

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
Michael Stjernholm	LC	NERI/AU

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: 2002-2006 Area: DK

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

2. Aerial orthophotos

No Yes Year: 2005?? Google Earth Area: DK

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

3. Very High Resolution satellite data

No Yes Year: Area: Please, select:

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):

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

- a. Short description of errors found (if any): The real dense urban fabric is in general found, however unexplainable variations occur in the urban fabric, where even large building may not be identified at all (soil sealing < 30 % , see illustrations no XXXXX)
- b. Open urban fabric (low density sub-urban) is generally under-mapped – the class 1- 30 % rarely used

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): Also non-sealed parking areas are mapped as soil-sealed.
- d. Like the urban fabric do even major buildings disappear

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
- e. Short description of errors found (if any): The wide infrastructure line are generally found. The degree of identification of roads narrower than 20 meters is however very heterogeneous, which make the interpretation/exploitation very dubious.

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
- f. Short description of errors found (if any): Leisure port areas are exaggerated – leisure piers with boats show up with 100% soil sealing, see fig 10 – Roskilde. In other areas- see Kalundborg SW are even a wide pier missing. The land part is mapped as with the industrial fabric – see above.

Airports:

- c. Did you check if built-up/non built-up areas in airports are correctly mapped (e.g. runways, buildings, etc)?
 Yes No Not possible
- d. How would you assess the quality?
 very poor insufficient acceptable good excellent
- g. 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
- h. Short description of errors found (if any): Almost all gravel pits are mapped as 100 % soil sealing – In general are all non-vegetated areas mapped as if sealed.

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
- i. Short description of errors found (if any): The evaluation of soil sealing in this LC type is tricky. The non-built-up areas of arable land are correctly identified as non-built-up. However the farmhouses and associated areas are mapped very heterogeneously and often to the negative side – buildings and associated land are not identified as partly built-up.

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

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

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
- l. Short description of errors found (if any): The areas are generally mapped correctly. Error may occur where summer hut are placed on heathland – see Fig Soendervig

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
- m. Short description of errors found (if any): Natural sand areas are generally mapped correct. Main deviations are gravel / sand pits – see above.

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

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

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

c. How would you assess the quality?

very poor insufficient acceptable good excellent

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

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

p. 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

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

Salines:

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

Yes No Not possible

d. How would you assess the quality?

very poor insufficient acceptable good excellent

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

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
 - t. Short description of errors found (if any):
3. Comments concerning thematic content check (if any). Please indicate which part of the data was verified (full coverage or partial coverage, etc.):

The content check has been performed on the full Danish coverage as also exemplified by the illustrations found below

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:

The major problem in providing EEA with an overall qualitative assessment of the data is the lack of uses cases to be verified. It is not clear at all for which analysis the data could be applied:

- a) Urban areas – only high density urban fabric can be seen – residential areas are heterogeneously mapped and occupied area can't be assessed.
- b) Land occupied by infrastructures can't be assessed as many roads are not captured.
- c) The mapping of gravel pits (100 % error) disturbs the interpretation – though it may not offset the overall target of 85 % error.

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:

a) We do unfortunately for the time being not have access to Aerial photos for this exercise – conversion to Google Earth – seem for the time being to be cumbersome – guidance most welcome.

b) We have made two simple statistical (histogram) assessments:

- assessing the overall distribution of soil sealing values (see figure 11)
The assessment is that values are missing values from 50% and down – not least below the threshold of 30 % soil sealing.
- selecting building with an area above > 1000 sq meter

