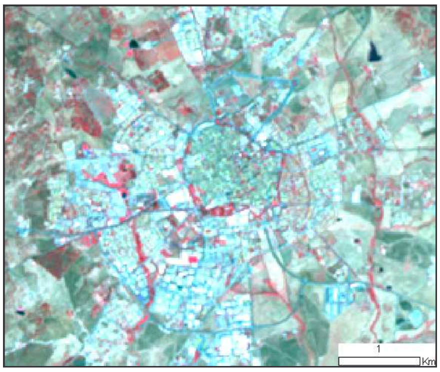
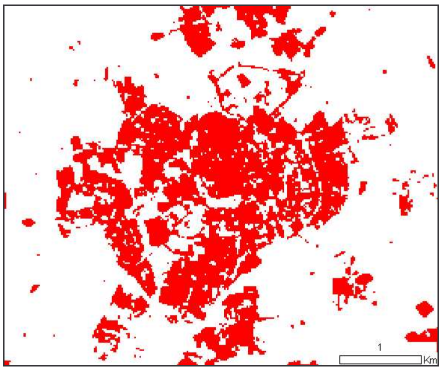
Portugal CORINE Land Cover 2006

Accuracy assessment of the High Resolution Built-up map for Continental Portugal





Mário Caetano, António Araújo, Vasco Nunes and Hugo Carrão

April, 2008

AGÊNCIA PORTUGUESA DO AMBIENTE
Ministério do Ambiente, do Ordenamento do Território e do Desenvolvimento Regional

INSTITUTO GEOGRÁFICO PORTUGUÊS

GRUPO DE DETECÇÃO REMOTA

Index

Table index

Figure index

1 Introduction

The Global Monitoring for Environment and Security (GMES) Fast Track Service Precursor on Land Monitoring is a joint initiative of the European Environmental Agency (EEA), European Space Agency (ESA) and European Commission (EC). The project consists on combining the CORINE Land Cover (CLC) update with the production of additional high-resolution data for built-up (soil sealing) and forest areas, with the objective of solving the shortcomings of a standard CLC update, which is not sufficient to meet the wide range of user needs (Maucha and Buttner, 2007).

In one of the project meetings, the Eionet workshop on quality control and validation of land cover data (Copenhagen, 12-13 November 2007), it was agreed that the National Reference Centres on Land Use and Spatial Information (NRC-LUSI) would support EEA in doing the validation of the high-resolution built-up (soil sealing) dataset. The Portuguese Geographic Institute (IGP), being the Portuguese NRC-LUSI, is responsible for the validation of the Portuguese product.

The present document describes the methodology to validate and determine the accuracy of the high-resolution soil sealing dataset for Continental Portugal, produced in the framework of the referred project. According to project requirements, the classification accuracy must be estimated per hectare (i.e. based on a 100 m x 100 m grid) for a soil sealing product of built-up and non built-up areas. Built-up areas are defined as areas where impervious surfaces account fro 80 to 100% of the total cover. Accuracy should be at least 85%, for the European product (Maucha and Buttner, 2007).

This document has five main chapters, including this introduction. The following chapter introduces the built up dataset that will be validated. The third main section presents the protocol that summarizes and explain step-by-step the proposed accuracy assessment strategy. The fourth section of this document presents and discusses the results of the validation. The last chapter presents the final considerations of the developed work.

2 High-resolution soil sealing map

A consortium of European service providers under contract with EEA produced the European high-resolution soil sealing land cover dataset. This map characterizes built-up areas, including degree of soil sealing, for the reference year 2006 (Sánchez and Kahabka, 2008).

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, 2005 *in* Maucha and Buttner, 2007).

According to the consortium definition, built-up areas are represented by a degree of soil sealing between 1 and 100%. These areas comprises pixels that are fully or partly covered by houses, roads, mines and quarries and any other facilities, including their auxiliary spaces, deliberately installed for the pursuit, of human activities. Built-up area does not include any fully vegetated pixels, even if they are closely related to these activities (such as city parks and gardens), or any other unvegetated non built-up opens spaces covered with bare soil, sand, glacier, bare rocks or water (Sánchez and Kahabka, 2008).

In the high-resolution soil sealing dataset, a per-pixel estimate of imperviousness (continuous variable from 0 to 100 percent) is provided as index for degree of soil sealing for the whole geographic coverage (Maucha and Buttner, 2007). The product accuracy will be estimated only for a binary product, i.e. soil sealing map of two classes: built-up (soil sealing degree equal or higher than 80%) and non built-up areas (soil sealing degree lower than 80%).

The high-resolution soil sealing map was produced in full spatial resolution, i.e. 20 m by 20 m, based on the supervised classification of the IMAGE2006 satellite data with following visual improvement of classification result and derivation of degree of soil sealing based on calibrated NDVI classification (Sánchez and Kahabka, 2008). IMAGE2006 corresponds to high quality orthorectified high-resolution satellite imagery (SPOT-4/5 and IRS LISS-3) obtained in two time windows selected by the countries for the years 2006+-1. These images were used both for the production of the high-resolution land cover datasets and for the CLC map update.

Table 1 summarizes the characteristics of the high-resolution soil sealing dataset for Continental Portugal.

Table 1 – Summarized characteristics of the high-resolution soil sealing map for Continental Portugal (Sánchez and Kahabka, 2008).

Data format	Raster
Spatial resolution	20 m
Coordinate system Projection Datum False E / False N Central Meridian Latitude of Origin Scale factor	 Transverse Mercator GRS80 0.00 / 0.00 -8°07'59.19'' 39°40'05.73'' 1.00
Raster coding	0 – Non built-up, water bodies inland 1-100 – sealing (imperviousness) values for built-up areas 254 – Unclassifiable areas (clouds, shadows, ets) 255 – No Data
Thematic overall accuracy (based on 100 m x 100 m grid) of built-up and non built-up areas	>85%

3 Methodology

The accuracy assessment of the high-resolution soil sealing map was made according to a procedure that involved the construction and analysis of contingencies tables, also called error or confusion matrixes, throughout a validation approach described in Figure 1.

