S E S 2 0 1 3
Ninth Scientific Conference with International Participation
SPACE, ECOLOGY, SAFETY
20 – 22 November 2013, Sofia, Bulgaria

VERIFICATION OF FOREST AND WATER BODIES HIGH RESOLUTION LAYERS 2012 FOR BULGARIA

Ventzeslav Dimitrov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: vdimitrov@stil.bas.bg

Keywords: *Land cover, satellite image, high resolution layers, verification*

Abstract: *Production of high resolution layers (HRLs) is an essential part of the pan-European component of Copernicus programme's land monitoring service. In the frame of the initial operations of the service five HRL were created by a number of service providers from multi-sensor and multi-temporal satellite images for the reference year 2012: imperviousness, forests, grasslands, wetlands and permanent water bodies. This paper is focused on the implementation of verification task for forest and permanent water bodies HRL products for Bulgarian territory. The purpose of the verification is to identify systematic classification errors suitable for further correction. It is carried out through visual inspection of stratified samples in the HRL using reliable reference data sets. Verification activity includes three steps, the first two – obligatory: general overview of data quality, look-and-feel verification and statistical verification. Stratification scheme and evaluation grades for Forest and statistical accuracy estimates for PWB HRL are presented. Results show that Forest and PWB HRL products meet the 85% accuracy requirements.*

ВЕРИФИКАЦИЯ НА СЛОЕВЕ С ВИСОКА РЕЗОЛЮЦИЯ 2012 ЗА ГОРИТЕ И ВОДНИТЕ ТЕЛА ЗА БЪЛГАРИЯ

Венцеслав Димитров

Институт за космически изследвания и технологии – Българска академия на науките
e-mail: vdimitrov@stil.bas.bg

Ключови думи: *Земно покритие, спътниково изображение, слой с висока резолюция, верификация*

Резюме: *Създаването на слоевете с висока разделителна способност (резолюция) (СВР) е съществена част от пан-европейски компонент на услугата за мониторинг на земята на програмата „Коперник". При началното функциониране на услугата няколко доставчика на услуги създават пет СВР от мултисензорни и разновременни спътникови изображения с базова година 2012: „Степен на непроницаемост"," Гори", „Пасища и ливади", „Влажни зони" и „Постоянни водни тела". Този доклад е фокусиран върху изпълнението на задачата за верификация на СВР "Гори" и "Постоянни водни тела" за територията на България. Целта на верификацията е да открие систематичните грешки от класификацията, поддаващи се на последваща корекция. Тя се извършва чрез визуална проверка на стратифицирана извадка проби от СВР с използването на надеждни референтни множества от данни. Работата по верификацията включва три стъпки, първите две – задължителни: общ преглед на качеството на данните, визуална проверка по принципа „гледай и чувствай" и статистическа верификация. Представени са схеми за стратификация и оценъчни степени за горите, както и статистически оценки за постоянните водни тела. Резултатите показват, че СВР "Гори" и "Постоянни водни тела" постигат изискваната точност от 85%.*

Introduction

The Copernicus programme, (formerly Global Monitoring for Environment and Security - GMES) combines satellite-borne earth observation, in-situ data and services to provide value added information, necessary for monitoring the environment. The initial operations phase of the land monitoring service (GIO land) produces more detailed complementary information in parallel with the

updating of CORINE Land Cover (CLC) for the reference year 2012. Five high resolution layers (HRL) describing main land cover characteristics are currently in production: imperviousness, forests, grasslands, wetlands and permanent water bodies [1] HRLs are created through applying sophisticated analysis algorithms at pixel level of multi-sensor and multi-temporal satellite imagery with spatial resolution of 20x20 m. These so called intermediate products set up the input for verification and enhancement, which are implemented by separate countries. As a result enhanced final products of countries are created.

Experience in working with one of the HRL already exists in Bulgaria, as in 2006-2008 the Soil Sealing HRL was evaluated as a part of the GMES Fast Track Service [2], [3]. The European validation of the soil sealing layer was further performed in 2011 using a dedicated methodology for statistical evaluation and calculation of quantitative accuracy figures for soil sealing data [4].

This paper is focused on the implementation of verification task for forest and PWB HRL products for Bulgarian territory. The verification has to identify systematic classification problems suitable for further correction checking for commission and omission errors against reliable in-situ spatial data sets.

Data used

Several data sets were involved in the verification process. By their origin they could generally be subdivided to space borne and in-situ data. The space borne group was formed from country-wide data sets representing part pan-European satellite image coverages or their derivatives, namely:

- satellite images from different sensors;
- HRL layers;
- CLC2006.

In-situ data comprises all non-space borne georeferenced data at European and national level used in the project activities.

Satellite images

Satellite image data are the basic source for HRL production. Overview of the satellite imagery used in the project is presented in Table 1.