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

Yes No

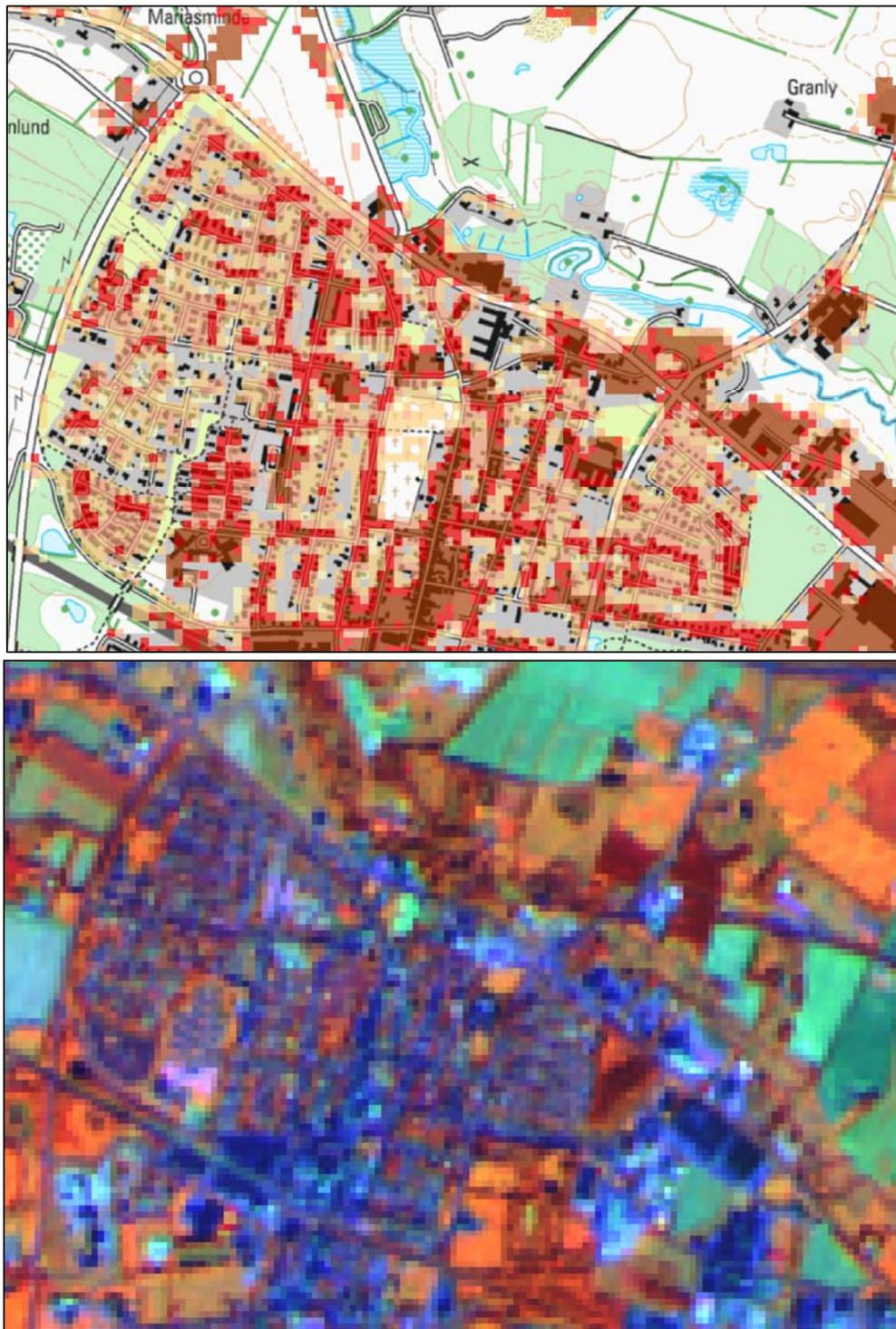
Filled in by Michael Stjernholm

Telephone number: +45 8920 1439

Email address: msh@dmu.dk

Date: Jan 7th 2008

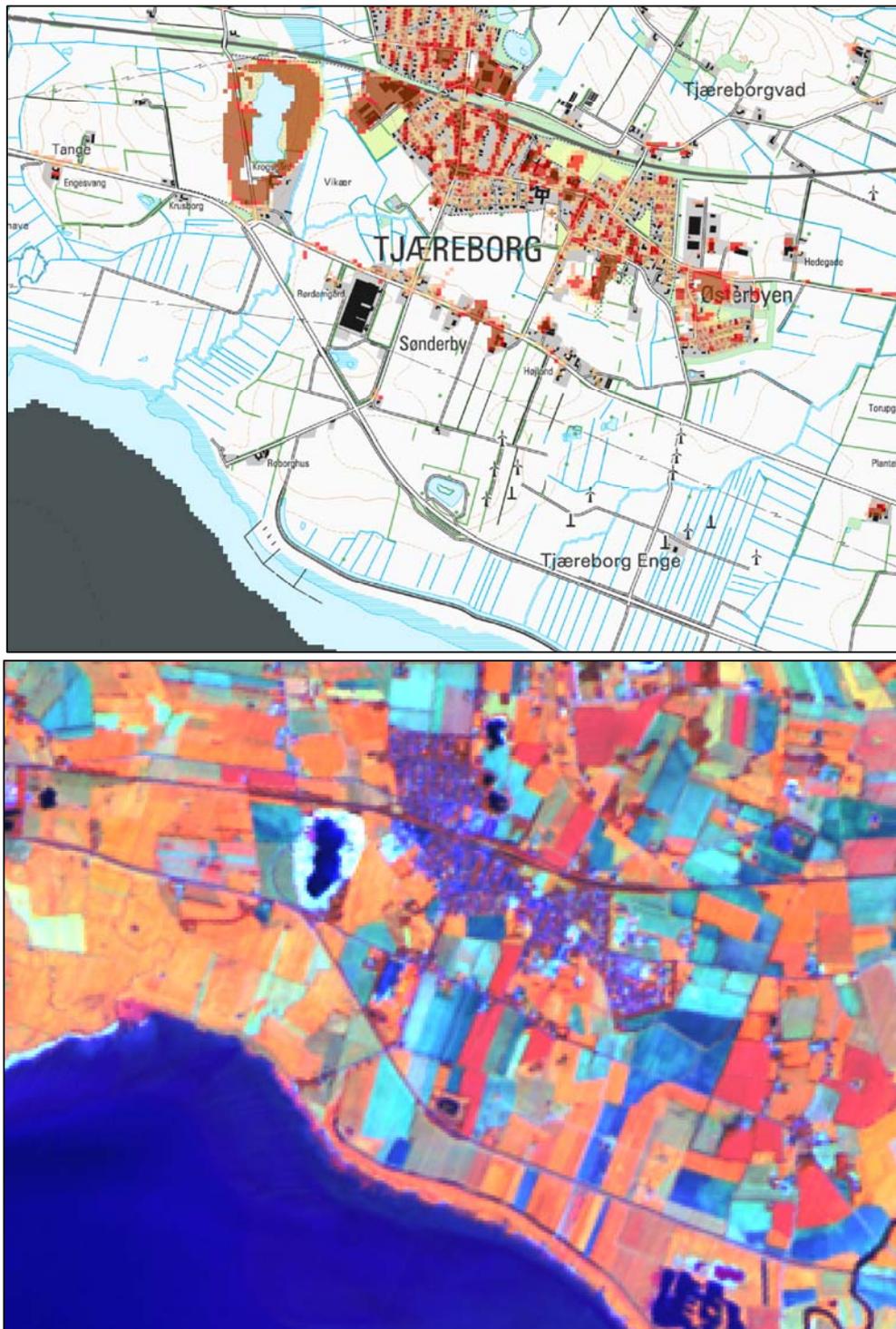
Thank you!