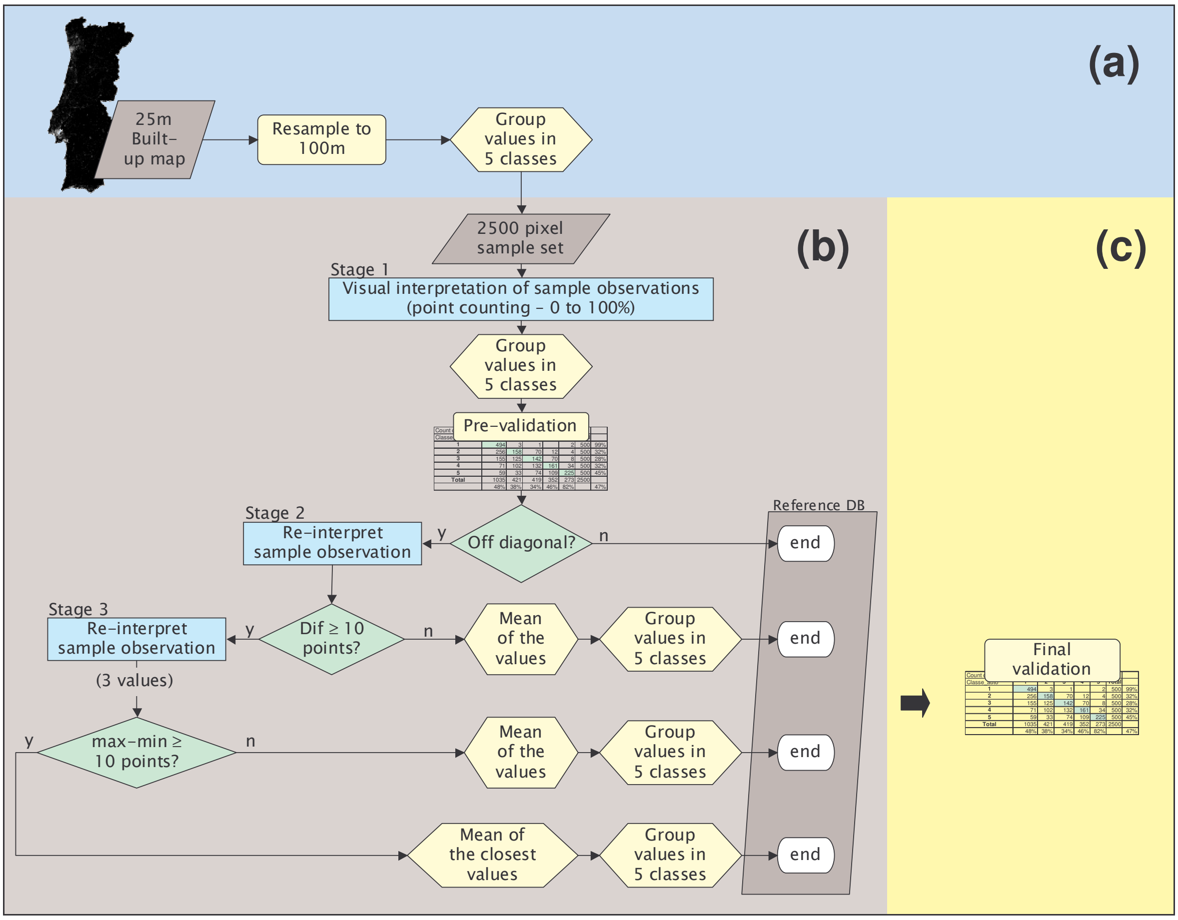


Figure 1 – Fluxogram explaining the developed validation methodology.

In brief, the methodology consisted in: (a) converting the original 20 m x 20 m soil sealing map into a 100 m x 100 m map, since according to project requirements classification accuracy must be calculated per hectare (i.e. based on a 100 m x 100 m grid); (b) developing a reference land cover database; and (c) comparing the reference land cover database with the soil sealing map for the production of an error matrix, to be used on the estimation of the specific and overall accuracy indexes (user, producer and overall accuracies).

Concerning the developed methodology, we should stress that in (a) the converted map values were group in 5 classes of soil sealing in order to assist the sampling design. In (b) several analyses were produced along 3 sequential intermediate stages in order to filter miss-interpreted reference sample observations. Regarding (c), the validation was done throughout a rigid approach and also a fuzzy approach.

In the following sections, the procedures developed for the soil sealing map validation are explained in detail and by the performed order.

3.1 Built-up map preparation

The first thing to do was to convert the 20 m x 20 m soil sealing map into a 100 m x 100 m grid. This was done by averaging the soil sealing values of 5 x 5 20 m grid cells (Figure 2).

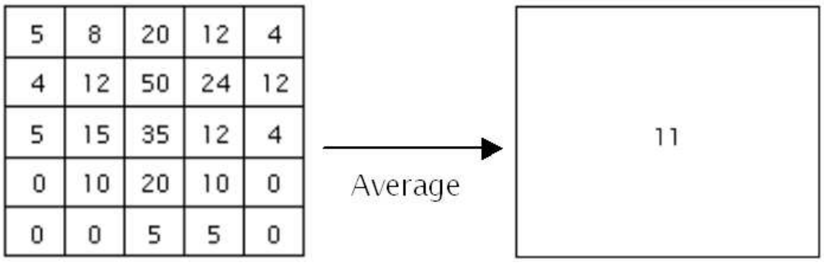


Figure 2 - Conversion of the original 20 m x 20 m soil sealing map (left) into a 100 m x 100 m soil sealing map (right). Cell values correspond to degree of soil sealing.

3.2 Reference database design

The reference database production process consisted in two main parts: 1) the survey sampling, namely the definition of the reference sampling observation and the selection of the most adequate sampling design for geographical random sample collection; and 2) the so-called response design (Stehman and Czaplewski, 1998), which consists on the evaluation of the variable of interest at each sampling observation, i.e. the degree of soil sealing, using for that purpose accurate reference data, e.g. aerial imagery.

Next we detail each of these components, explaining the main aspects taken into account during the reference database construction.

3.2.1 Reference sampling observation

The reference sample observation serves as the basic element of comparison between the map classification and the reference, or "true" classification. Because classification accuracy must be calculated per hectare (i.e. based on a 100 m x 100 m grid), the reference sample unit is a single pixel of the derived 100 m spatial resolution soil sealing map.

