



Integrating nature & biodiversity and land use data

Final Report

Version 2.1

Service Contract: No. 070307/2010/58/0092/SER/B2



in collaboration with



08.02.2012

The report was coordinated by:

Stefan Kleeschulte (GeoVille Environmental Services)

The following authors contributed to the report:

Alena Dostalova, Michael Hosek (AOPK)

Tomas Soukup (GISAT)

Maria José Ramos, Roger Milego (UAB)

Christian Schleicher (GeoVille Information Systems)

This report should be quoted as follows:

Kleeschulte, S. et al (2011): Integrating nature & biodiversity and land use data – Final Report. Output of the project: INBALUD - Integrating Nature & biodiversity and land use data. GeoVille Environmental Services, Luxembourg.

All opinions and views presented in this report are those of the authors, and do not necessarily represent views of the members of the Steering Group or European Commission.

Table of contents

1	Executive Summary	5
2	Introduction	6
3	Project objectives	7
3.1	EU biodiversity objectives and targets	7
3.2	The situation today	8
3.3	Project expectations	9
4	Project achievements	11
4.1	Part 1: Examine possibilities of how to bring existing data collections together	11
4.1.1	Assessment of potentially relevant data sets	11
4.1.2	Integration of LUCAS data	15
4.1.3	Indicator proposal.....	20
4.2	Part 2: Produce as useful first result a processed database.....	22
4.2.1	Data storage	22
4.2.2	Data integration and query	22
4.3	Part 3: Include follow-up work by producing an EU analysis on selected biodiversity parameter	25
4.3.1	Habitat Directive Species and Habitats Index Proposal.....	25
4.3.2	Sites designated under the EU Habitats and Birds Directives	27
4.3.3	High Nature Value Farmlands	29
4.3.4	Extracting biodiversity information from LUCAS	30
4.3.5	Characterising other data sets with LUCAS.....	33
4.3.6	Conclusions	37
4.4	Part 4: Make sure that metadata will be available for future data use requests	38
4.4.1	Meta data	38
4.4.2	Indicator Processing Documentation	39
4.4.3	Conclusions	41
4.5	Part 5: Feed in the future strategy on data storing and handling within European Institutions	41
4.6	Part 6: Formulate input to recommendations on EU biodiversity data policy	42
4.7	Part 7: Prepare conceptually the biodiversity element of in-situ surveys such as LUCAS	44
5	Conclusions	47
6	References	53

Annexes to the final report (provided as separate reports):

- Part 1: Examine possibilities of how to bring existing data collections together
- Part 2: Produce as useful first result a processed database
- Part 3: Include follow-up work by producing an EU analysis on selected biodiversity parameter
- Part 3a: The biodiversity element of in-situ surveys such as LUCAS
- Part 4: Meta data – Indicator Fact Sheets

- Part 5: Feed in the future strategy on data storing and handing within European Institution
- Part 6: Formulate input to recommendations on EU biodiversity data policy
- Part 7: Prepare conceptually the biodiversity element of in-situ surveys such as LUCAS
- OLAP cube user manual

1 Executive Summary

The objectives of the INBALUD project have been fourfold:

- a) to assess the added value of a better integration of land use (especially LUCAS) and biodiversity information,
- b) to contribute to the EU biodiversity objectives and targets through the provision of targeted information,
- c) to make the information accessible to interested stakeholders and finally
- d) to make recommendations for future biodiversity policies and biodiversity monitoring in Europe.

With respect to most of its goals the project has been successful. A new approach (the Habitat Diversity Index) to the assessment of existing information was proposed in support of "SEBI No 3 – Species of European interest" and "SEBI No 5 – Habitats of European interest". For SEBI No 8 – Sites designated under the EU Habitats and Birds Directives the project has provided important background information. In both cases the INBALUD information has been able to provide information at a more detailed spatial level than the currently existing information.

All project data and information have been stored in a spatial database, including meta data information of the input data as well as the results (indicators), indicator fact sheets with processing details and an interface to query the database contents in an easy, non-expert manner, without the need for complex GIS software.

The database together with its query interface has been provided to DG ENV and EEA. Once final details are clarified it is foreseen to make the database publicly available via the EEA respectively one of its European Topic Centres.

The project has carried out a gap analysis of the current data supporting European biodiversity policy and provided a series of recommendations how to overcome these gaps. Recommendations were done with respect to 2020 biodiversity targets with a specific aim on data required for indicators which provide important information on current status, achievements and challenges in the EU biodiversity conservation.