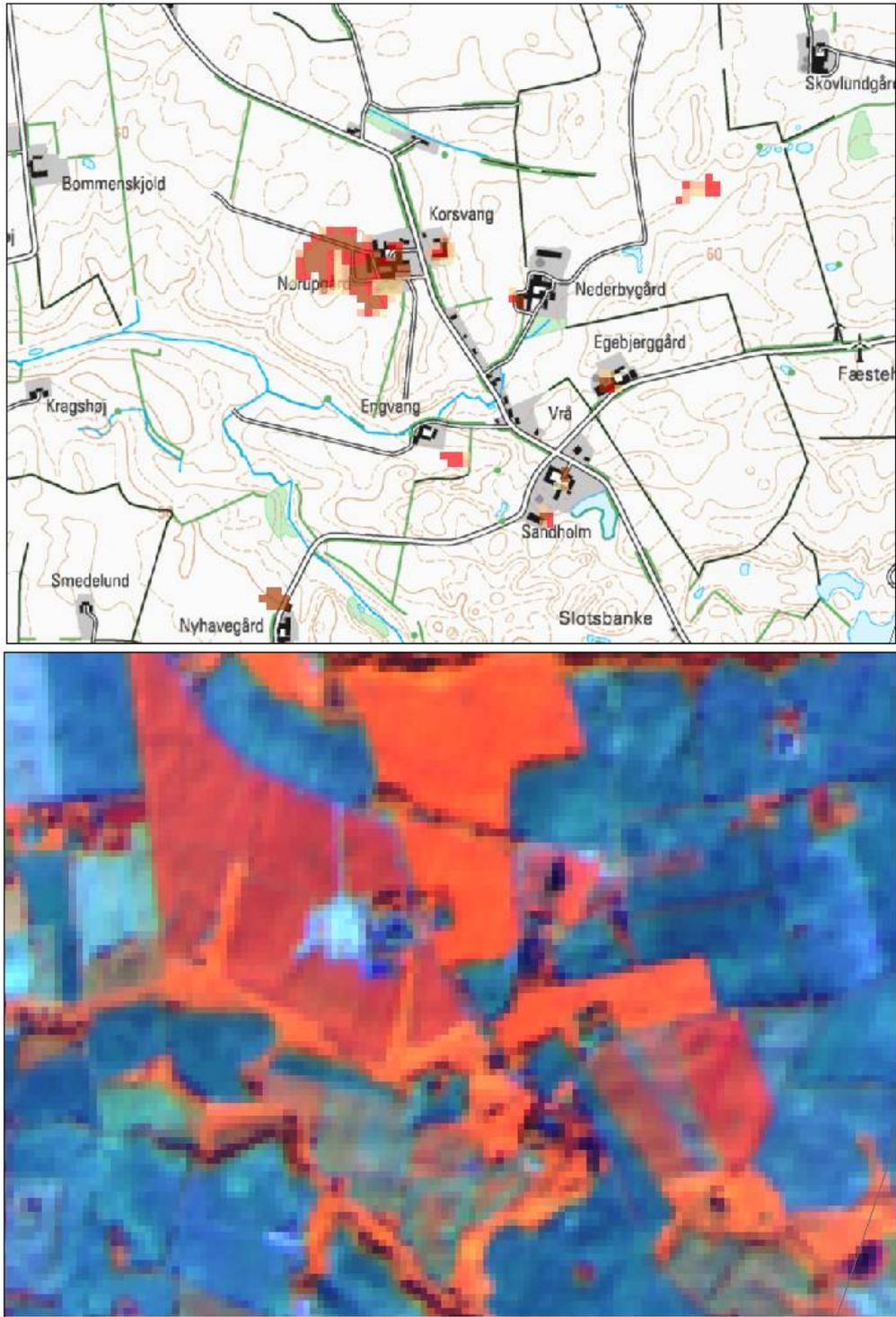
Figur 1a Vejen – Observe that big building in center is completely missing in soil sealing layer



