Final draft

European validation of GMES FTS Soil Sealing Enhancement data

Prepared by:
Gergely Maucha, György Büttner and Barbara Kosztra
Institute of Geodesy, Cartography and Remote Sensing (FÖMI)

29.06.2010

Project manager:
Markus Erhard with Hermann Peifer

Universitat Autònoma de Barcelona
Edifici C – Torre C5 4ª planta
08193 Bellaterra (Barcelona)
Spain, EU

Contact: etclusi@uab.cat

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Soil Sealing (or imperviousness) is the first high-resolution Land Monitoring layer of the EEA with European coverage. Its main use is the characterisation of the human impact on the environment. Multi-sensor and bi-temporal, orthorectified satellite imagery (IMAGE2006) was used to derive soil sealing data covering 38 countries of Europe. Production of the soil sealing database was implemented in two phases: (1) Initial Soil Sealing (ISS) and (2) Soil Sealing Enhancement data (SSE), which is the improvement of the ISS database on the basis of evaluation of ISS data by some Member States.

The main deliverable was a raster dataset of continuous degree of soil sealing ranging from 0 - 100% in full spatial resolution (20 m x 20 m) with the associated metadata. A derived product, a raster dataset of continuous degree of soil sealing ranging from 0 - 100% in aggregated spatial resolution (100 m x 100 m) in European projection was validated.

According to the descriptive statistics, **6.5 % of the European territory is** covered by 1 ha cells **including sealing** (any percent between 1-100), and the total **sealed surface is 1,8 %.** Built-up covers 0.5 % of Europe (if the sealing threshold is 80%) or 2.5% (with 30% threshold).

The purpose of this report is to introduce the method and results of validation of SSE data, more precisely the "built-up" class. Very High Resolution (VHR) Google Earth imagery was used as independent, higher resolution dataset. Reference sealing percentages were obtained using a 10x10 grid positioned around the sampling points. Checking **classification accuracy** was understood as estimating **omission error, commission error and overall accuracy.**

Due to the small size of the sealed layer (and consequently of the built-up class) the full error matrix could not be derived for the whole dataset with affordable efforts, therefore two experiments were devoted to estimate the accuracy of the built-up class:

1. In the 1st experiment only sealed cells were sampled;
2. In the 2nd experiment sampling was restricted to the CLC Artificial surfaces layer, where most of the sealed areas are found.

Using randomly selected samples, the SSE values and reference sealing values were compared. The mean difference between the two was close to zero i.e. neither systematic overestimation nor systematic underestimation occurred. The large standard deviation however shows the limitations of this technology. A number of factors might be responsible for the large standard deviation. Two of these were examined, one being the effect of data producer, second being the effect of source data quality. Differences between results of Service Providers were discovered. For a surprise, EO data quality as shown by the mitigation file did not prove to be a significant factor of this variability.

Main findings of the validation of the built-up class (builtup80):

- According to service specification (builtup80 class) **85% overall accuracy has been achieved** according to both experiments.
- However, **significant (>50%) commission error** was shown by both experiments. This means that lots of non built-up samples were erroneously classified as built-up by SP.
- **Calculated omission error is low**, below the 15% limit. This means that only few built-up samples were classified as non-built-up by SP (few built-up left out).

An important finding of both the "look and feel" comparison and the statistical analysis is that **setting the threshold of built-up class at 30% SSE provides more reliable estimation of built-up** area than setting the threshold at 80%. For the builtup30 class the calculated **overall accuracy is above 85%, and both omission and commission errors are below 15%**.

1 INTRODUCTION

Soil Sealing or imperviousness is an important environmental parameter. It can be used to characterise the human impact on the environment, e.g. the extent of built-up areas or the change in runoff due to human construction activity. The 2001 National Land Cover Database of the US includes an imperviousness layer [1] (in addition to the Land Cover and percent tree canopy layers).

The first soil sealing database for Europe was produced as part of the GMES Fast Track Service on Land Monitoring (Land FTS LM) in 2006-2008. Multi-sensor and bi-temporal, orthorectified satellite imagery – called IMAGE2006 - was used [2] to derive soil sealing data, the same as the CORINE Land Cover 2006 update.

Production of Soil Sealing database covering 38 European countries (32 EEA Member States and 6 West-Balkan countries) was implemented in two phases:

- Initial Soil Sealing (ISS) data based on EEA specification [3], and
- Soil Sealing Enhancement data (SSE) based on a new tender specification by EEA [5], which was prepared on the basis of evaluation of ISS data by some Member States. The SSE database is thus an improvement of the ISS database based on evaluation of users.

1.1 DEFINITIONS

Both publicly available definitions and definitions according to the Tender Specifications [5] are presented below.

Soil Sealing is the loss of soil resources due to the covering of land for housing, roads or other construction work [6].

Soil sealing layer (either ISS or SSE) is seamless raster layer containing continuous values ranging from 0 - 100% representing a degree of soil sealing. The original product has a full spatial resolution of 20 m x 20 m with the associated metadata [4].

Impervious surfaces are mainly artificial structures - such as pavements (roads, sidewalks, driveways and parking lots) that are covered by impenetrable materials such as asphalt, concrete, brick, and stone and rooftops [7]. Imperviousness (or degree of soil sealing) is estimated in relation to the pixel area [5].

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 [9]. Impervious surfaces of built-up areas account for 80 to 100% of the total cover [3]. Built-up areas at 1 ha level are defined by an average sealing degree per 1 ha unit above 80% [5].

According to FAO [11] industrial and urban areas are classified as follows based upon the occurrence of impervious surfaces compared to permeable surfaces:

- High density: more than 75 percent of the total surface consists of impervious surface(s).
- Medium density: 50 to 75 percent of the total surface consists of impervious surface(s).
- Low density: 50 to 30 percent of the total surface consists of impervious surface(s).
- Scattered: less than 30 to 15 percent of the total surface consists of impervious surface(s).

In this report the following definitions will be used to distinguish built-up from non-built-up and sealed from non-sealed areas.

Table 1 Definitions to separate built-up/non-built-up and sealed/non-sealed surfaces

Raster value	Production class	Validation class
0	Non-sealed cells	Non-built-up area
1-79	Sealed cells	
80-100	Sealed cells	Built-up area
254	Unclassifiable pixels	
255	No data	

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% [5].

2 SERVICE SPECIFICATION

2.1 INITIAL SOIL SEALING

Service specification is quoted from the Project Management Plan [4].

"The main deliverable is a raster dataset of built-up and non built-up areas[1] including continuous degree of soil sealing ranging from 0 - 100% in full spatial resolution (20 m x 20 m) with the associated metadata.

Input data provided by ESA:

Orthorectified satellite data coverage for Europe (Image2006), acquired primarily in the reference year 2006 (+/- 1 year), covering two dates, used sensors SPOT 4 and 5 (HRVIR) and IRS-P6 LISS-III:

- 20 m resampled (with cubic convolution interpolation)
- 4 spectral bands
- Max. 5% cloud coverage
- Covering 2 dates, at least 6 weeks apart from the respective scene selected for the first coverage
- Orthorectified towards national projection systems (used DTM unknown)
- Metadata to each scene

Input data provided by EEA:

- Dataset with national country borders (to be used for clipping the data at a national level) as defined and provided by the EEA
- European-wide reference grid (100 m x 100 m)

Ancillary input data:

- Image2000, in national projection
- VHR remote sensing data as provided by Google Earth

Methodology:

Supervised classification of built-up[2] areas following with visual improvement of classification result and derivation of degree of soil sealing based on calibrated NDVI.

Geometric resolution: Pixel resolution 20 m x 20 m

Coordinate Reference System:

- National projection systems for country data sets
- ETRS89 for seamless European data set.

[1] Misleading: Built-up areas mentioned here mean a binary mask as a result of a supervised classification used during production to separate areas for the calculation of soil sealing levels 1-100. This does not correspond to the later definition of built-up areas for the QC criteria (e.g. built-up areas are represented a degree of soil sealing of 80 - 100%, non built-up areas are represented a degree of soil sealing of 0 – 79%.)

[2] More precisely: sealed cells – see Table 1.

Geometric accuracy :

According to orthorectified satellite image base delivered by ESA.

Thematic accuracy (in %):

Classification accuracy per hectare (based on 100 m x 100 m grid) of built-up and non built-up areas[3] is > 85%.

Data type: Raster

Delivery format: IMAGINE Image (IMG)

Raster coding:

0 – Non-built up areas[4], water bodies inland;

1-100 - sealing values in percentage of the area;

254 – Unclassifiable areas (clouds, shadows, etc.);

255 – No Data (no thematic information)

Metadata: According to EEA metadata standards (EEA MSGI specification)"

2.2 Soil Sealing Enhancement

The tender describing SSE includes additional thematic specifications as improvements of the ISS data [5].

Removal of CLC class 1.3.x objects

Objects corresponding to CLC class 1.3.x (mines, quarries, dump and construction sites) must be removed from the 20m pixel layer. The functional outline of these objects should be removed, while buildings belonging to these units must be maintained and their degree of imperviousness must be assessed.

Closing of gaps in settlements

Gaps in settlements (errors of omission) must be corrected. Member States reported that special roof types were often not mapped as impervious areas. Furthermore, parts of discontinuous urban fabric were not mapped. These areas should be included in the new dataset.

Correction of airports and harbours

Misclassification (errors of omission and errors of commission) of airports and harbour areas should be corrected.

Misclassifications

Errors of commission related to beaches, sand and dunes (CLC class 3.3.1), bare rock (CLC class 3.3.2) and sparsely vegetated areas (CLC class 3.3.3) should be corrected.

2.3 Satellite Image Processing

In order to better understand the results of the validation an overview of satellite image processing is quoted from the Project Management Plan [4].

"As the main challenge, the derivation of a continuous degree of soil sealing had to be solved in a robust, reliable and reproducible way. The applied image processing approach was based on the fact that a reliable derivation of soil sealing degree was not possible directly from the vegetation index. Low vegetation index values, which are characteristic

[3] As defined in Table 1.

[4] Precisely: non-sealed area – see Table 1.

for densely built-up areas, are e.g. also found in bare soil areas of agricultural fields. Even when using multi-temporal satellite images with different acquisition dates in combination with bi-temporal, multi-spectral classification techniques the result may be improved, but the vegetation indices of two acquisitions are still too ambiguous.

Therefore, the applied image processing approach started with deriving a binary map of built-up areas[5] and then further subdivided this area into 100 degrees of soil sealing, ranging from totally sealed surfaces (100% degree of soil sealing) up to built-up areas with extensive vegetation cover (1% degree of soil sealing).