Table 1. Overview of the satellite data sets

Feature/Satellite	IRS LISS III (coverage-1)	RapidEye (coverage-2)	IRS AWiFS	SPOT-5 pan-sharpened
No. of bands delivered	4	5	4	3
Channels	green, red, NIR, SWIR	blue, green, red, red-edge, NIR	green, red, NIR, SWIR	green, red, NIR,
Ground sampling distance (m)	23.5	6.5	56-70	2.5 or 10
Bit depth	7	up to 12	10	8
Delivered resolution	20 m	20 m	60 m	2.5

Two pan-European satellite image coverages – 1 and 2 form the core image data set. Image data from IRS-P6 satellite are the same used by the service providers to produce the HRL multi-date image coverage of the country's territory is delivered ensuring two or more images at every location. Pan-sharpened SPOT 5 image data with three spectral bands were used instead of colour infrared aerial photos, needed for forest type layer. All images were acquired during the period 2011-2012.

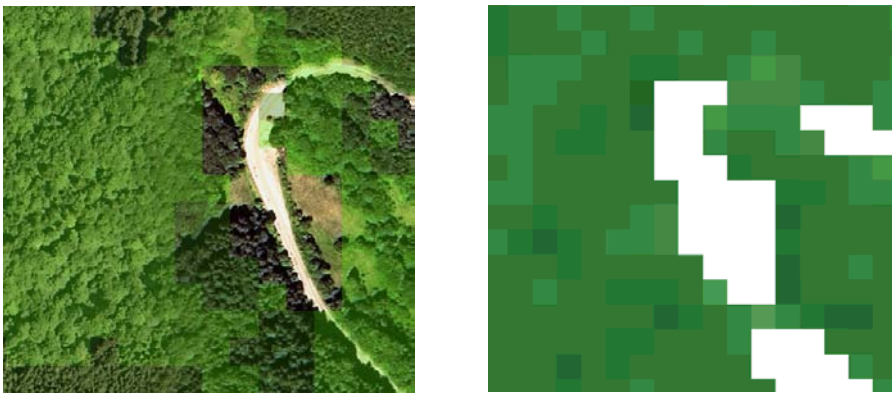
High Resolution Layers

In the following text only most important definitions and outline parameters of forest and water bodies HRLs are provided. More detailed description and technical specifications of all five high resolution layers can be found in [1]. The required classification accuracy for HRL is at least 85%.

HRL data set for the forest includes two products: Tree Cover Density and Forest Type. The latter consists of two raster layers: Dominant Leaf Type and Additional Support layer. Orchards, forest nurseries and transitional woodlands are included in the tree cover density product. On the other hand, non-forest trees are excluded from Forest Type product, following the FAO forest definition. This is achieved through the Support layer which maps trees under agricultural use and in urban context only, while the Dominant Leaf Type layer maps all the trees irrespective of their use.

Figure 1 gives an impression of forest HRLs through an example which represents a situation of a turn of the road in a forest area. On Figure 1 a) a semi-transparent Forest Type raster layer is overlaid on aerial photo with dark-green pixels corresponding to coniferous forest and light-green ones

corresponding to broadleaved forest. On the right-hand image (Figure 1 b), the Tree Cover Density layer of the same area is shown where darker colour means higher density.



a) Forest Type b) Tree Cover Density

Fig. 1. Example of Forest product HRLs

In the case of PWB HRL in order to detect the permanent presence of surface water image data for three reference years were used: 2006, 2009 and 2012. Seasonal AWiFS data were involved for the 2012 reference year to separate temporary water from permanent water bodies.

In-situ data

In-situ data comprise all non-space-borne data with a geographic dimension. They are divided to ancillary and reference data, at European and national level. Collecting national level in-situ data is the responsibility of the countries. Primary use of in-situ data the project is for providing a reliable reference for HRL evaluation. Working on the forest and water bodies HRLs the following in-situ digital data sets were used: forest inventory database, physical blocks of the Land Parcel Identification System, river network, digital orthophoto map, polygons of settlement, topographic maps, National cadastre layer, CLC2006 and RAMSAR sites layer. In addition, the web-based tools Google Earth with StreetView and OpenStreetMap were used when necessary.

Methods

The GIO Land tasks, assigned by project's terms of reference, were further detailed in comprehensive Guidelines for verification [5], which served as a methodological and technical basis in our work. On the other hand, the local technical team had the freedom to make decisions on the specifics of methods' implementation.

The work on forest and water bodies HRLs verification is based on three main methods:

- stratified sampling
- computer-aided visual interpretation of samples
- geospatial data editing.

The purpose of the verification is to identify systematic errors in a HRL that allow further correction. It is carried out through visual inspection of selected samples on in-situ data, checking for commission and omission errors. The verification for all HRLs includes three steps (the first two obligatory): general overview, look-and-feel verification and statistical verification.