Figur 1b Vejle – seen from Google Earth



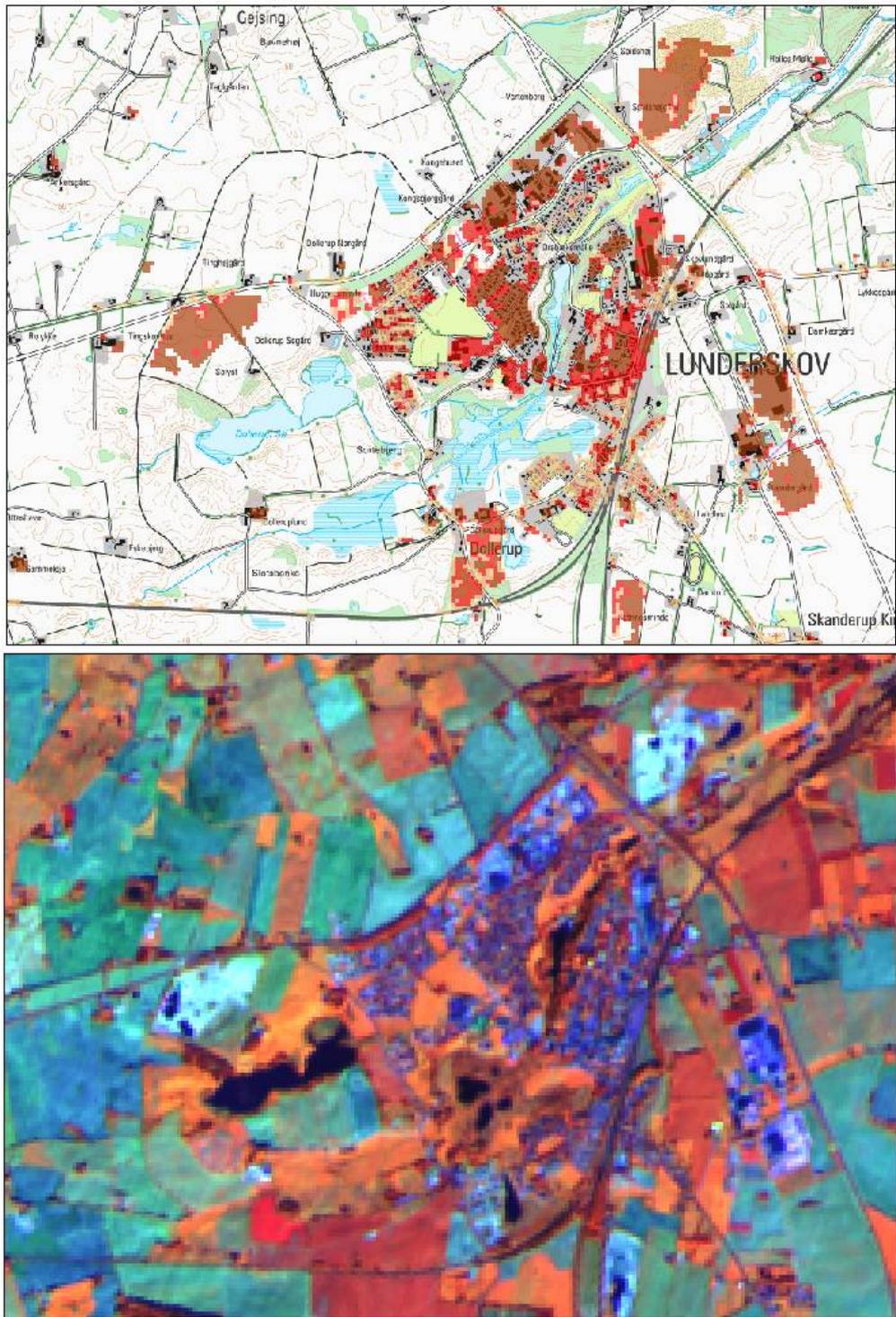
Figur 2 Tjæreborg - Observe gravel pit in upper left and major buildings missing



Figur 3a Farm area – quite some heterogeneity in what is recorded and omitted



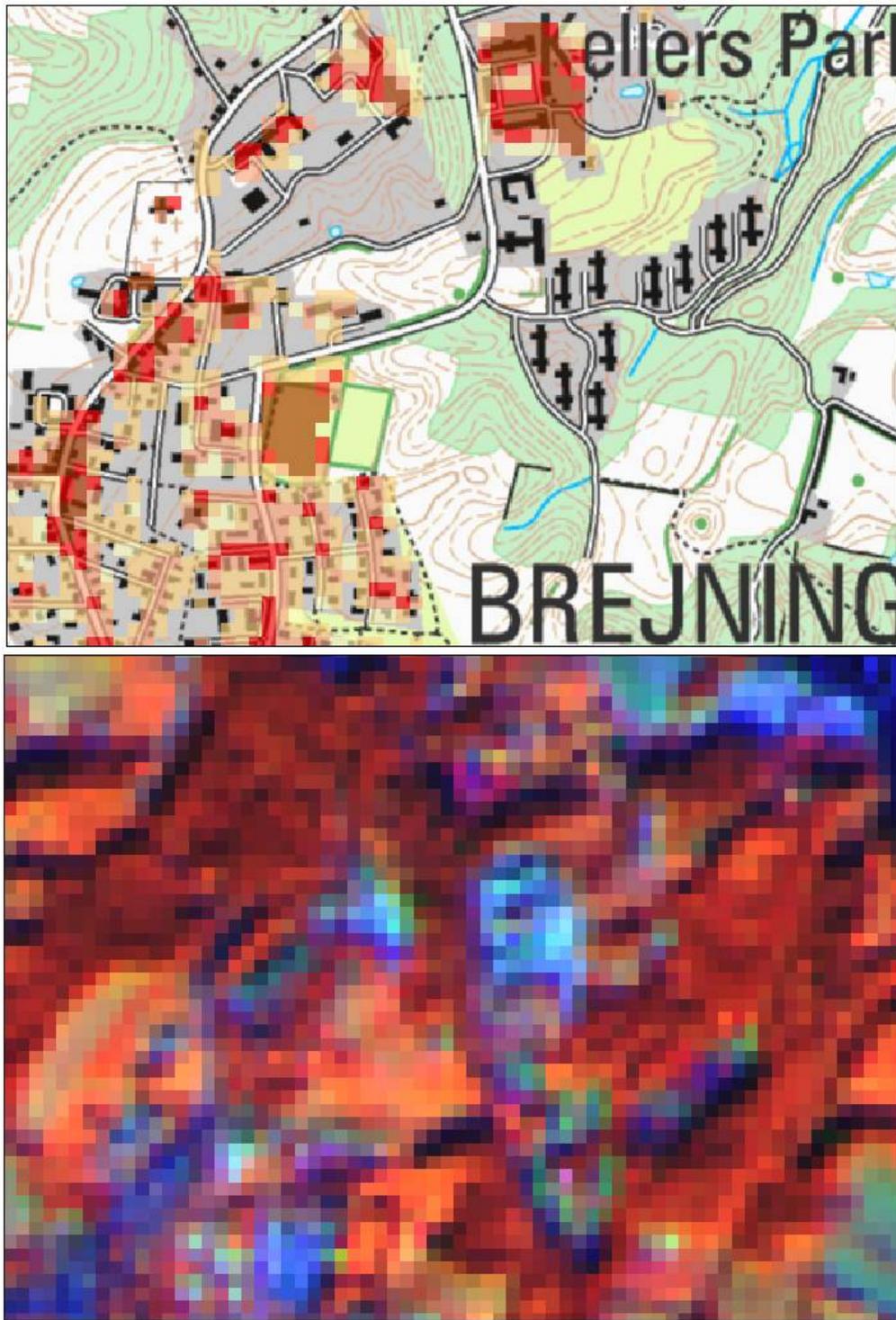
Figur 3b Farm area – From Google Earth



Figur 4a Lunderskov - Observe gravel pits upper left and upper right and new residential area in construction to the south



