Accuracy of Built-up Area Mapping in Europe at Varying Scales and Thresholds

Pavol Hurbanek
Peter M. Atkinson
Jeganathan Chockalingam
University of Southampton
Southampton, United Kingdom
pavalhurbanek@gmail.com
P.M.Atkinson@soton.ac.uk
j.chockalingam@soton.ac.uk

Robert Pazur
Slovak Academy of Sciences
Bratislava, Slovakia
g美貌@gmail.com

Konstantin Rosina
Comenius University
Bratislava, Slovakia
konstantin.rosina@gmail.com

Abstract — The paper provides an accuracy assessment of the Soil Sealing Layer (SSL), a map of impervious surfaces in most of Europe. It focuses on the extent of mapped built-up area and the accuracy of built-up area mapping as a function of the soil sealing threshold applied, the spatial resolution used, and the spatial configuration of built-up areas mapped. The results from the stratified random sampling comparison of SSL with aerial orthophotos derived for Slovakia are compared with those for other countries and complemented by "complete coverage" comparisons in three 6x6 km study areas in Slovakia.

Keywords: soil sealing; impervious surfaces; land cover; spatial resolution; threshold; Europe; Slovakia; population downscaling

I. INTRODUCTION

One of the main purposes of the Soil Sealing Layer (SSL) 2006 (EEA 2010) developed within the Global Monitoring for the Environment and Security (GMES) programme is to serve as a source of high spatial resolution land cover data for disaggregation of socioeconomic statistics (GSE Land 2008). An example of such disaggregation is the model-based estimation of population density in 100 m grid cells (Gallego 2009a) using Corine Land Cover (CLC) data (EEA 2009a) and census population data for communes (i.e. the smallest spatial units used in Europe-wide applications). However, because of the rather coarse 25 ha minimum mapping unit of CLC, there is no CLC Category 1 "Artificial Surfaces" reported in 29% of the approximately 114,000 communes in Europe in 2000 (Gallego 2009b). Therefore, an additional data source is required to downscale population more accurately in these communes. Because, for example, in Slovakia in 2006 (2,891 communes) there are about 15% communes with no "Artificial Surfaces" grid cells in CLC at 100 m, but less than 1% communes with no greater-than-zero soil sealing pixels at 20 m, SSL seems to fit this role of additional data source.

II. SOIL SEALING LAYER

Similarly to CLC2006, SSL2006 is also derived from IMAGET2006. It is based on satellite sensor imagery from SPOT 4/5 HRVIR (10/20 m ground sampling distance - GSU) and IRS-P6 LISS III-3 (23 m GSU) from 2006 (+/- 1 year) re-sampled to 20 m ground resolution using cubic convolution interpolation with geometrical correction towards national projection systems (EEA 2009b).

The soil sealing 0-100% integer values in SSL are available at 20 m spatial resolution as an intermediate product and at 100 m spatial resolution (averaged over 5x5 20 m pixels) as a final product. While the former is provided "as it is", the latter is validated with respect to the >85% thematic accuracy tender specification requirement imposed on its binary version: built-up §80% soil sealing, non-built-up <80% soil sealing. It can be assumed that the decision to validate the 100 m rather than the 20 m resolution was taken because the geometric accuracy of IMAGE2006 is an RMSE of about 10 m in each direction with respect to IMAGE2000 and USGS Land Cover reference datasets (Muller et al. 2009), and possibly also because the cubic convolution had been used in pre-processing (mentioned above).

The choice of 80% threshold in the tender specifications is backed by work of Di Gregorio (2005). However, it seems to be quite widely criticised by users, including the experts from individual European countries' National Reference Centres on Land Cover (EEA2009c). The reason is probably that this threshold is applied in a binary classification, but in the original document it is just the first out of three different thresholds (80, 50, and 20%) of a fourfold classification — i.e. built-up, built-up/non-built-up, non-built-up/built-up, and non-built-up — where the sequence in the mixed class names represents the dominance.

