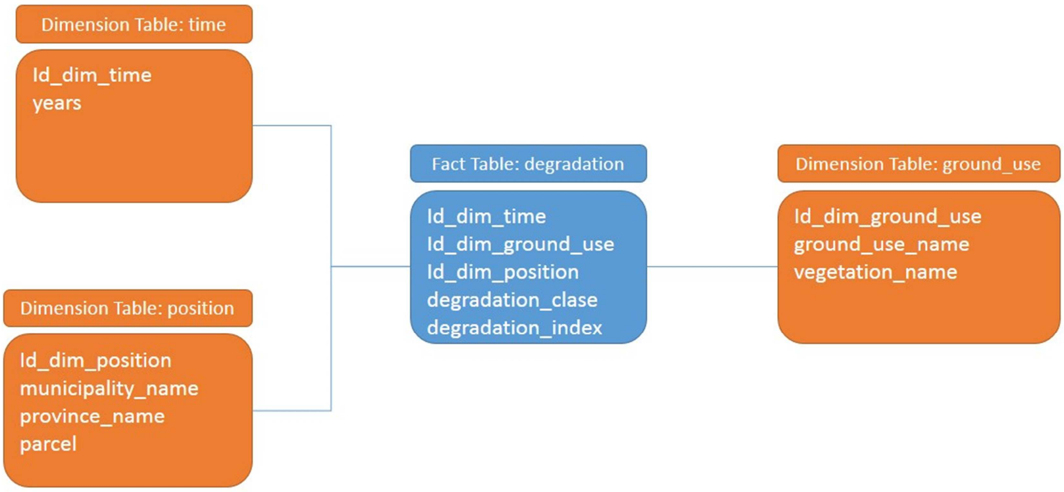

Figure 1: Data warehouse design.

The general architecture for the proposal BI/data-driven SDSS solution is depicted in the Figure 2 below.

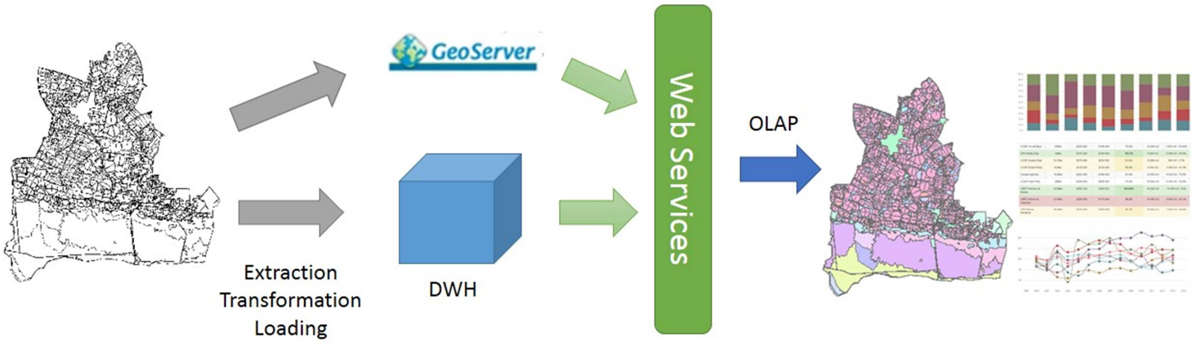

Figure 2: Architecture of the data-driven Spatial Decision Support System

Primary data source were the layers of degradation indexes obtained for each time period following the AHP technique explained above in this paper. ETL proceedings were developed to load data into the data warehouse, by means of Pentaho Data Integration 5.0.1. This powerful tool facilitates extracting, transforming and cleaning; as well as loading data into a data warehouse.

Extraction: During the first phase, data are extracted from the available sources (GIS for the example). The subsequent incremental extractions allowed update the data using new data that become available over time. The selection of data to be imported is based upon spatial database design, which in turn depends on the information needed by business intelligence analyses and decision support systems operating in the specific application domain.

Transformation: The goal of the cleaning and transformation phase is to improve the quality of the data extracted from the different sources, through the correction of inconsistencies, inaccuracies and missing values.

