

Final

Guidelines for verification of high resolution soil sealing layer

- Qualitative assessment -

Prepared by:

Chris Steenmans

Ana Sousa

03.12.2007

Version 1.0

European Environment Agency



Introduction

This document provides the guidelines for the verification of the high resolution soil sealing layer, based on a qualitative assessment of the mapped area. As agreed at the Eionet workshop on quality control and validation of land cover data (Copenhagen, 12-13 November 2007), these guidelines should help National Reference Centres on Land Cover (NRCs) to support EEA in doing the verification of the soil sealing layer that is being produced in the frame of GMES land monitoring fast track service precursor.

The soil sealing data is produced by a consortium of European service providers under contract with EEA and is based on the classification of the IMAGE2006 satellite data. The overall objective is the production of a seamless European high resolution core land cover dataset of built-up areas, including degree of soil sealing, for the reference year 2006. Built-up areas are characterized by the substitution of the original (semi)-natural cover or water surface with an artificial, often impervious, cover. This artificial cover is usually characterized by long cover duration (FAO Land Cover Classification System, 2005). Impervious surfaces of built-up areas account for 80 to 100% of the total cover. A per-pixel estimate of imperviousness (continuous variable from 0 to 100 percent) will be provided as index for degree of soil sealing for the whole geographic coverage. The data will be produced in full spatial resolution, i.e. 20 m by 20 m, which provides the best possible core data for any further analysis. The classification accuracy per hectare (based on a 100 m x 100 m grid) of built-up and non built-up areas should be at least 85%, for the European product.

The verification task will run from end November 2007 (when the first country deliveries are expected) until October 2008 (deadline for the last country to be delivered by the contractor) and should support EEA in accepting or rejecting the delivery of the country datasets produced by the service provider.

This qualitative assessment supported by NRCs is part of the grant agreement between EEA and participating countries in the GMES project land monitoring fast track service precursor/CLC2006.

NRCs are invited to carry out this assessment and to give feedback to the Agency within 4 weeks after reception of the data. If it is not possible to perform the verification task within these 4 weeks, it is expected that it will be completed before the end of the grant agreement, according to Article I.2 (Duration).

If countries would like to do additional checks or a quantitative assessment based on statistical validation, they are welcome to do so and to share the results with EEA.

Guidelines are provided for the preparatory work, the inventory of reference data that will be used, the description of the geometric and thematic quality and the overall qualitative assessment. NRCs should use this document template to report on the verification of the data, by filling in the grey boxes: insert free text in the “Text Form Fields” (); tick the “Check Box Form Field” (☐); and select from “Drop Down Form Field” (Please, select). Feel free to add additional text or illustrations (e.g. examples from screenshots).

A quantitative assessment or final validation of the European dataset will be carried out by EEA in collaboration with Eionet during late 2008-2009 (project details to be confirmed during the second half of 2008). This European validation will be based as much as possible on the results of national validations. NRCs are invited to inform EEA about planned activities (if any) at national level. Preliminary recommendations for such a statistical validation (quantitative assessment) are attached in annex for information.

Note: After filling in the template save it as a word document: filename: countryISOcode.doc (e.g. AT.doc).

1. Preparatory work

1. Upload the data that will be made available by EEA via ftp server or sent by mail. Please inform EEA on reception of the data;
2. Check for available reference data that will be used during the verification;
3. List the experts/expertise that are involved in the verification task:

Expert name	Field of expertise	Institution
Claire FitzGerald	Remote sensing	ERA Maptec

The average time needed for this verification is estimated at one person/day per 10.000 km². Please note that this time can vary depending on the experience of the interpreter, the availability of the reference data and the complexity of the landscape. The table below gives an indicative estimate for the EEA member countries.