III. ACCURACY ASSESSMENT

The objective of this paper was to provide an assessment of the accuracy of SSL before its further use in modelling, especially in population downscaling. Specifically, it focuses on:

- The extent of mapped built-up area as a function of the soil sealing threshold and the spatial resolution in Europe and Slovakia: Fig. 1 (a1-b6)
The accuracy of built-up area mapping at 80% soil sealing threshold and 100 m spatial resolution in Slovakia and other European countries: Table 1, Fig. 2

The extent of mapped built-up area and the accuracy of built-up area mapping as a function of the soil sealing threshold applied, the spatial resolution used, and the spatial configuration of built-up areas mapped in Dubnica (DB), Poprad (PP), and Myjava (MY): Fig. 1 (c1–g6), Fig. 3

For the purpose of the second objective, random stratified sampling design suggested by Maucha and Büttnner (2008) was applied to the 100 m grid cells of SSL in Slovakia with 500 samples drawn from Strata 1 (≥80% soil sealing in SSL) and 2,000 from Strata 2 (<80% soil sealing in SSL). A regular square grid of 100 points 10 m apart was overlaid on top of each corresponding sample plot in the reference dataset (Geodis Slovakia and EuroSens 2003 and Google Earth historical imagery) and the percentage of points overlapping with impervious surfaces in each plot was calculated. This revealed that 388 Strata 1 samples had soil sealing <80% in the reference dataset and none of the Strata 2 samples had soil sealing ≥80% in the reference dataset. The respective commission and omission errors for Slovakia were derived using the methodology developed by Maucha and Büttnner (2008) and compared with the results of quantitative assessments in other European countries (Table 1, Fig. 2) available from EEA (2009c).

Figure 1. Frequencies (%) of 12 soil sealing categories as a function of spatial resolution: (1) 0%, (2) 1-9%, (3) 10-19%, (4) 20-29%, (5) 30-39%, (6) 40-49%, (7) 50-59%, (8) 60-69%, (9) 70-79%, (10) 80-89%, (11) 90-99%, (12) 100%. Note that y axis has logarithmic scale and depicts the following tick marks: 0, 1, 10, 100%.

Figure 2. Frequencies (%) of 12 soil sealing categories at 100 m spatial resolution in the samples from Strata 1 (≥80% soil sealing in SSL) and Strata 2 (<80% soil sealing in SSL) in Slovakia and Austria. Labels of the 12 bars on the x axis (left to right): (1) 0%, (2) 1-9%, (3) 10-19%, (4) 20-29%, (5) 30-39%, (6) 40-49%, (7) 50-59%, (8) 60-69%, (9) 70-79%, (10) 80-89%, (11) 90-99%, (12) 100%. Tick marks of the logarithmic scale on the y axis (down to up): 0, 1, 10, 100%.

For the purpose of the third objective, three different 6x6 km study areas in Slovakia were selected: Dubnica (DB) with several smaller towns spread over the area, Poprad (PP) with one relatively compact city of more than 50,000 inhabitants, and Myjava (MY) with numerous but very small dispersed
rural settlements. Using the same reference datasets, a regular square grid of sample points 5 m apart was overlaid on top of each study area, resulting in 16 points in each 20 m pixel. This allowed for deriving of maximum 17 (65) different possible reference soil sealing values at 20 m (40 m) spatial resolution, and more than 100 values at all other coarser spatial resolutions analysed.

The frequencies and spatial configurations of these values compared with frequencies and spatial configurations of the SSL values in the same study areas are shown in Fig. 1 (e1-f6) and Fig. 3 (e1-f12).

Using a threshold of 60% or larger at a spatial resolution of 20 m and 40% or larger (except 80%) at a spatial resolution of 100 m led to larger commission errors in DB than in PP, i.e. the study area with the most compact / clustered and least dispersed / spread settlement pattern at all scales of the three study areas. When comparing omission errors in DB and PP at almost all combinations of thresholds 1% to 90% and spatial resolutions 20 m to 300 m, they were again larger in DB than in PP (see Fig. 3 a1-d6). This is in spite of that the underlying IMAGE2006 had failed one of the input data requirements (both acquisition dates within the country specific vegetation period) in PP, but none in DB.