The applied methodological approach consisted of the following main steps:

a) Data preparation & management: Provision of spatial database of bi-temporal satellite images and derived working sub-areas ("Working Units" = WU) to be processed in the following steps:

b) Core processing, containing the 3 main processing steps:

   a. Hybrid automated classification with supervised and unsupervised elements, leading to binary maps of built-up area:

      i. Automated supervised training of built-up areas using a region growing tool, to capture the variety of urban spectral signatures. Definition of ca. 50 AOIs (Areas of Interest) per WU, and clustering of the retrieved training pixels into 20 spectral classes via (e.g.) ISODATA.

      ii. Maximum Likelihood (ML) classification and thresholding based on spectral distance measures, to receive a preliminary built-up layer.

      iii. Unsupervised clustering of all pixels outside this preliminary built-up layer in order to automatically derive spectral signatures for non-built-up areas.

      iv. ML classification of the complete WU using all built-up and non-built-up spectral signatures.

      v. Additional spectral signature refinement of spectrally mixed classes via clustering of the respective areas. New classification of the complete image using these new signatures in addition to the initial signatures.

   b. Manual correction of the binary built-up map to obtain the required quantitative thematic accuracy (85%) as well as good qualitative results.

   c. Derivation of degree of soil sealing based on the NDVI (Normalised Difference Vegetation Index).

c) Generation of sub-country / country data sets.

d) Accuracy assessment.

e) Re-projection & mosaicking, generation of seamless European dataset.

Neither topographic normalisation nor atmospheric correction was applied to satellite imagery [4].

2.3.1 Mitigation file

A special file was created to show which parts of Europe were covered by suboptimal EO data in order to mark these areas as where "the Service Provider does not take the responsibility for data quality". Sufficient EO data quality means that the following criteria are fulfilled:

- availability of minimum two "proper" EO data coverage;
- acquisition dates are inside the user-specified image acquisition window;

[5] Precisely: sealed cells – see Table 1.

- minimum six weeks between acquisition days;
- no cloud coverage is present.

These characteristics are found as attributes in the mitigation vector file. Deriving these data from the file it was shown in [8] that in Nordic countries (FI, IS, NO) practically 100% of the area was “not OK” concerning EO coverage. There are 22 countries where the SP does not take the responsibility for more than 50% of the area according to the mitigation file.

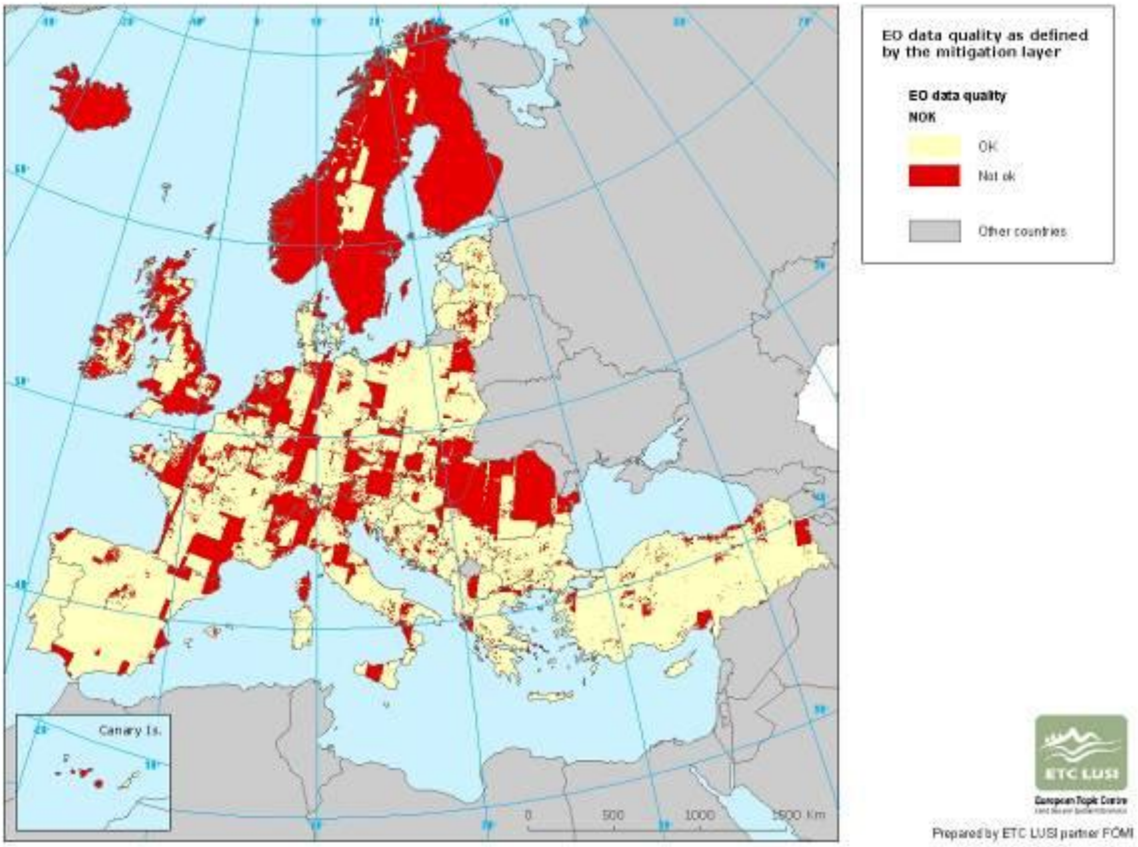

Figure 1 EO data quality as shown by the mitigation layer. Red area is where input EO data do not meet service specifications.

2.4 PRODUCTS

a) Primary product: A raster dataset of <u>continuous degree of soil sealing</u> ranging from 0 - 100% in full spatial resolution (20 m x 20 m) in national projection with the associated metadata. All additional datasets have been derived from the primary product.

b) Raster dataset of <u>continuous degree of soil sealing</u> ranging from 0 - 100% in aggregated spatial resolution (100 m x 100 m) in European projection with associated metadata. This dataset was the main target of validation.

c) LUT-s allowing the aggregation of the continuous values of degree of soil sealing into the following five soil sealing level classes:

• 0 - 29%

• 30 - 49% (30% threshold compatible with lower limit of CLC class 1.1.2, 49% threshold a median of CLC class 1.1.2)

• 50 - 79% (79% threshold compatible with ceiling of CLC class 1.1.2)

• 80 - 99% (80% threshold compatible with CLC class 1.1.1)

• 100%

European Soil Sealing data accessible at the EEA data server [12] corresponds to products derived from the SSE database (products b) and c) in the above list).

2.5 ACCURACY ASSESSMENT BY SERVICE PROVIDER

Classification accuracy is rather loosely defined in the tender specification [3] saying: "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%". According to project documentation Service Provider understands the accuracy assessment of the product in the following way [4].

1. Definition of 100 m x 100 m reference grid in national projection of the respective country assessed.
2. Stratification of the area based on CORINE Land Cover level I. To emphasize the accuracy assessment in the urban areas, 50 % of the sample plots are placed within CLC class Artificial Surfaces, the other 50% are placed in the remaining classes[6].
3. Cluster- based random sampling based on 100 m x 100 m reference grid, defined per single nation, number of samples adapted to nation size in km².
4. Re-projection of reference samples to allow overlay with Google Earth.
5. Estimation, if reference cell will be labelled as "built-up" according to EEA definition or not (80% threshold degree of soil sealing). Taking into account the visibility of objects in the satellite images used for the production of the raster product[7].
6. Calculation of overall accuracy to generate accuracy measure (overall accuracy, user accuracy (commission error), producer accuracy (omission error)), per single nation (for internal use & validation only) and for European dataset for publication by EEA.[8]
7. Adaptation of statistics with regard to the mitigation shape file. All sample plots falling within areas of the raster product, where the underlying IMAGE2006 data has been identified to fail the ITT's specifications, are not included in the final statistics[9]."

According to the specifications, the built-up raster product, which is subject to the accuracy assessment, is accepted if the final statistics indicate an overall accuracy above 85 %.

Accuracy assessment was performed for each country product for internal quality control.

---

[6] The stratification method raises questions, and causes bias in the final statistics: (1) CLC Artificial surfaces classes are not fully sealed, includes significant amount of non-sealed surfaces, (2) Results from both strata were used together in the same error matrix without any normalization.

[7] The estimation was probably performed by "educated guess", without using a point grid. The uncertainty of this method was not estimated.

[8] The overall accuracy is <u>not</u> a measure of class accuracy, it is the average user's accuracy of both (built-up and non built-up) classes, heavily dominated by the larger (non built-up) class. Commission and omission errors are the right measures of a single class accuracy, but because of the stratification (without normalizing the matrix) only the calculated commission error is valid.

[9] The consequences are twofold: (1) Usually only very few samples were left, not enough for a representative estimation. (2) Further bias in the statistics, not indicated how many samples were left by strata.

3 METHOD OF VALIDATION

In any validation exercise, two basic questions should be raised when planning or evaluating the QC of a product:

Question 1: How much the product corresponds to the specifications?

- Informs about the quality of the completed work.
- Usually provides the basis for accepting/refusing/revising the product.
- If a decision has to be made quantitative results are needed.

Question 2: How much the product corresponds to the reality?

- Informs about the usability of the product.
- Different uses often require different kinds of quality parameters.
- Qualitative as well as quantitative methods can characterise the data.

The two questions above concern overlapping issues. Both are valid questions usually with different answers. Product specifications and QC criteria might significantly influence the accuracy figure of the product.

3.1 CHARACTERISATION OF SSE DATA

As described in chapter 2.4, the SSE product is a seamless raster layer containing the degree of imperviousness ranging from 0 to 100%. The built-up class (according to the QC criteria) can be created by setting thresholds to form sealing levels (e.g. 0-79: Non built-up, 80-100: Built-up).

Because in tender document and the service specification the expression "built-up" (see Ch. 2.1) is used in two different meanings, we will use a definition introduced in Table 1 in order to avoid misunderstanding.

As seen in Table 1, sealed cells may have a sealing level 1-100. If we want to calculate the amount of estimated **sealed area** on the real earth surface for a 100 m x 100 m cell with 1% sealing level we have to write:

$$area_{sealed} = \text{[illegible]}$$

Correspondingly, we can calculate the amount of sealed area for a larger region by summing the sealed area of the "n" number of cells included:

$$A_{sealed} = \sum_{i=1}^{n} \text{[illegible]} = \text{[illegible]} \sum_{i=1}^{n} \text{[illegible]}$$

as the cell area is the same all over the database. Applying this equation we calculated the amount of sealed area for Europe based on SSE database.

Figure 2 shows the distribution of sealed cells in Europe, while Table 2 shows the calculated percentages for Europe using the defined categories.

