

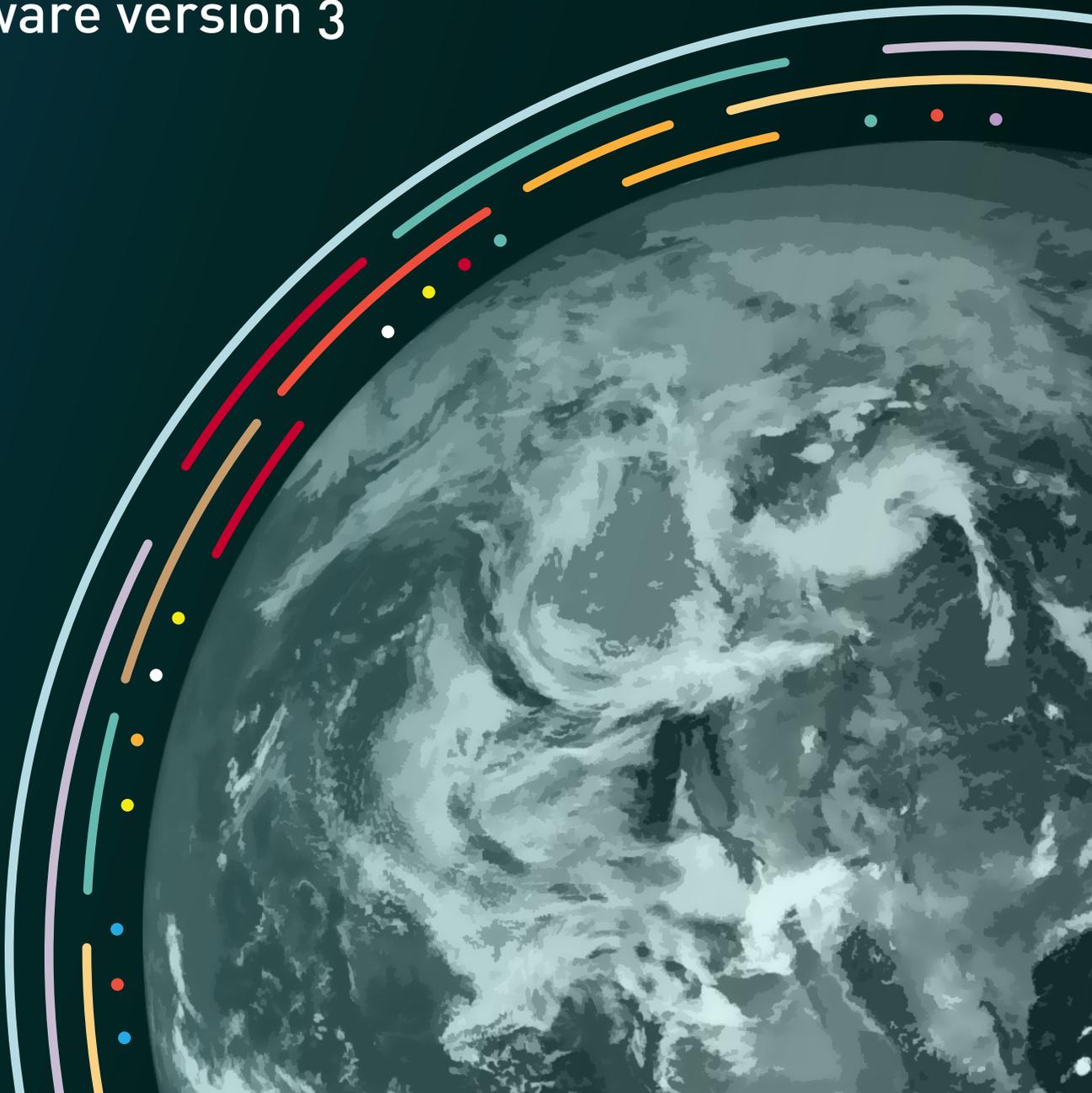


Food and Agriculture Organization
of the United Nations

Land Cover Classification System

Classification concepts

Software version 3



Classification Concepts



Land Cover Classification System

Software version 3

Author

Antonio Di Gregorio

Contributors

Matieu Henry
Emily Donegan
Yelena Finegold
John Latham
Inge Jonckheere
Renato Cumani

The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO) concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO in preference to others of a similar nature that are not mentioned.

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO.

ISBN 978-92-5-109017-6

© FAO, 2016

FAO encourages the use, reproduction and dissemination of material in this information product. Except where otherwise indicated, material may be copied, downloaded and printed for private study, research and teaching purposes, or for use in non-commercial products or services, provided that appropriate acknowledgement of FAO as the source and copyright holder is given and that FAO's endorsement of users' views, products or services is not implied in any way.

All requests for translation and adaptation rights, and for resale and other commercial use rights should be made via www.fao.org/contact-us/licence-request or addressed to copyright@fao.org.

FAO information products are available on the FAO website (www.fao.org/publications) and can be purchased through publications-sales@fao.org.

CONTENTS

An overview of the LCML / LCCS3. Basic Concepts	1
1 - Conceptual Background	3
1.1 - The context	3
1.2 - Classification and legends	4
1.3 - Shortcomings and problems of semantic interoperability with current systems	4
2 - The FAO LCCS approach	7
3 - The LCML General System Design	9
4 - An overview of LCML fundamentals	11
4.1 - The general principle	11
4.2 - The LCML “basic Elements”	11
4.3 - The LCML properties and characteristics	16
4.4 - The LCML Elements vertical relationship or vertical stratification	16
4.5 - Inter strata relationship	18
4.6 - The LCML Element presence type	18
4.7 - The horizontal pattern	18
5 - A new general framework for data production and management	21
6 - LCML versus LCCS3	23
7 - LCML / LCCS3 from the operational point of view	25
Option 1 - Translation of the existing national legend / nomenclature into LCML / LCCS	25
Option 2 - Upgrading of an existing national legend / nomenclature using the LCML / LCCS approach	26
Option 3 - Upgrading of field data	27
Option 4 - Build up a national legend / nomenclature using the LCML / LCCS approach	27
Option 5 - Expand the LCML based national legend / nomenclature to become the national reference system for upscale and future upgrading	28
Option 6 - Build up a national “relational database” using the LCML / LCCS approach	28

The authors wish to thank
the Food and Agriculture Organization of the United Nations and the UN-REDD programme
for funding the publication and translation of this manual.

“Whenever you find yourself on the side of majority it is time to pause and reflect”.
Mark TWAIN



An overview of the LCML / LCCS3 Basic Concepts

Introduction

Land cover (LC) can be defined as the “observed (bio) physical cover of the earth’s surface”¹, and is a synthesis of the many processes taking place on the land. It reflects land occupation (and its transformation) by various natural, modified or artificial systems and to some extent how these systems affect the land. LC is one of the most easily detectable indicators of human intervention on land. Because it can change quickly over time, it is also a good proxy for dynamics of the earth surface resulting from a variety of drivers and factors.

Data on LC and changes to that LC are important for satisfying the ever increasing demand for reliable data to support studies and research at regional and global scales. These types of data support comprehension and analysis of natural phenomena such as climate change; provide a means to assess carbon stock accountability; and help monitor agriculture development, disaster management, land planning, defence of biodiversity, etc. (Fig.1)

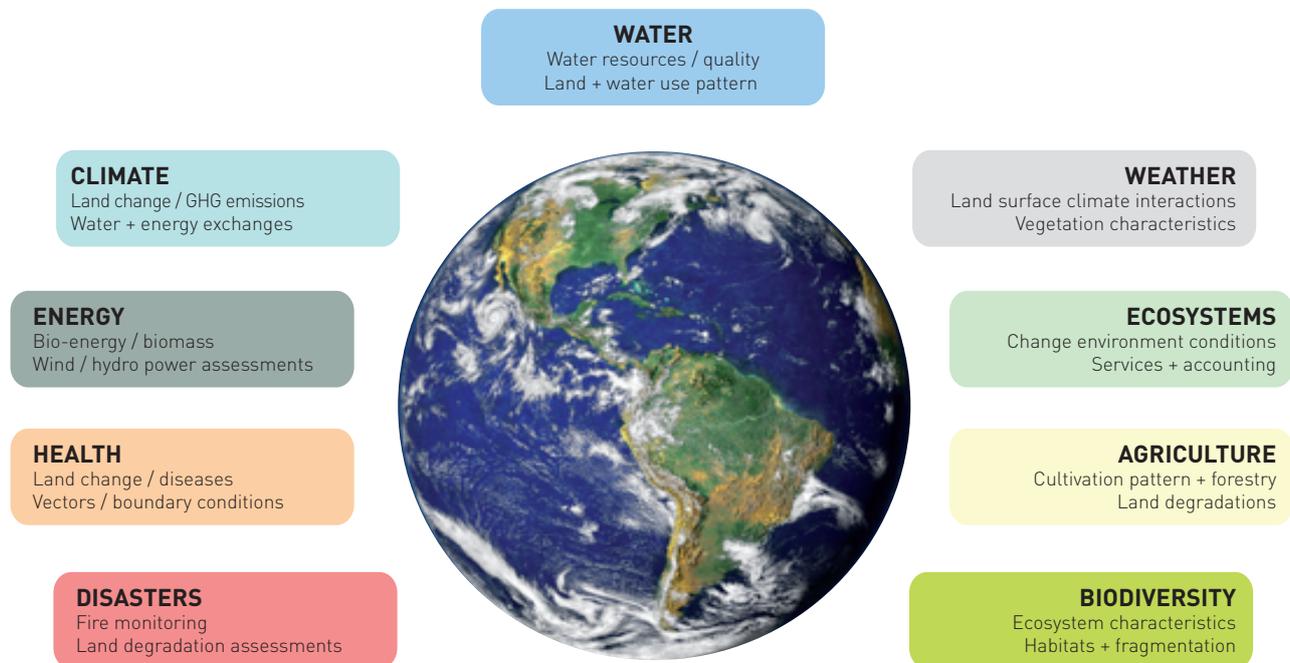


Fig.1 - GEO societal benefits and Land Cover/vegetation observations²⁻³.

Information on LC is critical for the implementation of environmental, food security and humanitarian programmes of the United Nations, international and national institutions. Despite this recognized need for information on LC, there are only limited efforts to improve the quality of the data, which often lacks sufficient accuracy or is collected using a variety of procedures, tools and thematic nomenclatures (classifications/legends) the incomparability of which has a negative effect on the utility of the data.

- 1 Di Gregorio A, LJM Jansen (2005) Land Cover Classification System Classification concepts and user manual Software version (2). In: 8 EaNRS (ed). Food and Agriculture Organization of the United Nations, Rome.
- 2 Herold M, C Schmullius, O Arino (2008) Building Saliency, Legitimacy and Credibility toward Operational Global and Regional Land-Cover Observations and Assessments in the Context of International Processes and Observing Essential Climate Variables. Proc of the 2nd MERIS / (A)ATSR User Workshop. ESA SP-666, Frascati, Italy 22-26 September 2008.
- 3 Herold M, L Kooistra, A van Groenestijn, P. Defourny, C Schmullius, V Kalogirou, O Arino (2012) Chapter 26, Building Saliency, Legitimacy, and Credibility toward Operational Global and Regional Land-Cover Observations and Assessments in the Context of International Processes and Observing Essential Climate Variables. In Giri CP (ed) Remote Sensing of Land Use and Land Cover, Principles and Applications. CRC Press pp. 397-414.

An integrated system for LC observation provides worldwide consistency and links local and global levels of observation. The concept assumes that there is observational continuity on all observation scales and that the data and information products are consistent and compatible, using common approaches to characterize, describe and compare LC information (standardization, harmonization, and validation) and to facilitate the joint application of mapping products⁴ to demonstrate the technical difficulty in differentiating.

4 Herold M, L Kooistra, A van Groenestijn, P. Defourny, C Schmullius, V Kalogirou, O Arino (2012) Chapter 26, Building Saliency, Legitimacy, and Credibility toward Operational Global and Regional Land-Cover Observations and Assessments in the Context of International Processes and Observing Essential Climate Variables. In Giri CP (ed) Remote Sensing of Land Use and Land Cover, Principles and Applications. CRC Press pp. 397-414.

1 CONCEPTUAL BACKGROUND

1.1 - The context

Land Cover (LC) can be considered a geographically explicit feature that can be used in different disciplines (geography, ecology, geology, forestry, land policy and planning etc.) as a geographical reference (e.g. for land-use, climatic or ecological studies). In modern geographic databases, due to its relatively easy spatialization, LC has become a sort of “boundary object” between different disciplines. This development, on the one hand enhances the value of the LC information, but on the other hand enlarges the base of potential users, posing new challenges for its harmonization and consistent use. New LC information is constantly being produced.

With the evolution of technology - including the operational availability of new earth observation satellites and digital data interpretation - the analyses and integration of those information has certainly started a new era in land mapping. Prominent examples are the variety of global and regional LC products (Fig.2) and their increasingly frequent use in studying land characteristics and dynamics. Any LC mapping activity can be defined as a process of information extraction governed by a process of generalization and abstract representation of the real world. Technology, like satellite earth observation and Geographical Information Systems (GIS), has helped to overcome many previous limitations in LC mapping on different scales, increasing the availability of LC information and the base for exchange and integration of such data.

