ERA-STAR Regions Project

GMES-DSL

GMES - Downstream Service Land: Austria-Slovenia-Andalusia.
Concept for a Harmonized Cross-border Land Information System

Executive Summary

submitted by

JOANNEUM RESEARCH, Institute of Digital Image Processing (JR)

GeoVille Information Systems GmbH

in co-operation with

Umweltbundesamt (UBA-AT)

GIS Styria (GISST)

Scientific Research Centre of the Slovenian Academy of Sciences and Arts, Institute for Anthropological and Spatial Studies (ZRC)

Slovenian Forestry Institute (SFI)

Slovenian Environmental Agency (ARSO)

Environment Ministry of the Regional Government of Andalusia (EGMASA)

ERA-STAR Project

GMES-Downstream Service Land (GMES-DSL)

Executive Summary

A) Project background

Land cover and land use information is essential for a row of applications – especially in public administration and the areas of spatial planning, forestry, agriculture, nature conservation or hydrology. The spatial explicit knowledge of the current status of land cover and land use in relation with the knowledge on their changes play an important role in many politic and administration areas.

On European level the issue of land monitoring is handled as priority area within the GMES (Global Monitoring of Environment and Security) programme. Currently, the usage of European wide datasets is limited for applications on regional level due to their coarse geometric resolution (e.g. CORINE Land cover with 25 ha minimum object size). Therefore some EU member countries (e.g. Spain, UK, Germany, and Sweden) have initiated national programmes that cover also these regional needs by producing more detailed land cover databases. Of major importance for the generation of regional land cover information is the compatibility and interoperability for multiple users and across different application domains, and their vertical compatibility to European wide harmonised data sets.

The objectives of the ERA-STAR Regions project GMES-DSL (Downstream Service Land) are twofold:

- to develop a concept for a cross-border (AU-SLO) land cover/land use (LC/LU) information system, and
- to integrate European-level GMES data, mostly derived employing spaceborne EO data, with data already existing in the regions (GIS & in-situ data, including aerial photographs).

As such GMES-DSL fills a major gap as appropriate and harmonized land cover inventories on national and regional level do currently not exist.

B) Organisation

The project consortium consists of research institutes (Joanneum Research – JR, Scientific Research Centre of the Slovenian Academy of Sciences and Arts – ZRC, Slovenian Forestry Institute – SFI), national and regional authorities (Umweltbundesamt – UBA, Slovenian Environmental Agency – ARSO, GIS Styria, Environment Ministry of the Regional Government of Andalusia) and a private company (GeoVille).

The project was divided into five major workpackages that represent the logical workflow in LC/LU mapping:

- Requirements – Definition of user needs & requirements
- Data Model – Definition
- Benchmarking – Benchmarking of processing chains in selected areas
- Mapping – LC/LU mapping on pilot study areas
- Evaluation – Evaluation and applicability control

C) Requirements

First, the existing data and user requirements were studied. In Austria and Slovenia no detailed area-wide land monitoring system exists. In Austria each federal state has its own geographic land information system with a vast array of GIS maps and datasets, but no coherent picture and time-series of entire Austria are available. Slovenia on the other hand has several datasets that are tailored to different users (e.g. Actual land use of agricultural and forest land of the Ministry of Agriculture, Forestry and Food, MKGP). The need for a uniform – area-wide – land information and monitoring system becomes obvious. Indeed, integration of existing data into a new developed concept system plays an important role for user acceptance.

It is inevitable to understand the user needs before developing a service. However, the development of land monitoring systems – specifically the object classes and service parameters – has to deal with two interrelated constraints:

- cost of service, and
- user requirements expressed in form of production parameters:
    - object classes (thematic detail),
    - minimum mapping unit (MMU, geometric detail),
    - time interval between updates

Having these in mind a special focused user questionnaire has been developed in the project to structure the discussion with selected users and to document these findings in a standardized procedure. In total 13 user interviews have been carried out pointing out the following main priorities for LC/LU mapping.

1) Thematic detail should be oriented towards CLC level 3, even going beyond to more thematic detail on an additional level 4 for national scales.

