A STUDY ON THE AUTOMATIC REVISION OF THE EUROPEAN COMMUNITY'S CORINE LAND COVER DATABASE USING SATELLITE DATA

G. Wilkinson, S. Folving, K. Fullerton and J. Mégier
Environmental Mapping and Modelling Group
Institute for Remote Sensing Applications
Joint Research Centre
Commission of the European Communities
21020, Ispra, Varese, Italy.

ISPRS COMMISSION IV

ABSTRACT: This paper describes work in progress at the European Commission's Joint Research Centre on the potential for automatic revision of land cover maps using remote sensing data for the whole of Europe at 1:100,000 scale. The research on the automatic updating is concerned with a number of important issues in thematic mapping such as high accuracy image classification, image data generalisation, correspondences between image classes and pre-defined mapping classes. Studies are reported which have been undertaken in Portugal on inter-comparisons between first edition and revised land cover maps which reveal considerable subjectivity problems in the conventional mapping procedures. Some conclusions are drawn on the implications of these subjectivity issues for automatic updating via remote sensing and some possible approaches, involving the possibility of knowledge-based decision-making are described.

KEY WORDS: Land Cover Mapping; Map Revision; GIS; Thematic Mapping.

1. INTRODUCTION

1.1. The CORINE Programme

The CORINE programme of the European Commission is concerned with the production of a large Geographic Information System containing 22 data layers representing the landscape of the twelve member states of the European Community (CORINE 1991). Examples of the data layers currently being produced are:

- Biotopes
- Water Resources
- Soil Erosion Risk
- Land Cover
- Soil Types
- Climate
- Administrative Units

These data layers are being assembled to provide an integrated geographical and statistical view of the European Community, although the cartographic scales of the datasets are not all the same. One of the most important requirements of the complete information system is accuracy both in the spatial and temporal sense. Since it will be used widely througout Europe and will be used to provide "value-added" products, absolute accuracy in the CORINE base maps is essential to avoid the problem of error accumulation in the derived products -a problem which can be very serious in using GIS data (Heuvelink, Burrough and Stein 1989).

1.2. The CORINE Land Cover Database

One of the most important layers in the overall CORINE system is the Land Cover database. This database will eventually contain a land cover map of the entire European Community at 1:100,000 scale as a digital polygon "coverage" plus attribute information. In order to produce such a database a uniform classification scheme - The Land Cover Nomenclature - has been devised to represent the landscape in all the member states. The Nomenclature is hierarchical and contains three levels with forty-four classes at the lowest (most detailed) level -see list below:

Corine Land Cover Class Nomenclature:-

1.1.1 Continuous Urban Fabric
1.1.2 Discontinuous Urban Fabric
1.2.1 Industrial or Commercial Units
1.2.2 Road and Rail Networks and Associated Land
1.2.3 Port Areas
1.2.4 Airports
1.3.1 Mineral Extraction Sites
1.3.2 Dump Sites
1.3.3 Construction Sites
1.4.1 Green Urban Areas
1.4.2 Sport and Leisure Facilities

2.1.1 Non-Irrigated Arable Land
2.1.2 Permanently-Irrigated Land
2.1.3 Rice Fields
2.2.1 Vineyards
2.2.2 Fruit Trees and Berry Plantations
2.2.3 Olive Groves
2.3.1 Pastures
2.4.1Annual Crops Associated with Permanent Crops
2.4.2 Complex Cultivation Patterns
2.4.3 Land Principally Occupied by Agriculture
with Significant Areas of Natural Vegetation
2.4.4 Agro-Forestry Areas

3.1.1 Broad-Leaved Forest
3.1.2 Coniferous Forest

3.1.3 Mixed Forest
3.2.1 Natural Grasslands
3.2.2 Moors and Heathland
3.2.3 Sclerophyllous Vegetation
3.2.4 Transitional Woodland-Scrub
3.3.1 Beaches, Dunes, Sands
3.3.2 Bare Rocks
3.3.3 Sparsely Vegetated Areas
3.3.4 Burnt Areas
3.3.5 Glaciers and Perpetual Snow

4.1.1 Inland Marshes
4.1.2 Peat Bogs
4.2.1 Salt Marshes
4.2.2 Salines
4.2.3 Intertidal Flats

5.1.1 Water Courses
5.1.2 Water Bodies
5.2.1 Coastal Lagoons
5.2.2 Estuaries
5.2.3 Sea and Ocean

Apart from the fixed Nomenclature and the common map scale employed in each member state, the CORINE land cover maps contain only land cover parcels with a minimum size of 25ha. Hence all mapped areas have to be generalised into polygons of such a size as to match this constraint.