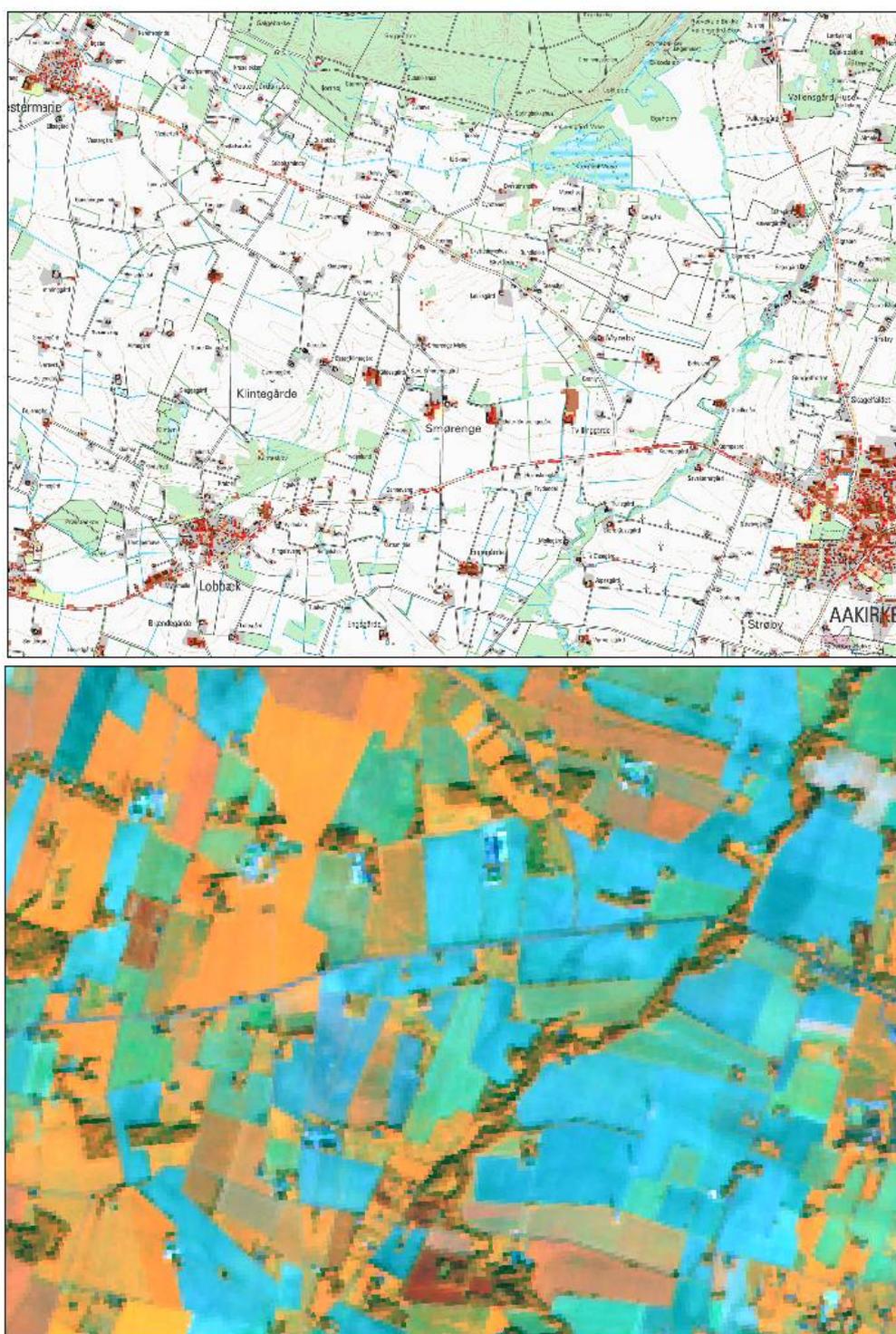
Figur 4b Lunderskov from Google Earth



Figur 5a Brejning. Observe buildings in center/right are large and intersected with major parking areas.