2) Another priority is to focus on time interval which should be between 3 and 5 years. Intermediate updates allowing for detection of hot-spots are sometimes requested after 1 to 2 years.

3) Geometric detail, i.e. the minimum mapping unit (MMU), depends on the thematic classes of interest, but should be approximately 0.1-1 ha in a map scale of 1:25.000, corresponding to the typical size of a land parcel or agricultural field. For specific objects, such as buildings, even smaller MMUs between 50 and 100m^2 are of interest, that can be represented an a map scale of 1:10.000.

The main challenge is to find the common denominator of these very heterogeneous requirements coming from various application domains (spatial planning, forestry, agriculture, hydrology, etc.) and to translate them into technical specifications for a national LC data model.

D) Data model definition

For LU/LC mapping two kinds of nomenclatures can be distinguished. Firstly, the well known hierarchical nomenclatures based on hierarchical class definitions and secondly, the – newer – object based nomenclatures that are gaining importance but have not yet been widely used due to their inherent complexity.

CORINE Land Cover is the best known hierarchical nomenclature, with a wide user acceptance, easy readability and simple technical implementation. However, drawbacks of hierarchical models include the missing comparability with other nomenclatures, the missing information for mixed classes and the absence of changes within classes if thresholds of the class definitions are crossed.

For regional applications an object oriented approach is more and more demanded because of its increased flexibility. A specific polygon is not assigned to a specific hierarchical class in the nomenclature (e.g. coniferous forest), but is described in detail and can be assigned to different classes (e.g. coniferous forest as forest or a park). The attributes of the polygon can be in a textual or quantitative manner (e.g. 85% coniferous, 15% deciduous forest). Basic land cover classes are defined as smallest entity for mapping ("simple" land cover classes), while differently composed, "complex" classes describe the mix-fraction as percentage of simple classes. With the availability of such detailed information, both land cover and land use can be obtained for each polygon. The definition of minimum mapping units plays an important role in the object based approach. A major advantage of object oriented nomenclatures is that they can be converted to standard hierarchical ones, but not vice-versa.

E) Definition of the nomenclature

The definition of the nomenclature for a harmonised LC/LU data model was done iteratively in several discussion rounds. Attributes were added to gain the advantages of the object based data model. The definition of a data model depends very much on the applied scaling; consequently 1:25.000 was chosen as appropriate scale for both regional and national applications. The users pointed out, that in urban areas there is a demand for higher geometric accuracy (i.e. delineation of individual buildings). Therefore, a second scale with 1:10.000, especially for settlements has been fixed. As an example the urban data model for the scale 1:10.000 is shown in Figure 1.

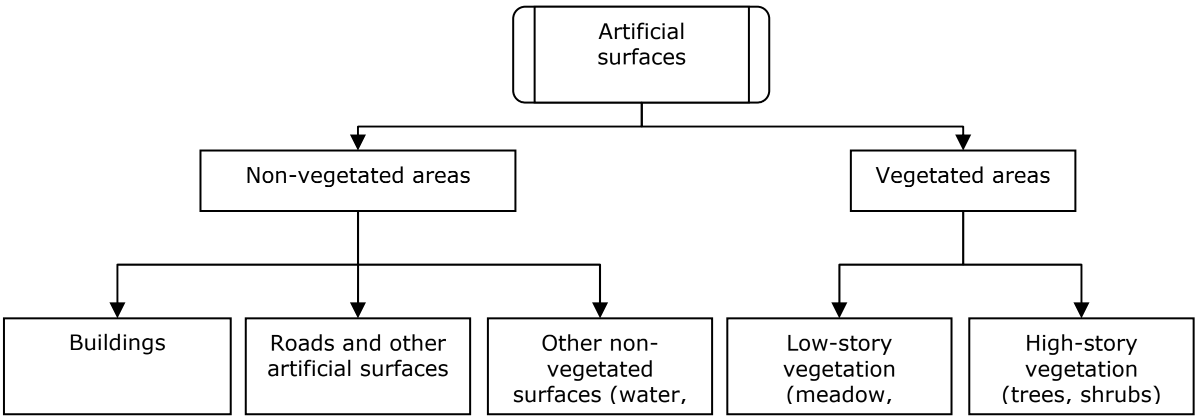


Figure 1: Simple Classes for Scale 1:10.000.