In principle the defined Nomenclature is used for each member state in the EC. Some states, however, have introduced a fourth level of detail for local purposes, though this does not form part of the CORINE standard.

The Nomenclature was originally designed with the objective of representing all types of landscape including northern European/Scandinavian landscapes (e.g. Ireland, Denmark) and Mediterranean landscapes (e.g. Portugal, Greece).

2. STATUS OF THE EC LAND COVER MAP

2.1 Current Situation

The total area of the EC Member States is approximately 2.25 Million hectares. The mapping for this area is being undertaken by project teams in the individual states under the direction of the CORINE Land Cover Team which is part of the European Environment Agency Task Force.

The work on the CORINE project commenced in 1985, and by the beginning of 1992 the full land cover maps had been produced only for Portugal and Luxembourg (and in addition the island of Corsica).

The initial mapping exercise is being undertaken primarily by manual photo-interpretation using hard copy Landsat Thematic Mapper images. In some areas additional information in the form of topographic maps, orthophotos and air photos have been used. At the present time the full database status is shown in table 1.

DB STATUS	COUNTRIES
Digital database complete	Portugal Luxembourg
Information at map sheet level digitised (but not yet integrated into database)	Spain Netherlands
Images selected/pre-processed and/or interpretation in progress	France Belgium Italy Greece Denmark
Information at national level in preparation but no interpretation yet carried out	Ireland Germany

Table 1. Status of the CORINE Land Cover Project, Spring 1992.

The United Kingdom, has followed a separate methodology which does not fit with the CORINE land cover definition and is therefore not included in the categories in table 1. However it is likely that the UK product will be integrated into the land cover database in the future. Also, several states outside the Community (e.g. Eastern European and North African states) have shown interest in following the CORINE Land Cover approach and it is likely that the eventual database coverage will be more extensive than just the 12 existing members of the EC.

2.2. The Need for Updating

For statistical accuracy it is necessary to revise the land cover maps on a regular basis to reflect real changes in the European landscape besides correcting any errors existing in the first edition maps. At present the methodology for revising the maps and the revision frequency have yet to be defined. However it is apparent that for economic reasons the revision procedure should be automated as much as possible and that the frequency of revision should be sufficient to observe significant land use changes which either require central intervention by the Community or which affect Community policy. For these reasons the Joint Research Centre (JRC) is investigating the potential role of automatic methods based on digital remote sensing data.

3. EXPERIMENTAL STUDIES ON FIRST EDITION MAPS

3.1 Experiments in Portugal

Our studies on possible updating approaches have so far been restricted to Portugal and Luxembourg as the only two member states with completed maps.

In 1991 an extensive field survey was carried out in both countries with the aim of understanding the characteristics of the first edition maps and to identify any constraints that would apply to the map revision exercise. A total of eleven test sites were selected all of which were within approximately 60km. of the city of Lisbon (see figure 1). Each test site was 3km. x 3km. in size. These test sites were selected to cover as much variety of land cover as possible and they ranged from concentrated urban areas (Sites 5, 6), to agricultural areas (Sites 3, 7, 8, 10, 11), forested areas (Sites 1, 2, 9, 10) and natural vegetation areas (Sites 2, 4).

On each site two separate types of survey were made. The first was concerned with producing a new land cover map following the convention of the CORINE project. The aim of this was mostly to understand the accuracy of the original (1985) map and its reproducability when a different mapping team is used. The second type of survey was concerned with recording areas of landscape homogeneity that could be used in experiments on image classification. For this type of survey very detailed information was recorded such as vegetation type, canopy conditions, stand height, soil type, soil colour, soil wetness, and percentage cover etc.