Load: Finally, after being extracted and transformed, data are loaded into the tables of the spatial data warehouse to make them available to analysts and decision support applications. [16]

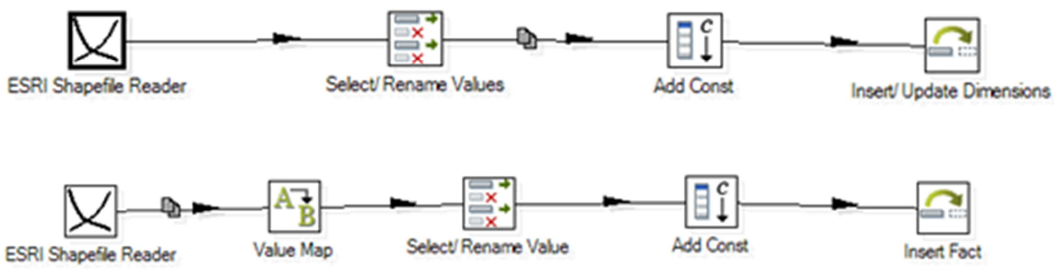

Figure 3: Some ETL proceedings used for load data in data warehouse

Part of this process corresponds to publish the layer of land degradation in the Map Server GeoServer. GeoServer has implemented the main OGC specifications [17-19]: WMS, WFS, WCS y WPS. GeoServer is also useful to visualize output maps, since it uses OpenLayers, an opensource viewer for OGC specifications.

With the view to develop MDX queries on the data warehouse, a specific Web Service was built. This allows the OLAP analyses and the unification of non-spatial and spatial information consumed from the Map Server GeoServer, compatible with OGC (Open Geodata Consortium). This Web Service includes logic for reports and analyses, making transparent for users work with the data warehouse.

A client application was developed as the BI user interface to support queries and visualization. This application invokes the Web Service developed to facilitate spatial and OLAP analyses regarding land degradation. JavaScript libraries like OpenLayers (to manage and visualize geospatial components), Jquery (to build graphs and tables) and JPivot (to support OLAP operations as Drill-down, drill-up, Pivoting, Slice) were included into the client interface.

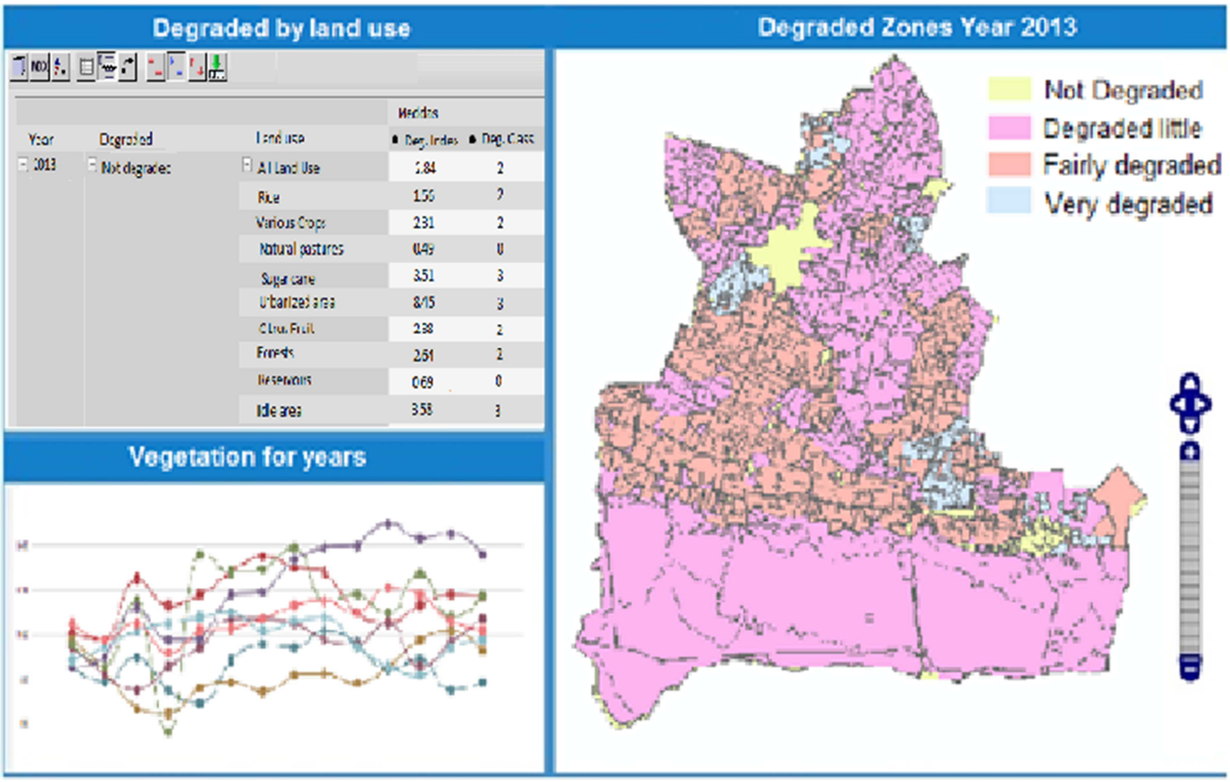

Figure 4: A view of the client application developed to visualize analysis of land degradation

Figure 4 shows the web application developed for the analysis of land degradation with a view of an OLAP query engine, statistical data and information on a map. Achieving calculate a degradation index per parcel or unit of land is very important to local authorities in Güira de Melena. Having this tool local governments can achieve greater effectiveness in actions to combat land degradation.