3.2.2 Sampling design

For the reference database construction we have chosen to use a stratified random sampling design, because:

- It can contribute to a scientifically defensible accuracy assessment of land cover products (Stehman and Czaplewski, 1998);

- It can guarantee a minimum sample size in each stratum to derive precise accuracy estimates for each map class (Stehman, 1999).

Because classification accuracy must be calculated for built-up (80-100% of soil sealing) and non built-up areas (0-80% of soil sealing), we decided that stratification should follow the scheme shown in Table 2 in order to achieve a better sample distribution over the non built-up map areas. Moreover, this stratification allowed us to produce a more detailed quality assessment of the soil sealing map, which can be observed in Annex I.

Table 2 - Map classes used as strata for the stratified random sampling.

Class (Code)	% degree of soil sealing	Designation
1	[0 - 20[	Very low
2	[20 - 40[	Low
3	[40 - 60[	Mean
4	[60 - 80[	High
5	[80 - 100]	Very High

Regarding the appropriate sample size to be collected per map class, in general, the larger the sample size the greater the confidence one can have in assessments based on that sample. However, extensively large sample sizes obligate to extensive efforts in sample interpretation and therefore are not cost effective. For that reason, several main authors (e.g. Hay, 1979; Carrão *et al.*, 2007a) have suggested and used an equation (1), based on the binomial approximation to the normal distribution, to estimate the required sample size for the accuracy assessment of land cover maps. Congalton and Green (1999) states that this approach is statistically sound for estimating the sample size needed to estimate the overall accuracy of a classification or the accuracy of a single land cover category.

$$N = \left(\frac{Z_{1-\alpha/2}}{d} \right)^2 p(1-p) \tag{1}$$

N - Sample size
α - Significance level
d - Absolute precision
p - Probability of belonging to the map class

Instead of using the equation (1) to estimate the appropriate sample size, we decide to follow the EEA recommendations (Maucha and Buttner, 2007) and use a sample size of 500 sample observations per map class. The equation (1) was used to estimate the maximum error interval (the obtained value was 3.67%,) regarding overall and specific accuracies, for a sample size of 500 observations, and at the confidence level of 90% (0.1 significance) required by EEA.

In the end, the sampling design resulted as shown in Figure 3.

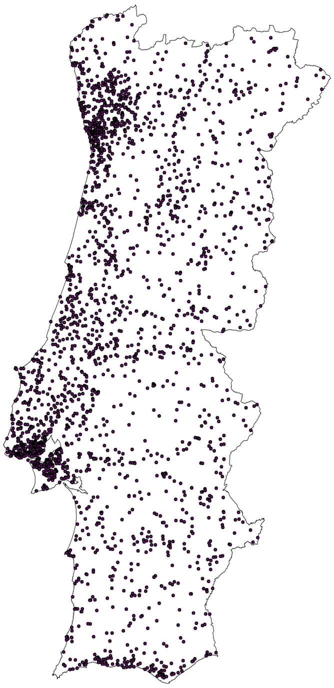


Figure 3 - Distribution of the 2500 samples for Continental Portugal.

3.2.3 Reference imagery data

Reference labels were derived by visual analysis of orthorectified aerial images acquired during the years of 2004, 2005 and 2006 and covering the whole Portuguese territory (Figure 4). These images have four spectral bands (blue, green, red, and near-infrared) and 50 cm of spatial resolution.

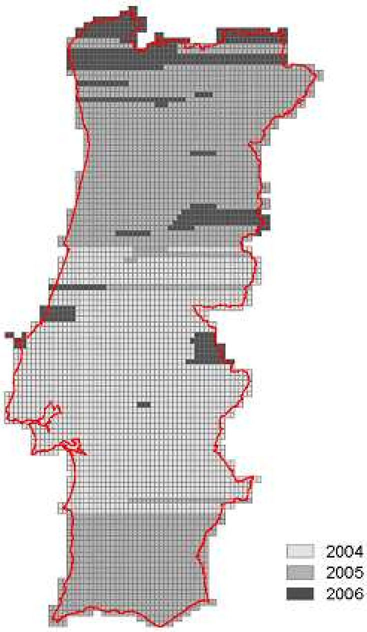


Figure 4 - Distribution of the orthorectified aerial images acquired during the years of 2004, 2005 and 2006 for Continental Portugal.

Because of the temporal difference between some orthorectified aerial images and the map, we also used, whenever needed, SPOT-4/5 and IRS LISS-3 imagery as reference imagery data.

3.2.4 Sample interpretation

The procedure for sample interpretation consisted in the visual interpretation, at a scale of 1:1000, of the reference imagery data for the set of 2500 sample observations. This was performed by a team of five interpreters, responsible for the interpretation and classification of the sample observations. The high skill of the interpreters was guaranteed, as they are working in the CLC update. Moreover, they were submitted to a training period in which they interpreted the same set of more than 100 sample observations, having results compared in order to harmonize interpretation criteria.

The sample interpretation procedure was done throughout 3 sequential intermediate stages in order to filter miss-interpreted reference sample observations (Figure 1 - b).

<u>1st Stage</u>

The whole set was randomly and equally distributed to the five image interpreters. At each sample observation, the interpreters classified the degree of soil sealing (0-100%) existing in the observation area. This was done by means of counting 100 equally distributed points, and for each of these 100 points the interpreters determined a binary imperviousness state, i.e. impervious or not impervious (Figure 5). This strategy was suggested and established in the 2nd meeting of the GMES Land FTS Precursor Steering Committee (29th January 2008) as proposed by the Austrian team (Banko, 2008).

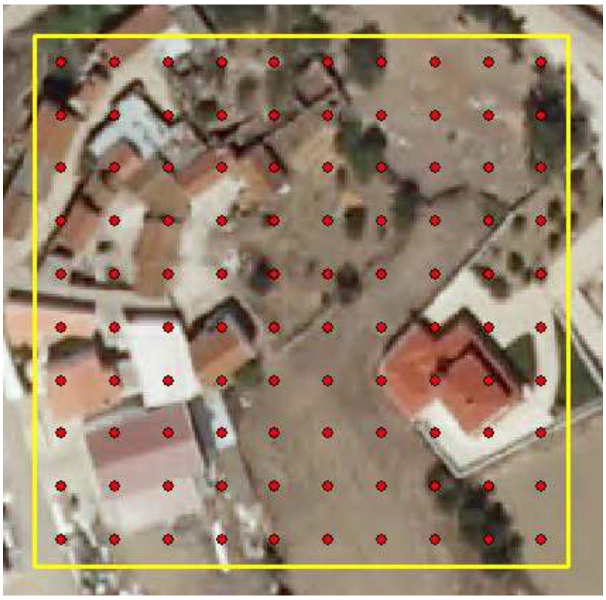


Figure 5 - Visual interpretation of a sample observation (100 m x 100 m) at the scale of 1:1000. By counting the impervious points out of the 100 points the interpreters classify the soil sealing degree of the sample observation.