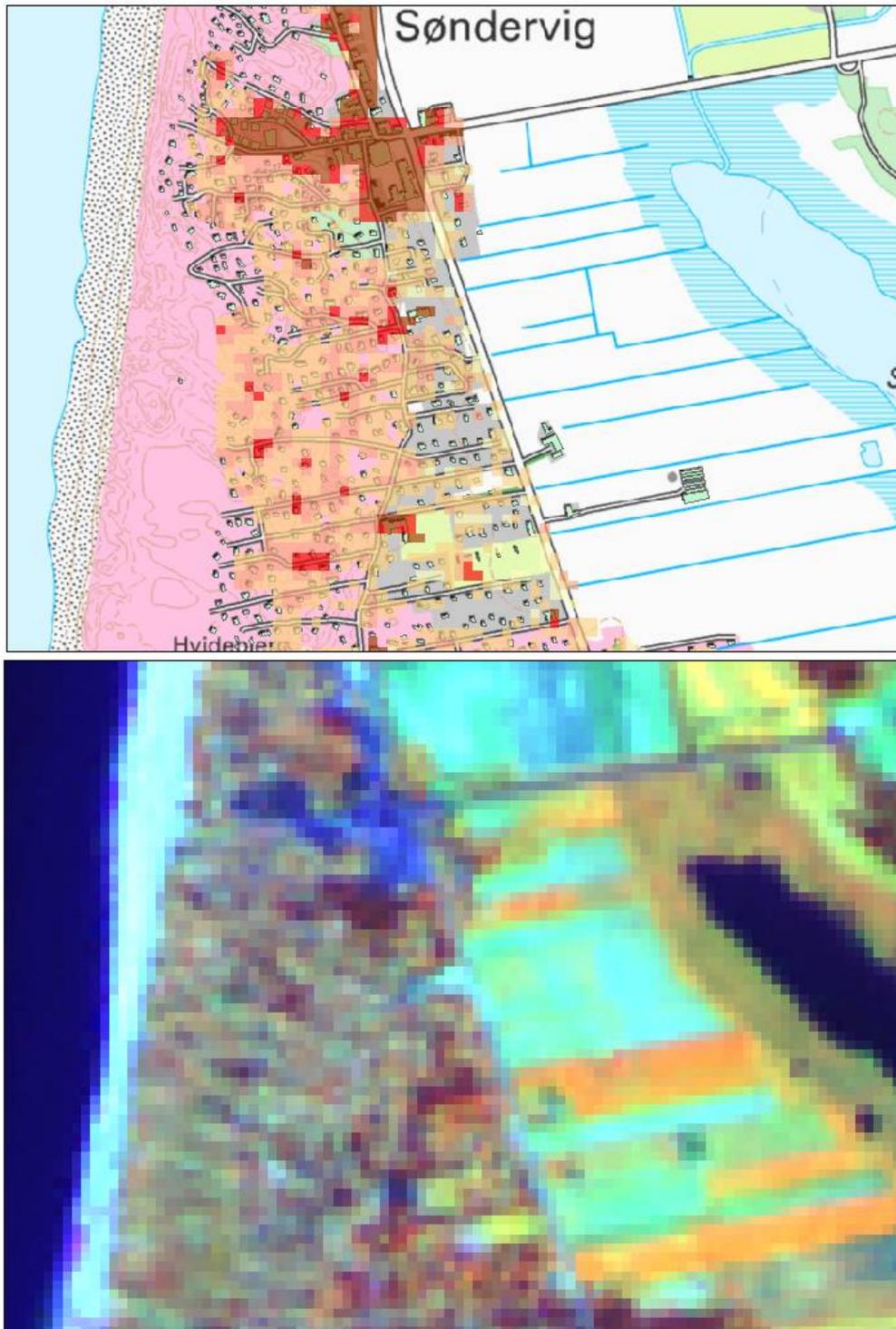
Figur 5b Brejning from Google Earth



Figur 6a Bornholm / Aakirkeby - an agricultural area showing some smaller roads being recognised in SS layer



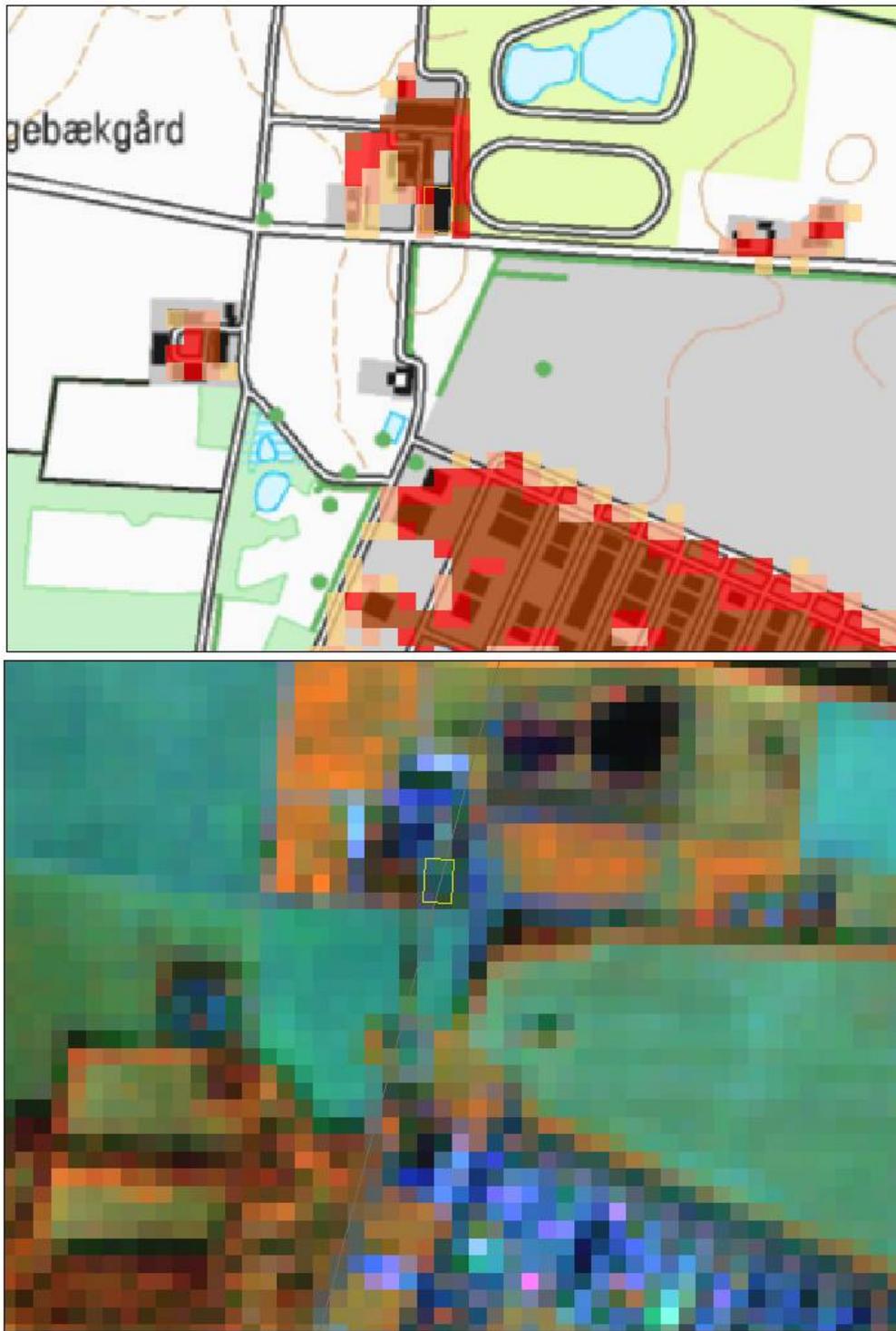
Figur 6b Bornholm/Aakirkeby from Google Earth