All data obtained in the field surveys were digitised and entered into an ARC/INFO system. The original Portuguese CORINE map was also available in ARC/INFO format allowing inter-comparisons to be made.

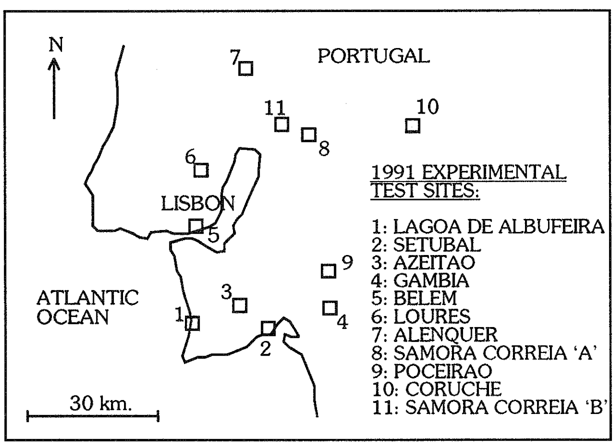


Figure 1. Test Sites Used in 1991 Field Experiments

Multitemporal sequences of digital Landsat Thematic Mapper imagery were also acquired from 1985 (i.e. at the same time as the original mapping exercise) and from 1991 (coincident with our own revised field survey). The images were rectified and co-registered with the digital map information in the UTM projection.

3.2 Land Cover Map Reproducability Studies

Following the assembly of all the above-mentioned map and image datasets with a 6 year time gap, it has been possible to undertake a variety of studies on change detection and map revision besides on original map accuracy. Most of the work to date has been concerned with the inter-comparison of the 1991 survey CORINE-type land cover map and the original 1985 Land Cover map product on the the eleven mapped test sites. Typical results from this are shown in figures 2 and 3.

Figure 2 shows such a comparison on a coastal zone test site which contains a lagoon (running from upper right to lower left of each box) and areas of forest, heathland and agriculture. In fact this area has seen little land cover change over the 85-91 time period and the lagoon and beach front (to the far left of each box) impose a basic similarity on the two maps. However in the zones