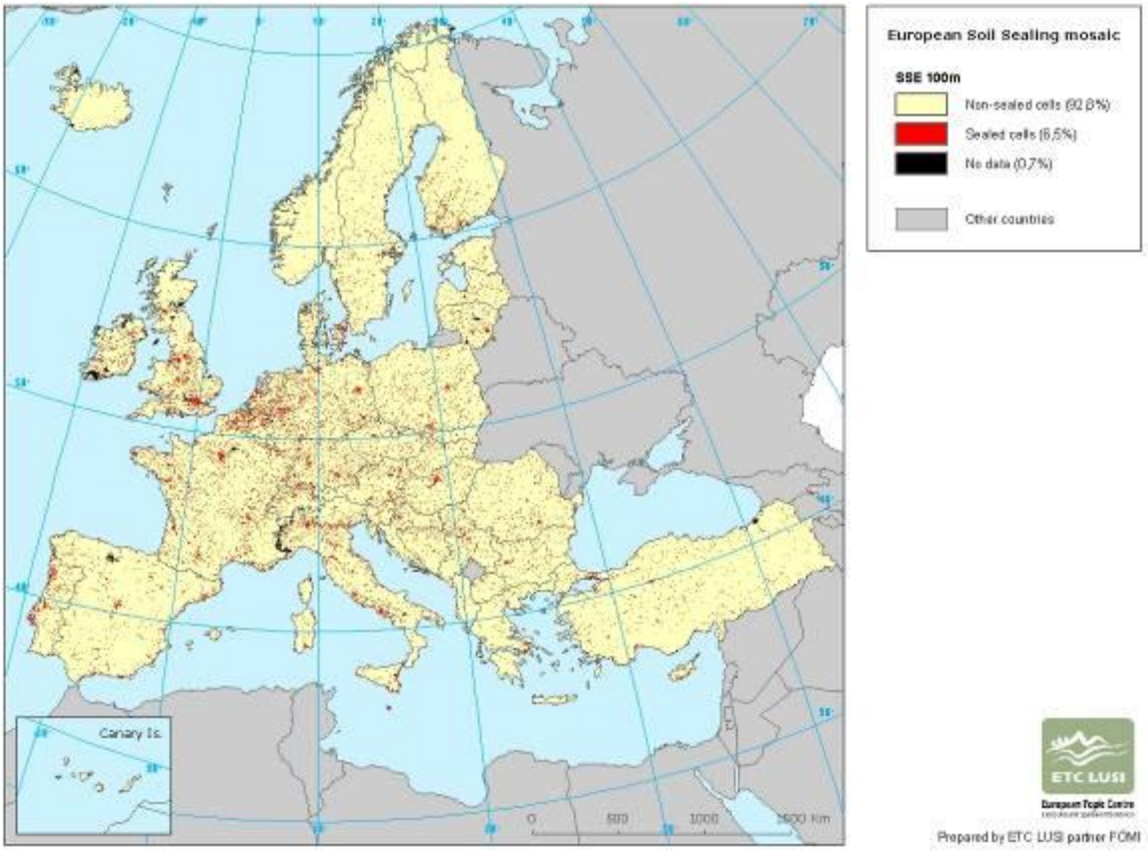

Figure 2 Distribution of sealed cells in Europe (all areas ≥ 1% soil sealing) based on SSE100 data.

Table 2 Soil Sealing statistics for Europe

Sealing level (raster value)	Category	Area (ha)	Area (%)
0	Non-sealed cells	568 636 267	92,8%
1-100	Sealed cells	39 829 049	6,5%
254-255	Unclassifiable pixels and no data	4 289 282	0,7%
Total:	**Total in participating countries**	**612 754 598**	**100,0%**
0-79	Non built-up area	605 401 543	98,8%
80-100	Built-up area	3 063 773	0,5%
254-255	Unclassifiable pixels and no data	4 289 282	0,7%
Total:	**Total in participating countries**	**612 754 598**	**100,0%**
Calculated by summing the non-sealed part from cells having sealing level between 0-99	Amount of non-sealed areas	597 357 664	97,5%
Calculated by summing the sealed part from cells having sealing level between 1-100	Amount of sealed areas	11 107 652	1,8%
254-255	Unclassifiable pixels and no data	4 289 282	0,7%
All	**Total in participating countries**	**612 754 598**	100%

Table 3 shows sealing statistics by countries. It shows that the most sealed area was found in Malta, Belgium and the Netherlands. On the other hand the least sealed area was observed in Iceland, Norway and Sweden.

Table 3 Sealing statistics for countries

Country	**Total area**	**Sealed area**	
	ha	**ha**	**%**
Albania	2 851 147	17 596	0,62%
Austria	8 393 512	160 871	1,92%
Belgium	3 066 326	226 020	7,37%
Bosnia and Herzegovina	5 120 432	56 039	1,09%
Bulgaria	11 095 596	203 642	1,84%
Croatia	5 655 489	119 343	2,11%
Cyprus	924 614	33 501	3,62%
Czech Republic	7 886 610	251 471	3,19%
Denmark	4 335 115	153 045	3,53%
Estonia	4 533 695	39 538	0,87%
Finland	39 091 136	201 024	0,51%
France	54 917 795	1 522 972	2,77%
Greece	13 203 750	177 718	1,35%
Germany	35 780 884	1 815 360	5,07%
Hungary	9 301 366	294 125	3,16%
Iceland	10 292 575	15 288	0,15%
Ireland	7 025 866	111 401	1,59%
Italy	30 147 542	848 561	2,81%
Latvia	6 459 914	71 998	1,11%
Liechtenstein	16 002	922	5,76%
Lithuania	6 529 252	130 480	2,00%
Luxembourg	259 571	12 728	4,90%
Macedonia	2 528 541	21 957	0,87%
Malta	31 546	4 186	13,27%
Montenegro	1 386 954	10 652	0,77%
Netherlands	4 152 908	304 372	7,33%
Norway	46 934 067	92 359	0,20%
Poland	31 265 769	738 132	2,36%
Portugal	9 203 156	285 443	3,10%
Romania	23 845 495	382 840	1,61%
Serbia	8 854 117	161 184	1,82%
Slovakia	4 901 098	115 372	2,35%
Slovenia	2 028 016	37 305	1,84%
Spain	50 598 509	719 528	1,42%
Sweden	53 177 470	198 356	0,37%
Switzerland	4 128 912	110 358	2,67%
Turkey	77 945 992	629 882	0,81%
United Kingdom	24 883 859	832 083	3,34%
Total in participating countries	**612 754 598**	**11 107 652**	**1,81%**

Table 4 shows sealing statistics by CLC classes. As expected most sealed surfaces are found in the artificial surfaces classes. Some sealing can be recognised in classes of agriculture, and less amount in classes of forests and semi-natural group, wetlands and water. The highest sealing percentage in Table 4 is associated to CLC class 111, value for sealing level 80-100% being 49.5%. This should be interpreted so that 49.5% of Continuous urban class area in Europe includes cells with ≥ 80% sealing.

Table 4 Sealing statistics for CLC classes

CLC code	Sealing level (%)				
	0	1-29	30-49	50-79	80-100
111	2,6%	10,0%	10,4%	27,5%	49,5%
112	17,0%	31,0%	18,9%	23,2%	9,9%
121	15,3%	17,3%	12,4%	21,7%	33,3%
122	15,6%	30,6%	18,9%	17,7%	17,2%
123	18,0%	15,6%	10,8%	17,3%	38,3%
124	40,3%	27,9%	12,8%	11,1%	7,9%
131	89,5%	5,6%	1,8%	1,8%	1,4%
132	85,6%	7,0%	2,6%	2,5%	2,3%
133	55,0%	18,5%	9,5%	9,8%	7,3%
141	48,3%	35,0%	8,8%	6,1%	1,8%
142	65,0%	20,9%	6,8%	5,6%	1,8%
211	95,0%	4,1%	0,6%	0,3%	0,1%
212	95,7%	2,9%	0,6%	0,5%	0,2%
213	97,3%	2,3%	0,2%	0,1%	0,0%
221	92,0%	6,4%	1,0%	0,4%	0,1%
222	91,7%	6,6%	1,0%	0,5%	0,1%
223	94,2%	4,9%	0,6%	0,3%	0,1%
231	92,5%	6,5%	0,7%	0,3%	0,0%
241	83,7%	13,6%	1,8%	0,7%	0,1%
242	85,1%	11,8%	2,0%	1,0%	0,2%
243	93,4%	5,7%	0,6%	0,3%	0,1%
244	98,7%	1,1%	0,1%	0,1%	0,0%
311	98,8%	1,1%	0,1%	0,0%	0,0%
312	98,4%	1,4%	0,1%	0,0%	0,0%
313	98,2%	1,6%	0,1%	0,0%	0,0%
321	98,9%	0,9%	0,1%	0,1%	0,0%
322	99,3%	0,6%	0,0%	0,0%	0,0%
323	98,7%	1,0%	0,2%	0,1%	0,0%
324	98,6%	1,2%	0,1%	0,0%	0,0%
331	96,9%	2,1%	0,5%	0,3%	0,1%
332	99,9%	0,1%	0,0%	0,0%	0,0%
333	99,6%	0,3%	0,0%	0,0%	0,0%
334	98,3%	1,5%	0,2%	0,1%	0,0%
335	100,0%	0,0%	0,0%	0,0%	0,0%
411	98,4%	1,3%	0,2%	0,1%	0,0%
412	99,5%	0,5%	0,0%	0,0%	0,0%
421	97,5%	2,0%	0,3%	0,2%	0,1%
422	94,1%	3,7%	1,3%	0,6%	0,2%
423	99,4%	0,5%	0,1%	0,0%	0,0%
511	95,1%	3,8%	0,7%	0,3%	0,1%
512	99,3%	0,6%	0,1%	0,0%	0,0%
521	99,4%	0,5%	0,1%	0,0%	0,0%
522	98,0%	1,5%	0,3%	0,1%	0,0%
Sum	**93,5%**	**4,0%**	**1,0%**	**1,0%**	**0,5%**

CLC data used: 100m raster version; CLC2006 in 36 countries; CLC2000 in CH, GR, UK

3.2 BASIC CONSIDERATIONS

Detailed SSE data specification is presented in Ch. 1.1 and Ch.2 (only briefly repeated here):

Built-up areas at 1 ha level: The average sealing degree per 1 ha unit above 80%.

QC criteria: Classification accuracy per hectare (100 m x 100 m) of the built-up and nonbuilt-up areas should be above 85%.

As seen above, the classification accuracy is rather loosely defined in the tender specification. Not a specific measure (like overall accuracy) was appraised, neither any indication how to measure the accuracy.

The following facts were taken into consideration when deciding how to measure the accuracy.

"Look and feel" checks have to be performed to yield an overall impression on the data.

The accuracy of the built-up class can be characterised by two parameters:

- **commission error**: classification: SSE ≥ 80%, reference: SSE < 80%
- **omission error**: classification: SSE < 80, reference: SSE ≥ 80%

Both errors are related to the area of the built-up class, and both have to be smaller than 15%. In other words: the producer's accuracy and the user's accuracy both have to be larger than 85%.

The advantage of the **overall accuracy** would be to provide a single value as accuracy measure, but the overall accuracy is always a measure of overall quality of a classification, including more classes (i.e. the average commission error of all classes). In our case overall accuracy is always dominated by the larger (non built-up) class. Still, this measure was also calculated in this validation.

Reference sealing values have to be determined for 100 m x 100 m grid cells with a help of a **point grid** containing 10x10 points as described in Ch. 3.3. This way SSE values are compared to a more precise reference measurement.

Stratified random sampling (see Ch. 3.2.1) has to be applied to be able to measure the relevant accuracy parameters effectively.