Figur 7a Søndervig - Coastal resort with vacation huts. In this area the density is shown with a higher (maybe even correct level of soil sealing - than found in most other summer hut areas. The major road from the east is not identified although it show up in satellite image.



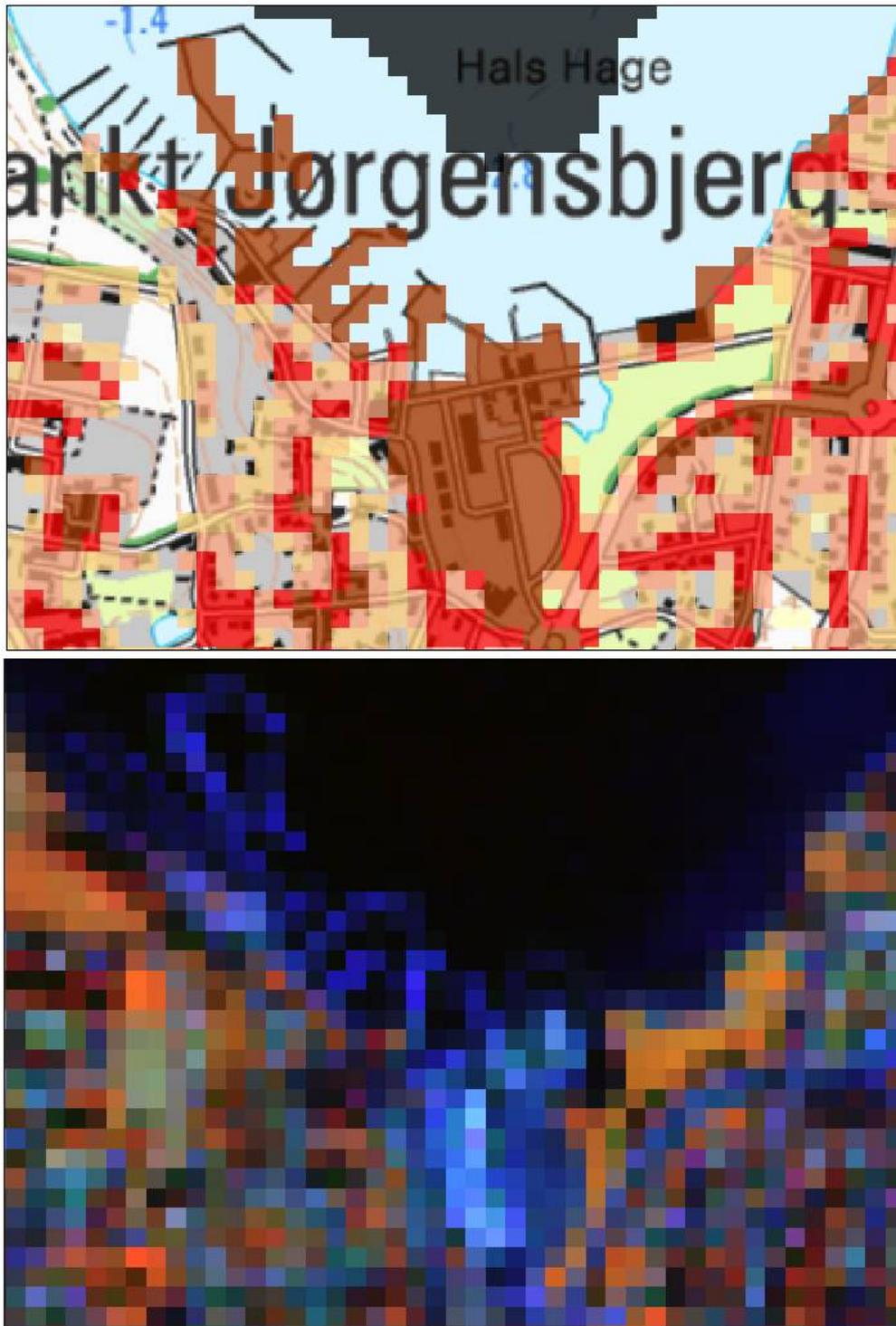
Figur 7b Søndervig from Google Earth



Figur 8a Tune - an odd combination in the top - the building (marked with yellow outline) is not identified, whereas some of the track / road is identified.