A comprehensive discussion about the final data models for settlement, forest, alpine, agriculture, water, and non-alpine nature areas is present in the corresponding chapters of the final report. It is important to note that the proposed nomenclature is tailored to the needs of regional users interviewed in this project and is probably not applicable for nation-wide realization due to cost limits. However, with the proposed detailed data model, such a regional nomenclature could be directly integrated and aggregated to future nation-wide mapping object models.

F) Pilot study areas

The data models were tested in three test areas including one cross border aree. The location of the test areas are shown in Figure 2.

Figure 2: Overview of the three test sites (Source: Google Earth)

The first test area is a cross border region with a large portion of urban areas and agricultural fields. It is located around Bad Radkersburg/Gornja Radgona. The region is mainly flat and beside the settlements and extensive agricultural areas, contains also water bodies (river, ponds). As a cross border region (divided by the river Mur) is a perfect spot for comparing and harmonizing cross border data. The second test area focuses mainly on alpine and forest environments. With respect to the availability of ancillary data and the diversity of nature, the area around Schladming was chosen. It covers forests, rocks, glacier, alpine lakes, touristic infrastructure with skiing areas, golf courses and so on. The third test area encompasses the town of Kobarid with its surroundings which mostly represent forests with overgrown pastures and alpine areas with rocks and screes. The following image and ancillary data sets have been used:

Table 1: Used image and ancillary data.

Data		Use, properties
GMES data	Image 2006 satellite data (SPOT, IRS)	Basic data, available for all test areas
	Mapping products (Soil sealing layer, GSE-FM forest map)	Involved in mapping and stratification procedure
Other data	SPOT 5	Acquired within the project for the cross-border area
	Orthophoto	High resolution, available for all test areas
	Ancillary data (digital cadastre, agricultural land use map, road network, topographic maps, forest stand maps etc.)	Integrated in all levels of classification

G) Stratification

The production can be split in thematic or geographic workflow:

- Thematic: The area is split (stratified) into thematic domains (urban, forest etc.) and one production partner is responsible for one or more specific thematic domains.
- Geographic: The area is split according to geographic criteria (e.g. administrative units) and one production partner is the sole responsible for the entire area.

Another dimension for the production layout is a centralised or decentralised production. For centralised production, some tasks are done by only one partner compared to a decentralised approach, where all tasks are done by every partner.

Internal discussions about the methodology lead to the decision to follow a thematic work split, i.e. utilizing the partners' expertise and domain knowledge for specific data models. The consortium is aware that stratification can be handled only as first approximation. In case of errors of the stratification or temporal changes, the stratification data set must be corrected before starting the classification or the classification system has to be able to eliminate these errors using several classification loops between the models.

Based on already existing data, the following sequence was chosen for stratification:

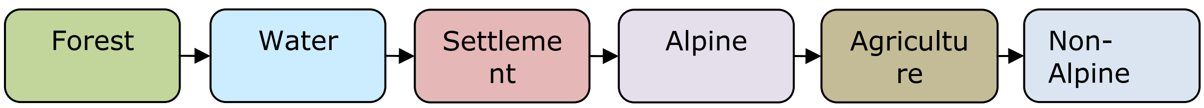


Figure 3: Stratification procedure.

- Forest is based on the forest layer from GMES Service Element Forest Monitoring (GSE-FM). The scale and MMU is appropriate for maps in 1:25.000.
- Water is stratified next from existing data sets from public administration and added to the stratification map
- In the next step urban areas are extracted. This stratification step is the most labour-intensive one as no up-to-date and accurate representation of the urban outline exists.
- The remaining part of the surface consists of agricultural and natural areas located in alpine terrain or lowlands. For the separation of these two strata, a DSM and the forest layer are used in a semi-automated processing. The

agriculture stratum at this point still includes non-alpine natural areas. However, the whole area is processed with the agricultural data model and not classifiable objects are then in the end summarized into a non-alpine nature class.