General overview of data quality aims to provide a general impression on the HRL and in-situ data and to reveal problematic areas for detailed checking in the next step. During this phase considerations for stratification to be applied in the look-and-feel verification phase are collected.

Both look-and-feel and statistical verification rely on sampling. Samples are evaluated by an expert using relevant in-situ data. Look-and-feel samples should be chosen in such a way which ensures areas of potential classification errors to be checked. Knowledge about the territory and classifier's performance is necessary to be able to define such problematic zones. Due to all the above conditions, look-and-feel sampling could not be random. Sufficient number of samples should be provided both for commissions and omission errors, therefore appropriate stratification is applied. As a result of sampling and interpretation of HRL, evaluation on five grades is provided for each stratum: excellent, good, acceptable, insufficient and very poor.

Statistical verification provides quantitative accuracy results that are comparable between HRLs and countries. The sampling is random. The problem of ensuring reasonably low number of samples for the omission strata is solved through adequate stratification. In the case of commission error the estimation uncertainty, expressed as percent of the commission error, depends on the number of samples, but is independent from the area of the HRL class. Thus, if 250 valid samples are taken in the commission stratum the maximum uncertainty is ±3.16% [5]. The commission stratum is the mapped HRL itself. The commission error is calculated by the ratio (1):

$$(1) \quad E_{comm\ (HRL)} = \frac{Area\ _{error}}{Area\ _{valid}},$$

where
$E_{comm(HRL)}$ – commission error,
$Area_{error}$ - area of erroneously classified samples,
$Area_{valid}$ - area of valid samples.

The omission error is estimated through the commission error of the non-HRL class, formula (2). If stratification is applied for omissions, then instead of "non-HRL" we should consider the term and entity "omission stratum".

$$(2) \quad E_{omis\ (HRL)} = E_{comm\ (non-HRL)} = \frac{Area\ _{total} - Area\ _{HRL}}{Area\ _{HRL}},$$

where
$E_{omis(HRL)}$ – omission error
$E_{comm(non-HRL)}$ – commission error of the non-HRL class
$Area_{total}$ – the whole country area classified.
$Area_{HRL}$ – area of the mapped HRL.

Results of the verification are presented in the form of Verification Report and data set to be used in the enhancement task.

Results and discussion

Verification of Tree Cover Density and Forest Type

For the look-and-feel verification of forest layers maximum 15 thematic strata and 5–10 samples in each are required, which means 75-150 samples in total. The list of strata, number of samples selected and evaluation grades per stratum for the Tree Cover Density layer are presented in Table 2.

Table 2. Sampling results of look-and-feel verification for Tree Cover Density layer

№	Name of the stratum	No. of samples	Evaluation grades
	COMMISSION		
1.	Transitional woodland-shrub	10	*acceptable*
2.	Land principally occupied by agriculture, with significant areas of natural vegetation	10	*acceptable*
3.	Moors and heathland	9	*insufficient*
4.	Wetlands	10	*insufficient*
5.	Grasslands	9	*acceptable*
	OMISSION		
6.	Urban vegetation (trees in parks, cemeteries...)	10	*good*
7.	Trees in sport and recreation areas	9	*acceptable*
8.	Orchards	10	*insufficient*
9.	Lowland forests, broadleaved	10	*good*
10.	Lowland forests, coniferous	7	*good*
11.	Mountain forests, broadleaved	10	*good*
12.	Mountain forests, coniferous	10	*good*
13.	Forest along rivers & lakes	10	*good*
14.	Coastal forests	5	*good*
15.	Agricultural areas with scattered small forest patches (if ≥ 0.5 ha)	8	*good*
Overall evaluation			*good*

As a result of the analysis of errors found during the look-and-feel verification the following comments are formulated the Tree Cover Density layer:

- Commission errors occur in areas of grassland with shrubs in CLC 324 polygons, as well as in CLC 243 polygons with forested patches.
- Large areas with commission errors are present in alpine parts of Rila and Pirin mountains, mostly occupied by dwarf pine. Some commission error cases occur in treeless grassed zones, as well.
- Within the Wetland stratum marshlands or grassed parts of wetland area (mostly near the Danube river and seashore) are incorrectly mapped as forest.
- Commission errors over grasslands occur predominantly within or near forest areas.
- Many of CLC2006 222 based orchards, checked on ortho-photo, look abandoned and/or forested. Polygons with LPIS land use codes 020 and 022 were also involved to evaluate the Tree Cover Map over orchard areas. A lot of orchards are not mapped, probably, due to insufficient tree cover density of orchard fields. Palmette orchards are often missed, as well.
- In forest areas commission errors occur, mapping included grassland areas as forest. Classifier's performance degrades in case of appearance of non-forest features in large forest areas, e.g. road turns, rocks, leading to minor local omission errors.