3.2.1 Stratification

Sealed cells (as well as built-up) cover just a small fraction of Europe (Ch.3.1). If we drew random samples from the entire area, most of the samples would fall on non-sealed areas. Consequently the representativeness of the sealed / built-up samples would be low. (Increasing the total number of samples so much, as it would be required by the small classes would yield unnecessarily large number of samples for large classes.) Therefore stratification should be used to increase the efficiency of sampling. Selecting the right way of stratification should be done especially carefully. Ad-hoc applied stratification strategies make nearly impossible to puzzle out the meaning of the final statistical results. Additionally, if stratification is applied the resulting error matrix has to be properly normalised (see Annex 1).

In this study two stratification strategies were applied:

- In order to estimate the commission error, random samples were selected from the population of the sealed cells, which also includes the class to be validated (built-up class). The resulting error percentage will directly provide the commission error (see results in Ch. 4.1). With the direct comparison of SSE and reference sealing level, a calibration of the sealing levels could be performed.
- Estimating omission error for a small class is nearly impossible, because errors (omitted sealed area) have to be searched for inside the large non-sealed class.

In our case this area is 92,8% of the total (see Table 2). To provide reliable results with limited efforts a restricted area was defined and examined. The CLC2006 Artificial surfaces layer[10] was used as a mask for a second random sampling. The methodology provided omission and commission errors for the Artificial surfaces stratum, and overall accuracy was calculated as well.

[10] Table 4 shows that the CLC Artificial surfaces class includes the large majority of sealed surfaces

3.3 LOOK AND FEEL CHECKS

Figures in this chapter intend to characterise SSE data quality by simple visualisation. SSE data are overlaid on IMAGE2006 data (IRS LISS III imagery).

Figure 3 SSE 20m data (Budapest airport). Sealed surfaces (runway, road, built-up area) are very well indicated by SSE. The bright linear feature was a road construction in 2006 (not yet sealed).

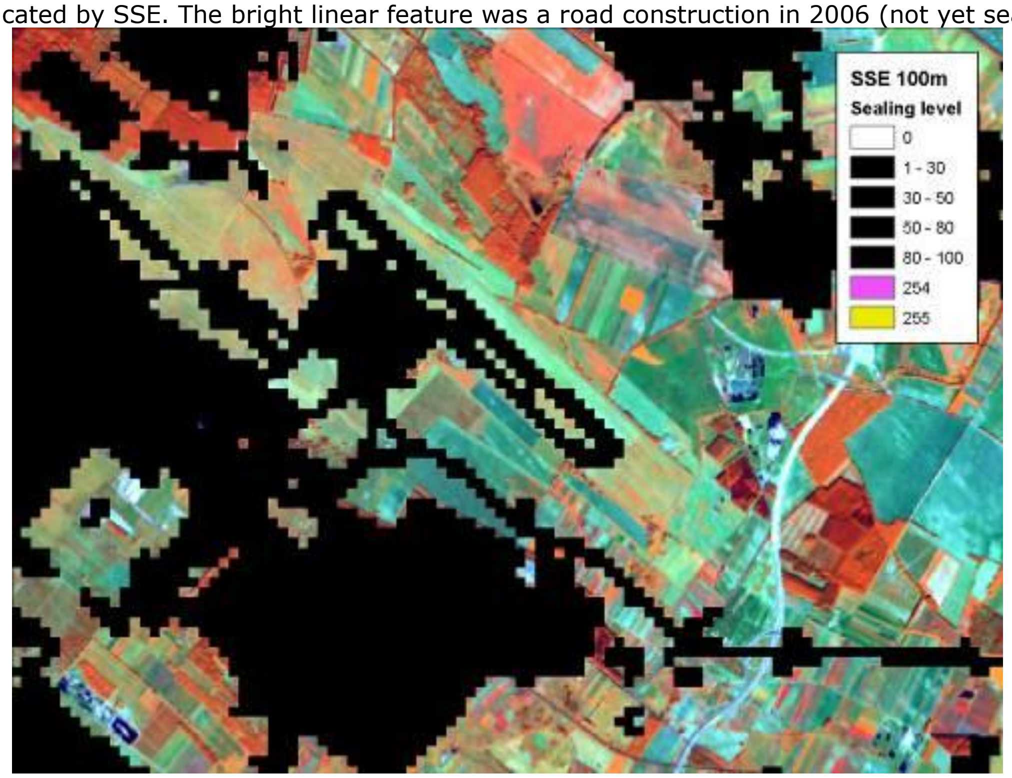

Figure 4 SSE 100m data (Budapest airport). Note the widening of linear features and disappearance of fine structure inside built-up area south of the airport as consequences of the aggregation.

Figure 5 Discontinuous built-up areas with lots of green surfaces between houses. The grey rectangles highlight those 20m pixels where SSE exceeds 80%. If we set the built-up limit to 80% sealing, we can expect omissions in such urban areas. Yellow dots represent the 10m x 10m grid, where the reference sealing was estimated (see Fig. 9)

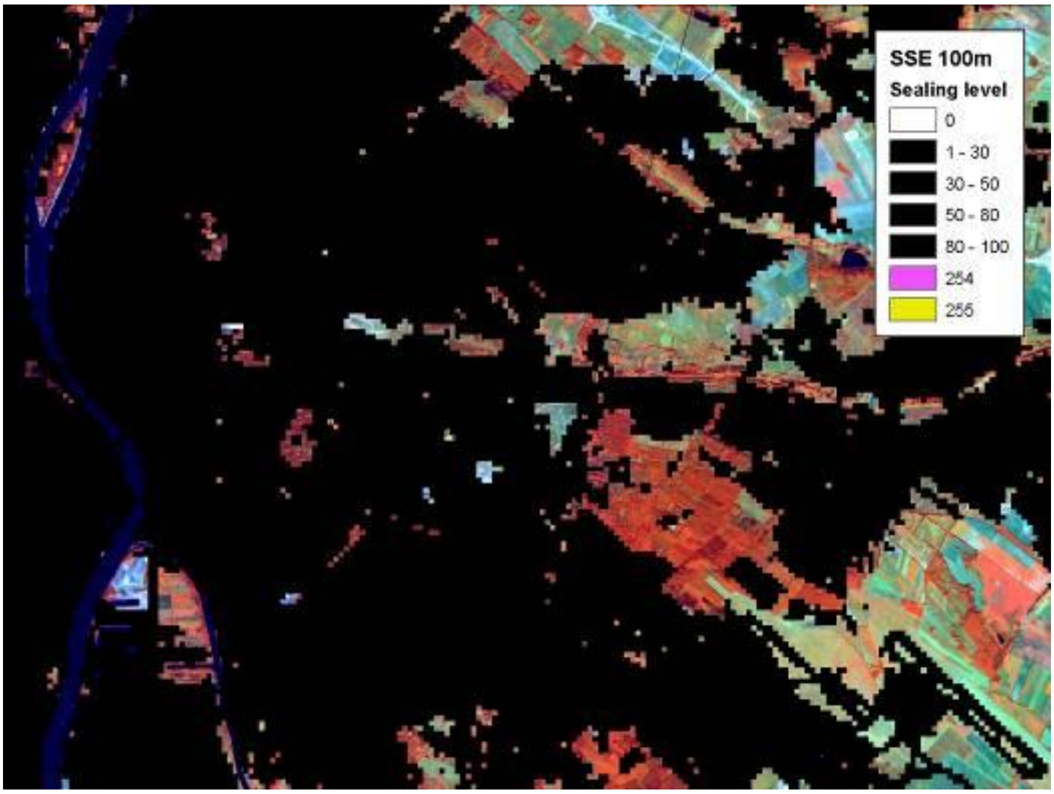

Figure 6 Black mask covers all sealed cells (1-100%). Unmasked areas show all non-sealed cells (water, forest, agriculture, etc). Bright man-made features were under construction in 2006.

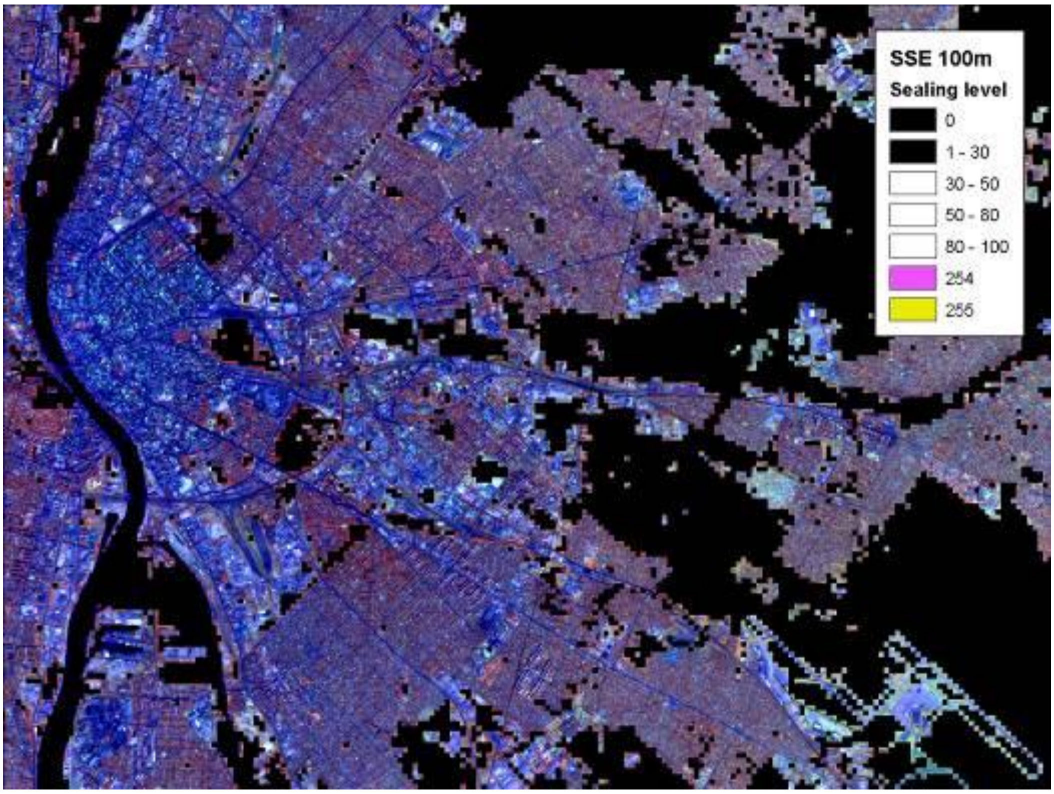

Figure 7 All cells having a sealing level less than 30 are masked for the Eastern part of Budapest. Surfaces with sealing levels equal to or larger than 30 (corresponding to the class called "builtup30") are shown in the colours of the satellite image colour composite.