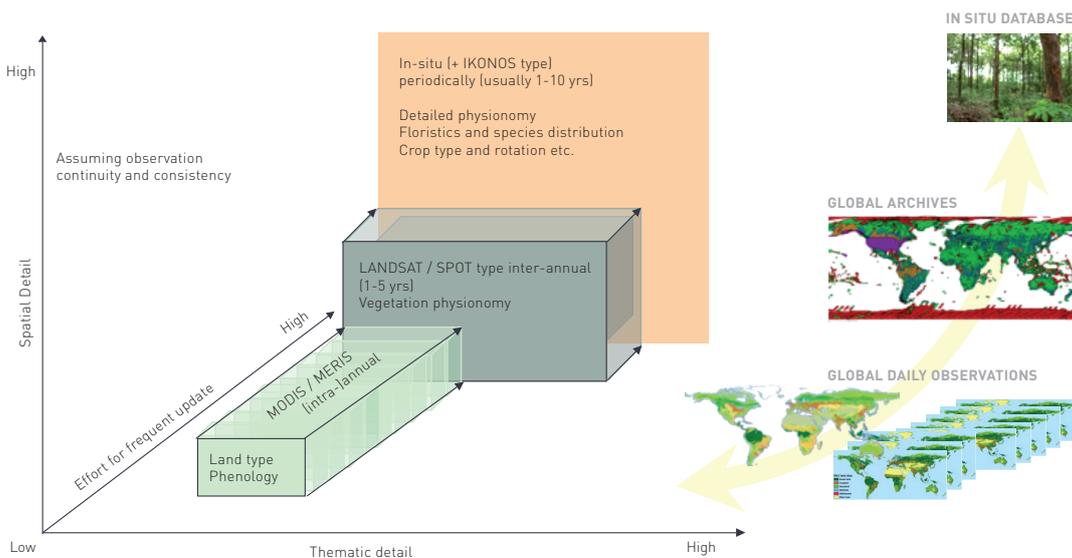


Fig.2 - Variety of land observations from local to global. Framework for integrated global observations of LC and vegetation⁵.

A side effect is that technology driven mapping often reflects more technological capabilities rather than the needs of end users. Many global and regional land mapping products are tailored to serve too specific or too generic applications and tend to exist mainly as independent datasets. Their development is driven by different national or international initiatives; the subsequent mapping standards adopted reflect the varied interests, requirements and methodologies of the originating programs. It is challenging to derive ‘efficient’ maps that are interoperable and potentially harmonized to satisfy the growing requirements of a wider end user community; often they lack flexibility and interoperability with other information. Thus, this situation is very similar to the pre –GIS era where each discipline was producing its own LC paper maps, the same geographic area was mapped several times, at different scales, for different purposes, with different types of data, and with varying accuracy.

⁵ Herold M, C Schmullius, O Arino (2008) Building Saliency, Legitimacy and Credibility toward Operational Global and Regional Land-Cover Observations and Assessments in the Context of International Processes and Observing Essential Climate Variables. Proc of the 2nd MERIS / (A)ATSR User Workshop. ESA SP-666, Frascati, Italy 22-26 September 2008. Modified by Antonio Di Gregorio.

The associated LC legends were discipline-specific with sufficient detail for the specific applications but with poor information for others. The exchange and/or use of natural resources data between disciplines, organizations and countries was absent or very limited. This situation fostered discussion and review for developing internationally accepted LC standards is a crucial task in minimizing current inadequacies and responding to the requests and needs of the international community. This is particularly evident when dealing with the ontology of a specific database and how its formalization of the meaning is defined. Formalization of the meaning of a map's thematic content and its semantic interoperability with other existing data set has become the bottle-neck that the advance of science has to solve.

1.2 - Classification and legends

The necessity to catalogue/categorize different aspects or activities of the real world is an old and common practice. We live in a regulated world; our lives are surrounded with systems of classification, regulated by standards, formats, etc. Classification, standardization/harmonization and hierarchical thinking do not come as complementary responses to our technological age where the use of computers implies a certain organization of the information, but is deeply related to human nature and is related to our instinctive ability to classify/order things to cope with the complexity (or the continuum nature) of the real world. The oldest method to communicate knowledge is, no doubt, human language and conversation, where specific language elements or specialized terms are created to exchange particular types of information⁶. Each discipline has developed its own specialized vocabulary. Beginning with the work of Carl Linnaeus, the concept of a hierarchical and standardized method to describe aspects of the real world has become the most efficient and concrete way to communicate and share information and knowledge in contemporary science. In the case of spatial information, classification is usually a predefined list of categories that try to signify an abstract representation of the real world. Sokal⁷ defines this process of classifying as: “the ordering or arrangement of objects into groups or sets on the basis of their relationships”, and Bowker and Star⁸ as: “a spatial temporal or spatio-temporal segmentation of the world”. They, therefore, define a ‘classification system’ as “a set of boxes (metaphorical or literal) into which things can be put in order to then do some kind of work (bureaucratic or knowledge production)”. A classification, therefore, describes a systematic and logical framework, populated by “a priori defined” number of classes, with a description of criteria used to distinguish them and the relationship between classes themselves. Classification thus requires the definition of class boundaries, which should be clear, precise, and possibly quantitative and based upon objective criteria. In an abstract, ideal sense a classification system should exhibit the following properties:

- use of consistent, unique and systematically applied classificatory principles;
- adapted to fully describe the whole gamut of feature types and their details;
- the system is complete, providing total coverage of the world it describes;
- the classes derived from it are all unique, mutually exclusive and unambiguous.

Classification therefore aims to address an entire information domain and subdivide it according to a set of rules to produce a set of classes and sub-classes allowing for all of the possibilities in the logical space.

1.3 - Shortcomings and problems of semantic interoperability with current systems

Categorization has always been a useful method to minimize the complexity of the real world and it is familiar to many aspects of our life. The process of categorization, however, is far from a perfect approach in modern information management. Its limitations are partly inherent with its intrinsic nature of grouping the real world phenomena in a certain number of artificial categories and partly related to the methods with which its “formalization of the meaning” is ratified. Reality is by its nature a “continuum”, and any partition of this virtual continuum into categories is intrinsically arbitrary and often reflects specific needs on the part of the data producer and not necessarily reflecting the varied needs of individual end users. Threshold parameters, for instance, produce arbitrary and artificial differences in values in the real world (Fig.3). Generalization, as well as the creation of a class itself, is often an arbitrary process. In addition very little effort is generally devoted to the way we explain the thematic content (“formalization of the meaning”) of each class.

6 Ahlqvist O (2008) In search of classification that supports the dynamics of science - The FAO Land Cover Classification System and proposed modifications. *Environment and Planning B: Planning and Design* 35:169-1996.

7 Sokal R (1974) Classification: purposes, principles, progress, prospects. *Science* 185:111-123.

8 Bowker GC, SL Star (1999) *Sorting Things out: Classification and its consequences*, Cambridge, MA:MIT Press.

Formalization of the meaning is the means by which the author of a classification system (legend) makes official and manifests the ontology (intend as “meaning” or “significance” of the things) of the categories and explains their relationship. The persistent use of simplified text descriptions (a class name with class description) exacerbates the inherent problems of any “generalization” process introducing supplementary constraints that increase the fuzziness of the data and create huge interoperability problems. This process implies rigidity in the transfer of information from the data producer to the end user community. End users have limited if any possibility to interact with the data, and must therefore accept them ‘as is’.

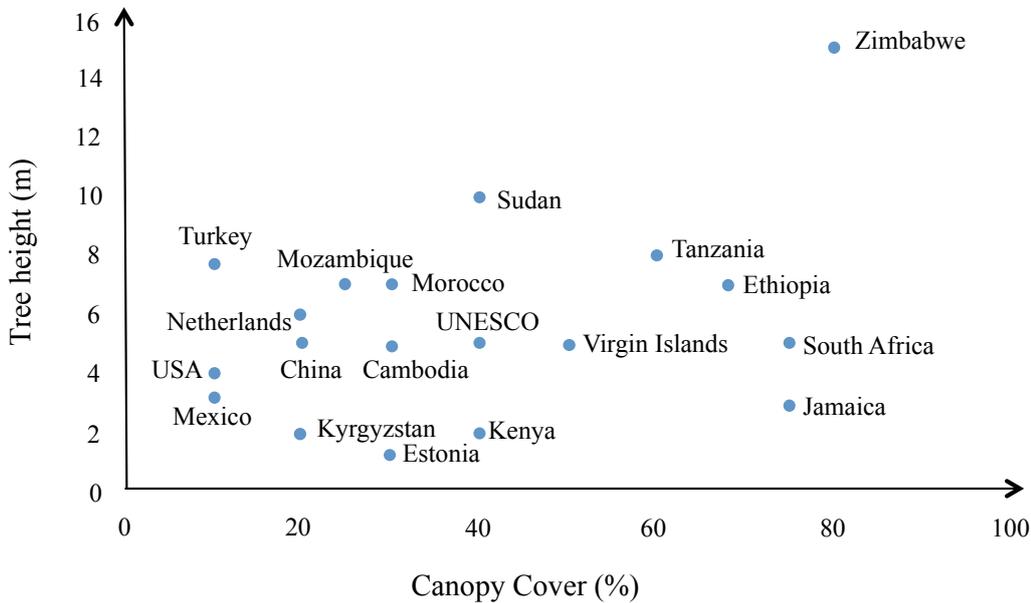


Fig.3 - Visualization of different threshold parameters used in the forest definition that use LC as core information⁹.

Class definitions can be often imprecise, ambiguous or, sometimes, absent. The build-up of the definition in the form of a narrative text is unsystematic (many diagnostic criteria forming the system are not always applied in a consistent way) and in any case do not always reflect the full extent, amount and type of the information, and is not logically structured. In addition, this structure of a database makes it very difficult to manage with modern GIS techniques. Semantic interoperability is actually the main challenge in Spatial Data Infrastructures. The representation of the multiplicity of the aspects summarizing a specific feature of the real world is drastically reduced or partially lost. Often some vagueness in the class definition is artificially included by the map producer to hide some ‘technical anomalies’ when reproducing a certain feature on the map. Moreover, vagueness or extreme complexity in the class definition makes it difficult to correctly assess the accuracy of the data set. Semantic interoperability is the most evident and tangible result of this situation. Interoperability is defined as “the ability of systems to operate in conjunction on the exchange or re-use of available resources according to the intended use of their providers”¹⁰. In the case of ‘semantic interoperability’, we refer to the understanding of the ‘meanings’ of different classes and relations among concepts. On these aspects, current classification and legends shows severe limitations that risk affecting the practical use of LC information. The list below shows the most common problems encountered when dealing with semantic interoperability of classification systems¹¹:

- different terms used for concepts (Synonymy);
- different understanding of homonymous concepts (Polysemy) (e.g. the various meanings of the term ‘forest’ for forestry, environmental modelling);
- different understandings of the relations among common concepts;
- common instances across databases assigned to different concepts in different ontologies;
- common instances allocated to a more general concept in one hierarchy than in other;
- equivalent concepts formalized differently;
- equivalent concepts explicated differently.

9 Ahlqvist O (2008) In search of classification that supports the dynamics of science - The FAO Land Cover Classification System and proposed modifications. *Environment and Planning B: Planning and Design* 35:169-1996.

10 Kavouras M, M Kokla (2002) A method for the formalization and integration of geographical categorizations. *Int J Geogr Inf Sci* 16:439-453. doi: 10.1080/13658810210129120.

11 Kavouras M, M Kokla (2002) A method for the formalization and integration of geographical categorizations. *Int J Geogr Inf Sci* 16:439-453. doi: 10.1080/13658810210129120.

The graded and fuzzy nature of common LC categories derived from traditional classifications/legends has been recognized for a long time by the remote sensing community, however limited efforts to sort out semantic uncertainty in land cover studies has been proposed. Several studies to ascertain classes semantic similarity have been done and several techniques have been elaborated. However, especially when using a feature matching process to solve semantic discrepancies, would be more effective to concentrate the efforts promoting and testing new ways to formalize the semantic of a classification system rather than continue to implicitly support obsolete paradigms (class names and narrative text description) that make the efforts for semantic interoperability difficult, time consuming and, in many cases, ineffective. The present work wants to demonstrate that alternative solutions of LC data representation and functional management of their semantic interoperability exist, this is particularly important with the increasing need to capture as much as possible the diversity of information of “real world” features.

We start from the following assumptions:

- mapping (or conceptual representation of a particular geographic feature) is a local activity, so at one level it can be understood why there is the tendency to establish unique classification systems to fit local conditions;
- any land surface is at a certain level (or scale of observation) heterogeneous and the standards to represent and generalize those land characteristics can be as diverse as the land itself;
- in geographic information truth as a distinct, incontrovertible and correct fact cannot exist a classification (legend/nomenclature) of geographic phenomena is inherently subject to indeterminacy and relativism mostly reflected in its ontology;
- it should be recognized that no classification system can reflect the social and/or the natural world completely accurately;
- classification (categorization) is a highly dynamic process related to geographic areas, time and culture;
- it should be recognised that there is a “balancing act” inherent in the process of classifying that must strike a balance between the huge complexity of the “continuum” nature of the real world and the necessity to represent or utilize it in a database with a finite boundary;
- there are and always will be multiple ways to categorize (segment) real world phenomena, and all are equally legitimate;
- in the process of classifying (categorizing) both standardization and harmonization efforts are needed. The effectiveness of a classification process depends at which level standardization and harmonization are used.

Despite these serious constraints, we should recognize that:

- despite the obvious constraints, categories are useful means whereby we cope with the “continuous” nature of the real world and its multiplicity of information;
- categorization is also a powerful method for sharing knowledge;
- not necessarily a categorization process should be the base for an “a priori” generation of a database legend the categorization can be used as rules for database interrogation, rather than for a database generation.