For the Dominant Leaf Type product 12 thematic strata and 85 samples respectively were selected. Overall evaluation grades for both layers are "good". There are some differences in the strata list for commission errors due to the specifics of expected forest – non-forest errors (Table 3).

Table 3. Dominant Leaf Type strata, samples and grades

№	Name of the stratum	No. of samples	Evaluation grades
	commission		
1	Major cities	10	*good*
2	Sport and recreation areas in urban context	7	*insufficient*
3	Orchards	10	*insufficient*
4	Moors and heathland	8	*insufficient*
5	Wetland	10	*insufficient*
	omission		
6	Lowland forests, broadleaved	7	*good*
7	Lowland forests, coniferous	7	*good*
8	Mountain forests, broadleaved	7	*good*
9	Mountain forests, coniferous	5	*good*
10	Forest along rivers & lakes	8	*good*
11	Coastal forests	6	*good*
12	Agricultural areas with scattered small forest patches (if ≥ 0.5 ha)	8	*good*
Overall evaluation			*good*

Major problems, which occurred during the look-and-feel verification of this product, are revealed by the following comments:

- Most of problems were found in commission strata.
- Many of CLC2006 class 222 based orchards are abandoned and/or forested and does not exist in LPIS physical blocks. In such cases Support layer commissions occur.
- Only 30% of LPIS orchards overlap with CLC 222 polygons.
- Large areas in alpine parts of Rila and Pirin mountains, covered by dwarf pine, are mapped as forest causing commission errors. Some commission error cases also occur in treeless grassed zones.
- Treeless grassed and/or wet parts of wetland areas are often mapped as forest.
- Dominant leaf type is generally correctly mapped. Young coniferous forests were found mapped as broadleaved. Open area patches surrounded by coniferous forest could be mapped as broadleaved.
- The requirement for minimum width of 20 m for forest fields to be mapped is generally not covered by the classifier. Windbreaks, more than 30 m wide, are simply missed or later filtered out because of 4 connectivity requirement.

Statistical verification of PWB layer

Statistical evaluation procedure is non-obligatory one and was applied to the PWB layer only. Two strata were defined – one for commission and one for omission errors. The commission stratum,

as it was mentioned before, is the mapped HRL itself. The omission stratum was defined by merging polygons related to water classes from LPIS and agricultural cadastre data sets. As recommended by the guidelines, 280 random samples were taken in both commission and omission stratum (Table 4).

Table 4. PWB sampling results for commission and omission error

Measured quantity	Value
COMMISSION	
All valid samples	280
Samples showing error	19
User's accuracy	93.21 %
Uncertainty	1.50 %
Commission error	6.79 %
OMISSION	
All valid samples	280
Samples showing error	28
Error ratio	10 %
Accuracy	90 %
Producer's accuracy	94.02 %
Uncertainty	1.07 %
Omission error	5.98 %

Sampling results show low uncertainty of accuracy values, i.e., low standard deviation of error values, which fact is a guarantee for a reliability of the estimations.

Conclusions

Overall results of this study show the good general ability of automated analysis of multi-temporal satellite imagery to extract forest and permanent water bodies land cover data. In general, the tree cover is overestimated by the classifier causing a prevalence of commission errors. The attempt to extract orchards information using the above approach comes across with serious problems. Both the qualitative and the statistical evaluation show that the Forest and PWB HRL products meet the accuracy requirements.

References:

1. GIO land High Resolution Layers (HRLs) – summary of product specifications - Prepared by: T. Langanke, EEA, 2013, http://land.copernicus.eu/user-corner/technical-library/gio-land-high-resolution-layers-hrls-2013-summary-of-product-specifications/at_download/file .
2. S t e e n m a n s, C., A. S o u s a. Guidelines for verification of high resolution soil sealing layer - Qualitative assessment, EEA, 2007.
3. S t o i m e n o v, A., R. V a t s e v a, V. D i m i t r o v. Soil Sealing Part of CORINE Land Cover 2006 Bulgaria Project, International Conference Fundamental Space Research, Sunny Beach, Bulgaria, September 21-28, 2008, Proceedings, pp. 110-113, ISBN 978-954-322-316-9, 2008b.Maucha, 2011.
4. Guidelines for verification of high-resolution layers produced under GMES / Copernicus Initial Operations (GIO) land monitoring 2011 – 2013 - Prepared by: G. Büttner, EEA, 2013.
5. M a u c h a, G., G. B ü t t n e r, B. K o s z t r a. European Validation of GMES FTS Soil Sealing Enhancement Data, In Remote Sensing and Geoinformation not only for Scientific Cooperation, pp. 223-238, EARSeL, 2011.