The current LUCAS survey approach was reviewed and recommendations were made how to better incorporate biodiversity issues in the survey. Two options were presented a) the stricter implementation of current survey guidelines and b) a biodiversity targeted survey to be carried out by expert surveyors aimed at habitats and selected groups of species. For the suggested biodiversity survey a scheme was drafted and a first cost estimation was provided.

With respect to the first goal of the project – the integration of LUCAS and biodiversity data – the results are mixed. Due to the underlying differences in concept (statistics vs. mapping) and the design of LUCAS as an agricultural survey in the beginning, the value of LUCAS data for biodiversity assessments is limited. Better results can be achieved when attaching information from other data themes to the LUCAS sample points.

2 Introduction

The present report is the final report of the INBALUD project (Integrating nature & biodiversity and land use data) which has been financed by DG Environment under the service contract number 070307/2010/58/0092/SER/B2.

The report summarises in short the main achievements of the project. For all of the main parts of the project (except part 4), individual task (part) reports are available. For part 4 (meta documentation of the project data) meta data sheets exist and are included with the data itself.

The interested reader is referred to these detailed reports in case the summary information provided here is not sufficient.

3 Project objectives

The present project has been executed in the framework and context of the EU 2020 biodiversity strategy, the definition of new and measurable targets for the assessment of specific operational biodiversity objectives and sub-targets and the need to better explore the use of existing land cover and land use data for understanding the effects of nature and biodiversity policy in order to provide (spatial and thematic) data for further investigation of land repartition and quality trends in Europe.

Concrete INBALUD aims included:

- Integration of land cover and land use data, including a technical solution;
- Gap-analysis of present data sets and recommendation on EU biodiversity data policy;
- Suggestion on biodiversity in-situ element which would overcome detected gaps.

3.1 *EU biodiversity objectives and targets*

The EU biodiversity objectives and targets have been set to be in line with Strategic Plan for Biodiversity of Convention on Biodiversity (CBD). The former EU 2010 biodiversity target – “to halt biodiversity loss within the EU by 2010” – has not been achieved¹ in accordance to the global biodiversity targets².

On 3 May 2011, the European Commission adopted an ambitious new strategy³ to halt the loss of biodiversity and ecosystem services in the EU by 2020. Six main targets with 20 actions were agreed to help Europe reach its goal.

The six targets cover:

- Fully implement the Birds and Habitats Directives;
- Maintain and restore ecosystems and their services / promote green infrastructure;
- Increase the contribution of agriculture and forestry to maintaining and enhancing biodiversity;
- Ensure the sustainable use of fisheries resources;
- Combat invasive alien species;
- Help avert global biodiversity loss.

The strategy is in line with global commitments made in Nagoya in October 2010, in the context of the Convention on Biological Diversity, and it is based on the two commitments made by EU leaders in March 2010. The first is the 2020 headline target:

¹ Anonymous 2009: Progress towards the European 2010 biodiversity target – indicator fact sheet

² Global Biodiversity Outlook 2 and 3

³ <http://ec.europa.eu/environment/nature/biodiversity/comm2006/2020.htm>

"**Halting the loss** of biodiversity and the degradation of ecosystem services in the EU by 2020, and **restoring** them in so far as feasible, while stepping up the **EU contribution** to averting global biodiversity loss"; the second is the 2050 vision: "By 2050, European Union biodiversity and the ecosystem services it provides – its natural capital – are protected, valued and appropriately restored for biodiversity's intrinsic value and for their essential contribution to human wellbeing and economic prosperity, and so that catastrophic changes caused by the loss of biodiversity are avoided."

Operational objectives

Measurable thresholds are supposed to be set for evaluation of the specific targets' attainment. Because these thresholds are still under discussion, they are not included in this text.

3.2 The situation today

The vision of the 2020 biodiversity strategy and the objectives of this project need to be contrasted with the reality of today's situation of available information which can be explored for better understanding the effects of nature and biodiversity policy.

Most biodiversity related assessments for the EU 2010 biodiversity baseline were based on two main widely used data sources:

- Reporting under the Article 17 of the Habitats Directive and thus the range or distribution of certain species or habitats of Community Importance and the conservation status of species and habitat types of European interest (assessed for the entire area of the EU Member States for the period 2001-2006, published in 2009). The importance of this data is strengthened by the fact, that they cover distribution and conservation status for the whole territory and not only in Natura2000 Sites or protected areas;
- Corine Land Cover (CLC) and derived data sets (e.g. land accounting, ecotones, HNV farmlands). However, knowing the specifications of CLC (nomenclature, mapping scale, minimum mapping unit, minimum change unit) it is to be concluded that CLC is not really adapted to biodiversity needs.