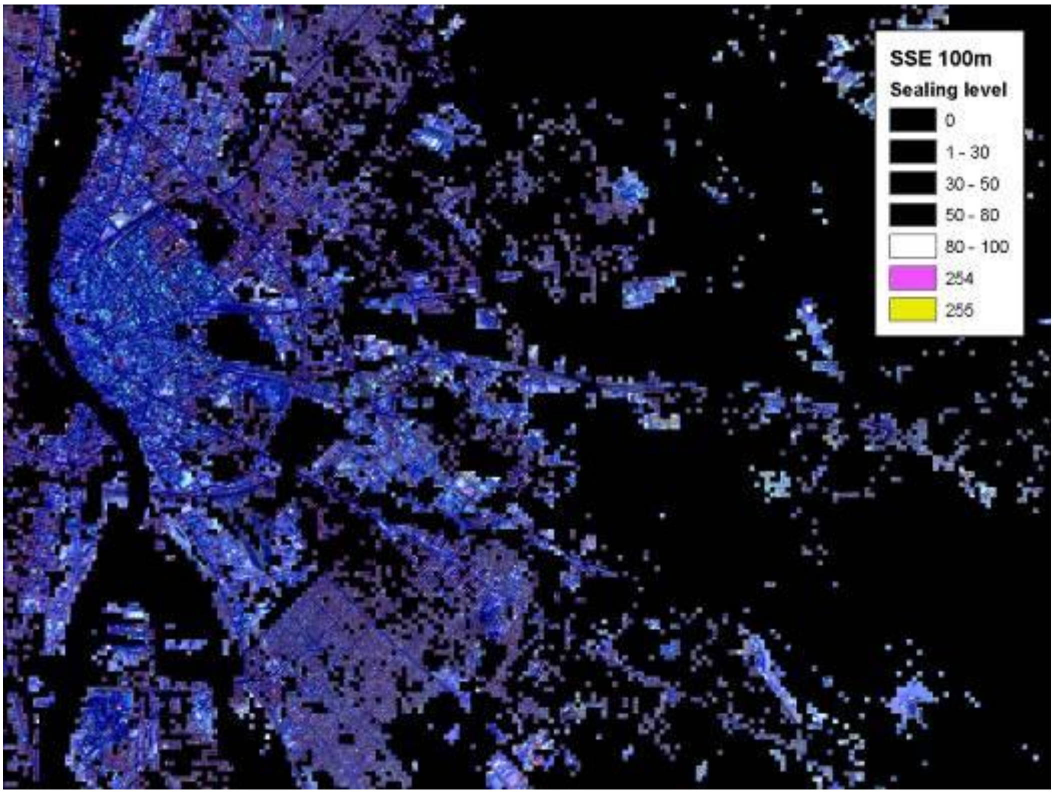

Figure 8 All cells having a sealing level less than 80 are masked for the Eastern part of Budapest. Surfaces with sealing levels equal or larger than 80 (corresponding to the class called "builtup80") are shown in the colours of the satellite image colour composite.

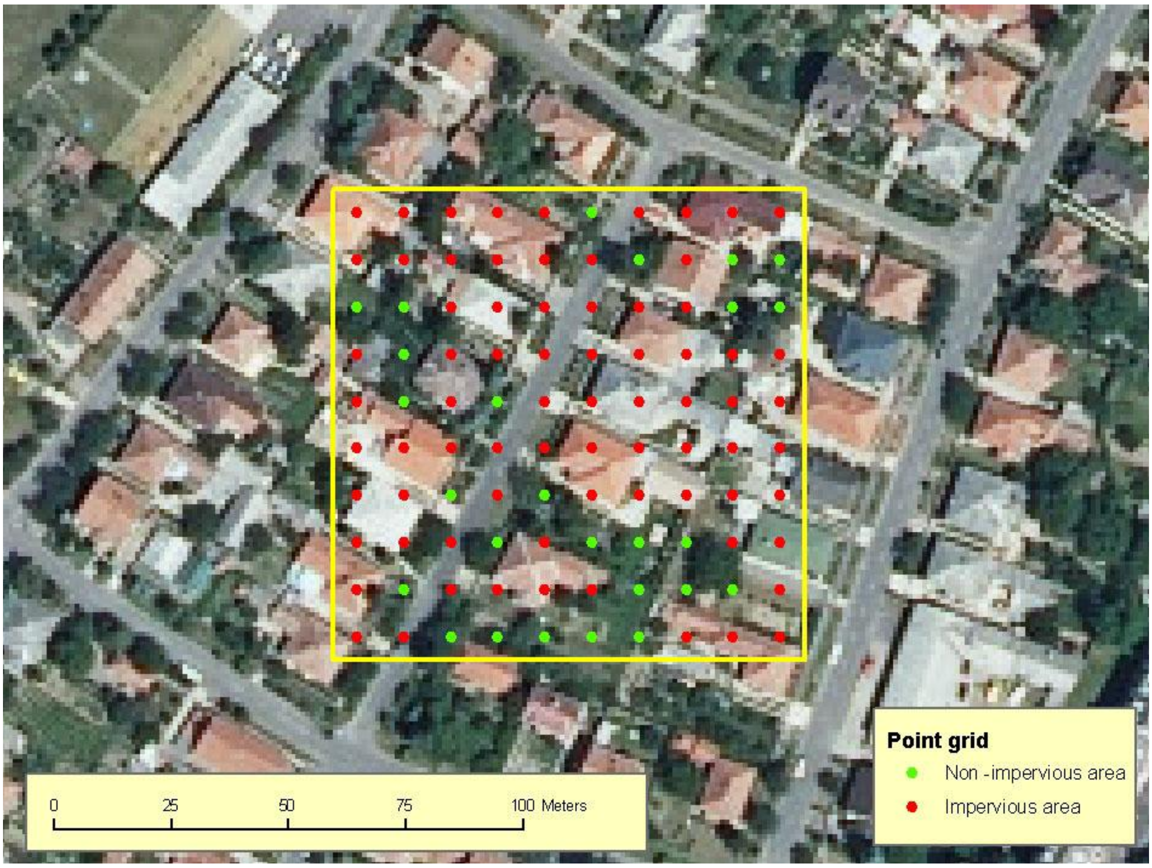

Figure 9 Validation methodology of European SSE data. By counting the number of impervious points inside the 100x100 m grid cell (total of 100 points), the interpreter estimates sealing degree of the sample cell.

3.4 ESTIMATING SOIL SEALING LEVELS

European Soil Sealing Enhancement (SSE) mosaic data at 100 m resolution were validated.

Google Earth (GE) provided valuable reference data to be compared with SSE data. The date of GE imagery was always considered in order to fit to the 2006 ± 1 year date of SSE data. Unfortunately, the spatial coverage of GE imagery was not optimal (see Fig. 10).

The essence of the validation methodology is the estimation of soil sealing level inside the 100 m x100 m sample cells based on available VHR imagery. The interpretation is "blind" in the sense that the interpretation should be carried out without having access to the SSE product.

A specific ArcGIS based tool has been developed for the purposes of the validation of soil sealing. The 100 m x 100 m sample cells were complemented by a 10 m x 10 m point grid inside each sample cell. By counting the impervious points out of the 100 points, the interpreter estimated the sealing degree of the sample observation (Fig. 9).

The potential error sources of the method are:

- Geometric shifts caused by improper georeferencing of HR/VHR imagery. If differences in geometry are in the range of pixel size (20 m), they will probable be levelled out in case of the 10 m x 10 m grid.
- Short term changes between the date of HR imagery and reference imagery. The validating expert can handle this in most of the cases, based on pure logic.

- Statistical uncertainty. In case of 50% sealing the statistical error of the estimation is ±5%, while in case of 80% sealing the error is around ±4%[11].

3.5 "CALIBRATION" OF SSE DATA

SSE data have been derived by analysing high-resolution satellite images. No atmospheric or topographic correction has been done by Service Providers (SPs). An obvious checking of the data is to compare SSE with a more precise estimation of soil sealing, based on VHR imagery. This process could be called calibration.

Google Earth (GE) was used to provide reference material for the calibration exercise. In case GE would not have provided sufficient amount of sampling, the idea was to consult the EIONET to provide orthophoto samples for selected locations. However, to ask large number of countries to provide relevant VHR information in acceptable time would have been a challenge. In a later phase of the work, looking at the sufficient coverage (Fig. 10) and quality of GE imagery the idea of requesting contribution from the EIONET was dropped.

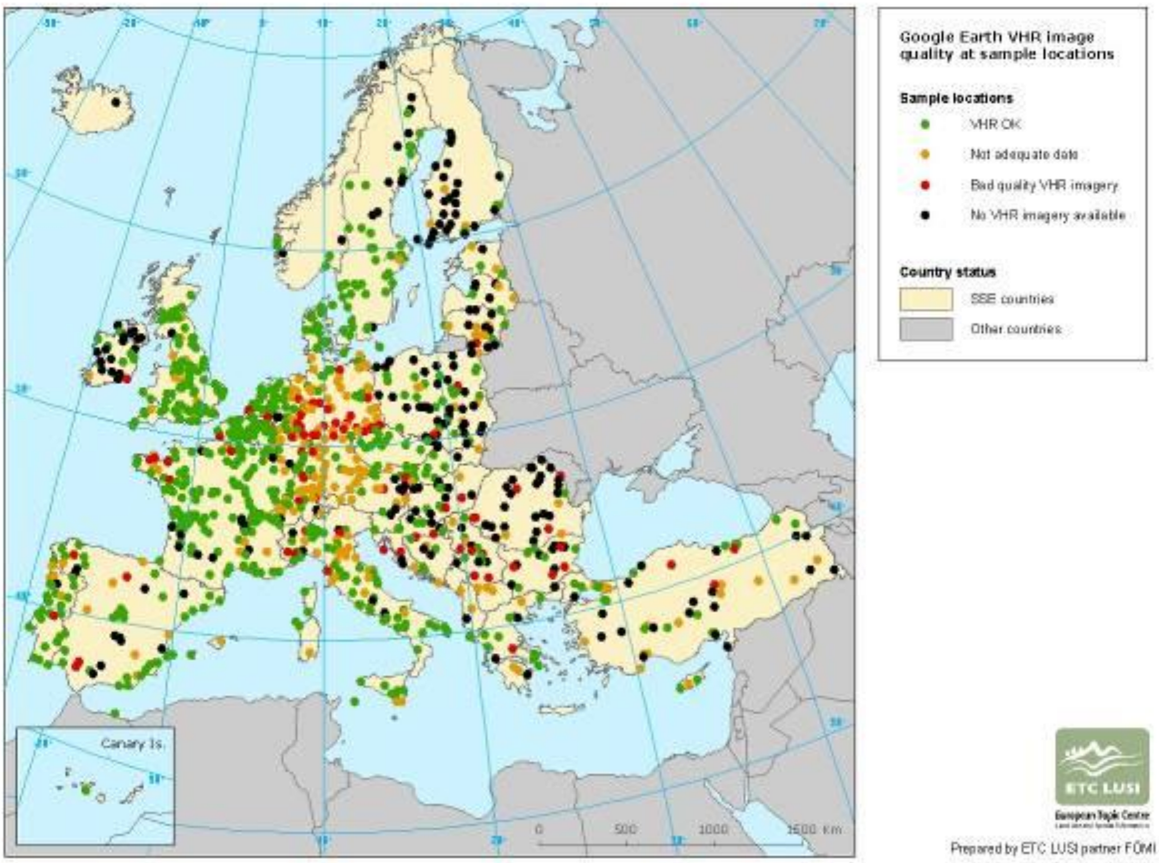

Figure 10 Distribution of sampling points used for the calibration exercise. Colours mean the quality of Google Earth imagery - green: good hi-res image (601 samples), yellow: bad date (201); red: bad quality (80), black: missing hi-res image (224).