A classification system therefore should be able to support the many different ways a database can be generated. The recent advances in many fields of information technology and data management impose a modern and advanced approach to support evolving standards and in general the dynamic of science. In paragraph 3 the basic characteristic of an ideal classification system are listed, and in addition to those conventional ones new properties are needed, therefore a modern system should:

- be potentially able to converse with other systems. This inherent harmonization property should not rely (only) on expert judgment but the harmonization process should be automated as much as possible;
- recognize the balancing act inherent in classifying (Bowker and Star, 1999). A classification will never be able to fully represent all the aspects of the real world, therefore it must be clear it reflects (just) a specific scope for which has been developed;
- render voice retrieval¹² by allowing users to detail and compare classes using a detailed class description (systematically organized with a list of explicit measurable diagnostic attributes), thus avoiding the risk of systems being impermeable to the end users;
- standardization process should focus on the rules and conditions how a feature is conceptualized rather than acting just on the class name;
- the “formalization of the meaning” of the system and its components should be formulated with the most modern methods of modelling;
- a modern classification should not be considered an isolated structure but more a functional component of a rather complete system for data management.

12 Bowker GC, SL Star (1999) *Sorting Things out: Classification and its consequences*, Cambridge, MA:MIT Press.

2 THE FAO LCCS APPROACH

In 1996, FAO made a contribution to upgrade the data categorization (mainly for Vegetation and LC) by starting to develop a new approach. A new set of classification concepts were elaborated and were discussed and endorsed at the meeting of the International Africover Working Group on Classification and Legend¹³. The system was developed in collaboration with other international initiatives on classification of LC, such as the U.S. Federal Geographic Data Committee (FGCD) – Vegetation Subcommittee and Earth Cover Working Group (ECWG); the South African National Land Cover Database Project¹⁴; and the international Geosphere-Biosphere Programme (IGBP) – Data and information System (DIS) Land Cover Working Group and Land Use Land Cover Change (LUCC) Core Project.

After a test period in the FAO Africover project in 1997–1999, the first official release of LCCS (1) was published in 2000 (Di Gregorio and Jansen, 2000). A second version was developed based on international feedback involving a large global community, and published in 2005 (LCCS2)¹⁵. LCCS adheres to the concept that it is deemed as more important to standardize the attribute terminology rather than the final categories. LCCS works by creating a set of standard diagnostic attributes (called classifiers) to create or describe different LC classes. The classifiers act as standardized building blocks and can be combined to describe the more complex semantics of each LC class in any separate application ontology (classification system)¹⁶.

The creation of or increase in detail in the conceptualization and description of an LC feature is not linked to a text description of the classifier but to the choice of clearly defined diagnostic attributes. Hence the emphasis is no longer on the class name but on the set of clearly quantifiable attributes. This follows the idea of a hybrid ontology approach, with standardized descriptors allowing for heterogeneous user conceptualization¹⁷. During the practical use of the LCCS through the years, there has been an unexpected trend in the utilization of the system by the international user community. In addition to the creation of specific legends for specific applications, the system has also been used as a reference bridging system to compare classes belonging to other existing classifications.

In 2003, FAO submitted the LCCS to the ISO Technical Committee 211 on Geographic Information as a contribution toward establishing an international standard for LC classification systems. This was the first time that this ISO committee had addressed a standard for a particular community of interest within the general field of geographical information. All of its previous standards had been higher level or abstract standards that established rules for application schema, spatial schema or similar concepts. There was some initial difficulty in initiating the standardization activity due to this more specific focus could leave this out. The result was that a standard was first developed to address classification systems in general (ISO 19144-1 Classification Systems) and then one to address LC (ISO 19144-2 Land Cover Meta Language). Both have been approved in 2012 as official international ISO standards. The LCML (Land Cover Meta Language) is therefore a powerful tool to characterize geographic features using a more modern approach with respect to the conventional classification methods. Its intrinsic structure of an open object oriented system allows not only an unambiguous description of real world features more consistent with the logic and structure of modern databases but also enlarges the capability of the system to describe phenomena related to inputs and activities peoples undertake on a certain LC feature typical of agriculture. A more detailed description of the LCML schema and a summary of how it works is illustrated in the next chapter.

13 Di Gregorio A, LJM Jansen (1996) FAO Land Cover Classification System: A Dichotomous, Modular-Hierarchical Approach. The Federal Geographic Data Committee Meeting - Vegetation Subcommittee and Earth Cover Working Group 15-17 October 1996, Washington D.C.
Di Gregorio A, LJM Jansen (1997) A new concept for a land cover classification system. Proceedings of the Earth Observation and Environmental Information 1997 Conference. Alexandria =, Egypt, 13-16 October 1997.

14 Thompson M (1996) A standard land-cover classification for remote-sensing applications in South Africa. South African Journal of Science 92:34-42.

15 Di Gregorio A, LJM Jansen (2005) Land Cover Classification System Classification concepts and user manual Software version (2). In: 8 EaNRS (ed). Food and Agriculture Organization of the United Nations, Rome.

16 Ahlqvist O (2008) In search of classification that supports the dynamics of science - The FAO Land Cover Classification System and proposed modifications. Environment and Planning B: Planning and Design 35:169-1996.

17 Ahlqvist O (2008) In search of classification that supports the dynamics of science - The FAO Land Cover Classification System and proposed modifications. Environment and Planning B: Planning and Design 35:169-1996.

3 THE LCML GENERAL SYSTEM DESIGN

LCML acts as a method to bring the LC community together to create a common understanding of LC nomenclatures with the aim to produce global, regional and national data sets able to be reconciled at different scales, level of detail and geographic location. The LCML provides a general framework of rules from which more exclusive conditions can be derived to create specific legends. It is a language based on physiognomy and stratification of both biotic and abiotic materials. The system may be used to specify any LC feature anywhere in the world, using a set of independent diagnostic criteria that allow correlation with existing classifications and legends.

LCML Land cover classes are defined by a combination of a set of LC elements. These LC meta-elements are divided in two categories: “basic elements” the elements that constitutes the main physiognomic aspects of biotic and abiotic cover features, for instance for biotic features trees, shrubs, herbaceous vegetation etc.; and “element properties” that further define the physiognomic/structural aspect of the basic objects.

The fundamental idea of LCML is that a predefined set of LCML basic elements (BIOTIC and ABIOTIC) and their properties enriched in their semantic significance with “element” and “ class” characteristics can be arranged in different types of vertical and horizontal patterns to describe a wide variety of distinctive and detailed land cover situations.

THE OBJECT ORIENTED METHOD TO MAP LAND COVER ACCORDING TO FAO/GLCN

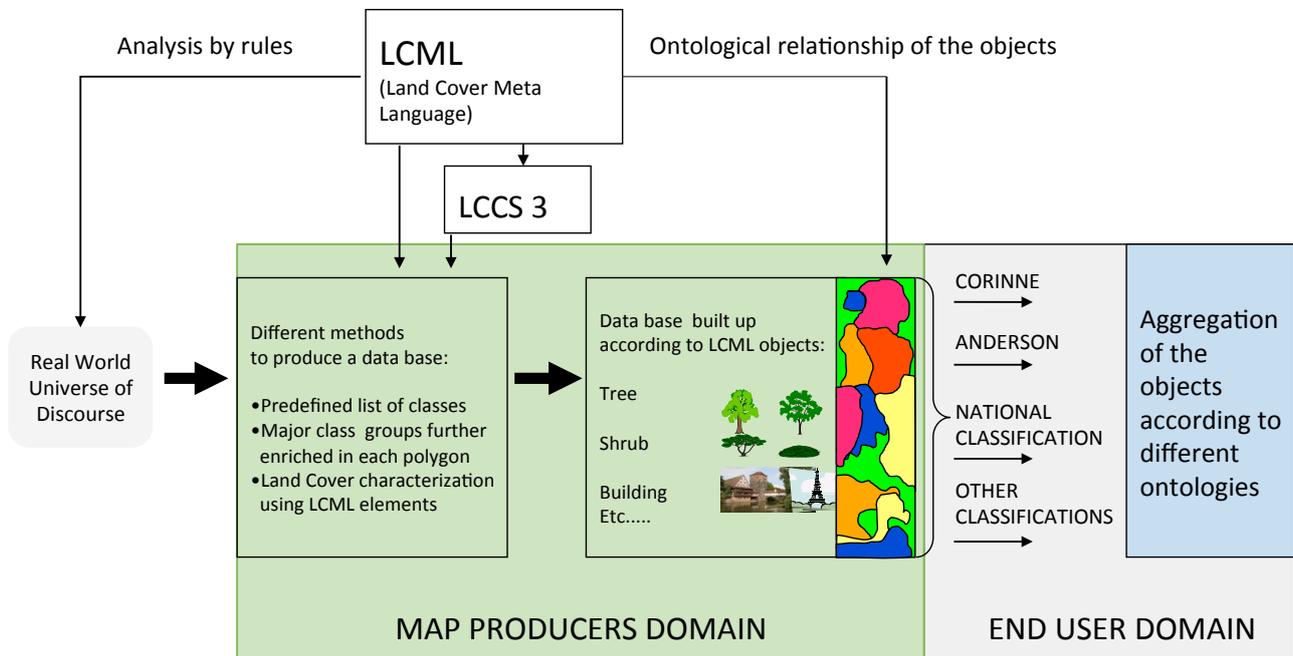


Fig.4 - The object oriented method to map LC, according to FAO/GLCN.

4 AN OVERVIEW OF THE LCML FUNDAMENTALS

4.1 - The general principle

The system utilize a physiognomic-structural approach to describe the “real world” features. This method is derived from vegetation studies were an ample evidence exist in applying physiognomic-structural criteria for the classification of plant communities. On this framework:

- physiognomy is defined as the overall appearance of a vegetation element;
- structure is defined as the spatial distribution pattern of growth forma in a plant community¹⁸. The structure, then, defines the horizontal and vertical arrangements of vegetation elements (growth forms etc) and it is usually characterized by horizontal macro patterns, vegetation layering, height and growth form density.

This principle is extended not only to the vegetated features but to the whole LC features populating the “real world”.

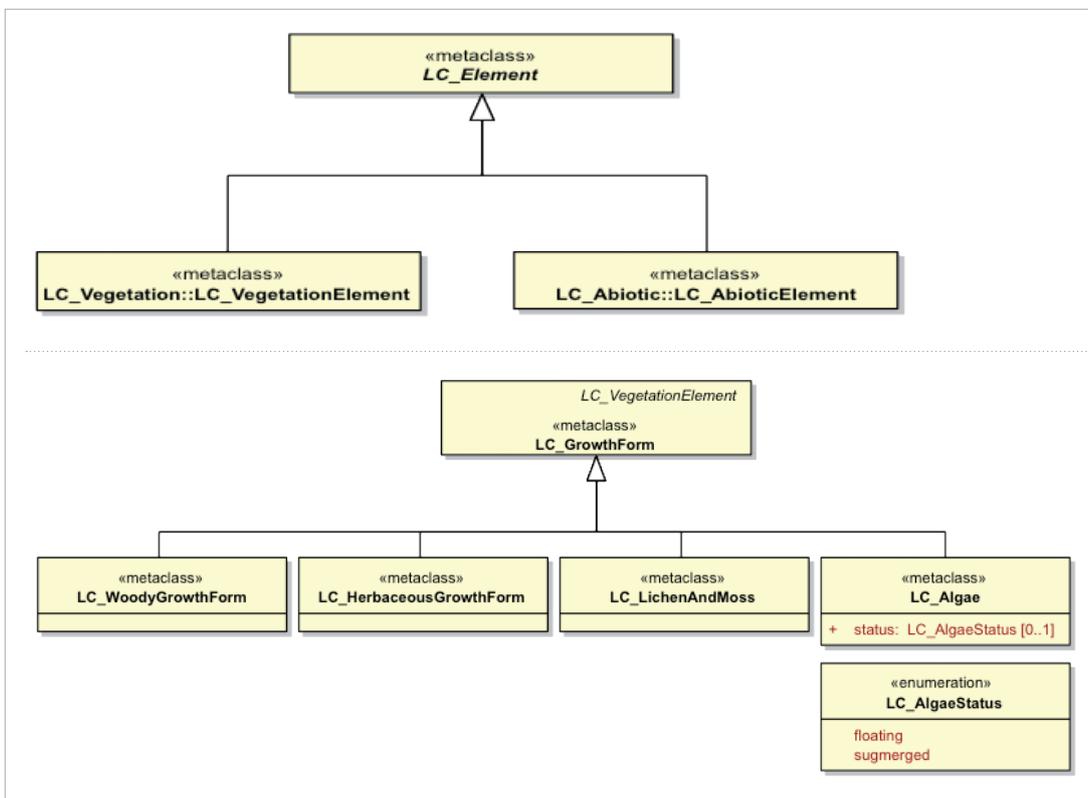


Fig.5A
The LCML
Element major
groups.

Fig.5B
Growth forms
subtypes.