2nd Stage

In the 2nd stage, the off-diagonal sample observations that resulted from the 1st reference vs map comparison were submitted to re-interpretation by an interpreter not involved in the first stage, while the diagonal observations were considered final.

After re-interpreted, the sample observations that presented a very different soil sealing degree (difference superior to 10%), when comparing with their own label in the 1st stage, were sent to re-interpretation (3rd stage). The sample observations that presented a soil sealing degree difference equal or inferior to 10% were considered final, and their soil sealing degree label is an average of the soil sealing defined by the two interpreters.

3rd Stage

In the 3rd stage, the sample observations that were not considered final in the previous stages were re-interpreted. This was done by an interpreter not involved in the first two phases. Those observations that, after re-interpreted, presented a very different soil sealing degree (difference superior to 10%), when comparing with their own label in the 1st and 2nd stage, were considered final and their soil sealing degree label is the average of the two closest values coming from the third stages. The sample observations that presented a soil sealing degree difference equal or inferior to 10% were considered final and their soil sealing degree label is the average of the values estimated in the three stages.

3.3 Accuracy assessment

The accuracy assessment of the high-resolution soil sealing map was made through the construction and analysis of a contingency table, also called error or confusion matrix.

Error matrixes are used to compare the information obtained in a classification process with the reference data. They are produced trough the intersection of map classes with the classes attributed to each sample observation in the reference database. In the error matrix, usually, lines represent the map and columns represent the reference labels. From it, we can estimate the specific and overall accuracy indexes, i.e. user, producer and overall accuracies, commission and omission errors (Jensen, 1996).

The overall accuracy is the sum of the diagonal elements of the matrix (correctly classified elements) divided by the total number of pixels in the sample. The producer's accuracy is the number of pixels correctly classified in class (i) divided by the total number of pixels in the reference class (i). It traduces the probability of a reference observation being correctly classified in the map. The omission error is its complementary. The user accuracy corresponds to the number of pixels correctly classified in class (i) divided by the total number of pixels in the map class (i). This measure is the probability that a pixel classified on the map actually represents that category on the ground (reference). The commission error is its complementary (Jensen, 1996).

These indexes, calculated this way, should be considered as sample statistics not calibrated for the map, since, among others, they do not consider the area of each map class. Card (1982) points out that, for the stratified sampling case, the overall proportion of correctly classified individuals in the map, given the reference land cover categories, should not be simply estimated by the diagonal entry divided by the row sum of the contingency table, because of the bias introduced by possible differential sampling rates within map categories. Therefore, the overall and per class accuracy estimations that will be used to evaluate the land cover map should include the known areas of each map category to improve the estimation of the proportion of correctly mapped individuals.

Taking into account the above, we decided to do the validation considered as sample statistics not calibrated for the map and also, to develop an accuracy assessment of the high-resolution soil sealing map considering the area of each map class. For this latter purpose we supported our methodology on Cochran (1977) work, later adopted by Card (1982). The equations (2, 3, 4, 5) used to derive the specific and overall accuracy indexes were those proposed by Carrão *et al.* (2007b), derived from Cochran (1977). In accordance with Carrão *et al.* (2007b), given an error matrix A (Figure 6) and considering that:

Reference land cover category

Land cover map category (Stratum)		ϖ_1	ϖ_2	$\cdots$	ϖ_G	$RowTotal$
	ϖ_{11}	n_{11}	n_{12}	$\cdots$	n_{1G}	n_{1+}
	ϖ_2	n_{21}	n_{22}	$\cdots$	n_{2G}	n_{2+}
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	ϖ_H	n_{H1}	n_{H2}	$\cdots$	n_{HG}	n_{H+}
	$ColumnTotal$	n_{+1}	n_{+2}	$\cdots$	n_{+G}	n

Figure 6 - Error matrix A (Carrão *et al.*, 2007b).

N - number of pixels in the map;

N_h - number of pixels in the map category ϖh;

N_{h1} - number of pixels in the map category ϖh that correspond to same reference category;

N_g - number of pixels in the reference category ϖg;

N_{hg} - number of pixels in the map category ϖh that intersect reference category ϖg;

n_h - number of pixels in the sample collected in map category ϖh;

n_{hg} - number of pixels in the sample collected in map category ϖh that intersect reference category ϖg;

n_{h1} - number of pixels in the sample collected in map category ϖh that correspond to same reference category.

The overall accuracy estimation for the map, i.e. the proportion of correctly classified pixels in the map, is

$$\hat{P}_c = \frac{1}{N}\sum_{h=1}^{H}\frac{N_h}{n_h}n_{h1} \tag{2}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_c) = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{h1}}{n_h}\left(1 - \frac{n_{h1}}{n_h}\right)\right] \tag{3}$$

The user's accuracy estimation for each land cover class in the map is

$$\hat{P}_{h,c} = \frac{n_{h1}}{n_h} \tag{4}$$

and the respective estimated variance

$$\hat{V}(\hat{P}_{h,c}) = \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{h1}}{n_h}\left(1 - \frac{n_{h1}}{n_h}\right)\right] \tag{5}$$

The estimation of producer's accuracy under a stratified random sampling survey must be calculated using an unbiased estimation of the population unknown quantities $Nh1$ and Ng. Their unbiased estimators are:

$$\hat{N}_{h1} = \frac{N_h}{n_h}n_{h1} \tag{6}$$

$$\hat{N}_g = \sum_{h=1}^{H}\frac{N_h}{n_h}n_{hg} = \sum_{h=1}^{H}\hat{N}_{hg} \tag{7}$$

and their respective estimated variances

$$\hat{V}(\hat{N}_{h1}) = {N_h}^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{h1}}{n_h}\left(1 - \frac{n_{h1}}{n_h}\right)\right] \tag{8}$$

$$\hat{V}(\hat{N}_g) = \sum_{h=1}^{H}{N_h}^2 \frac{N_h - n_h}{N_h n_h}\left[\frac{n_{hg}}{n_h}\left(1 - \frac{n_{hg}}{n_h}\right)\right] \tag{9}$$