In the 1st experiment 1106 random samples were selected for the calibration exercise from the "sealed" cells (sealing level > 0; covering 6.5% of Europe). If no GE/VHR imagery was available, the point was excluded from the analysis. If VHR reference imagery existed, its acquisition date was recorded. It was noticed, if the geometry of the

[11] Considering binomial distribution. Error values correspond to about 1 sigma.

reference image was questionable. Based on visual inspection of each sample, 601 samples proved to have good quality VHR/GE image, taken between 2004-2008. Photointerpretation (i.e. counting of sealed percentage, see Ch.3.4, Fig.9) on these samples provided the reference data to be compared with SSE data.

Differences of SSE and reference sealing level values have been calculated and analysed statistically as presented in Ch.4.1.

Commission error for built-up areas has been computed by means of the relevant subset of these samples. Thresholds were set to SSE data for this purpose. Due to the special sampling design (samples selected from the sealed layer) omission error was not possible to estimate. Results are presented in Ch.4.2.

3.6 SAMPLING WITHIN THE CLC ARTIFICIAL SURFACES CLASS

The aim of the 2nd sampling design was to provide an estimation of the full error matrix inside the CLC Artificial surfaces layer. Within this mask the ratio of non-sealed / sealed (and the non built-up / built-up) areas are favourable, and we could get a representative number of samples to calculate omission errors as well.

Because we have already had samples for sealed cells within the CLC Artificial surfaces from the first sampling experiment, we decided to sample only non-sealed cells within the CLC Artificial surfaces mask and to combine intelligently the two sets of samples:

- 274 of 601 valuable "sealed" samples from the first sampling fell into artificial areas;
- We have found 294 valuable samples out of 516 samples selected within the CLC Artificial surfaces "non-sealed" cells.

Altogether 568 samples were available within the CLC Artificial surfaces layer. The two sample subsets were combined following a proper normalisation (see details in Annex 1).

4 RESULTS

4.1 RESULTS OF THE "CALIBRATION" EXPERIMENT

In an ideal situation the points in a cross-plot of SSE and reference sealing data would distribute along a straight line from 0,0 to 100,100 coordinates, crossing the zero point at the origin of the coordinate system. However, due to the complexity of the real world and the limited capabilities of 20 m resolution EO data and that of the processing methods deviations are expected from the theoretical situation.

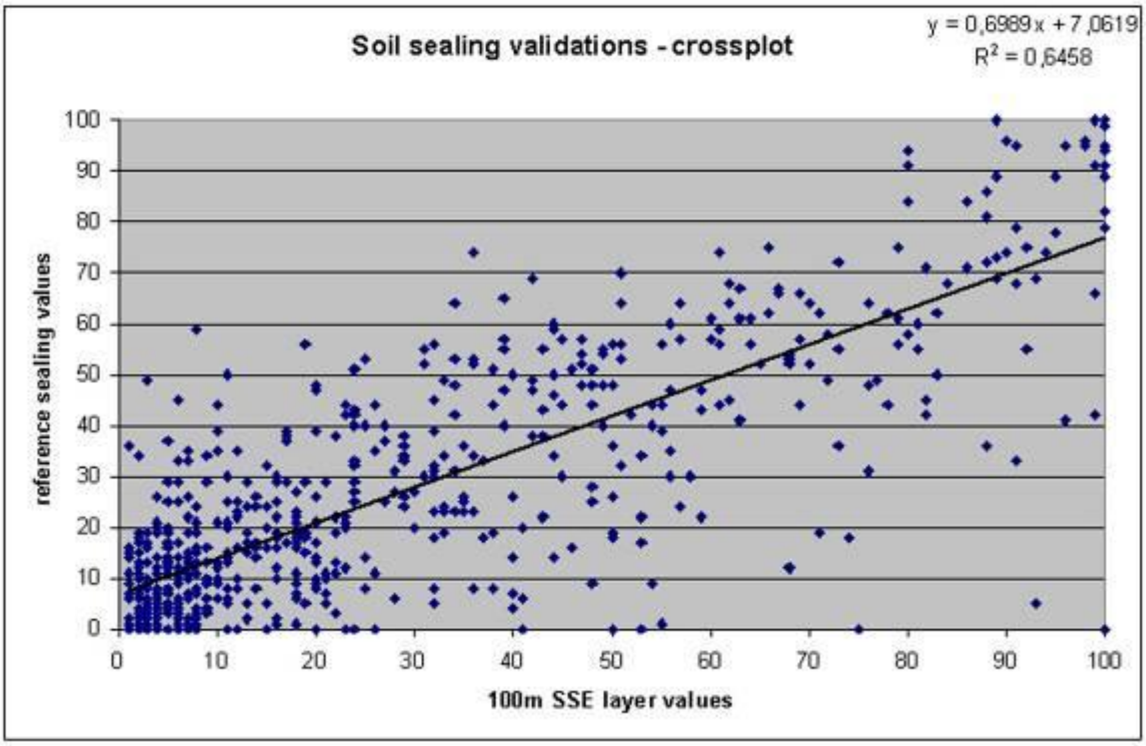

Figure 11 Cross-plot of SSE and reference sealing values

As shown on Fig. 11, there are significant deviations from the trend-line calculated for the crossplot, and the trend-line itself is less steep, than it would be in the theoretical case. On the bottom of the plot we see many different SSE values with a measured reference of zero. At the right side of the plot however we have different reference values measured for an SSE value of 100 (saturation).

By drawing the histogram of the differences of SSE and reference sealing values we have received an almost symmetric distribution around a value close to zero (Fig. 12):

Mean difference = (1,8 ±16,8) sealing levels

The mean difference is close to zero, meaning that neither systematic overestimation nor systematic underestimation occurs. The large standard deviation however shows the limitations of this technology.

Some of the possible reasons to explain the large spread seen on Fig. 12:

- Heterogeneity caused by IMAGE2006 (different EO sensors, large number of scenes, variable atmospheric conditions, topographic effects, etc)
- Heterogeneity of processing methodology (see Ch. 4.1.2)
- Seasonal differences between IMAGE2006 and GE imagery.

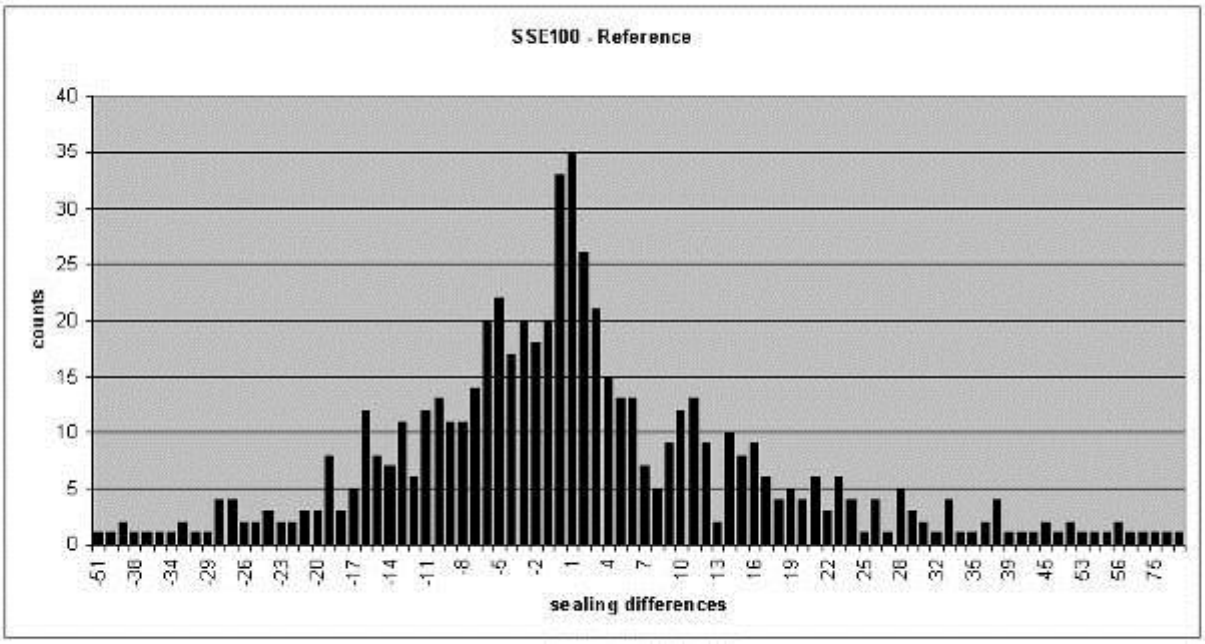

Figure 12 Distribution of the differences between SSE and reference sealing

4.1.1 Effect of Service Provider

In order to investigate the reasons of the large spread in differences between SSE values and observed sealing values, SSE data were separated according to responsible SPs. Six SPs worked on the project and the work was distributed mostly on a country base (Fig. 13). Regions are named after the corresponding SP. In case of Turkey, which was shared between two SPs and the cut line was not available the area is named after the two SPs: Planetek/Geoville. Mean and standard deviation of the differences between SSE and reference data have been computed (Table 5 and Fig. 14.) It was expected that different SPs would have similar distribution of differences between SSE and reference sealing.

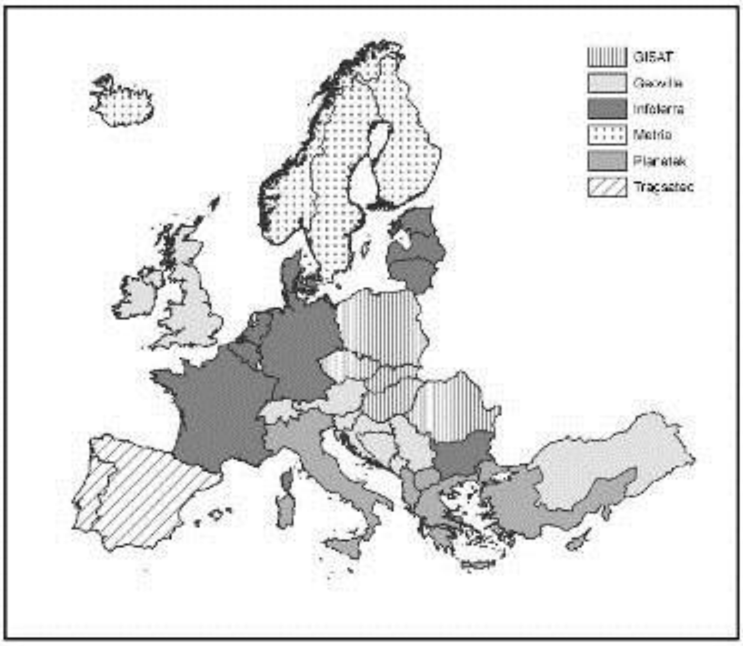

Figure 13 Geographic distribution of the work between Service Providers [10]

Table 5 Mean and standard deviation of differences between SSE and reference sealing by Service Providers

SP	Count	Minimum	Maximum	Mean	SD
GISAT	51	-16	75	11.73	20.58
Geoville	120	-30	53	-1.68	13.26
Infoterra	241	-38	100	3.20	16.40
Metria	42	-27	28	-4.57	11.54
Planetek	70	-37	56	0.29	17.18
Planetek/Geoville	18	-32	41	9.00	18.84
Tragsatec	59	-51	50	-0.98	19.46
ALL	**601**	**-51**	**100**	**1.82**	**16.84**

Table 5 and Fig. 14 show that two SPs (GISAT and Planetek/Geoville) have a strikingly different mean difference between SSE and estimated soil sealing compared to other SP's mean. These values are coupled with large standard deviation. Although the reason of difference is unknown, it will definitely increase the standard deviation (spread) of the total distribution. Lowest standard deviation was obtained for METRIA and Geoville data.