4.2 - The LCML “basic Elements”

The meta language operates by representing LC features in terms of the subtypes of LCML “basic elements” called LC-Element(s) (or a set of them) with associated attributes and characteristics (see below). The LCML “basic Elements” (LC-Elements) form the basic building blocks of any Land feature representation. In order for the LCML to operate it is always necessary to define at least one LCML Element. This ensures (at least at the minimum level) a basic level of different class compatibility, because all of them, irrespective of their semantic definition, have to list the LCML Element(s) that populate each LCML subtype (class). The LCML Elements are subdivided in two major groups, VEGETATED and ABIOTIC (Fig.5A). The LCML VEGETATED Element comprise the all Growth Forms of vegetation plant community. The Growth Forms are further divided in four subtypes (Fig.5B). Each of the four subtypes is further subdivided into the final higher detail level subtypes (Figs. 6,7,8).

18 Kuchler AW and Zonneveld IS (1988) Vegetation Mapping. Handbook of Vegetation Science. Kluwer Academic Publishers, Dordrecht, The Netherlands

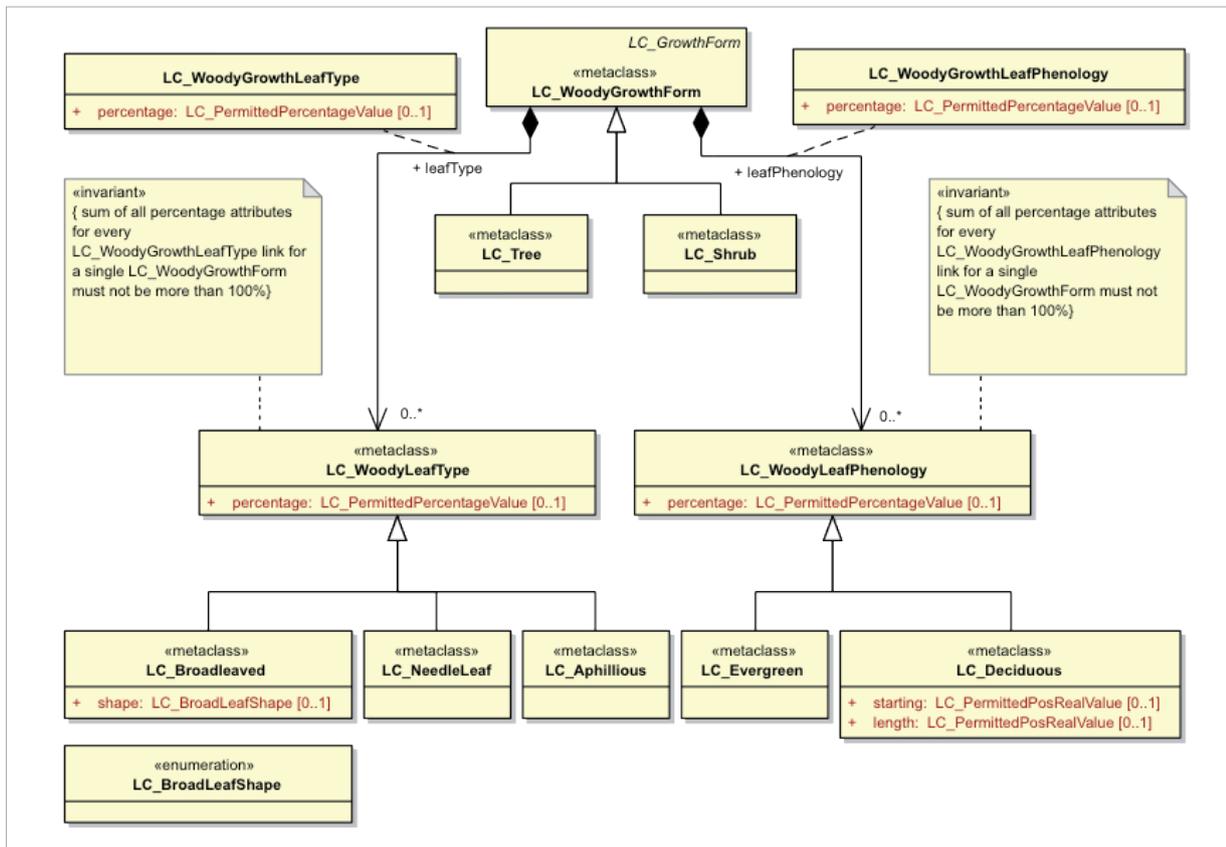


Fig.6 - Higher detail level of the growth form Woody further subdivided in Trees and Shrubs with the attributes leaf type and Leaf phenology.

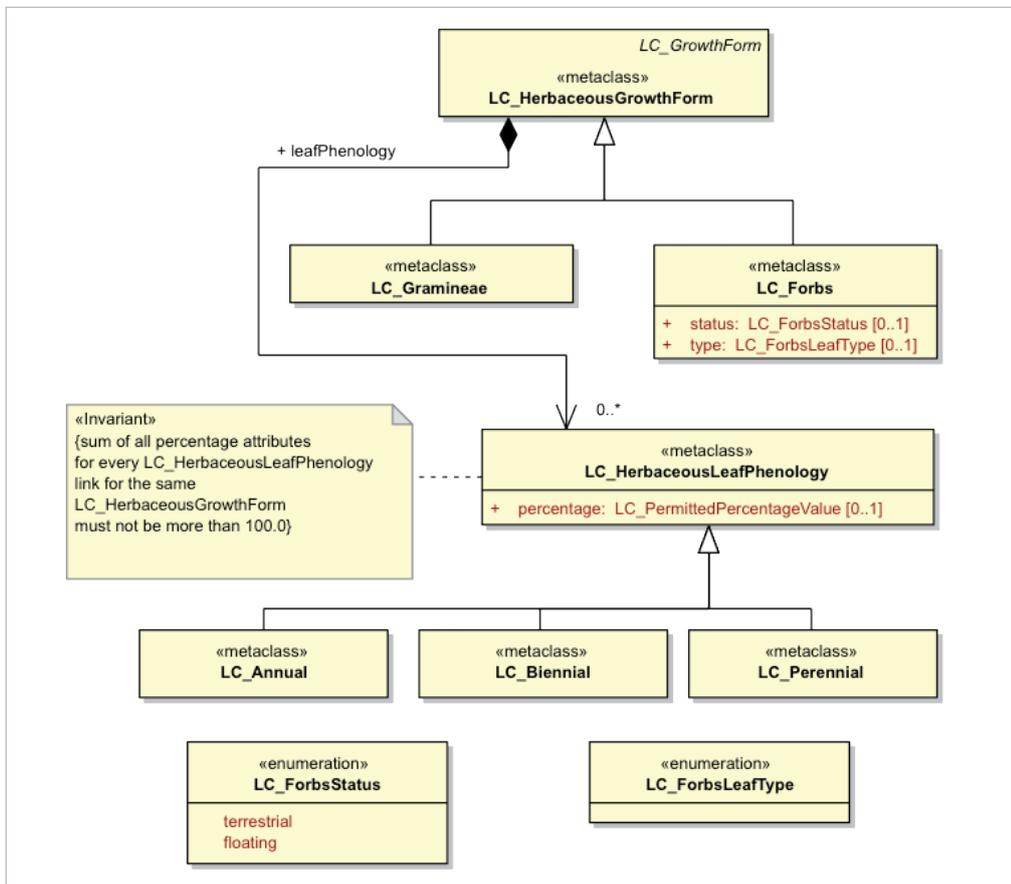


Fig.7 - Higher detail level for the growth form herbaceous further subdivided in Graminae and Forbs with the attribute leaf phenology.

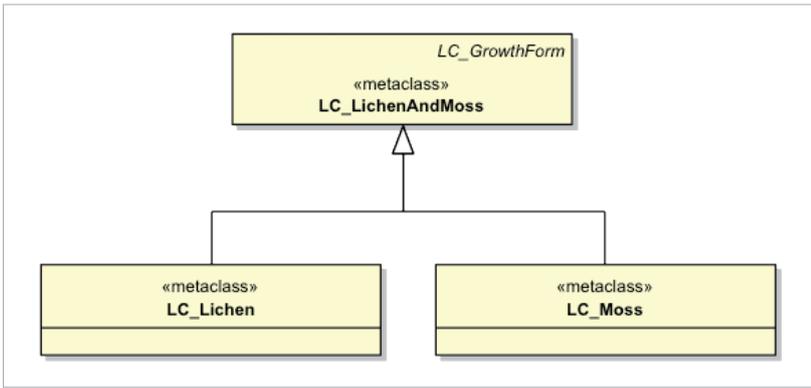


Fig.8 - Higher detail level for the growth form Lichen and Mosses further subdivided in Lichen and Mosses.

The LCML ABIOTIC elements comprise the whole real world objects not included in the LCML Vegetated elements group. They are subdivided in a first level in three basic types of abiotic surface elements:

1. Artificial Surface Elements;
2. Natural Surface Elements;
3. Water Body and Associated Surface Elements (Fig.9).

Each Abiotic Element subtype is further in two additional sublevels (see Figs 10, 11A-B for the Artificial surfaces sublevels, Figs.12, 13A-B for the Natural surfaces sublevels and Figs.14, 15 for the Water and associated surfaces sublevels).

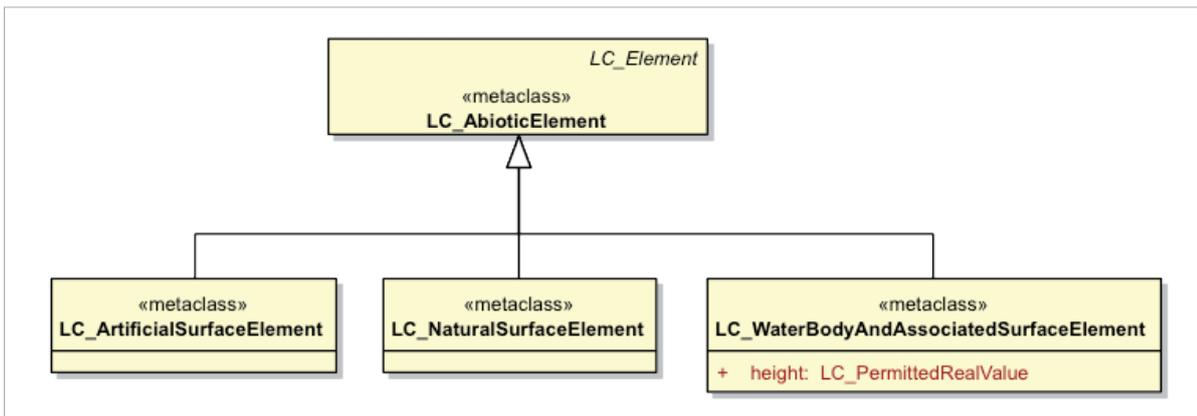


Fig.9 - Abiotic Elements subtypes.

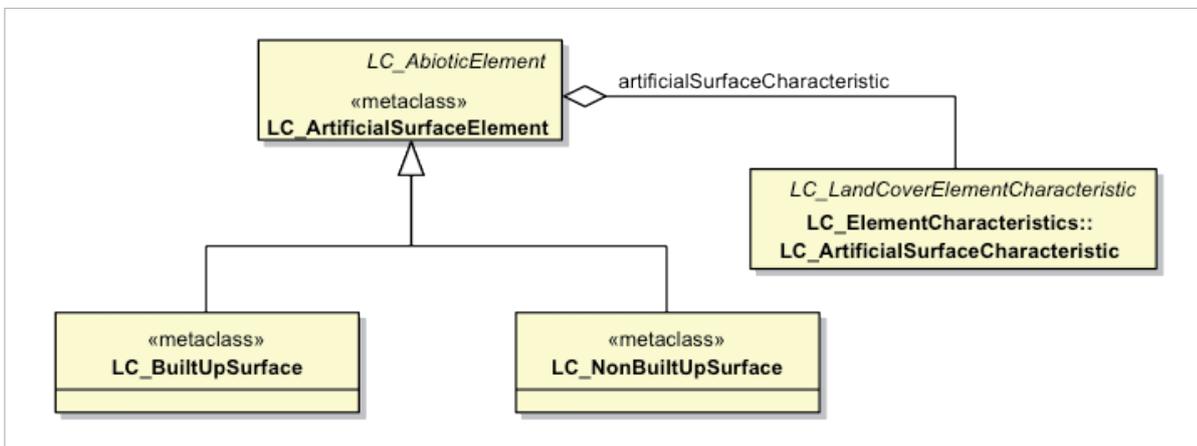


Fig.10 - First subdivision of the Artificial surface element.