![Figure 3](image.png)

*Figure 3.* The extent and spatial configuration of mapped soil sealing values (e1-f4, e4-f10) and built-up areas (a1-d2, a4-d5) and the errors of soil sealing mapping (e3-f6, e6-f12; with light gray to white underestimations and dark grey to black overestimations) and built-up area mapping (a3-d3, a6-d6; with white omission error pixels and black commission error pixels) at different soil sealing thresholds and spatial resolutions in Dubnica (DB) and Poprad (PP), i.e. 6x6 km study areas in Slovakia.
TABLE I. COMMISSION AND OMISSION ERRORS IN SSL IN SOME OF 
THE EUROPEAN COUNTRIES (EEA 2009c)

<table>
<thead>
<tr>
<th>Country</th>
<th>B (%)</th>
<th>C (%)</th>
<th>D (%)</th>
<th>E (%)</th>
<th>F (%)</th>
<th>G (%)</th>
<th>H (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slovakia</td>
<td>500</td>
<td>388</td>
<td>74.52</td>
<td>80.64</td>
<td>77.58</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>0</td>
<td>0.00</td>
<td>18.31</td>
<td>9.15</td>
<td>8.55</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>496</td>
<td>123</td>
<td>21.81</td>
<td>28.20</td>
<td>25.01</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>3</td>
<td>12.24</td>
<td>79.56</td>
<td>45.90</td>
<td>93.67</td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>498</td>
<td>278</td>
<td>51.84</td>
<td>59.16</td>
<td>55.50</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1991</td>
<td>6</td>
<td>14.39</td>
<td>53.95</td>
<td>34.17</td>
<td>94.77</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>500</td>
<td>134</td>
<td>23.74</td>
<td>30.26</td>
<td>27.00</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>11</td>
<td>75.66</td>
<td>202.50</td>
<td>139.08</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>500</td>
<td>88</td>
<td>15.03</td>
<td>20.65</td>
<td>17.84</td>
<td>95.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>24</td>
<td>20.88</td>
<td>40.56</td>
<td>30.72</td>
<td>99.89</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>500</td>
<td>231</td>
<td>42.60</td>
<td>49.80</td>
<td>46.20</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>6</td>
<td>1.20</td>
<td>5.80</td>
<td>3.50</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

A: Country; B: Error (Commission error); C: All samples; D: Misclassified samples; 
E: 95% confidence interval lower bound (%); F: 95% confidence interval upper bound (%); 
G: Mean error (%); H: Probability of having more than 15% error (%)

The tendency observed in the comparison of DB and PP, i.e. that the more dispersed (less compact) is the settlement pattern, the more erroneous is its representation in SSL, was further supported by the case of MY. Although the commission and omission errors at 100 m spatial resolution and 80% threshold in this study area were practically 0%, there are only 30 non-zero 20 m pixels (with mean soil sealing value of 67%) in SSL, but 5,250 pixels (with mean soil sealing value of 24%) in the reference dataset in this study area (Fig. 1 g1-g6).

IV. CONCLUSION

This preliminary accuracy assessment of SSL suggests that it forces values of soil sealing out to the extremes (0% and 100%) by overestimating (usually but not only) medium and higher soil sealing values (e.g. Fig. 2 a3-b4) and underestimating (usually but not only) medium and smaller soil sealing values (e.g. Fig. 2 a1-b2). This potentially results in overestimation of share of impervious surfaces in areas with more compact settlement pattern (usually urban areas) and underrepresentation or complete omission of small and dispersed rural settlements.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 220832.

REFERENCES


specifications-eea-ses-09-003)


GSE Land (2008). Brochure InfoSheet impervious areas and sealing levels. GSE Land Information Services website, Communications, Promotion Material. GSE (GMES Service Element) Land project, 2 p. (http://www.gmes-gseland.info/com Prom/ImperviousAreas_Sep08.pdf)