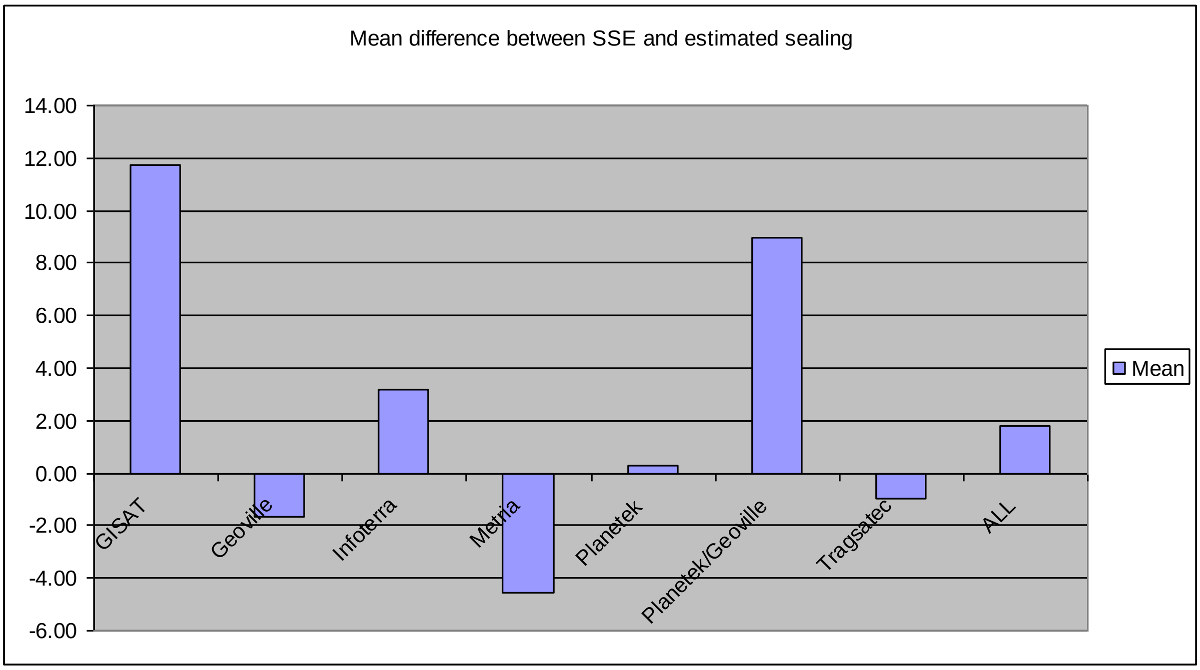

Figure 14 Mean differences between SSE and the reference sealing values by Service Providers

4.1.2 Effect of satellite image quality

In order to test the effect of image quality on SSE data quality, the area was cut into two parts based on mitigation data (location where the quality of satellite imagery is not optimal for deriving SSE – see Fig.1.):

- Area with optimal IMAGE2006 coverage;
- Area with non-optimal IMAGE2006 coverage.

Mean differences between SSE and reference sealing level were computed for these two strata (Table 6). It was expected that the two distributions would differ, as better estimation was expected from optimal IMAGE2006 coverage.

Table 6 Effect of satellite image data quality

EO image quality	Number of samples	%	Mean	Standard deviation
Optimal	354	59%	3,2	16,0
Non-optimal	247	41%	-0,1	17,8
Sum	**601**	**100%**	**1,8**	**16,8**

Figures show that no significant difference was found for the SSE validation results in these two strata regarding EO data quality.

4.2 ACCURACY ASSESSMENT FOR THE BUILT-UP CLASS

The accuracy of built-up / non-built-up separation was measured in two experiments:

- With samples drawn from the sealed area (only the commission error could be estimated, see Ch 3.5).
- Using samples derived from the CLC Artificial surfaces class the full error matrix could be derived (omission error, commission error and overall accuracy, see Ch. 3.6).

In both cases two different specifications of built-up were checked:

- Builtup80, which is the Service Specification, meaning that 80% SSE is the limit of the built-up.
- Builtup30, meaning that 30% SSE is the limit of the built-up.

Both the "look and feel" comparison (see Figure 7-8) and the calibration experiment showed that lots of built-up areas are characterised by lower than 80% SSE. FAO [11] and CORINE Land Cover [13] classification also puts the threshold of discontinuous built-up class to 30% sealing level.

4.2.1 Sealed area stratum

Tables 7 and 8 show accuracies obtained with samples drawn from the sealed area stratum. Due to this sampling strategy, omission errors could not be derived. Table 7 includes and evaluates the accuracy figures. Confidence values have been computed as standard deviations assuming a binomial distribution.

According to the error matrix, the commission error of the class is very high: 53,7% ± 6,8%. This means that more than half (53,7%) of the samples coded as built-up (>80% sealing) by the SPs is not built-up in reality (i.e. they have less than 80% sealing).

The very high overall accuracy value (in contrast to the high commission error) seen in Table 7 can be explained by the dominance of the class representing sealing level between 1-79 (i.e. non built-up).

Table 7 Accuracy matrix for builtup80 class derived from European SSE data (omission error was not possible to estimate due to the special sampling strategy)

		Reference				
		0-79	**80-100**	**Sum**	***User's accuracy***	***Commission error***
Classification	**0**	No data	No data	**-**	-	-
	1-79	547	0	**547**	*100,0%*	*0,0%±0,0%*
	80-100	29	25	**54**	*46,3%*	*53,7%±6,8%*
	Sum	**576**	**25**	**601**		
	Producer's accuracy	-	-			
	Omission error	-	-			

Overall accuracy: 95,2%±*0,9%*

Table 8 Accuracy matrix for builtup30 class derived from European SSE data (omission error was not possible to estimate due to the special sampling strategy)

		Reference				
		0-29	**30-100**	**Sum**	***User's accuracy***	***Commission error***
Classification	**0**	No data	No data	**-**	-	-
	1-29	327	46	**373**	*87,7%*	*12,3%±1,7%*
	30-100	48	180	**228**	*78,9%*	*21,1%±2,7%*
	Sum	**375**	**226**	**601**		
	Producer's accuracy	-	-			
	Omission error	-	-			

Overall accuracy: 84,4%±*1,5%*

The matrix in Table 8 is more balanced, the overall accuracy value being more in harmony with the commission errors. Commission error figure shows that only about 1/5 (21,2%) of the built-up samples is not correctly coded by SPs.

Table 9 compares results obtained for builtup80 and builtup30.

Table 9 Comparison of accuracy figures obtained for sealed area stratum

	Omission error, built-up	Commission error, built-up	Overall accuracy	Comment
Builtup30	-	21,1%	84,4%	Overall accuracy is almost fulfilled; commission error is close to 15%.
Builtup80	-	**53,7%**	**95,2%**	Overall accuracy is fulfilled; commission error is much higher than 15%.
Accuracy target	15%	15%	85%	

In case of builtup80 commission error is rather high. At the same time overall accuracy is also high, which is the result of the dominance of the non built-up (1-79) class among the samples.

If the built-up limit is set to 30% sealing (builtup30), commission error becomes significantly lower with a modest decrease in overall accuracy. These figures are more in harmony with each other as the two classes (built-up / non built-up) are represented by approximately the same number of samples.

4.2.2 CLC Artificial surfaces stratum

Tables 10 and 11 show accuracies obtained with samples drawn from the CLC Artificial layer. The full, normalized[12] error matrix (omission error, commission error and overall accuracy) has been computed. Results are valid in the CLC Artificial surfaces stratum only. Table 12 compares and evaluates the accuracy figures.

Table 10 Accuracy matrix for builtup80 class derived from SSE data (valid for CLC Artificial classes only)

		Reference				
		0-79	**80-100**	**Sum**	***User's accuracy***	***Commission error***
Classification	**0-79**	301,7	0,3	**302,0**	*99,9%*	*0,1%±0,2%*
	80-100	29	23	**52**	*44,2%*	*55,8%±6,9%*
	Sum	**331,7**	**23,3**	**354**		
	Producer's accuracy	*91,2%*	*98,8%*			
	Omission error	*8,8%±1,6%*	*1,2%±2,3%*			

Overall accuracy: **91,7%±*1,5%***

The very high overall accuracy value (in contrast to the high commission error) seen in Table 10 is explained by the dominance of the non built-up class (sealing levels between 0-79) among the samples.

Table 11 Accuracy matrix for builtup30 class derived from SSE data (valid for CLC Artificial classes only)

		Reference				
		0-29	**30-100**	**Sum**	***User's accuracy***	***Commission error***
Classification	**0-29**	144,6	27,4	**172,0**	*84,0%*	*15,9%±2,8%*
	30-100	23	159	**182**	*87,4%*	*12,6%±2,5%*
	Sum	**168,6**	**186,4**	**354**		
	Producer's accuracy	*86,3%*	*85,3%*			
	Omission error	*13,7%±2,7%*	*14,7%±2,6%*			

Overall accuracy: **85,7%±*1,9%***

[12] Because of the stratified sampling design the original error matrix had to be normalized, this explains the non-integer values in the matrix. See further explanation in Annex 1.

The matrix seen in Table 11 is more balanced, the overall accuracy value being more in harmony with the commission errors. This is due to balance between built-up and non built-up class in the stratum (number of samples is similar).

Table 12 Comparison of accuracy figures obtained for the CLC Artificial surfaces stratum

	Omission error, built-up	Commission error, built-up	Overall accuracy	Comment
Builtup30	**14,7%**	**12,6%**	**85,7%**	**All accuracy criteria are fulfilled.**
Builtup80	**1.2%**	**55,8%**	**91,7%**	Overall accuracy is fulfilled; commission error is much higher than 15%.
Accuracy target	15%	15%	85%	

In case of builtup80 commission error is rather high, but overall accuracy is also high (like in 4.2.1). If the built-up limit is set to 30% sealing (builtup30), commission error is below the 15% limit, and overall accuracy still remains above the required 85%.