Then, the producer's accuracy can be viewed as a ratio of these two estimators as follows

$$\hat{P}_{g,c} = \frac{\hat{N}_{h1}}{\hat{N}_g} \tag{10}$$

and the respective estimated Mean Square Error (MSE)

$$M\hat{S}E(\hat{P}_{g,c}) = \left(\frac{N}{\hat{N}_g}\right)^2 \sum_{h\neq g}\left[\left(\frac{N_h}{N}\right)^2\left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{hg}}{n_h}\left(1 - \frac{n_{hg}}{n_h}\right)\hat{P}^2_{g,c}\right] + \left[\left(\frac{N}{\hat{N}_g}\right)^2\left(\frac{N_h}{N}\right)^2\left(\frac{N_h - n_h}{N_h n_h}\right)\frac{n_{h1}}{n_h}\left(1 - \frac{n_{h1}}{n_h}\right)\left(1 - \hat{P}_{g,c}\right)^2\right] \tag{11}$$

The approximate confidence intervals for each generic estimate $\hat{P}$ are determined according to $\hat{P} \pm Z_{\alpha/2}\left[V(\hat{P})\right]^{1/2}$, where $Z_{\alpha/2}$ is the standard normal deviate for the desired confidence level 1-α. $Z_{\alpha/2}\left[V(\hat{P})\right]^{1/2}$ represents the absolute precision d of the estimate at the 1-α confidence level.

The accuracy assessment methodology here described was complemented with the implementation of fuzzy intervals in the breaking values of soil sealing classes (Figure 7), both in the 5 classes and in the 2 classes situations.

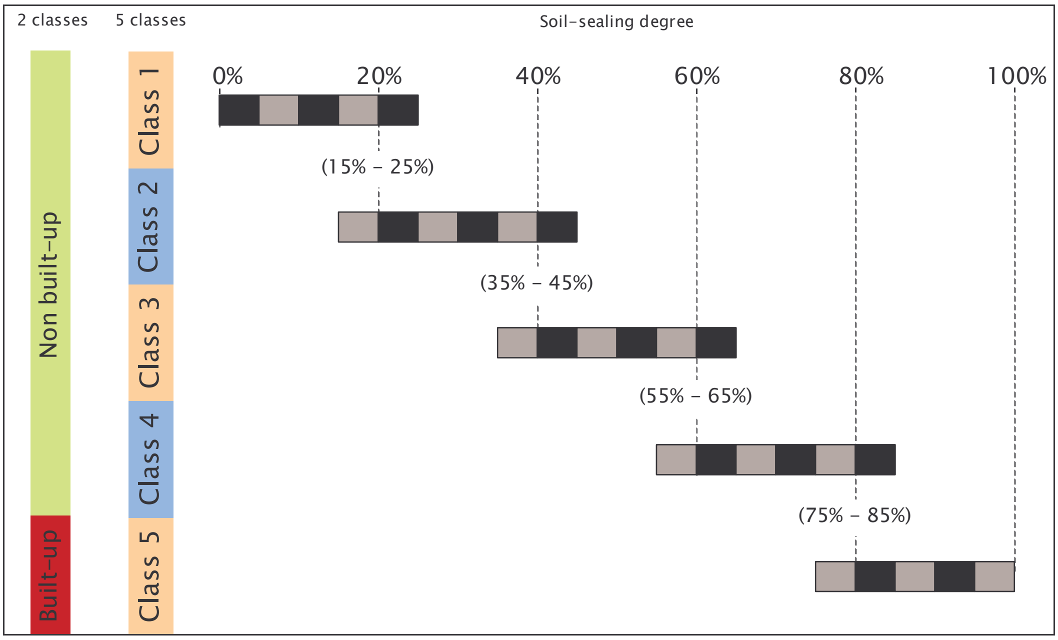


Figure 7 - Illustration of fuzzy intervals in the breaking values of soil-sealing classes.

These fuzzy intervals prevent situations of slight differences between soil sealing values from reference data and map to be considered as different classes, and therefore accounted as errors. For example, if the map has a value of 82% and the reference is 79%, this method considers the map correct, unlike the rigid intervals method would do.

4 Results

Results of the built-up map validation are described in two different sections, in accordance to the use of two different approaches in the error matrix construction. One approach, where the area of each map class is not accounted in the validation (sample error matrix), and the other, where it is (estimated map error matrix). The results for the 5 classes soil sealing map are presented in Annex 1.

4.1 Sample error matrix

This approach, as previously described, compares the values of soil sealing from the 100m built-up map with the visual interpretation of the 2500 sample set. This comparison was made accordingly to soil sealing threshold of 80%, meaning that the built-up map and visual interpretation values were grouped in two classes: less than 80% and equal or greater than 80%. Table 3 and Table 4 present these results for rigid and fuzzy methods, respectively.

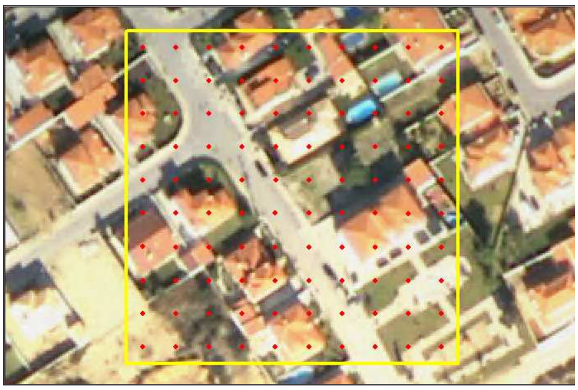
Table 3 – Sample error matrix (rigid), where values represent number of sample observations. PA is the producer's accuracy (%); UA is the user's accuracy (%); and OA is the overall accuracy (%). Reference is represented in columns while map is represented in lines.