Country	Area (km ²)	Person days	Country	Area (km ²)	Person days
Austria + Liechtenstein	83.855	9	Lithuania	65.200	7
Belgium	30.520	3	Luxembourg	2.586	<1
Bulgaria	110.994	11	Malta	316	<1
Cyprus	9.251	1	Netherlands	41.526	4
Czech Republic	78.864	8	Norway	323.878	33
Denmark	43.075	4	Poland	312.683	31
Estonia	45.200	5	Portugal	88.935	9
Finland	338.145	34	Romania	237.500	24
France	543.965	55	Slovakia	20.251	5
Germany	357.028	36	Slovenia	49.035	2
Greece	131.957	13	Spain	504.782	51
Hungary	93.030	9	Sweden	449.964	39
Iceland	102.820	10	Switzerland	41.293	4
Ireland	70.282	7	Turkey	789.452	79
Italy	301.245	30	United Kingdom	244.082	25
Latvia	63.700	6			

2. Reference data

Please list the reference data that is used for this verification:

1. Topographic maps

☐ No

☒ Yes

Year:

Area: Full country

If only a subset, then please specify the area(s):

2. Aerial orthophotos

☐ No

☒ Yes

Year: 2004-2006

Area: Full country

If only a subset, then please specify the area(s):

3. Very High Resolution satellite data

☐ No

☒ Yes

Year: 2005-2006

Area: Full country

If only a subset, then please specify the area(s):

4. CLC2000

☐ No

☒ Yes

5. Other

Name:

Year:

Area: Please, select:

If only a subset, then please specify the area(s):

Name:

Year:

Area: Please, select:

If only a subset, then please specify the area(s):

Qualitative assessment HR soil sealing layer

Name:

Year:

Area: Please, select:

If only a subset, then please specify the area(s):

Name:

Year:

Area: Please, select:

If only a subset, then please specify the area(s):

Comments concerning the reference data used (if any):

Image dates are variable, but mostly 2005 and 2006, with a lesser amount of 2004. All are 1m resolution orthos.

B. Geometric quality

Please provide your qualitative assessment of the geometric quality of the data. The objective of this task is to perform a visual analysis of the soil sealing dataset concerning its co-registration when put in overlay with other reference datasets.

1. Check geometric accuracy:

Is there a visible shift? ☐ Yes ☒ No

If yes:

a. Is there a systematic shift? ☐ Yes ☐ No

b. Is there a local shift? ☐ Yes ☐ No

Where?

Please indicate the region, place name, coordinates or other description of location:

2. Is the used projection correct? ☒ Yes ☐ No

3. Comments concerning geometric issues (if any), or in case the geometric quality could not be checked, please provide a short explanation:

C. Thematic quality

Please provide your qualitative assessment of the thematic quality of the data. The objective of this task is to perform a visual comparison between available reference data and the soil sealing dataset. You are requested to verify for a number of land cover classes (similar to the CLC classes at levels 2 or 3) to check if any errors in the data can be identified. Please note that many land cover classes can include sealed surfaces, especially for features <25 ha.

For this part of the verification, it is recommended to use a binary mask (built-up/non-built-up area) that can be used in overlay with the reference data:

1. Apply a lookup table to map all pixels > 80% degree of soil sealing as built-up area;
2. Perform the checks on pixels > 80% degree of soil sealing by screening for each of the land cover classes if built-up or non built-up areas are correctly mapped. Feel free to add screenshots with examples to illustrate the quality judgement.

For your qualitative assessment, following examples of check boxes can be ticked:

- ☐ “excellent” meaning that you expect that the accuracy of the built-up data is reaching almost 100%; no errors could be found in the areas that were verified.
- ☐ “good” meaning that you are confident that the classification results are at least 85 % correct; only sporadic errors were encountered in the areas that were verified.
- ☐ “acceptable” meaning that you estimate that in most of the verified areas the classification results will probably reach an accuracy of 85 %; some minor errors could be detected in the areas that were verified.
- ☐ “insufficient” meaning that you do not expect that the classification results will reach the minimum of 85 % accuracy; you encountered several errors in different regions.
- ☐ “very poor” meaning that you are confident that the classification results are bad with regard to presence of built-up area; most of the areas verified are wrongly mapped.

Urban fabric:

- a. Did you check if built-up/non built-up areas are correctly mapped within urban fabric (e.g. houses, buildings, streets, etc.)?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality of the mapped built-up area within the urban fabric?

☐ very poor ☐ insufficient ☐ acceptable ☐ good ☒ excellent

- c. Short description of errors found (if any): The accuracy of the built-up data is excellent with only one point not reaching almost 100% (this is due to confusion between the spectral signatures of bare soil and built-up areas).