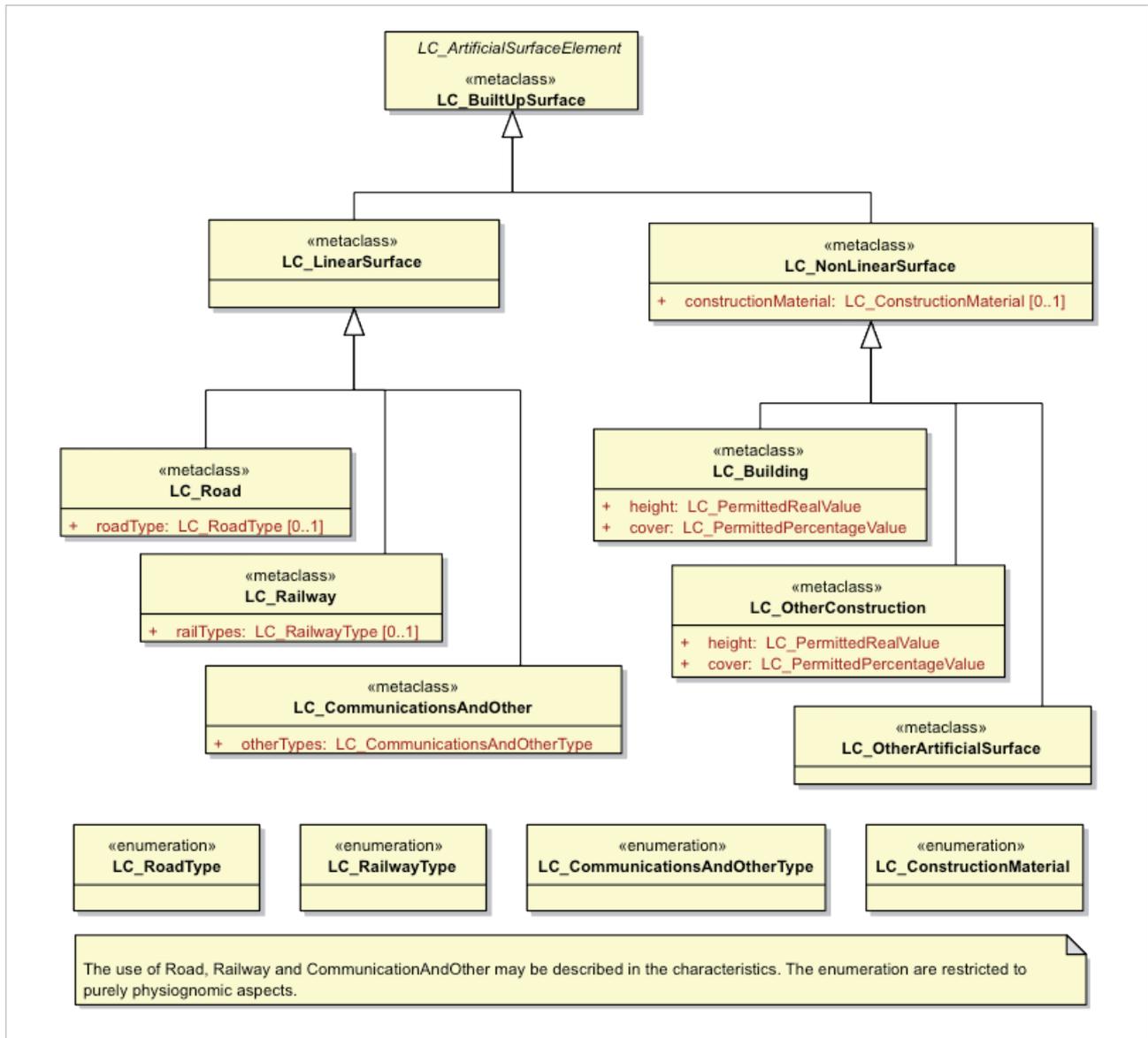


Fig.11A - Further subdivision of the element Built up surface.

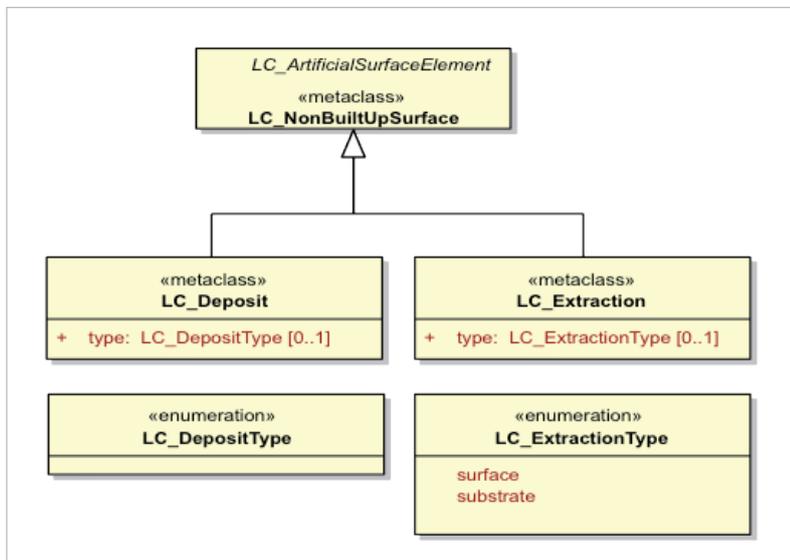


Fig.11B - Further subdivision of the element Non built up surface.

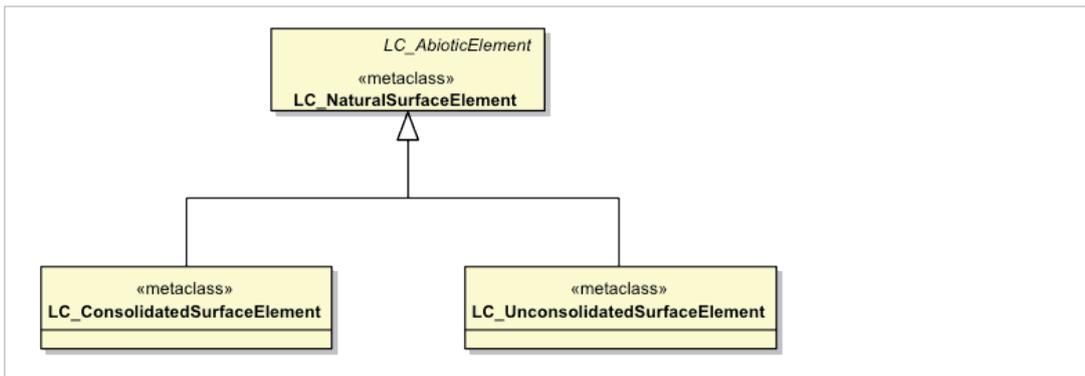


Fig.12
First subdivision of the element Natural surface.

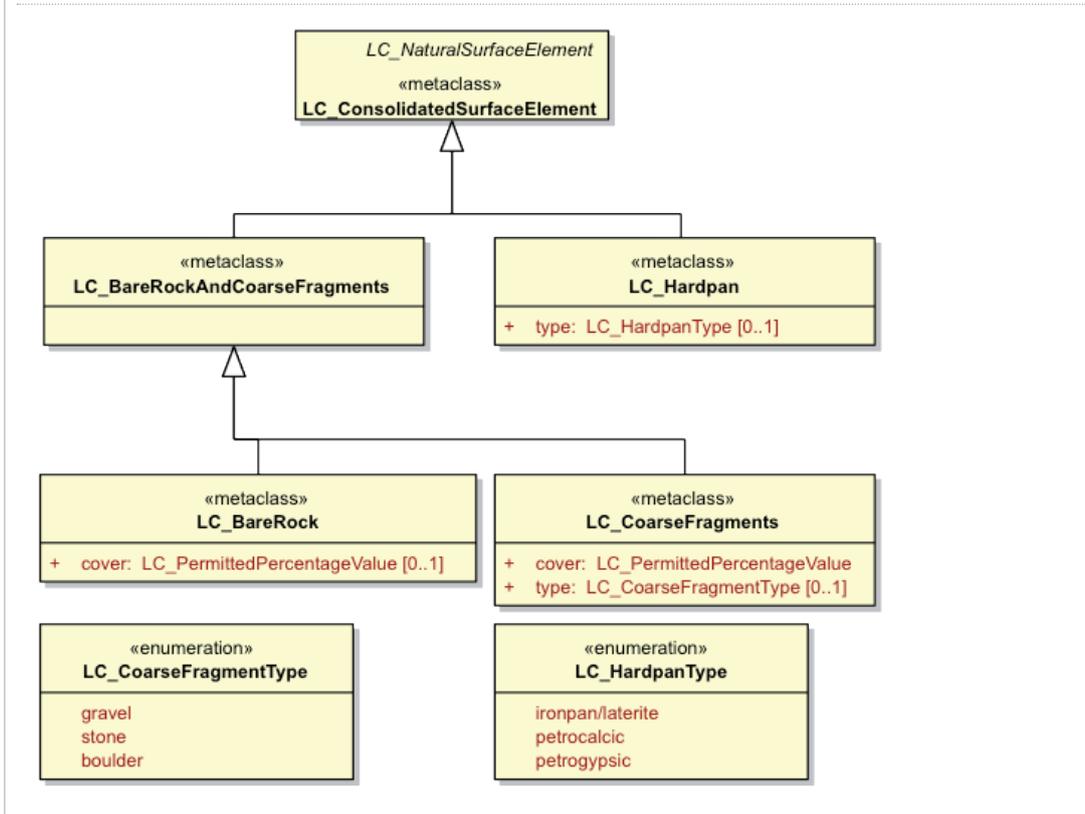


Fig.13A
Further subdivision of the element Consolidate surface.

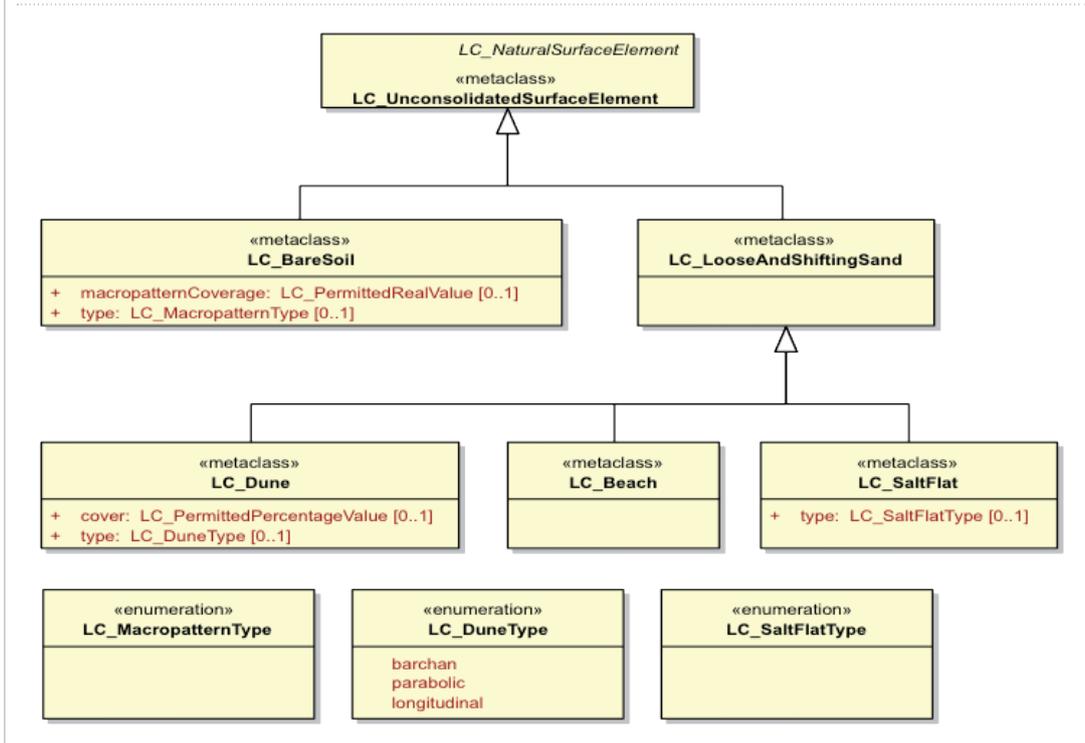


Fig.13B
Further subdivision of the element Un-consolidate surface.

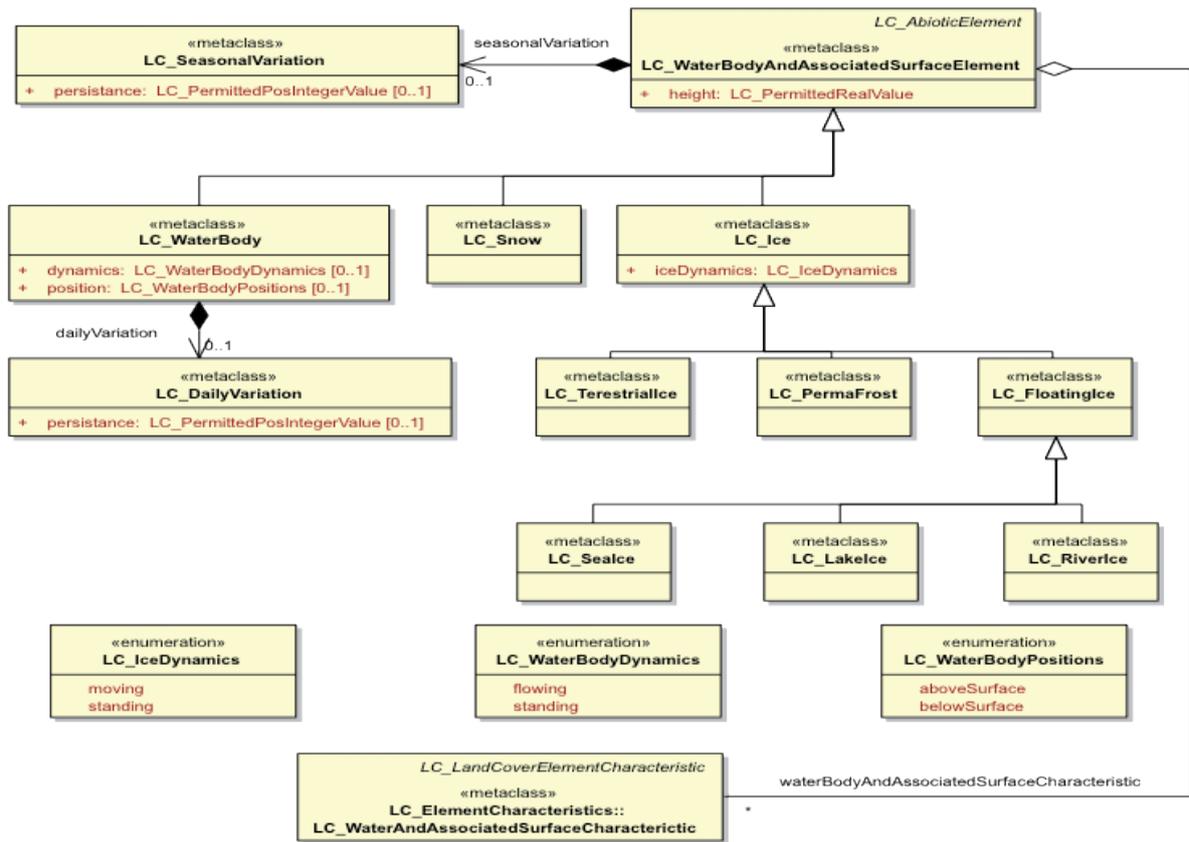


Fig.14 - Further subdivision of the element Water body and associated surface with a list of properties as seasonal or daily water variation, position and ice and water dynamic.