4. Conclusions

A land degradation index has been calculated in a GIS based on multi-criteria AHP techniques. The new map of land degradation temporary disaggregated serves as input data to a Business Intelligence solution.

OLAP analyses on the data warehouse built around the degradation allow diversify queries considering different dimensions as time, location and land use. A powerful spatial visual tool which integrates mapping of degradation using different views by means of OLAP analyses will allow supporting decision making processes on sustainable agriculture in Güira de Melena.

Climate Change is a driving force or indirect cause for land degradation. A spatio-temporal study of land degradation and its correlation with similar future studies of climate variability will support decision makers in order to create new policy actions regarding food security.

This work is part of a bigger Project entitled “Environmental bases for local sustainability of food production” (BASAL) and constitutes a first approximation to a data-driven Spatial Decision Support System to manage Climate Change effects on agriculture.

5. References and Bibliography

[1] Z. Mikolajuk and A. G.-O. Yeh, "Sustainable Development and Decision Support Systems," in *DECISION SUPPORT SYSTEMS FOR SUSTAINABLE DEVELOPMENT: A Resource Book of Methods and Applications*, G. E. Kersten, Z. Mikolajuk, and A. G.-O. Yeh, Eds., ed. New York: Kluwer Academic Publishers, 2002, p. 16.

[2] C. Vercellis. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making (Wiley ed.)*.

[3] H. K. Bhargava, D. J. Power, and D. Sun, "Progress in Web-based decision support technologie.pdf," *Decision Support Systems*, vol. 43, p. 13, 2007.

[4] C. Imhoff, N. Galemmo, and J. G. Geiger. (2003). *Mastering Data Warehouse Design*.

[5] W. H. Inmon. (2005). *Building the Data Warehouse*. *Fourth Edition*.

[6] G. F. R. Sprague, "Decision support systems:Issues and challenges," in *Proceedings of an International Task Force Meeting* vol. XI, P. Press, Ed., ed. UK: International Institute for Applied Systems Analysis, 1980, p. 202.

[7] A. Mahadevan, K. Ponnudurai, G. E. Kersten, and R. Thomas, "Knowledge Discovery in Databases and Decision Support," in *DECISION SUPPORT SYSTEMS FOR SUSTAINABLE DEVELOPMENT: A Resource Book of Methods and Applications*, G. E. Kersten, Z. Mikolajuk, and A. G.-O. Yeh, Eds., ed. New York, Boston, Dordrecht, London, Moscow: Kluwer Academic Publishers, 2002, p. 24.

[8] D. S. Frankel. (2003). *Model Driven Architecture Applying MDA to Enterprise Computing*.

[9] S. C. Lin and E. Yen. (2010). *Data Driven e-Science*.

[10] V. L. Sauter. (2010). *Decision Support Systems for Business Intelligence*.

[11] UNPD, "Project Document: Environmental Bases for Local Alimentary Sustainability (BASAL, for their initials in Spanish)," ed. La Habana, Cuba: UNPD, 2008, p. 79.

[12] T. L. Saaty. (1980). *The analytic hierarchy process*.

[13] T. L. Saaty, "Decision making with the analytic hierarchy process " *Int. J. Services Sciences*, vol. 1, p. 16, 2008.

[14] G. M. Morales, M. P. Moro, R. O. C. Díaz, A. V. Cause, R. P. Capote, and R. D. Telles, "USING OF THE GIS AND DISCREET METHODS IN THE DETERMINATION OF VULNERABLE AREAS TO DESERTIFITION " in *VI Congreso Internacional de Geomática* La Habana, 2009, p. 9.

[15] F. Adam and P. Humphreys, "Encyclopedia of Decision Making and Decision Support Technologies," in *Encyclopedia of Decision Making and Decision Support Technologies*, S. Reed, Ed., ed. United States of America by Information Science Reference: Information Science Reference, 2008, p. 1020.

[16] Spatialytics. (2013). *Introduction*. Available: http://docs.spatialytics.com/doku.php?id=en:spatialytics_etl:001_introduction

[17] OGC. (2006). *OpenGIS, Web Map Service Implementation Specification. Version 1.3.0*. Available: http://portal.opengeospatial.org/files/?artifact_id=14416

[18] OGC. (2006). *OpenGIS, Web Feature Service Implementation Specification. Version 1.1*. Available: https://portal.opengeospatial.org/files/?artifact_id=8339

[19] OGC. (2006). *OpenGIS, Geography Markup Language (GML) Implementation Specification. Version 3.1.1*. Available: https://portal.opengeospatial.org/files/?artifact_id=4700