Industrial or commercial units:

- a. Did you check if built-up/non built-up areas are correctly mapped within industrial or commercial units (e.g. parking lots, buildings, etc.)?
- ☒ Yes ☐ No ☐ Not possible
- b. How would you assess the quality?
- ☐ very poor ☐ insufficient ☐ acceptable ☐ good ☒ excellent
- c. Short description of errors found (if any): Bare soil misclassified as built-up.

Road and rail networks and associated land:

- a. Did you check if built-up/non built-up areas within road and rail networks and associated land are correctly mapped (e.g. railway stations, highways >20 m width, etc.)?
- ☒ Yes ☐ No ☐ Not possible
- b. How would you assess the quality?
- ☐ very poor ☐ insufficient ☐ acceptable ☒ good ☐ excellent
- c. Short description of errors found (if any): Where bare soil is misclassified as built-up.

Port areas:

- a. Did you check if built-up/non built-up areas in port areas are correctly mapped (e.g. installations, dykes, etc.)?
- ☒ Yes ☐ No ☐ Not possible
- b. How would you assess the quality?
- ☐ very poor ☐ insufficient ☐ acceptable ☒ good ☐ excellent
- c. Short description of errors found (if any): Bare soil and water misclassified as built-up.

Airports:

- a. Did you check if built-up/non built-up areas in airports are correctly mapped (e.g. runways, buildings, etc.)?
- ☒ Yes ☐ No ☐ Not possible
- b. How would you assess the quality?

☐ very poor ☐ insufficient ☐ acceptable ☐ good ☒ excellent

c. Short description of errors found (if any):

Mine, dump and construction sites:

a. Did you check if built-up/non built-up areas in mine, dump and construction sites are correctly mapped (e.g. buildings, infrastructure, etc)?

☒ Yes ☐ No ☐ Not possible

b. How would you assess the quality?

☐ very poor ☒ insufficient ☐ acceptable ☐ good ☐ excellent

c. Short description of errors found (if any): Areas of bare soil and water misclassified as built-up.

Arable land:

a. Did you check if built-up/non built-up areas in arable land are correctly mapped (e.g. bare soil, large farm houses, roads >20m width, etc)?

☒ Yes ☐ No ☐ Not possible

b. How would you assess the quality?

☐ very poor ☐ insufficient ☒ acceptable ☐ good ☐ excellent

c. Short description of errors found (if any): Similar spectral signals from built areas and land cover at this time of year has lead to confusion between the classes. Also some problems at the edges of fields.

Heterogeneous agricultural areas:

a. Did you check if built-up/non built-up areas in heterogeneous agricultural areas are correctly mapped (e.g. buildings, roads >20m, etc)?

☒ Yes ☐ No ☐ Not possible

b. How would you assess the quality?

☐ very poor ☐ insufficient ☒ acceptable ☐ good ☐ excellent

c. Short description of errors found (if any): Spectral confusion between crop type and artificial surfaces.

Forest:

a. Did you check built-up/non built-up areas in forests are correctly mapped (e.g. clear-cuts, roads, etc.)?

☒ Yes ☐ No ☐ Not possible

b. How would you assess the quality?

☒ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Bare soil and cleared forest misclassified as built-up, transitional woodland classified as built area, residential gardens identified as built-up areas.

Scrub and/or herbaceous vegetation associations:

- a. Did you check if built-up/non built-up areas in scrub and/or herbaceous vegetation areas are correctly mapped (e.g. dry vegetation, rock outcrop, etc.)?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☐ very poor ☐ insufficient ☒ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Bare soil and crop rotation leading to confusion among spectral signatures.

Beaches, dunes and sands:

- a. Did you check if built-up/non built-up areas in beaches, dunes and sand areas are correctly mapped?

☐ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☒ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Sand, sea and rock wrongly identified as built-up

Bare rocks:

- a. Did you check if built-up/non built-up areas in bare rock areas are correctly mapped?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☒ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Rock misclassified as built-up.

Sparsely vegetated areas:

- a. Did you check if built-up/non built-up areas in sparsely vegetated areas are correctly mapped?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☐ very poor ☒ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): The nature of the class leads to confusion between bare soil and artificial surfaces.