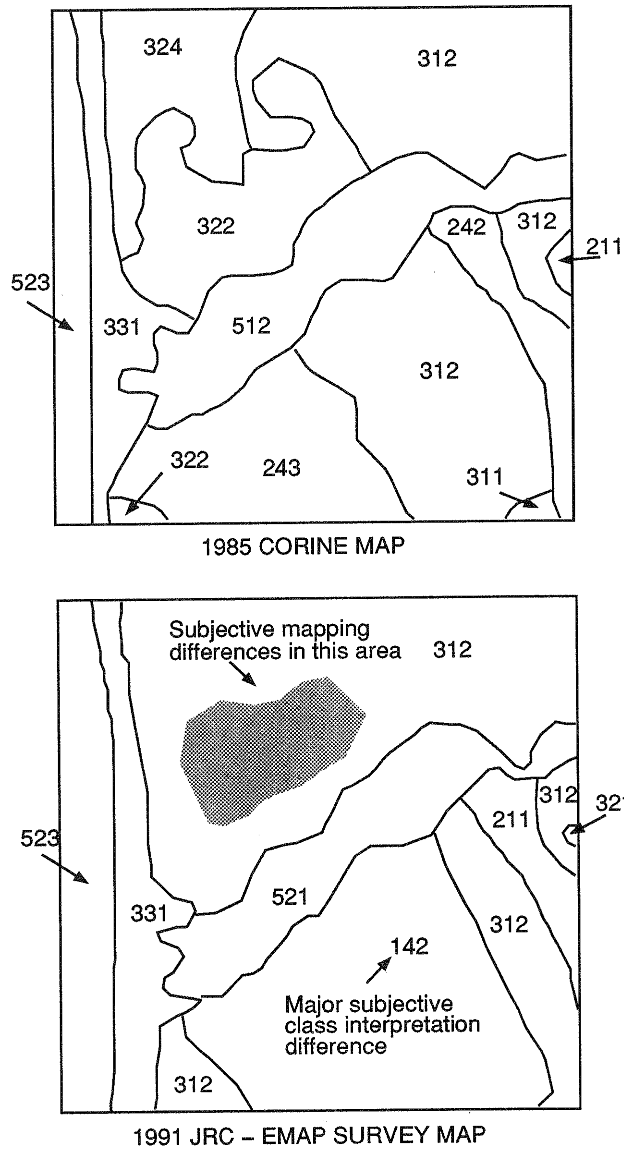
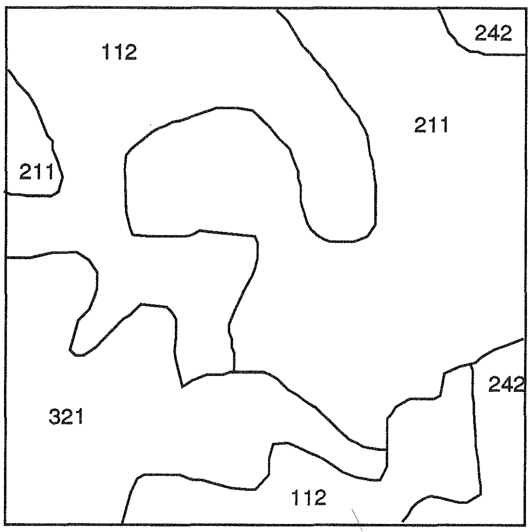
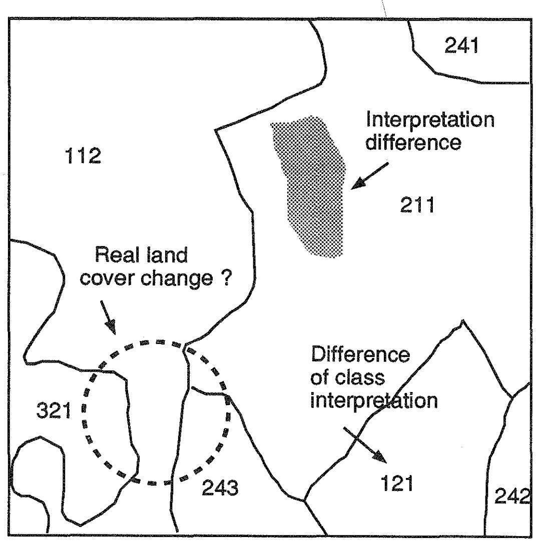


Figure 2. Intercomparison of Original Land Cover Map and JRC 1991 Survey - Lagoa de Albufeira Test Site

above and below the lagoon there are considerable mapping differences. These differences involve both different parcel boundaries and different land cover class labels. In general these differences arise from the subjectivity which is inherent to the CORINE mapping process. In many areas the European landscape, especially in the Mediterranean zone, contains very mixed terrain which is not easily categorised by the 44 classes of the standard nomenclature. Hence a considerable degree of subjective judgement comes into play when humans are forced to make the landscape conform to this model. This subjectivity is readily revealed by the intercomparison of figure 2. One of the most striking differences between the two maps of figure 2 concerns the parcel which was labelled class 243 (land principally occupied by agriculture with significant areas of natural vegetation) in the 1985 survey. This is in fact an area of holiday homes with large gardens and natural vegetation areas as ground inspection reveals. The original map class is not a completely valid description of the area. In the new survey of 1991 the 'recreational' nature of the area was



1985 CORINE MAP



1991 JRC – EMAP SURVEY MAP

Figure 3. Intercomparison of Original Land Cover Map and JRC Survey 1991 - Loures Test Site

recognised and the parcel (with a different extent) was labelled with class 142 ("sport and leisure facilties"). Again this is not an entirely appropriate description. In fact a class discrepancy of this magnitude is quite serious -the parcel is within the group of "agricultural" classes in the 1985 map and in the group of "artificial" or man-made classes in the 1991 survey. An area of between 1 and 2 square km. thus changes between major class groups without any real change on the landscape. The repetition of this kind of discrepancy at a European scale between map revisions would lead to highly erroneous land cover change statistics.