Various attempts and methods were tested to define the optimal workflow for the stratification. Basically the strata were obtained from ancillary data although also automatic procedures were tested. The processing was different for the Austrian and Slovenian part due to the availability of diverse ancillary datasets. Thus, the stratification for Slovenia was mainly straightforward because of the use of the agricultural land use map dataset from which the forest, water, alpine and partially settlement stratum were extracted. The stratification for Austria involved the use of various datasets (cadastre, GSE forest layer, orthophotos) and subsequent manual refinement of the domains, especially for stratification of the urban outline.

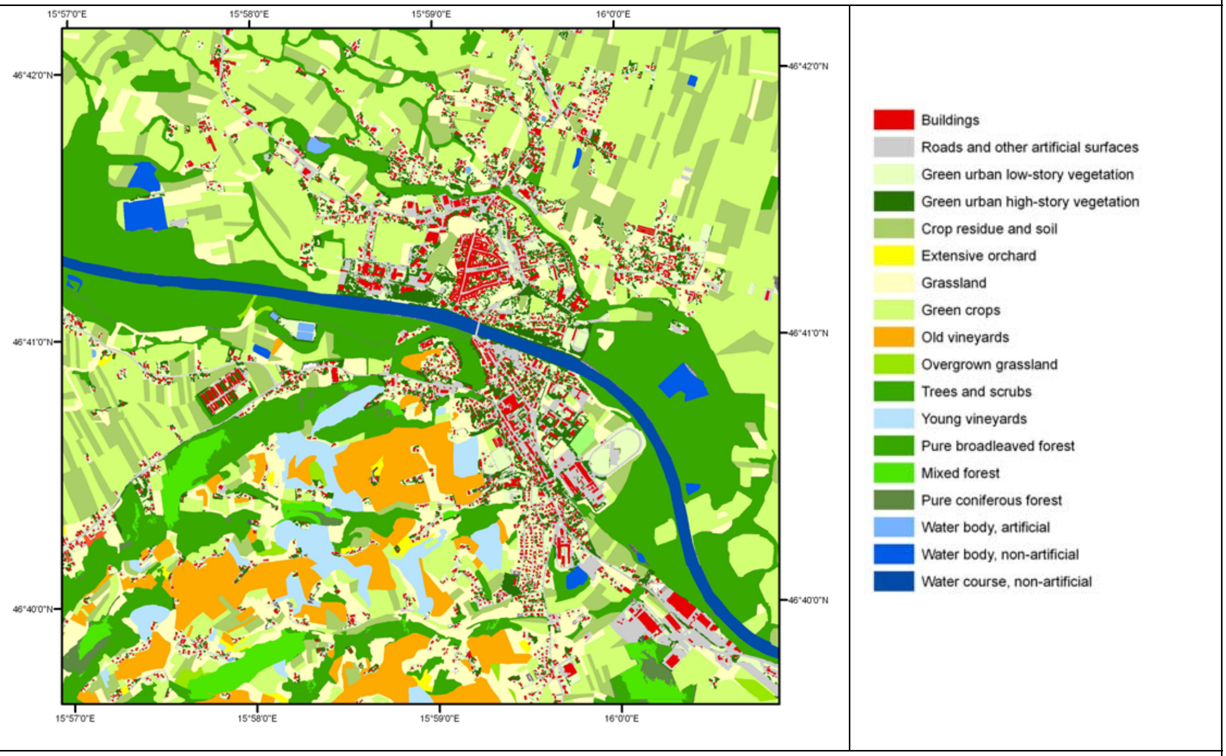
H) Mapping results

Mapping was performed using previously stratified areas as a mask. Because of this approach the mapping could be divided between partners, each handling a specific area containing only the respective classes. The starting classes were defined by the nomenclature but were still subject to change in the ongoing processing.

The methods used for classification are different in the various thematic domains. Often automatic image segmentation followed by automatic (rule based, machine learning etc.) and visual (manual) classification were used. Visual interpretation was mostly used for classification refining. The processing line and results for the domains forest, water, settlement, alpine and agriculture are provided in the corresponding chapters of the report.