	Non built-up	Built-up	Total	UA
Non built-up	1973	27	2000	98.65
Built-up	233	267	500	53.40
Total	2206	294	2500	
PA	89.44	90.82		OA = 89.60

It can be observed that about 10% of all sample set is inadequately classified in the built up map, according to the reference data. The most part of this misclassification (233 pixels) corresponds to sample observations with less than 80% of soil sealing that were classified as built-up (i.e. greater or equal to 80%) in the built-up map. Another important result is that only 53.4% of the sample set classified as built-up in the map is effectively built-up.

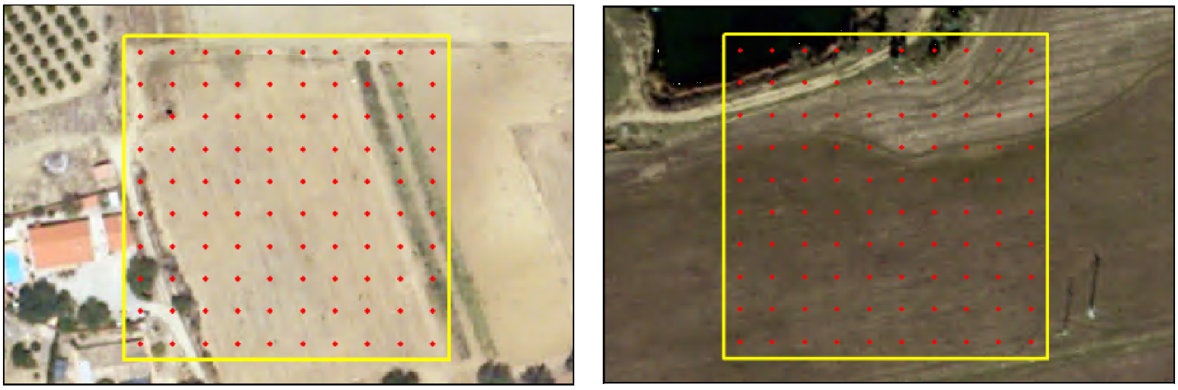
Most of the situations of disagreement can be due either to the over-estimation of the density of urban areas, where gardens, unproductive land and other non-sealed soils are only perceptible with higher spatial resolution data; or to wrong classification of bare soil areas into urban, like the examples of fallow land, beaches or sparse vegetation in dry soils. The following examples (Figure 8, Figure 9 and Figure 10) show some of these misclassifications, with the indication of soil sealing values in the built-up map and by visual-interpretation (i.e. reference data).

Visual interpretation: 59%
Built-up map: 82%



Visual interpretation: 53%
Built-up map: 82%

Figure 8 – Examples of over-estimation of the density of urban areas



Visual interpretation: 2%
Built-up map: 81%

Visual interpretation: 0%
Built-up map: 96%

Figure 9 – Examples of misclassification of agriculture land (fallow land) as built-up areas.

Visual interpretation: 47%
Built-up map: 100%

Visual interpretation: 2%
Built-up map: 91%

Figure 10 – Examples of misclassification of beaches, bare soil and sparse vegetation in dry soils.

Other type of errors that were found is related to the under-estimation of the soil sealing degree. These situations, however, are not as representative as the over-estimations, and can be generally explained by the presence of urban greenery, in some cases with concrete, asphalt or other soil sealing materials under tree cover (Figure 11).

Visual interpretation: 84%
Built-up map: 35%

Visual interpretation: 97%
Built-up map: 77%

Figure 11 – Examples of under-estimation of soil sealing.

Some other misclassifications can be due to slight differences in the values of soil sealing, like situations of reference data with 78% and built-up map with 82%. These situations, in which a fuzzy tolerance of 5% in class intervals would make reference data and built-up map agree, were detected in 67 sample observations, increasing Built-up UA from 53.4% to 64.8%. This means that, when a fuzzy validation with 5% tolerance (previously described) is applied, the overall accuracy of the built-up map improves to 92% (Table 4).

Table 4 – Sample error matrix (fuzzy), where values represent number of sample observations. PA is the producer's accuracy (%); UA is the user's accuracy (%); and OA is the overall accuracy (%). Reference is represented in columns while map is represented in lines.

	Non built-up	Built-up	Total	UA
Non built-up	1983	17	2000	99.15
Built-up	176	324	500	64.80
Total	2159	341	2500	
PA	91.85	95.01		OA = 92.28

4.2 Estimated map error matrix

In Table 5 we introduce the percentage area occupied per mapped land cover class (Nh), the percentage area occupied per reference land cover class (Ng), the overall accuracy (P), the user's accuracy (Ph), the producer's accuracy (Pg), and the absolute precisions, estimated for a 90% level of confidence, for Ng, P, Ph and Pg ($d(Ng)$, $d(P)$, $d(Ph)$ and $d(Pg)$, respectively.

Table 5 – Percentage area occupied per mapped land cover class (N_h); percentage area occupied per reference land cover class (N_g) plus 90% absolute precision $d(N_g)$; user's accuracy (P_h) plus 90% absolute precision $d(P_h)$; producer's accuracy (P_g) plus 90% absolute precision $d(P_g)$; overall accuracy (P) plus 90% absolute precision $d(P)$.

		N_h	N_g	$d(N_g)$	P_h	$d(P_h)$	P_g	$d(P_g)$
Non built-up		99.14	99.48	0.89	99.94	0.02	99.60	0.02
Built-up		0.86	0.52	0.04	53.40	3.65	89.40	3.34
P	99.54							
$d(P)$	0.037							

In Table 6 we introduce the proportion of matches between map and reference classes as a percentage of the total study area.

Table 6 – Estimated map error matrix (rigid), where values represent estimated proportion (N_{hg}) of matches between map and reference classes as a percentage of the total study area. Reference is represented in columns while map is represented in lines.

	Non built-up	Built-up	Total
Non built-up	99.08	0.06	99.14
Built-up	0.40	0.46	0.86
Total	99.48	0.52	100

Overall accuracy of the final map was estimated at 99.54% with an absolute precision of 0.037% at the 90% confidence level. This extremely high value of overall accuracy can only be explained by the major difference in map class areas (non built-up represents N_g = 99.48% of the map area while built-up only represents N_g = 0.52%) and by the fact that the largest class, non built-up, presented almost 100% user's accuracy. Comparing with the previous approach regarding the error matrix construction (sample error matrix), the sample observations that were found misclassified are the same. Again, most part of the misclassification corresponds to sample observations with less than 80% of soil sealing that were classified as built-up (i.e. greater or equal to 80%). However, overall accuracy was affected (from 89.60% to 99.54%, regarding sample error matrix and estimated map error matrix respectively) because the preponderance of the built-up class accuracy in the overall value was greatly reduced, since this class occupies a very small area in the whole analysed region.