Figur 8b Tune – from Google Earth



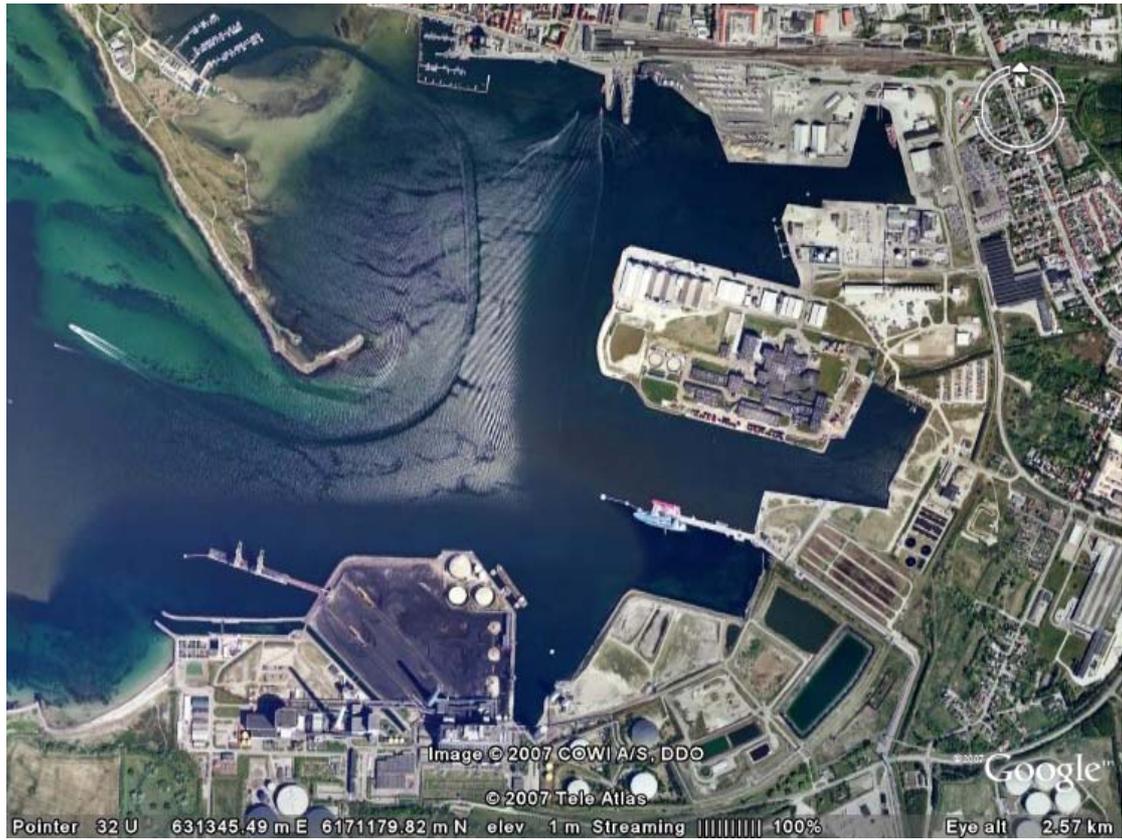
Figur 9a Roskilde - Observe leisure port piers include boats and are exaggerated



Figur 9b Roskilde from Google Earth



Figur 10a Kalundborg - Observer pier in lower left corner include a wide part not being identified, similar the large building in center /right with tarmac roofing



Figur 10b Kalundborg from Google Earth

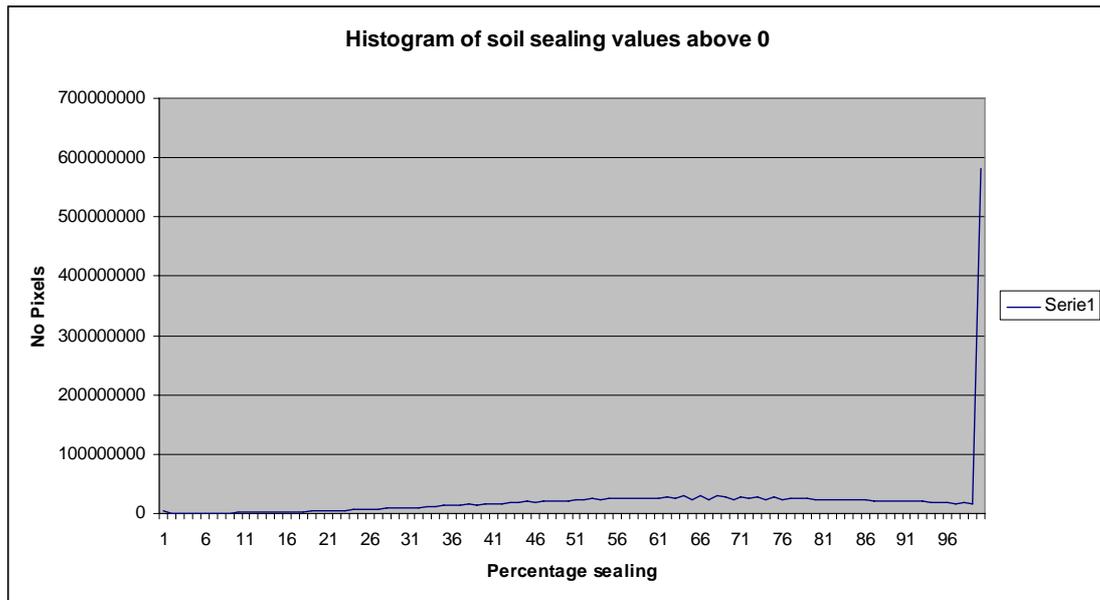


Figure 11 Histogram of all soil sealing values above 0