The forest domain was processed by SFI and JR. SFI used automatic image segmentation followed by machine learning decision trees method on aerial and satellite imagery; on the other hand JR processed SPOT and Landsat data with two different pixel-based approaches (maximum likelihood vs. regression estimator). The water and settlement domains were mapped by GeoVille. Water classes were extracted from ancillary datasets and orthophotos by means of visual interpretation. The extraction of settlement classes required extensive work due to the mapping in two scales. For the 10:000 scale the orthophotos were first segmented and automatically classified using a knowledge-based decision tree (rule based). For the smaller scale various ancillary data (soil sealing, road network, topographic maps, etc.) were used to segment and classify the images. As the 1:25.000 scale mainly deals with land use, visual interpretation is needed to achieve a reliable separation. The alpine stratum was classified with visual interpretation by JR. The remaining agricultural stratum was processed by ZRC and JR. The involved procedures ranged from automatic segmentation and classification by selecting examples and subsequent manual refining to completely visual segmentation. The final result of mapping in the Bad Radkersburg / Gornja Radgona test site is shown in figure 4.

Land use map 1:25.000 (with urban areas in scale 1:10.000)



Land use map 1:25.000 (with urban areas in scale 1:25.000)

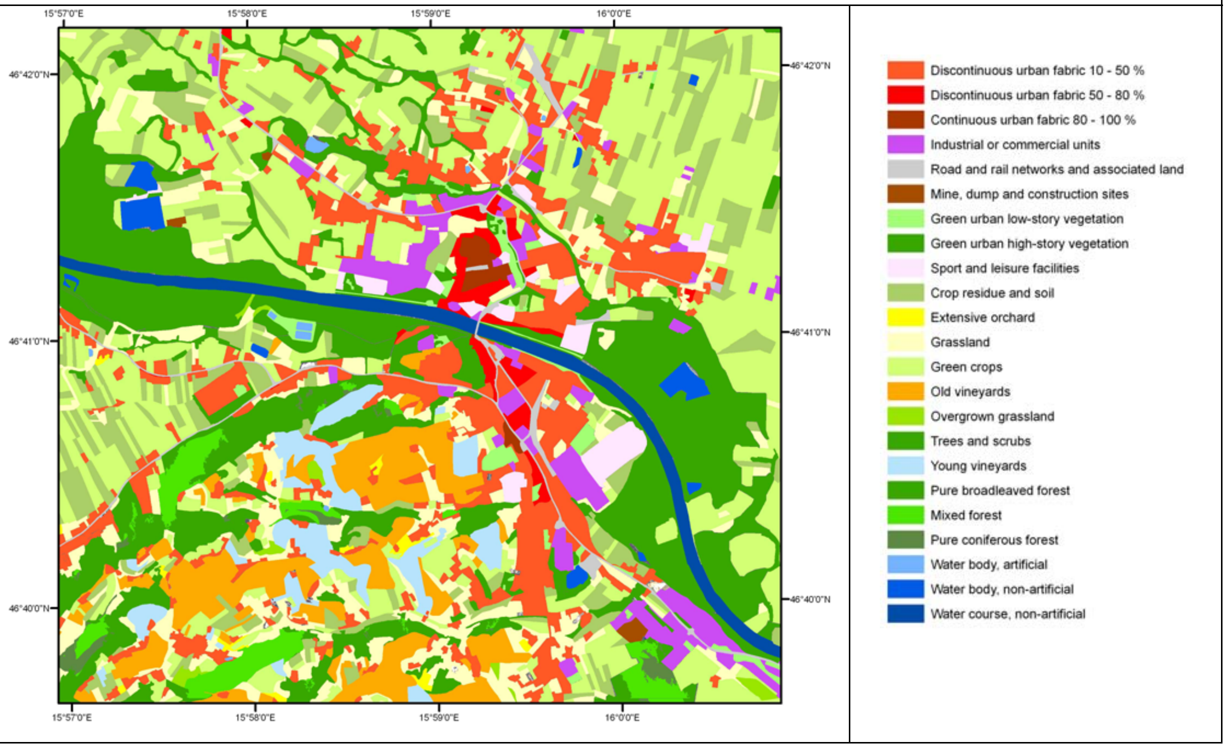


Figure 4: Mapping results for the Bad Radkersburg / Gornja Radgona test site.