When applying the fuzzy method in this accuracy assessment (Table 7 and Table 8), the differences encountered are mainly the same that in the sample error matrix. On the other hand, overall accuracy was not so affected because, as mentioned before, the preponderance of the built-up class accuracy in the overall value is small, when accounting map area proportions.

Table 7 – Percentage area occupied per mapped land cover class (N_h); percentage area occupied per reference land cover class (N_g) plus 90% absolute precision $d(N_g)$; user's accuracy (P_h) plus 90% absolute precision $d(P_h)$; producer's accuracy (P_g) plus 90% absolute precision $d(P_g)$; overall accuracy (P) plus 90% absolute precision $d(P)$.

		N_h	N_g	$d(N_g)$	P_h	$d(P_h)$	P_g	$d(P_g)$
Non built-up		99.14	99.40	0.78	99.96	0.02	99.69	0.02
Built-up		0.86	0.60	0.03	64.80	3.49	93.49	2.71
P	99.66							
$d(P)$	0.03							

Table 8 – Estimated map error matrix (fuzzy), where values represent estimated proportion (N_{hg}) of matches between map and reference classes as a percentage of the total study area. Reference is represented in columns while map is represented in lines.

	Non built-up	Built-up	Total
Non built-up	99.10	0.04	99.14
Built-up	0.30	0.56	0.86
Total	99.40	0.60	100

5 Conclusions

This report describes the validation procedures of the high-resolution built-up map for Continental Portugal, produced in the context of GMES Fast Track Service Precursor on Land Monitoring.

A reference database of 2500 sample observations was made by visual-interpretation of orthoimages of mainland Portugal, and a validation with two different approaches regarding error matrix construction, was then undertaken.

The results were affected by the different methods. When considering the sample error matrix, the high-resolution built-up map overestimates in about 47% the built-up areas (53% of user accuracy). Regarding estimated map error matrix, these over-estimations are also indicated even though the values were normalized by the representativeness of built-up areas in the total map. Therefore, the overall accuracies with sample error matrix and estimated map error matrix are respectively 89.6% and 99.5%.

Regardless of these differences, these values indicate the overall quality of the product is very high. However, when considering user's accuracy there is a clear over-estimation of the built-up areas, as indicated by both methods.

Also, the error matrixes were built considering the use of standard rigid methods and enhanced fuzzy methods. When analysing these results, the conclusion is that the fuzzy approach improves the specific and overall accuracies by tolerating slight differences in the values of soil sealing, thus promoting the quality of the map in a fair way.

Table 9 summarizes the results obtained for all accuracy assessment methods developed. This table shows that fuzzy results are always better than the rigid results. Regarding the comparison between sample error matrix and estimated map error matrix, overall accuracy results were higher for the latter. This is a consequence of the major difference in the proportion of the mapped area for each class combined with the good results achieved for the non built-up classification (UA > 98% for all cases). An important conclusion that arises from the table is the low user accuracy of the built-up class considering all the undertaken approaches, meaning that there is a clear over-estimation of the built-up areas in the map.

Table 9 – Summary of the results obtained for all accuracy assessment methods.

		Rigid		Fuzzy	
		Non built-up	Built-up	Non built-up	Built-up
Sample error matrix	UA	98.65	53.40	99.15	64.80
	PA	89.44	90.82	91.85	95.01
	OA	89.60		92.28	
Estimated map error matrix	UA	99.94	53.40	99.96	64.80
	PA	99.60	89.40	99.69	93.49
	OA	99.54		99.65	

Concerning the 5 classes evaluation (Annex 1), the differences between the 2 error matrix methods were equivalent, as well as when applying fuzzy and rigid approaches i.e., proportionally the methodology had the same effect on results with 5 classes that it had with 2. However comparing the results *per se*, the user's, producer's and overall accuracies are lower in the 5 classes evaluation.

References

Banko, G., 2008. AUSTRIA: verification of high resolution soil sealing layer - Qualitative assessment.

Card, D.H., 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy, Photogrammetric Engineering and Remote Sensing, 48:431-439.

Carrão, H., A. Araújo, C. Cerdeira, P. Sarmento, L. Capão, and M. Caetano, 2007a. A reference sample database for the accuracy assessment of medium spatial resolution land cover products in Portugal, Proceedings of the 2007 International Geoscience and Remote Sensing Symposium (IGARSS'2007), 23 – 27 Julho 2007, Barcelona, Espanha.

Carrão, H., M. Caetano, and P. Coelho, 2007b. Sample Design and Analysis for Thematic Map Accuracy Assessment: An Approach Based on Domain Estimation for the Validation of Land Cover Products, Proceedings of the 32nd International Symposium on Remote Sensing of Environment, 25 - 29 Junho de 2007, San Jose, Costa Rica.

Cochran, W.G., 1977. Sampling Techniques, Wiley, New York, 428 p.

Congalton, R. and K. Green, "Assessing the Accuracy of Remotely Sensed Data: Principles and Practices", Boca Raton, FL: CRC/Lewis Press, 1999.

Hay, A., 1979. Sampling designs to test land-use map accuracy, Photogrammetric Engineering and Remote Sensing, vol. 45, pp. 529–533.

Jensen, J.R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd Edition, Prentice Hall, New Jersey.

Maucha, G., and G. Büttner, 2007. Recommendations: Quantitative assessment high-resolution soil sealing layer, European Environment Agency.

Sánchez, B., and H., Kahabka, 2008. 2006 GMES fast track service precursor on land monitoring: high resolution core land cover data - built-up areas including degree of soil sealing - delivery report Portugal, European Environment Agency.

Stehman, S., 1999. Basic probability sampling designs for thematic map accuracy assessment, International Journal of Remote Sensing, vol. 20, pp. 2423–2441.

Stehman, S., and R., Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: fundamental principles, Remote Sensing of Environment, vol.64, pp. 331-344.