Figure 3 shows the same kind of comparison on a test site near the town of Loures. Here there is a very marked difference between the parcel boundaries since there are no dominating landscape features to impose a clear structure to the map. In fact the area is a dormitory town to the north of Lisbon. The mapped area contains some large modern housing estates intermingled with older low density housing, some light industry, some areas of natural vegetation, and some agriculture and abandoned land. When the area is generalised to 25 hectare parcels (as a minimum) it is extremely difficult to decide where the boundaries should be optimally placed. The result is two very different maps. Certainly there have been some real land cover changes in this area between 1985 and 1991. However the statistical significance of these changes will be lost in the general subjective confusion. The over-riding conclusion from such comparisons is that a land cover map revision procedure relying on the repetition of a subjective human process is unlikely to lead to a statistically useful result. The requirement for a more automatic, repeatable, procedure is paramount.

4. TOWARDS AUTOMATED LAND COVER MAP REVISION

4.1 Implications of the Portugal Study

One of the most important conclusions arising from the 1985-1991 map intercomparisons is that an attempt to revise the CORINE land cover maps by undertaking a re-mapping exercise (following the original procedure) will not yield a result which has a good statistical relationship with the original map on account of the subjectivity inherent to the photo-interpretation and spatial generalisation processes. Hence a change detection approach is more suitable using the original maps as the base-line against which changes should be observed. This effectively perpetuates the subjectivity of the first edition maps which may lead to the perpetuation of certain kinds of errors in the generalisation. However by following this change detection approach the statistical continuity is ensured allowing important studies to be undertaken on changes in land use over time. The use of the original map as a base-line for change detection also facilitates easier use of remote sensing data in the updating procedure since imagery can be partitioned according to the original parcel boundaries and decisions on revising class labels and/or re-drawing parcel boundaries are local rather than global.

4.2 Stages of an Automatic Revision Procedure

At the present time, the procedure for carrying out "automatic" updating is still under investigation. However, some of the key steps can be outlined.

One of the most important tasks is initially to determine which of the original land cover map parcels have changed in a significant way. This can be done in principle by detecting changes in the imagery within each of the original parcels. However this is a non-trivial problem because it is not possible to compare satellite radiances in old and new imagery on a pixel-by-pixel basis alone. The use of old and new images on slightly different days of year, with different illumination conditions, slightly different vegetation conditions, soil wetness etc. can create significant pixel radiance differences even though the main land cover features have not changed. For this reason it is much more acceptable to work in the classification domain -i.e. to initially classify pixels and then to make comparisons between old and new to detect changes. However this procedure will not be useful if conducted on a pixel-by-pixel basis since small scale changes may not be meaningful for parcels of at least 25 ha. Also small errors in old and new image co-registration (of the order of 1 - 2 pixels) may lead to large differences being detected between pixels in very patchy inhomogeneous zones at the pixel scale e.g. large houses with gardens, small-holding farms or horticulture with regularly interspersed buildings and cultivated plots. For this reason we are investigating a possible two-component approach to change detection in imagery as follows:

(i) within each parcel boundary from the original map produce class histograms from old and new imagery, then look for significant differences in class composition

(ii) within each parcel boundary from the original map perform a cross-correlation of the old and new pixel radiance data and then to threshold the result in order to detect major spatial differences.

A combination of (i) and (ii) then permits parcels with significant land cover change to be flagged. However it is axiomatic that once a land cover parcel is flagged as having changed its immediate neighbours will also be potential change candidates because of the possible requirement to re-draw boundaries. The posible parcel change detection approach is shown in figure 4.

4.3 Development of an Image Classifier with Good Generalisation

One of the most important requirements of the change detection procedure is a classification of the old and new imagery into as many classes as possible with high accuracy over large geographic areas. A potential method for doing this is to use an artificial neural network classifier which has been demonstrated in experiments undertaken at the JRC to yield results considerably better than conventional statistical image classifiers with large numbers of classes (Kanellopoulos et al. 1992). Work on further developing these classifiers is continuing at the JRC. The neural network approach has a distinct advantage in that it is appropriate for parallel processing. Studies are currently being undertaken at the JRC on implmenting neural networks for remote sensing data analysis on parallel transputer networks.