I) Benchmarking

Benchmarking implicates comparisons of the various processing methods and datasets in order to find out the best technique to materialize the theoretic data model into a land use/cover map. Combinations of different remote sensing data as well as pixel- and segment-based methods are tested and special attention is given to the workflow procedure. The results are summarized in the following table.

Table 2: Tested classification methods.

Domain	Method	Description
Urban	Pixel-based	Use for land use mapping is very limited, but useful for classification of soil sealing degree
	Segment-based	Better accuracy for mapping of "simple land cover objects", used with road network for mapping of land use objects
Water	Visual interpretation	Based on segments, extracted from mapping products
Agriculture	Pixel-based	Inaccurate, only with satellite data
	Segment-based	Better performance with orthophoto, supervised classification
	Visual interpretation	Only segmentation, time consuming
Forest	Pixel-based	supervised classification or regression analysis, information on tree types, crown coverage and growth
	Segment-based	supervised classification using machine-learned decision tree, good results

J) Evaluation

The evaluation is composed of feedback from the involved users and complemented by internal statistical evaluation of the mapping results. Regarding urban and water, a classification accuracy of 95% (scale 1:10.000) and 97% (1:25.000) was achieved by comparing the result with 600 visually interpreted points. Regarding forest, feedback from three users was obtained. JR presented the results to two forestry experts from "Land Steiermak – Forstwesen 10C", Austria. To get feedback from Slovenia, SFI had a meeting with the head of forest management planning at the Slovenian Forestry Service. Information of crown cover, tree type composition, growth class and clearances on a raster basis of 10m x 10m, as well as smaller forest lines in e.g. agriculture provides an excellent data source for many forestry projects. Scale of 1:25.000 is applicable for broader environmental applications, but is of restricted use for detailed silvicultural planning. Pixel-based classification is appreciated by Austrian users, whereas in Slovenia segment-based results were regarded as useful. The flexible classes without sharp class definition as suggested in the nomenclature were very much appreciated. Statistical evaluation of the agricultural classification in Slovenia led to an average accuracy of 87.1% for six classes. In order to improve the classification accuracy for different crop types, the use of multi-temporal image data has been recommended.

K) Main Conclusions

In the project the feasibility of a regional cross-border land cover/land use information system (Austria and Slovenia) has been demonstrated. The developed object oriented data model has been tested in different study areas and the most suitable processing methods have been extensively evaluated for the classification of settlements, forests, alpine areas, agriculture, water, and non-alpine nature areas. The developed concept is a GMES-compatible downstream solution, as it benefits from existing GMES data and demonstrates how GMES data from European level can be integrated with in-situ data (e.g. aerial orthophotos) available in the regions. This has been shown for example for urban areas with the existing GMES soil sealing layer and also in the area of forestry by using the GSE-FM forest mask.

Since the needs of users are very diverse a compromise has to be achieved between user requirements and the technical and economic feasibility. Some classes can be mapped automatically from remotely sensed data, while others may require extensive visual interpretation and detailed ancillary data. Thus both,

- a hybrid data model bringing together the advantages of object-oriented with hierarchical data models, and
- a combined mapping approach, i.e. a well-balanced combination of automated classification methods and supportive visual interpretation using ancillary data,

is considered the optimal solution.

The first data model and its evaluation w.r.t. technical and economic feasibility provide the fundament for the development of Austrian Land Information System LISA. However, this project had a regional (Styria-Northern Slovenia) focus with test areas allowing very detailed assessments. For a large area roll-out the individual methods and results have to be evaluated especially with respect to the connected costs.

The project has demonstrated the successful implementation of a LISA precursor and is an important contribution to improve sustainability of GMES downstream services. The multi-user and multi-purpose data model has been developed based on a first round of user interviews with Austrian and Slovenian users and experiences gained from Spanish SIOSE. The three test sites in Austria and Slovenia (incl. one cross-border site in Bad Radkersburg/Gornja Radgona) have been used as a joint feasibility test with project partners. A first estimation of technical and economic feasibility (i.e. cost matrix) has been elaborated and the proposed methodology successfully passed the internal and external evaluation. Last but not least the joint project has strengthened the co-operation at a national and regional level between users, researchers and companies in Austria and Slovenia in the field of the GMES Fast Track Service Land.