4.3 - The LCML properties and characteristics

Properties are specific attributes that further characterize the physiognomic aspect of an LCML element (for instance for natural vegetation height, cover, leaf type and leaf phenology). They are always related to one LCML element. Their list is fixed in order to maintain a stable basic vocabulary to assure a backward compatibility of the fundamental route with which any land feature is characterized in the LCML language. Changes are only allowed by the amendment of the LCML ISO standard (ISO 19144-2).

Characteristics, contrary to properties, are descriptive attributes not directly related to the physiognomic aspect of an LCML element (for instance for vegetation seeding time, grazing, vegetation management practices etc.). Two different types of characteristics exist: Element Characteristic and Class Characteristic. The first one, as the properties, is directly attached to one of the existing LCML elements, while the second one refers to the whole LCML elements used to create a particular class (examples of class characteristics are: climate, landform, topographical aspects etc.). Characteristics, unlike properties can be extended, providing a simple route to extend the descriptive aspect of the meta-language without having to modify its core structure (LCML elements and properties). A new set (or extension of an existing set) of characteristics can be done to the reference model proposing a new (or updated) registry of characteristics according to the rules described in Clause 9 of the ISO document. For the users of LCCS3 the extension of the characteristics list can be done utilizing the “user defined characteristic” module.

4.4 - The LCML Elements vertical relationship or vertical stratification

Individual or groups of LCML elements (vegetated and/or non-vegetated) may be assembled into strata (or layers). Several strata may be combined in an ordered sequence to model specific LC features. An example is a savannah or woodland composed by three separate layers of trees, shrubs and herbs with different cover of the woody component of the vegetation types (Fig.15A). The cover % of the elements in each stratum may overlap since each stratum is independent from each other.

The ordering of LCML element(s) into a single strata has the following rule: the strata can be populated by a single or a multitude of LCML elements (Vegetated and/or Abiotic). In case more than one LCML element populates the strata the space occupied by each single LCML element “partitioning” cannot be more than 100%.¹⁹



Fig.15A - This example describes the formation of a land cover class using different vegetation layers. Savannah vegetation is usually composed by a combination of sparse trees and shrubs over a grassland area. Three separate layers of Trees, Shrubs and Herbs with different cover of the woody component type.

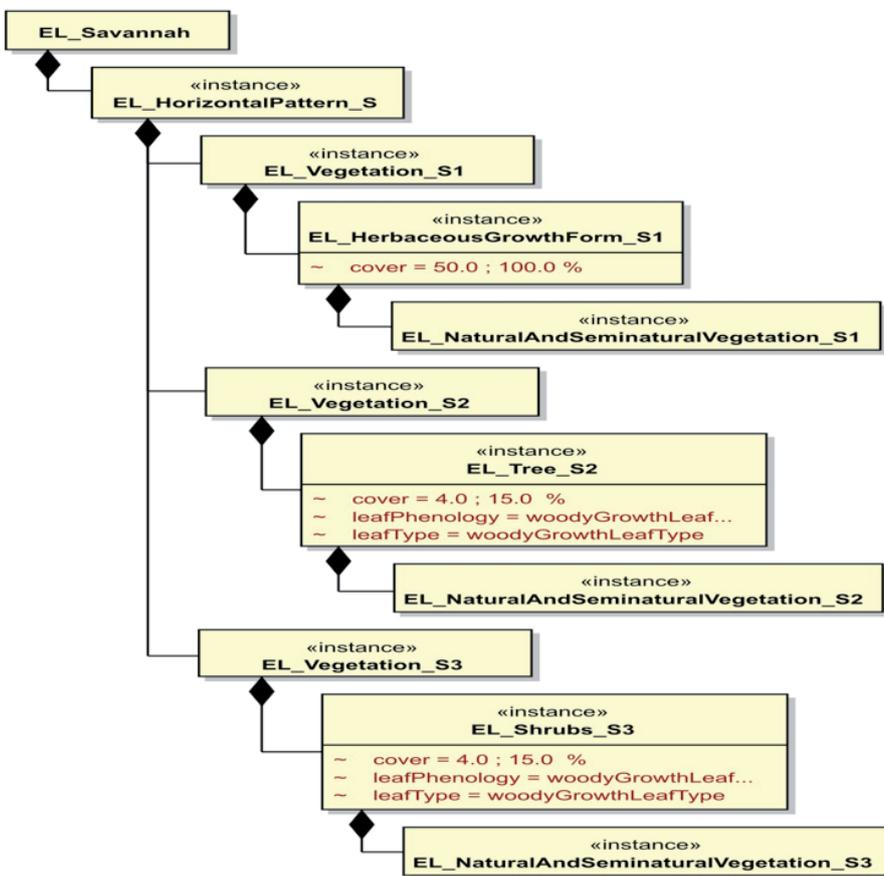


Fig.15B - The resulting LCML model for the class Savannah.

The model below in Fig. 15B shows how the class Savannah is modeled using the LCML rules. The class has been built up with three separate vegetation layers: a layer of herbaceous vegetation with a cover ranging from 50 to 100%, a layer of trees with cover from 4 to 15 % and a layer of shrubs with open cover from 4 to 15%.

- By default the “origin” of two or more LCML elements forming two or more strata is always the same. In **Fig.15**, for instance, both LCML elements Trees, Shrubs and Herbaceous vegetation all originate from the same soil background;
- The on top attribute indicate a relationship between strata were one stratum is defined in reference to another;
- The presence type indicate whether a stratum is fixed (mandatory) or is optional (conditional).

¹⁹ The introduction of the “partitioning” constraint has been done recently, therefore it is present in LCCS3, but not yet properly shown in the reference LCML UML schema.

4.5 - Inter strata relationship

It defines the rules governing the relationship between different strata. Two types of main condition exist:

- by default the “origin” of two or more LCML elements forming two or more strata is always the same. In Fig.15 for instance, both LCML elements Trees, Shrubs and Herbaceous vegetation all originate from the same soil background;
- the on top attribute indicate a relationship between strata were one stratum is defined in reference to another;
- the presence type indicate whether a stratum is fixed (mandatory) or is optional (conditional).

4.6 - LCML Element presence type

it defines the type of occurrence a specific LCML element has in a specific strata. It can be modelled in four different ways:

- mandatory, the LCML element must exist in the stratum;
- optional, the LCML element may exist in the stratum. It indicates a certain level of uncertainty of the information. It can be used when the LCML system is used to translate/represent specific legend classes characterized by an ambiguous definition, or to indicate that a further refinement of a specific class is possible;
- exclusive, one of two LCML elements may exist in a stratum (XOR relationship), this implies that between two LCML Elements A or B exist;
- sequential temporal, two LCML elements exist in the stratum but their presence is related by a temporal relationship. This type of syntax is very useful, for instance, when is necessary to describe a specific agriculture field with two or more alternating crops during the year.

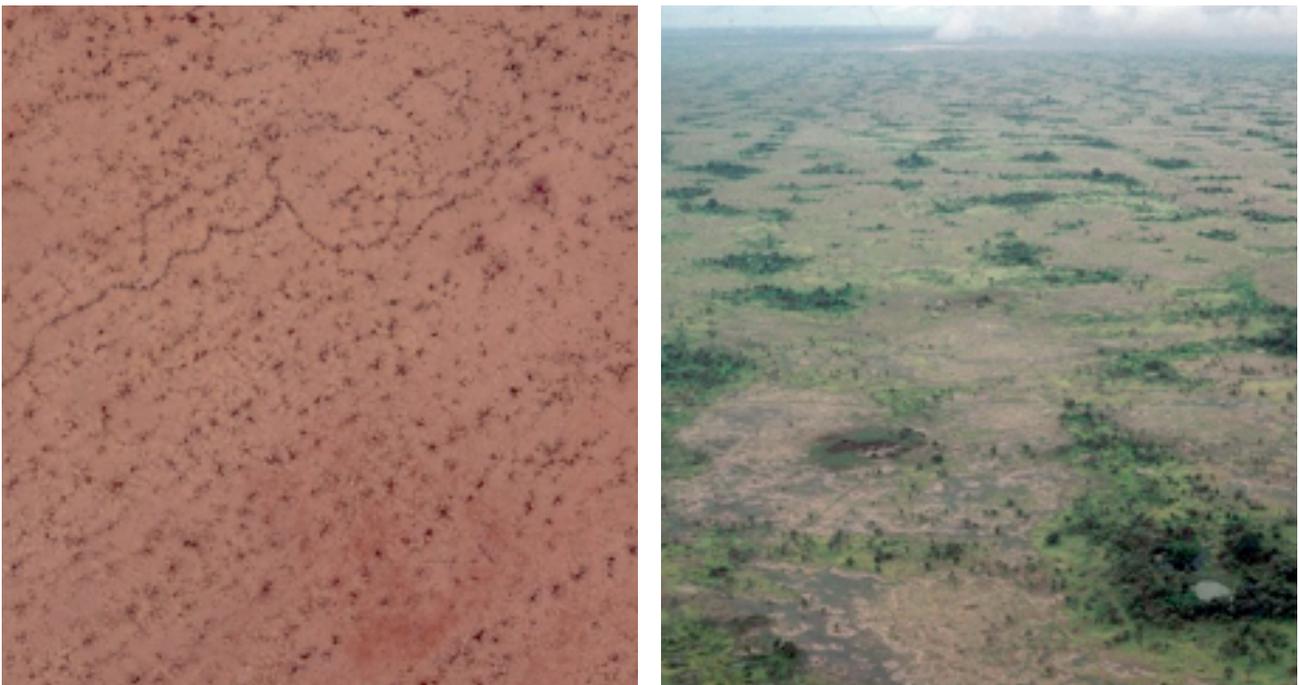


Fig.16 - Left: oblique view of a typical “Tiger Bush” formation. Right: the same area observed from a satellite image.

4.7 - The horizontal pattern

Allows the ordering of different strata consisting of single or groupings of LCML Elements, into one or more horizontal pattern(s). A horizontal pattern may be used for the description of a “functional unit” a complex Land Cover region composed by two or more distinct land cover features that will be handled as a “unicum” independently from scale constraints. A “functional unit” can have those peculiar characteristics because specific “ecological reasons” or a particular anthropic influence. An example is the vegetation formation called “tiger bush” which consists of a combination of patches of open shrubs and patches of open grassland in a specific multi horizontal pattern combination (Fig 16-17).

Note - A horizontal pattern to define a land feature is distinct from the representation of heterogeneous areas where two or more classes are combined just because of the actual given scale, and would otherwise be distinct at another scale. In this case the “cartographic standards” (the standards that regulate the representation of a geographic feature in a map according to M.M.U. constraint) should define if a mixed coding procedure must be applied (see for instance the FAO GLCN cartographic standards)

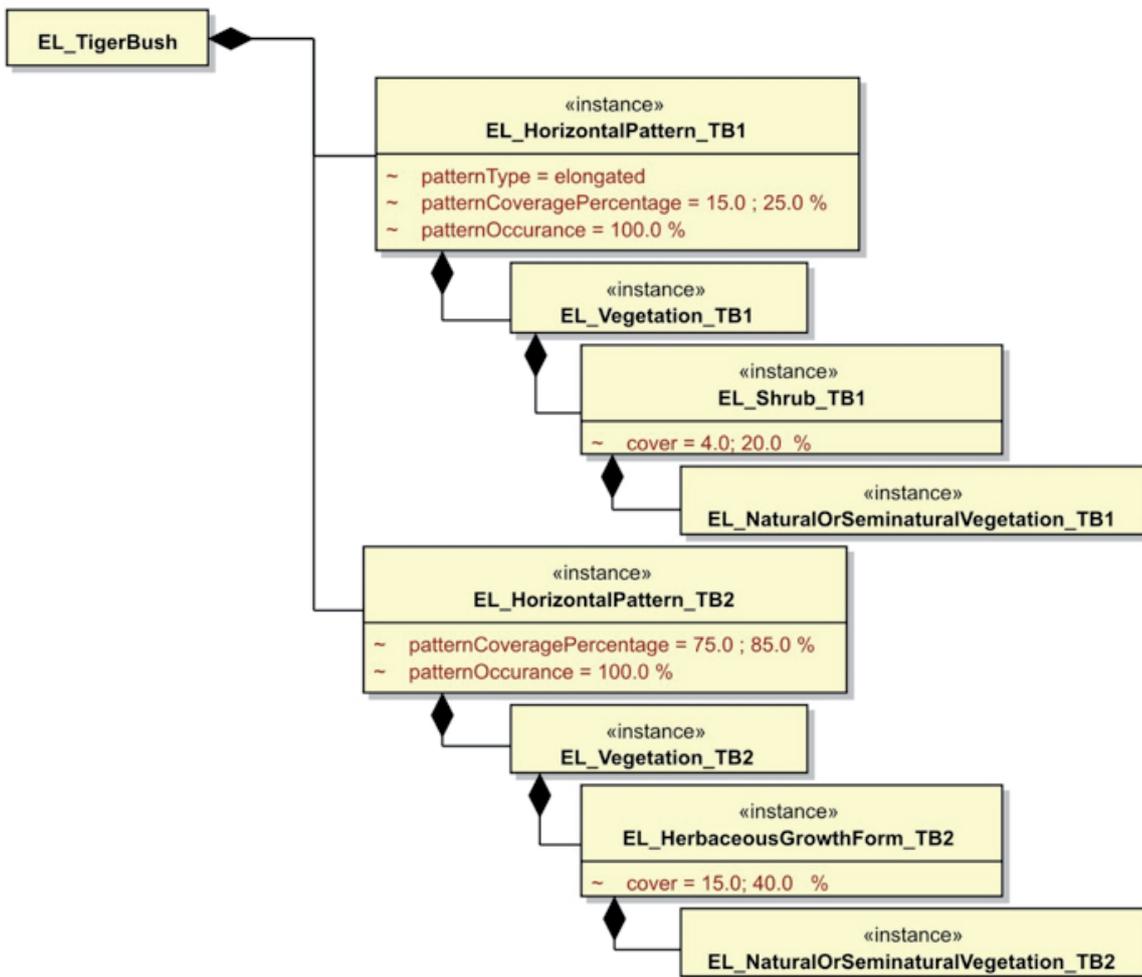


Fig.17 - The LCML model of the tiger bush.