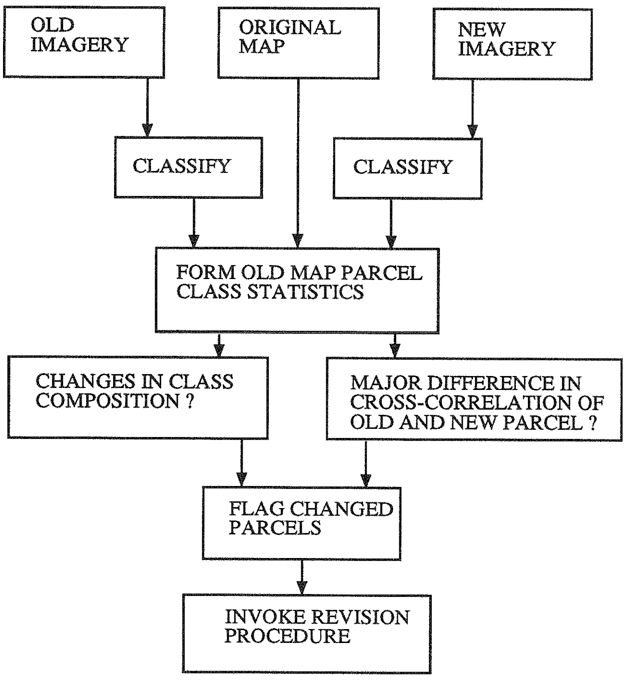


Figure 4. Proposed Procedure for Parcel Change Detection

4.4 Development of Rule-Based Boundary and Class Change

Perhaps the most difficult stage of "automatic" CORINE map revision is the decision making involved in (a) changing a class label for a parcel which has been detected as 'changed' and/or (b) re-drawing the boundaries for a changed parcel and its immediate neighbours for which there are topology-preserving and topology non-preserving possibilities.

We believe a rule-based approach is indispensable in this context and that a detailed parcel-by-parcel evaluation has to be made where the significant image class changes have occurred. Essentially these changes have to be isolated and their impact on boundaries and CORINE class labels carefully evaluated. This requires also that a useful relationship can be established between detectable image classes and CORINE land cover classes. There is a possibility of building up statistics of image class membership within the CORINE land cover nomenclature clases by examining the image classes detected (on old imagery) inside each parcel in the original land cover map. It will then be possible to determine the most likely new CORINE class label(s) for a changed parcel using its new image class composition on a probabilistic basis. However this procedure can not be separated from the spatial analysis since the image class composition of a parcel is dependent on its boundary. Procedures are thus suggested which involve, for example,: (a) identifying which pixels have changed, (b) deciding if these represent a growth of some existing landscape feature or a completely new feature (c) deciding whether simply to move an existing boundary or to create a new polygon and (d) deciding if new CORINE class labels are appropriate given the new proposed parcel geometry. Such complex inter-related decision-making in our view requires an expert system approach if it is to be automated. Such a system would require many rules and a good probabilistic evidential reasoning model.

These approaches are currently being investigated at the JRC and are likely to form an important component of our work on land cover map revision.

5. DISCUSSION

The revision of a land cover map with continental coverage is a formidable task by any approach. We believe remote sensing offers a possibility to do this automatically but there are many technical problems to be solved. Many of these problems are common to a multitude of applications of remote sensing -i.e. how to accurately extract large numbers of classes from satellite imagery, how to get good performance over wide geographical areas, how to spatially generalise products in a meaningful way. Work is underway on all these issues at the JRC and we hope to exploit techniques such as neural networks and expert systems in helping to solve the technical problems. The automatic updating of the CORINE land cover map makes an extremely interesting test case for such possibilities.

REFERENCES

CORINE, 1991. Corine - Examples of Use of the Programme 1985-1990. Office for Official Publications of the European Communities, Luxembourg, 57pp.

Heuvelink, G. B. M., Burrough, P. A. and Stein, A. 1989. Propagation of Errors in Spatial Modelling with GIS. Int. J. Geographical Information Systems, 3(4): 303-322.

Kanellopoulos, I., Varfis, A., Wilkinson, G. G. and Mégier, J. 1992. Land Cover Discrimination in SPOT HRV Imagery Using an Artificial Neural Network -A 20 Class Experiment. Int. J. Remote Sensing, 13(5): 917-924.