ANNEX I

Sample error matrix

Table 10 compares the values of soil sealing from the 100m built-up map with the visual interpretation of the 2500 sample set. This comparison was made in accordance with the soil sealing classes used for the built-up map stratification (class codes and respective soil sealing degree intervals: 1 [0-20%[; 2 [20-40%[; 3 [40-60%[; 4 [60%-80%[; and 5 [80-100%]).

Table 10 - Sample error matrix (rigid), where values represent number of sample observations. PA is the producer's accuracy (%); UA is the user's accuracy (%); and OA is the overall accuracy (%). Reference is represented in columns while map is represented in lines.

Code	1	2	3	4	5	Total	UA
1	496	3	1			500	99.20
2	250	179	59	10	2	500	35.80
3	153	119	166	57	5	500	33.20
4	73	109	121	177	20	500	35.40
5	33	34	69	97	267	500	53.40
Total	1005	444	416	341	294	2500	
PA	49.35	40.32	39.90	51.91	90.82		OA = 51

Considering the 51% overall accuracy, we may say that approximately half of the sample set is misclassified in the map. The major part of the confusions is due to the soil sealing degree overestimation, as it is demonstrated in the matrix, where the error is mainly located under the diagonal. Another interesting conclusion is that the classes better classified are the ones that represent the soil sealing degree extremes (class 1 and class 5). This may be due to the fact that both classes can only be confused with one other class, while all the others may be confused with two classes. This suggests that most of the error between consecutive classes may be related to slight differences of soil sealing degree near the breaking values of the classes. We also believe that extreme situations, like the ones represented in classes 1 and 5 (i.e. non imperviousness and total imperviousness) are easier to be discriminated both in the map and in the reference database.

The introduction of fuzziness into the accuracy assessment improved the specific and overall results (Table 11).

Table 11 – Sample error matrix (fuzzy), where values represent number of sample observations. PA is the producer's accuracy (%); UA is the user's accuracy (%); and OA is the overall accuracy (%). Reference is represented in columns while map is represented in lines.

Code	1	2	3	4	5	Total	UA
1	497	2	1			500	99.40
2	215	249	24	10	2	500	49.80
3	153	59	258	25	5	500	51.60
4	73	109	55	253	10	500	50.60
5	33	34	69	40	324	500	64.80
Total	971	453	407	328	341	2500	
PA	51.18	54.97	63.39	77.13	95.01		OA = 63

Estimated map error matrix

In Table 12, we introduce the percentage area occupied per mapped land cover class (Nh), the percentage area occupied per reference land cover class (Ng), the overall accuracy (P), the user's accuracy (Ph), the producer's accuracy (Pg), and the absolute precisions, estimated for a 90% level of confidence, for Ng, P, Ph and Pg ($d(Ng)$, $d(P)$, $d(Ph)$ and $d(Pg)$, respectively), regarding the rigid accuracy assessment of the 5 class map.

Table 12 – Percentage area occupied per mapped land cover class (Nh); percentage area occupied per reference land cover class (Ng) plus 90% absolute precision $d(Ng)$; user's accuracy (Ph) plus 90% absolute precision $d(Ph)$; producer's accuracy (Pg) plus 90% absolute precision $d(Pg)$; overall accuracy (P) plus 90% absolute precision $d(P)$.

Code	N_h	N_g	$d(N_g)$	P_h	$d(P_h)$	P_g	$d(P_g)$
1	94.48	95.58	0.63	99.20	0.65	98.06	0.11
2	2.59	2.03	0.54	35.80	3.52	45.62	12.32
3	1.29	1.23	0.32	33.20	3.45	34.82	9.27
4	0.78	0.64	0.05	35.40	3.50	43.09	3.98
5	0.86	0.52	0.04	53.40	3.65	89.40	3.34
P	95.81						
$d(P)$	0.63						

In Table 13 we introduce the proportion of matches between the 5 class map and reference classes as a percentage of the total study area.

Table 13 – Estimated map error matrix (rigid), where values represent estimated proportion (Nhg) of matches between map and reference classes as a percentage of the total study area. Reference is represented in columns while map is represented in lines.

Code	1	2	3	4	5	Total
1	93.72	0.57	0.19	0.00	0.00	94.48
2	1.29	0.93	0.31	0.05	0.01	2.59
3	0.40	0.31	0.43	0.15	0.01	1.29
4	0.11	0.17	0.19	0.28	0.03	0.78
5	0.06	0.06	0.12	0.17	0.46	0.86
Total	95.58	2.03	1.23	0.64	0.52	100

When comparing these results with the ones obtained from the sample error matrix validation, we may see that results improved. The major difference exists in the overall result once, as explained before, the preponderance of the class 1 (UA of 94%) outperforms largely the other classes due to the proportion of its area in the whole map.

The introduction of fuzziness into the accuracy assessment improved even more the specific and overall results (Table 14 and Table 15).

Table 14 – Percentage area occupied per mapped land cover class (N_h); percentage area occupied per reference land cover class (N_g) plus 90% absolute precision $d(N_g)$; user's accuracy (P_h) plus 90% absolute precision $d(P_h)$; producer's accuracy (P_g) plus 90% absolute precision $d(P_g)$; overall accuracy (P) plus 90% absolute precision $d(P)$.

Code	N_h	N_g	$d(N_g)$	P_h	$d(P_h)$	P_g	$d(P_g)$
1	94.48	95.59	0.55	99.40	0.57	98.24	0.11
2	2.59	2.05	0.45	49.80	3.67	62.90	13.60
3	1.29	1.18	0.32	51.60	3.66	56.25	14.99
4	0.78	0.58	0.05	50.60	3.66	68.15	4.68
5	0.86	0.60	0.03	64.80	3.49	93.49	2.71
P	96.82						
$d(P)$	0.55						

Table 15 – Estimated map error matrix (fuzzy), where values represent estimated proportion (Nhg) of matches between map and reference classes as a percentage of the total study area. Reference is represented in columns while map is represented in lines.

Code	1	2	3	4	5	Total
1	93.91	0.38	0.19	0.00	0.00	94.48
2	1.11	1.29	0.12	0.05	0.01	2.59
3	0.40	0.15	0.67	0.06	0.01	1.29
4	0.11	0.17	0.09	0.40	0.02	0.78
5	0.06	0.06	0.12	0.07	0.56	0.86
Total	95.59	2.05	1.18	0.58	0.60	100