4.3 Summary of accuracy assessment

Validation experiments described under Ch. 4.2.1 and Ch 4.2.2 show that

- Concerning service specification (i.e. builtup80 class) **85% overall accuracy has been achieved** according to both tests.
- **Significant (>50%) commission error** is shown by both experiments in case of builtup80 class. This means that about half of non built-up samples were erroneously classified as built-up by SP.
- **Omission error is low**, under the specified limit (valid for the CLC Artificial stratum only). This means that only few built-up samples were classified as non-built-up by SP.
- Setting the threshold of built-up at 30% sealing in SSE data provides more reliable estimation of built-up area than setting the threshold at 80%. In this case overall accuracy is above 85%, omission and commission errors are below 15%.

5 CONCLUSIONS AND RECOMMENDATIONS

Soil Sealing is the first high resolution Land Monitoring layer of the EEA with European coverage. Its main use is the characterisation of the human impact on the environment.

Multi-sensor and bi-temporal, orthorectified satellite imagery (IMAGE2006) was used to derive soil sealing data covering 38 countries of Europe. Production of the soil sealing database was implemented in two phases:

- Initial Soil Sealing (ISS), and
- Soil Sealing Enhancement data (SSE), which is the improvement of the ISS database on the basis of evaluation of ISS data by some Member States.

The main deliverable was a raster dataset of continuous degree of soil sealing ranging from 0 - 100% in full spatial resolution (20 m x 20 m) with the associated metadata. A derived product, a raster dataset of continuous degree of soil sealing ranging from 0 - 100% in aggregated spatial resolution (100 m x 100 m) in European projection was validated.

According to the descriptive statistics, 6.5 % of the European territory is covered by 1 ha cells including sealing (any percent between 1-100), and the total sealed surface is 1,8 %. Built-up areas cover 0.5 % of Europe (if the sealing threshold is 80%) or 2.5% (with 30% threshold).

The main purpose of this report is to introduce the method and results of validation of SSE data or more precisely the "built-up" class. Very High Resolution Google Earth imagery (VHR/GE) was used as independent, higher resolution dataset. Reference sealing percentages were obtained by using a 10x10 grid positioned around the sampling points. Checking classification accuracy was understood as estimating omission error, commission error and overall accuracy.

Due to the small size of the sealed layer (and consequently of the built-up class) the full error matrix could not be derived for the whole dataset with affordable efforts, therefore two experiments were devoted to estimate the accuracy of the built-up class:

1. In the 1st experiment only sealed cells were sampled;
2. In the 2nd experiment sampling was restricted to the CLC Artificial surfaces layer, where most of the sealed areas are found.

Using randomly selected samples, the SSE values and reference sealing values were compared. The mean difference between the two was close to zero i.e. neither systematic overestimation nor systematic underestimation occurred. The large standard deviation however shows the limitations of this technology. A number of factors might be responsible for the large standard deviation. Two of these were examined, one being the effect of data producer, second being the effect of source data quality. Differences between results of Service Providers were discovered. For a surprise, EO data quality as shown by the mitigation file did not prove to be a significant factor of this variability.

Main findings of the validation of the built-up class (builtup80):

- According to service specification (builtup80 class) **85% overall accuracy has been achieved** according to both experiments.
- **However, significant (>50%) commission error** was shown by both experiments. This means that lots of non built-up samples were erroneously classified as built-up by SP.
- **Calculated omission error is low**, much below the limit. This means that only few built-up samples were classified as non-built-up by SP (few built-up left out).

An important finding of both the "look and feel" comparison (see Figure 7-8) and the statistical analysis is that **setting the threshold of built-up class at 30% SSE provides more reliable estimation of built-up** area than setting the threshold at 80%. For the builtup30 class the calculated **overall accuracy is above 85%, and both omission and commission errors are below 15%**.

Main conclusions of the study are as follows:

- If overall accuracy is the only criteria, SSE has fulfilled the requirements (>85%).
- Concerning omission error it is under the 15% limit (valid only in the CLC Artificial layer).
- Concerning commission error this strongly exceeds the 15% threshold in case of builtup80 class.
- If built-up is defined as >30% SSE, all three accuracy criteria are fulfilled. Visual inspection also supports this definition of built-up. Builtup80 refers to just the most densely built-up areas according to FAO as well as CLC land cover classifications, while the builtup30 includes the discontinuously built-up areas as well according to these classification systems.
- In any further production of Soil Sealing datasets it is recommended to use either a 30% threshold or a multi-level threshold.

6 REFERENCES

[1] C. Homer et al.: Development of a 2001 National Land-Cover Database for the United States. Photogrammetric Engineering and Remote Sensing Viol 70, No. 7. pp. 828-840, 2004.

[2] R. Müller et al.: Automatic production of a European orthoimage coverage within the GMES Land Fast Track Service using SPOT 4/5 and IRS-P6 LISS III data; ISPRS Commission IV, WG3

[3] Tender Specification. GMES fast track service on land monitoring. High resolution core land cover data for built-up areas, including degree of soil sealing, 2006 EEA/IDS/07/001.

[4] M. Tinz, H. Kahabhka and H. Schrader: Project Management Plan, Issue 2, ITD0490-EEA-FTS-PMP Date Issued: 10.08.2007.

[5] Tender Specification. GMES Fast Track Service Precursor on land monitoring. High resolution core land cover data for built-up areas, including degree of soil sealing. Data enhancement in support of Member States requirements. EEA/SES/09/003.

[6] [http://eusoils.jrc.ec.europa.eu/library/themes/Sealing/].

[7] Wikipedia

[8] Technical conformity checking of GMES FTS Soil Sealing Enhancement data, ETC-LUSI Report, FÖMI, Project manager: Stefan Jensen, January 2010.

[9] EEA-FTSP-Sealing-Enhancement_CountryDeliveryReport-AL_F1v0.pdf . Delivery report for 38 countries.

[10] EEA-FTSP-Sealing-Enhancement_FinalReport, Issue 1.0 (22.12.2009)

[11] Land Cover Classification System. Classification concepts and User Manual. FAO and UN Rome, 2005.

[12] http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1

[13] Y. Heymann, C. Steenmans, G. Croissille., M. Bossard, CORINE Land Cover. Technical Guide. EUR12585 Luxembourg, Office for Official Publications of the European Communities, 1994

ANNEX 1: NORMALIZATION OF THE ERROR MATRIX

Ideally, a Simple Random Sampling design should be applied for thematic validation purposes. Sum of the corresponding rows/columns of an error matrix drawn based on a Simple Random Sampling well reflects the areal proportions of the classes. However, this way only a very few samples hit the small classes, thus statistics, which are calculated based on a few samples are not representative. Increasing the total number of samples so much, as it would be required by the small classes would yield unnecessarily large number of samples for large classes.

To increase the effectiveness of the sampling design stratification can be applied. Often the solution is to put the same number of samples into all classes. However, if we mechanically draw an error matrix from the results based on this kind of stratified sampling, some figures of the statistics (omission error) will have strongly biased false values. The purpose of the normalization described here is to transform the matrix to be similar to as it would be without stratification, in order to be able to calculate unbiased omission error and overall accuracy from the matrix.

As described in Ch 3.6 the aim of the second sampling design was to provide an estimation of the full error matrix for the artificial surface as defined by CLC Artificial surfaces class. Within this mask the ratio of non-sealed / sealed (and the non built-up / built-up) areas is favourable, consequently we can get a representative number of samples to calculate omission errors as well.

Because we already had validated samples for sealed cells within artificial surfaces from the first sampling experiment, we decided to sample only non-sealed cells within the CLC Artificial surfaces mask and to combine intelligently the two sets of samples:

- 274 of 601 valuable "sealed" samples from the first sampling fell into artificial areas;
- We found 294 valuable samples out of 516 initial samples selected within CLC Artificial surfaces "non-sealed" cells.

As seen above, altogether 568 valuable samples were found within Artificial surfaces. We could create a matrix based on this Stratified Random Sampling as seen in Table 1.1.

		Reference							
		0	**1-29**	**30-49**	**50-79**	**80-100**	**Sum**	***User's accuracy***	***Commission error***
Classification	**0**	123	162	6	2	1	**294**	*41,8%*	*58,2%*
	1-29	5	62	20	5		**92**	*67,4%*	*32,6%*
	30-49	1	17	21	23		**62**	*33,9%*	*66,1%*
	50-79		3	24	41		**68**	*60,3%*	*39,7%*
	80-100	1	1	6	21	23	**52**	*44,2%*	*55,8%*
	Sum	**130**	**245**	**77**	**92**	**24**	**568**		
	Producer's accuracy	*94,6%*	*25,3%*	*27,3%*	*44,6%*	*95,8%*			
	Omission error	*5,4%*	*74,7%*	*72,7%*	*55,4%*	*4,2%*			

Overall accuracy: 47,5%

Table 1.1 Error matrix for artificial areas without normalization. Red values indicate bias as a consequence of the stratified sampling.

The matrix shown in Table 1.1 is biased, sum of rows does not reflect the real proportions. We can calculate valid commission errors form the rows of the matrix, but the omission errors and overall accuracy values are false. To yield unbiased values we have to normalize the matrix.

Calculated from the CLC-SSE comparison (Table 4) the ratio of non-sealed / sealed areas within CLC Artificial surfaces area is 22,6% / 77,4%. In order to have well-proportioned number of samples we reduced the sum in the first row to 80 samples and calculated values in the 1st row respectively. As a consequence of the normalization, we have got non-integer values in the 1st row (and the sum). The normalized matrix is shown on Table 1.2.

		Reference							
		0	1-29	30-49	50-79	80-100	Sum	***User's accuracy***	***Commission error***
Classification	**0**	**33,5**	44,1	1,6	0,5	0,3	**80,0**	*41,8%*	*58,2%*
	1-29	5	**62**	20	5		**92**	*67,4%*	*32,6%*
	30-49	1	17	**21**	23		**62**	*33,9%*	*66,1%*
	50-79		3	24	**41**		**68**	*60,3%*	*39,7%*
	80-100	1	1	6	21	**23**	**52**	*44,2%*	*55,8%*
	Sum	**40,5**	**127,1**	**72,6**	**90,5**	**23,3**	**354**		
	Producer's accuracy	*82,7%*	*48,8%*	*28,9%*	*45,3%*	*98,8%*			
	Omission error	*17,3%*	*51,2%*	*71,1%*	*54,7%*	*1,2%*			

Overall accuracy: 51,0%[13]

Table 1.2 Normalized error matrix for artificial areas. Blue values indicate changed values as a consequence of normalization.

Results for the two built-up classes (builtup80 and builtup30) can be found in Ch.4.2.

[13] Note that overall accuracy values calculated for both error matrices containing 5 thematic classes are significantly lower than for a 2 class matrix shown in Ch.4.2. In contrary to overall accuracy values, the corresponding commission and omission error values are the same.