Glaciers and perpetual snow:

- a. Did you check if built-up/non built-up areas in glaciers and perpetual snow areas are correctly mapped?

☐ Yes ☐ No ☒ Not possible

- b. How would you assess the quality?

☐ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any):

Inland wetlands:

- a. Did you check if built-up/non built-up areas in inland wetlands are correctly mapped ?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☐ very poor ☐ insufficient ☒ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Bare soil is misclassified as built-up.

Salines:

- a. Did you check if built-up/non built-up areas in salines are correctly mapped?

☐ Yes ☐ No ☒ Not possible

- b. How would you assess the quality?

☐ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any):

Intertidal flats:

- a. Did you check if built-up/non built-up areas in intertidal flats are correctly mapped?

☒ Yes ☐ No ☐ Not possible

- b. How would you assess the quality?

☒ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent

- c. Short description of errors found (if any): Shadow, sand and water all lead to confusion in the classification.

Coastal lagoons:

- a. Did you check if built-up/non built-up areas in coastal lagoons are correctly mapped?
☒ Yes ☐ No ☐ Not possible
- b. How would you assess the quality?
☒ very poor ☐ insufficient ☐ acceptable ☐ good ☐ excellent
- c. Short description of errors found (if any): Sand and water misclassified as built-up areas.
3. Comments concerning thematic content check (if any). Please indicate which part of the data was verified (full coverage or partial coverage, etc.):

Thematic Content Check

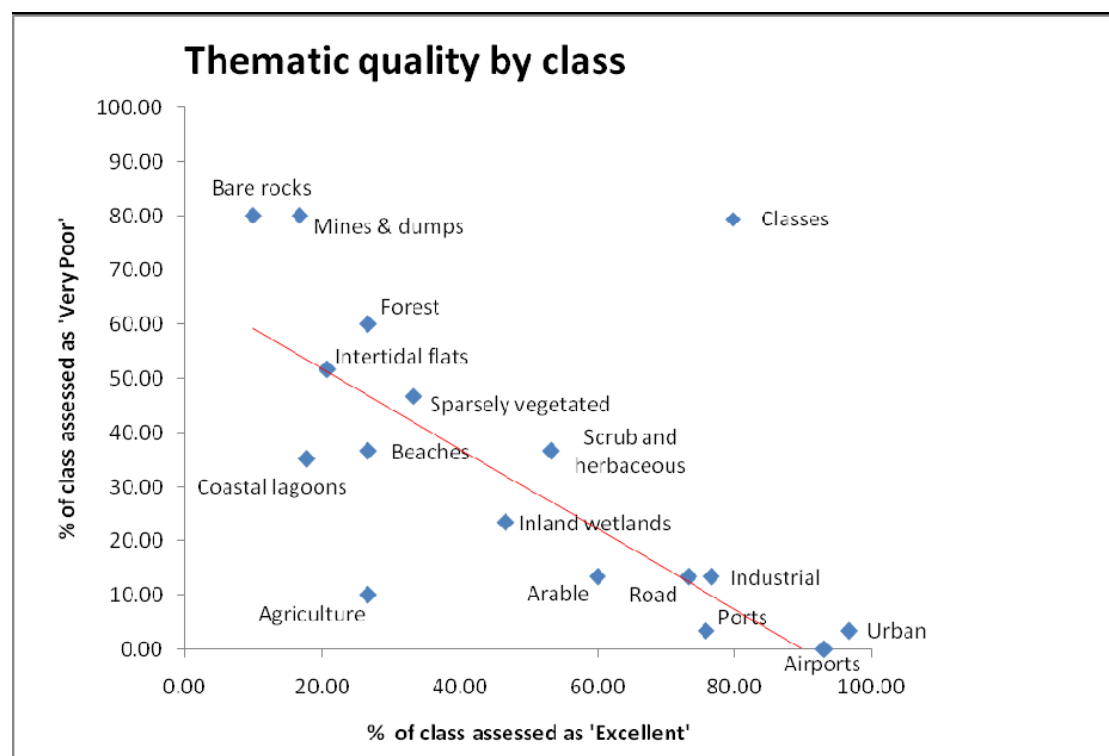
Please note that this is a qualitative assessment. The information provided below is by way of explaining the assessment. This is not a quantitative assessment and does not follow the guidelines for such an assessment.