. The Horizontal Pattern has three optional attributes pattern type, pattern percentage and pattern occurrence:

- the attribute pattern type is described in free text in a character string;
- the attribute pattern percentage describe the % value of each horizontal pattern.
- the attribute pattern occurrence describe the % value of the occurrence of a specific horizontal pattern

The model illustrated in Fig.17 shows how the “Tiger Bush” area is modelled in LCML:

- the class is composed of two separate Horizontal Patterns, Horizontal Pattern_TB1 and Horizontal Pattern_TB2;
- the Horizontal Pattern_TB1 is characterized by a pattern type: elongated, a pattern coverage: 15-20 %, a pattern occurrence: 100% and it is composed by the LCML element Shrub that has a cover of 4-20%;
- the Horizontal Pattern_TB2 is characterized by a pattern coverage: 75-85 %, a pattern occurrence: 100% and is composed by the LCML element Herbaceous Vegetation that has a cover of 15-40%.

5 A NEW GENERAL FRAMEWORK FOR DATA PRODUCTION AND MANAGEMENT

The LCML derived classification scheme cannot be seen as an isolated system but rather a functional component of a more composite framework of modern data production and management. Without this functional integration into the entire process of data generation it can have only a limited impact on the complex context of data standardization and harmonization.

An advanced data production and management system should be composed of different components, some of them already operative and others in a dynamic evolving process. The foundation of this system is the idea that complex real world features can be represented in an object oriented language with well-defined and quantifiable attributes and characteristics. A big step in this direction has been the accomplishment of the FAO LCML (LC Meta-Language) that has become an ISO standard in 2012. However other components are necessary to fully apply this object oriented logic.

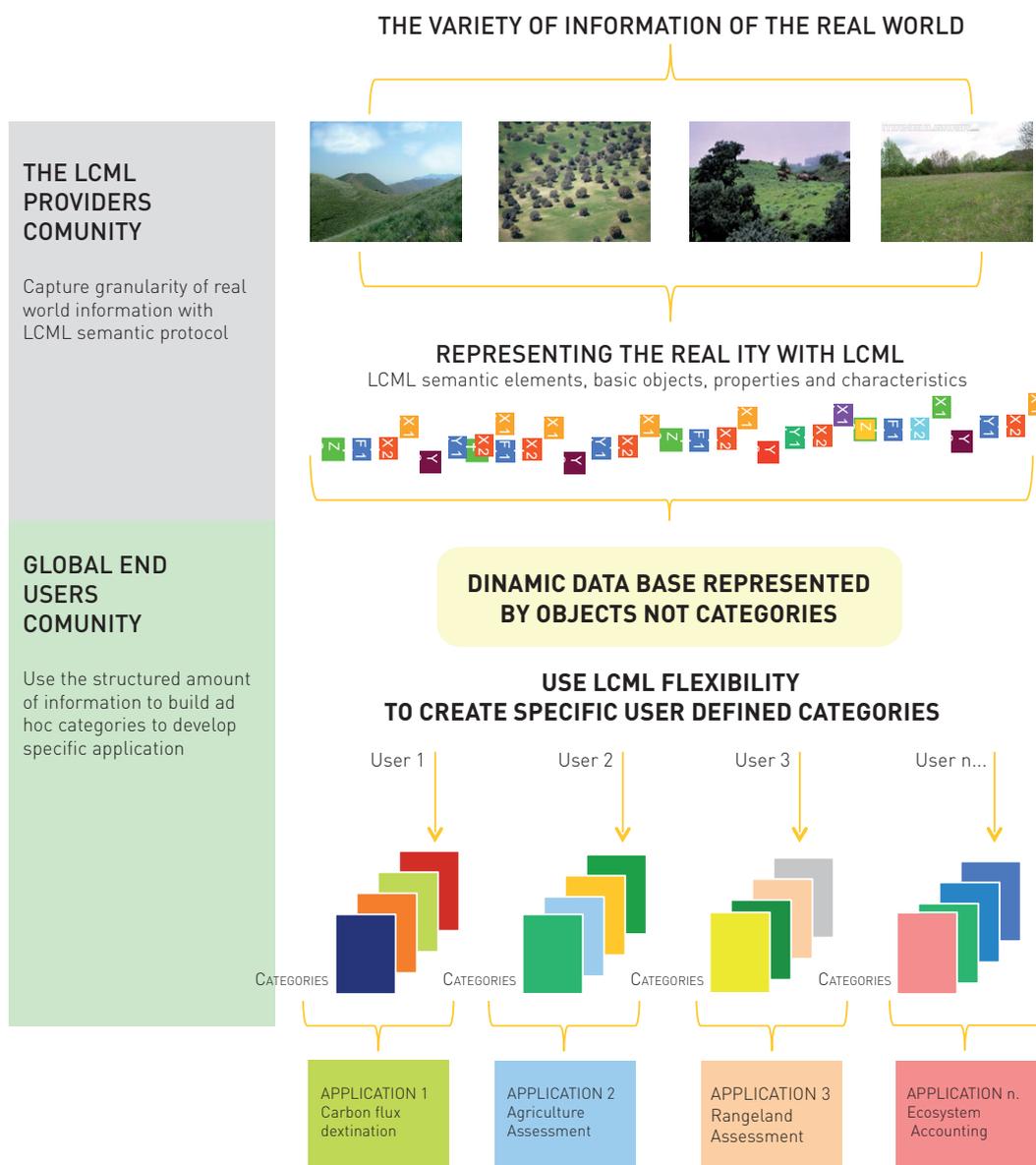


Fig.18 - The LCML flux of information.

New methods and devices to capture the variety of information of the real world features are of vital importance. Important potential advances in this sector are the launch of new sensors (Sentinel 1-2 etc.), the operational combination of different remote sensing devices, and the fast development of “citizen science” devices to upgrade with field observation remote sensing derived data. Directly linked with an “object oriented” description of the world feature is the development of a new way to store spatial information. The use of a “relational database” structure (Post GIS, Oracle) seems to be the most ergonomic solution to enhance LCML capabilities. The intrinsic logic of LCML seems to be fully compatible with this solution. The flux of information in this advanced system should first be the capability to capture in an extensive and detailed way the variety of information of the real world. These dynamic sets of information could be represented with the LCML logic into a relational database. At the end of this process, there is the development of specific categorization systems of which the primary function should be the formulation of rules for the interrogation of a dynamic database to give answers to specific applications (Fig.18).

The figure shows a schematic representation of the flux of information from the real world to a dynamic database. The “object oriented” nature of the LCML system allows the user to potentially capture the variety of information of the real world. At this stage the database can be populated by simple “objects” (enriched in their semantic meaning with specific quantifiable and measurable attributes and characteristics) and not necessarily by categories. However if categories are used at this stage of data information storage, they must have the quality to be decoded/converted in any moment into “objects” (“rendering voice retrieval” as suggested by Broker and Stars).

The conversion of the spatial data into a modern “relational” database should allow an easy and unambiguous (mostly done in an automatic way) translation of the data set into specific nomenclatures (categories) created always with LCML rules and customized to respond to a specific application. The different nomenclatures are perfectly and understandably harmonized between themselves.

6 LCML VERSUS LCCS3

It is indisputable that the basic concept/idea of LCML originates from the previous work done for the concept development of LCCS 1-2. The use of LCML elements organized in strata and enriched in their semantic meaning with properties and characteristics directly derive from the conceptual approach already developed before the year 2000, in the first elaboration of the LCCS concepts. In addition many LCML Elements, Properties and Characteristics both for their denomination and description are directly acquired from the LCCS 1-2 glossary. This not only to maintain consistency between the two systems but also to utilize terms that during the years of world- wide utilization of LCCS 1-2 have gained a considerable consensus and agreement. However despite these evident analogies the two systems show considerable differences:

- LCCS is a classification system, therefore its capability to describe the real world depends from the classification criteria adopted. A perfect example is the use of fixed value ranges for a series of parametric values (cover, height etc.). LCML on the contrary is conceived to be at a more general level of a specific classification system being a Meta-Language able, in theory, to support many classification criteria. This implies that LCML is far more flexible and efficient in the representation of the real world features respect to the old system;
- in LCCS the overall formalization of meaning (the overall logic) of the system is embedded in the proprietary software and synthetically explained in a text book. On the contrary in LCML the whole system logic is instantiated into an UML model that clearly documents the characteristics of the whole system;
- the LCML has gone through a more consistent and rigorous process of revision (the ISO process).

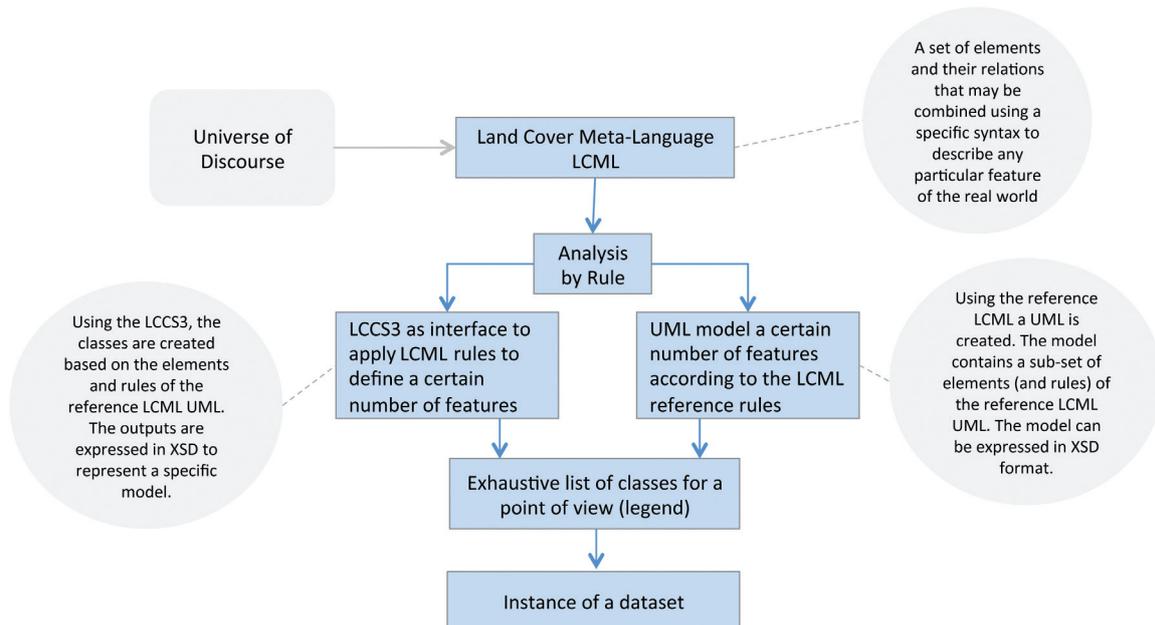


Fig.19 - Functional relationship between the LCML schema and the LCCS3.

In this new framework the new version of LCCS, the version 3, is intended to be a tool to apply in a “user friendly” way the overall concepts theorized in the LCML UML. Not all the users are conversant with the UML language, and it is also not functional to create manually an UML model (as sub-model of the reference LCML model) for each land feature to be represented into a database. LCCS3 has therefore the function to make easy, fast and intuitive the application of the LCML rules for the description of a given land feature. In addition LCCS3 has the capability to export through a specific application called “LCCS tool” the results in XSD format that can be converted into a proper UML schema or ingested into a system for advanced queries of the derived database. The Fig.19 shows the overall functional relationship between LCML and the new LCCS3.