In most cases 30 locations were sampled for each class.

Class	No. Points	Percentage (%)					Quality
		Excellent	Good	Acceptable	Insufficient	Very Poor	
Urban fabric	30	97	0	0	0	3	Excellent
Airports	29	93	7	0	0	0	Excellent
Industrial/Commercial	30	77	0	7	3	13	Good
Ports	29	76	7	3	10	3	Good
Road & Rail	30	73	0	7	7	13	Good
Arable	30	60	23	3	0	13	Acceptable
Scrub & Herbaceous	30	53	7	3	0	37	Acceptable
Inland Wetlands	30	47	13	7	10	23	Acceptable
Agriculture	30	27	30	17	17	10	Acceptable
Sparsely Vegetated	30	33	10	10	0	47	Insufficient
Forest	30	27	10	3	0	60	Very poor
Beaches	30	27	13	13	10	37	Very poor
Intertidal Flats	29	21	10	3	14	52	Very poor
Coastal Lagoons	17	18	18	24	6	35	Very poor
Mines & Dumps	30	17	0	0	3	80	Very poor
Bare Rocks	30	10	7	3	0	80	Very poor

Table 1: Points were randomly selected for each class and assessed using the categories given below.

- ☐ “excellent” meaning that you expect that the accuracy of the built-up data is reaching almost 100%; no errors could be found in the areas that were verified.
- ☐ “good” meaning that you are confident that the classification results are at least 85 % correct; only sporadic errors were encountered in the areas that were verified.
- ☐ “acceptable” meaning that you estimate that in most of the verified areas the classification results will probably reach an accuracy of 85 %; some minor errors could be detected in the areas that were verified.
- ☐ “insufficient” meaning that you do not expect that the classification results will reach the minimum of 85 % accuracy; you encountered several errors in different regions.
- ☐ “very poor” meaning that you are confident that the classification results are bad with regard to presence of built-up area; most of the areas verified are wrongly mapped.



Classification errors are highest when looking at land cover composed of exposed rock and bare soil (see chart above), and lowest amongst the classes composed of artificial surfaces, with the classes containing mixed parcels of permanent crops and scattered housing having a large range of accuracy.

D. Overall qualitative assessment of the dataset

The overall qualitative assessment is meant to support EEA in our contractual procedures with the service provider regarding the acceptance of the dataset. While the previous thematic quality assessment was looking at class by class, this section should provide your assessment of the quality for the whole territory.

How would you assess the overall quality of the mapped built-up/non built-up areas for the dataset provided?

☐ very poor ☐ insufficient ☒ acceptable ☐ good ☐ excellent

Please provide your final comments and additional remarks concerning overall qualitative assessment (e.g. difference in quality between regions e.g. mountains, agglomerations, coastal zones, etc), if any:

Areas of bare soil, exposed rock, water bodies and sand have all been misclassified as built areas, with coastal areas being the worst affected by this. Urban areas are correctly identified but anomalies in the classification occur at the edges of these areas.

We feel that the soil sealing layer will be useable in urban areas only in Ireland.

We suggest that the data set could be improved by the use of ancillary data, such as the high water mark (to remove coastal errors); quarries database and forestry database. All these ancillary data are available in digital format in Ireland

E. Quantitative validation

Are you planning to carry out a statistical validation (quantitative assessment) of the national dataset?

☐ Yes ☒ No

If yes, it would be helpful to provide us information about the timing, methodological approach or any other additional information which might be available:

Although we do not at present intend to undertake a formal quantitative assessment, two agencies (the Urban Institute and the national agricultural research organisation Teagasc) have offered to carry out a more targeted qualitative verification.

Are you willing to contribute to the final validation of the European dataset (actions scheduled from the second half of 2008 onwards)?

☒ Yes **in principle!** ☐ No

This is not budgeted for in 2008, but could be supported early in 2009 if sufficient information were available in advance to enable a budget to be provided.

Filled in by George McHugh

Telephone number: +353 539160612

Email address: g.mchugh@epa.ie

Date: 23 May 2008

Thank you!