7

LCML/LCCS3 FROM THE OPERATIONAL POINT OF VIEW

LCML/LCCS3 is an innovative object oriented meta-language that offers a new perspective for the semantic representation of land features. The characterization of land features results in a specific model that represent the objects, attributes and rules. In this respect, the proposed method is ergonomic, dynamic and allows an adequate management of geographic data. The proposed method is very different from traditional methods. However, the impact of the adoption of this methodology into a national system is much smoother and easier of what could be expected. It can be modulated in many different ways to fit different levels of expertise or willingness of the potential users to innovate their data production and management. The experience shows that the transition between two traditional legends is more complicated and difficult than the introduction to the LCML/LCCS method into a national system.

The different modalities of introduction of the LCML/LCCS concept into a national system have different levels of complexity. Therefore, the practical adoption of LCML/LCCS methodology requires different levels of expertise at different time periods for its full implementation. Each step, even the simplest one, is crucial for the adequate adoption and implementation of the system. Land mapping changes over time depending on objectives, national capacities, available technical means and progresses. In consequence, land mapping activities are assumed to improve over time. The proposed method allows any present choice to be fully ergonomic with any future (more technically advanced) options. Any action actually taken produces a solid technical foundation for the future ones.

Still, at country level, it is important to evaluate the most adequate way to introduce the LCML/LCCS approach. The objective of the proposed activities listed below, is to support/identify the best way to introduce LCML/LCCS approach into a national or sub-national context. The different options are listed with an increasing level of complexity and modification of the actual traditional way to generate the semantic meaning of a database. To facilitate the identification of the best option, each potential activity is described according to:

- general description/achievement;
- level of technical complexity;
- constraints.

Option 1 - Translation of the existing national legend / nomenclature into LCML/LCCS

General description

It consists in translating national legend classes, traditionally represented with a class name and a general text description, into the “object oriented” LCML/LCCS language. In this case, the semantic structure (formalization of meaning) of the legend remains the same with a class name and a text description. The LCCS schema will be an additional (more structured and systematic) description of the legend classes. The Fig. 20 shows an example of the integration of the LCML/LCCS schema as part of the conventional description of a traditional class legend (in this case a class of the Corinne legend has been considered).

Level of technical complexity

The only required expertise is a basic knowledge of the LCML/LCCS concept and software, together with a detailed understanding of the original national legend classes.

Constraints

Using LCML/LCCS, the overall structure of the legend, the semantic meaning, details of each classes, the structure of the derived spatial database remain untouched. However, the transition from a traditional way of legend description to a more structured language (where each class detail must be described with measurable and unambiguous attributes) is usually a quality control to test the consistence of a legend, both from a structural and thematic point of view. Very often, experts doing the translation will discover contradictions/inconsistencies in the original class definition. Different problems

can have different origins and can affect the database preparation in different ways. Some of these contradictions/inconsistencies are listed below:

- the legend was developed in other part of the world, under different climatic or geographic conditions, and doesn't fully reflect the local situation;
- the level of details of the original classes is much lower with respect to what could be gathered from advanced satellite data and combined field work;
- some legends (especially the ones describing natural vegetation) have been developed a long time ago. Usually they have been designed for field analysis and some key parameters are difficult to spatialize using remote sensing. Potentially, they can be substituted with parameters that are easier to measure and identify (e.g. Altitude values versus geographic characterizations such as low land, montane etc.);
- new classes, which were not in the original legend need to be mapped.

213 - RICE FIELDS

Original Description: Land prepared for rice cultivation. Flat surfaces with irrigation channels. Surfaces periodically flooded.
 Translation notes and LCCS graphical schema: one layer of "herbaceous growth forms" "cultivated" with specific floristic name "rice" and another of "natural water" "fresh" covering the field for the whole cultivation period have been considered.

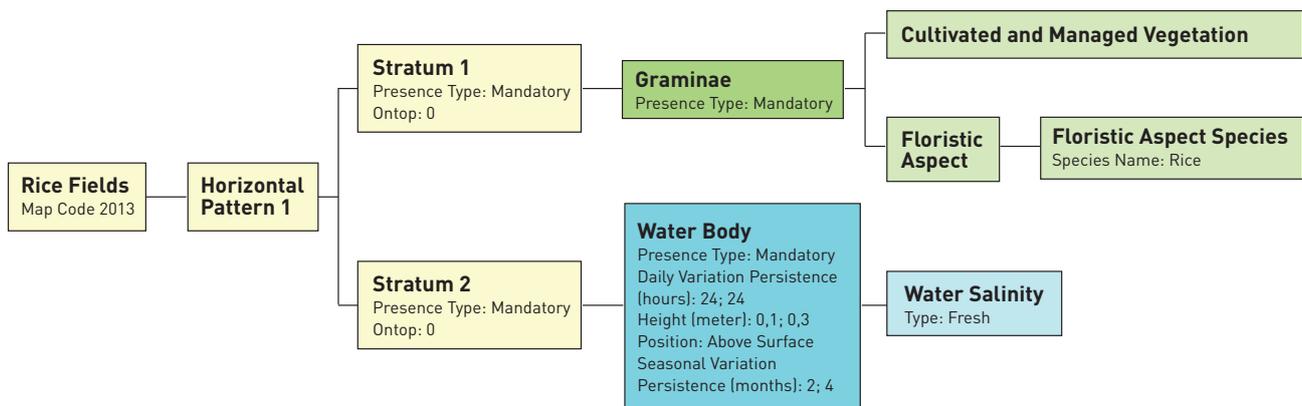


Fig.20 - Translation of the Corine "rice field" class into the LCML schema using LCCS3.

Option 2: upgrading of an existing national legend/nomenclature using the LCML/LCCS approach

General description

It consists in upgrading a national LC system, both in term of class number, extension and rationalization of class meaning. The transition from a traditional legend to a LCML/LCCS approach allows to evaluate the coherence, completeness and semantic content of a traditional legend. Very often, experts identify the limitation of the actual system and need to improve the system. This passes through the introduction of new classes or upgrading the level of information. Sometimes, the introduction of a new class, or the modification of the semantic meaning of some of them, is necessary to provide more information to emerging demands (e.g. biodiversity, climate change etc.).

In this context, LCML/LCCS is a perfect tool to ensure the consistent, systematic and functional upgrading of a previous legend structure. In this context, the general structure of the legend would be partially modified ensuring continuity with previous mapping activities. However, the rationale for upgrading the legend is based on the LCML logic. The LCML structure can become a parallel configuration to improve the national LC classification system. The LCML approach would gradually become the reference structure for upgrading all classes. This would ensure smooth transition between past and future legends. It would become obvious that increasing the semantic meaning of a class requires additional well defined parameters instead of a simple modification of the class name (or class description).

Level of technical expertise

Theoretically, the level of expertise will not be very different for the one required for the action 1. However, the modification (even if only partial) of a reference national legend requires a certain overview of the system and a specific knowledge of each features (and/or of the extra information that should be linked to the class). Any modification of the national legend requires the involvement of technicians and decision makers. An adequate documentation should be produced to explain

the reasons and the logic of the modifications/upgrading. The LCML approach and semantic structure will allow an easy and clear description of any modification.

Constraints

The overall hierarchical structure of the system will remain basically untouched. Those traditional “legend” structures, based on a rigid hierarchical arrangement of classes, can, however, be very rigid and static. Problems can occur, for instance, when it will be necessary to enhance the level of information (or to expand the number of class) only in a portion of the legend. Sooner or later these types of rigid hierarchical structure will become limited and a new architecture will become necessary.

Option 3: upgrading of field data

General description

It complements option 2. When the LCML/LCCS semantic structure is used to support the upgrading of the classification system, it becomes useful and ergonomic to modify the approach for the field data collection. This action implies the transfer of the LCML “object oriented” logic into the field data collection methodology. Basically, the field team will describe the objects of a given sample without assigning a class name. This approach simplifies the field survey avoiding the confusion between different class names and making the results from different teams comparable. In addition, it allows non-experts (which received basic training) to collect data and contribute to the built up of system. The management of field data and confirmation/upgrading of the legend classes will be done in the office by the mapping experts. In theory, they can delegate this activity to local officers once the consistency and transparency of information collected is ensured.

Level of technical expertise

It requires the production of a coherent ergonomic field form, together with a small guideline on how to collect this field information. Some field forms already exist and a conversion of them to fit national needs it is a rather easy process.

Constraints

Few days training may be necessary at local level (in the case local officers to collect the data).

Option 4: build up a national legend/nomenclature using the LCML/LCCS approach

General description

It consists of supporting the implementation of a (new) national LC system. The overall structure of the system (number and type of classes) can be tailored according to the country circumstances and reflecting the required level of information for the different applications/objectives. The overall system structure should be ergonomically linked with an “object based” approach to collect field data both for validation and information upgrade purposes. The overall structure of the database should be functional with a modern dynamic storage and query of the data. The software ADG would be used to support the “object oriented” query of the data. The full system will ensure compatibility and semantic interoperability with any other regional, national or sub-national LCML/LCCS based systems.

Level of technical expertise

Two different levels of expertise are required:

- good knowledge of LCML theoretical approach and practical use of LCCS3 software;
- extensive knowledge of the different LC features of the country.

Knowledge and practical use of the LCCS software can be handled by few experts. On the other hand, the deep knowledge of different LC aspects of the country is usually managed by a wider group of experts with different background (agriculture, forestry, rangeland etc.). Deeper is the knowledge on LC aspects of a country, higher will be level of details integrated in the national legend (and consequently in the derived database). An ergonomic combination of field, ancillary and satellite data, usually increases the level of information. The “object oriented” nature of LCML will help to make transparent those information directly derived from satellite observation from the others (e.g. rainfall, use of fertilizer, specific cultural practices etc.). This distinction is essential to set up a credible and scientifically sound procedure for the accuracy assessment (not all the class attributes can be properly assessed). It will be essential to support the legend system with an appropriate technical document.

Constraints

The transition from a traditional system to a new system may not be adequately understood by all stakeholders. The absence of pre-arrangement of the classes can complicate the setting up of the general system structure. At national level, some time will be required for users to fully understand the logic of the new system. On the other hand, allowing the representation of each class using both the old (class name and its text description) and new approach would facilitate bridging the gap and migration to a new system.

Option 5: expand the LCML based national legend/nomenclature to become the national reference system for upscale and future upgrading

General description

It consists of upgrading the activities described in Action 4. The major technical efforts will be (1) to determine a reference structure that local entities (or specific projects or NGO's) can follow to increase the thematic content of their databases. The aim of this reference structure should be to allow the production of local database with a higher degree of information and class flexibility while still remaining compatible with the overall structure of the national system. This functional system should ensure an efficient umbrella to integrate the land mapping activities from different scales and different sub-national locations. Such a system would ensure the consistency for activities at different levels and future activities that will use new sensors and/or new information extraction techniques.

Different methods can be adopted to support a National Land Cover Reference System, for sure a solid, clear and technically sound reference document is essential. However the use of a UML schema could give a more functional and comprehensive overview of the overall system structure. An alternative (or in addition) to the UML schema, a more comprehensive general graphical display can be used. The modelling capabilities of an UML schema are, however, much more effective than a simple graphical scheme.

Level of technical expertise

A panel of experts should agree on the list of attributes for the extension of the national system. An extended technical report should complement the official one (as foreseen in action 4). If the production of an UML schema is foreseen, some technical competence on the use of UML schema is required.

Constraints

A lot of attention and expertise should be provided when listing the extra attribute of the reference system. A possible constraint would be the risk to neglect some important parameters. However, the dynamic nature of attribute upgrading in LCML should allow a flexible adjustment of eventual problems.

Option 6: build up a national “relational database” using the LCML/LCCS approach

General description

Action 6 is not immediate because the LCML/LCCS language must be integrated into the structure of the most common relational databases (Oracle, Postgres etc.). The structure of LCML is already prepared for this. Therefore, this (sooner or later) will be an obvious future development of the system. However, the adoption of this approach will involve structuring the data and identifying different ways to collect the information.

The database will be structured/populated by objects and their relative attributes (see Fig.18). The conventional classes will be used later, only in the phase of data interrogation. Indeed, several possible legends will exist, according to the type of information that needs to be generated. This new approach will require a new procedure for extracting information from remote sensing and field data.

This innovative way to store and retrieve data is already partially feasible. The use of LCML/LCCS concept as foreseen in options 1 to 5 allows the end user to interrogate the resulting database by “objects” as in a relational database. Each legend classes, obtained from the translation of existing classes or directly generated from LCML/LCCS, are decoded back into objects (and their characteristics) using the FAO proprietary software ADG (Advance Database Gateway).

Level of technical expertise

Three different level of expertise will be required:

- good knowledge of LCML theoretical approach and practical use of LCCS3 software;
- extensive knowledge of a relational database;
- technical expertise to set up procedures to extract information from remote sensing data avoiding the use of predefined classes (continuous field, etc.)

Once (1) the incorporation of the LCML structure into Oracle or Postgres is done and (2) the libraries distributed and documented, this option will become much more ergonomic than the actual traditional methods for “remote sensing derived” information extraction. Most of the information we actually derive from remote sensing are continuous records transformed in discrete values to which we relate a semantic meaning (category or class).

It will be sufficient to avoid the process of creation of classes and store the information using the LCML logic with their “continuous values range” in the database. The process of categorization will be done later at the level of data interrogation.

Constraints

The transition from a conventional use of categories to populate a database to a fully object oriented system, can be difficult for the average user to understand. The extended use of ADG to interrogate existing traditional databases populated with LCML/LCCS derived classes will help understanding this new approach.





**Food and Agriculture Organization
of the United Nations**

I5232E/1/02.16