Several other additional data sets were used for SEBI 2010 (e.g. data on bird species) and for the analyses published in the EEA State of Environment Report 2010 and in the EEA Technical Report 12/2010: EU 2010 biodiversity baseline, which are cited in a list of references in the end of each chapter. However, most of them are spatially or temporary limited, i.e. they do not cover the whole EU and/or miss repeated measures using the same methods in the same sites or between time periods (for more see Part 1 Report, chap. 5; Part 6 Report)

The most important limiting factors for the data analyses are as follows:

- temporal limitation (see Part 6 Report) – time series not available for a majority of the data sets;

- difficult comparability of information between countries, due to differences in methods used for data collection (e.g. thematic, spatially resolution; for more see Part 6 Report);
- reporting under the Article 17 reporting has only been done once so far and thus change information is not available, only main trends are indicated. Moreover, some discrepancies in methods cause difficulties in data interpretation (see Part 1 Report, chap. 5.1);
- information in the Standard Data Forms (SDF) is not to be periodically updated. Information on quality of species and habitat types suffers of lack of common methodology across the EU level;
- most of the in-situ surveys, like LUCAS, collect EU-wide potentially relevant information which has not yet been sufficiently considered, but which unfortunately for the most part also lacks time series. Moreover, these EU-wide in-situ surveys were not designed for collection of biodiversity data, and therefore interpretation and analyses are limited (for more information see Part 3 Report);
- a general lack of information on biodiversity (specifically spatial distribution of biotopes and species) especially outside of protected areas at EU-level.

Given this background, the current project aimed at improving the availability of information based on the combination and integration of existing land use / cover data with biodiversity information and gap-analyses of present data sources including potential suggestions on data collection improvement and/or data processing for higher quality results obtained from these analyses.

3.3 *Project expectations*

The Commission wishes to ensure adequate use of existing land cover and land use data for understanding the effects of nature and biodiversity policy, and to provide (spatial and thematic) data for further investigation of land repartition and quality trends in Europe. The work foreseen under the present contract is divided into seven parts, each of which will have contributed to the achievement of these general objectives.

- Part 1: Examine possibilities of how to bring existing data collections together
- Part 2: Produce as useful first result a processed database
- Part 3: Include follow-up work by producing an EU analysis on selected biodiversity parameter
- Part 4: Make sure that meta data will be available for future data use requests
- Part 5: Feed in the future strategy on data storing and handing within European Institution
- Part 6: Formulate input to recommendations on EU biodiversity data policy
- Part 7: Prepare conceptually the biodiversity element of in-situ surveys such as LUCAS

One particular focus of the project was the exploration of LUCAS data. LUCAS originally designed as an agriculture focused survey has limited value for direct biodiversity and ecosystem assessments due to its spatial, temporal and methodological set up. Therefore, one of the special goals of this project was to explore the contribution of LUCAS for biodiversity and ecosystem assessments.

4 Project achievements

In the following chapters the main achievements of the INBALUD project are summarised; where relevant more elaborated task reports have been prepared. All separate reports are submitted together with this report.

4.1 Part 1: Examine possibilities of how to bring existing data collections together

The objectives of part 1 were defined as follows i) to create a list of potentially relevant data sets for assessing information on biodiversity, ii) to examine the integration and joint assessment of LUCAS data with other biodiversity relevant data and iii) to propose a number of biodiversity indicators for improvement.

The main conclusions are summarised in the project inception and task 1 reports.

4.1.1 Assessment of potentially relevant data sets

The following data sets were taken into consideration for integration into the project database.

Data set	Description and relevance for 2020 target
Corine Land Cover	<p>Only European-wide land cover database with time series (1990, 2000, 2006).</p> <p>Potentially relevant information for the INBALUD project: all information highly relevant.</p> <p>Basis of numerous derived data sets, used also for SEBI Ecosystem coverage..</p>
Natura 2000	<p>Sites designated under Directive 92/43/EEC (Habitat Directive): Sites of Community Importance (SCI) / Special Areas of Conservation (SAC)</p> <p>Sites designated under Directive 2009/147/EC (Birds Directive): Special Protection Areas (SPAs)</p> <p>Potentially relevant information for the INBALUD project: site spatial delineation, qualitative and qualitative parameters of the target species/habitats, list of negative impacts and activities on the site, general site character.</p> <p>Support to target 1 of the 2020 biodiversity strategy.</p>
CDDA	<p>Database of nationally protected areas.</p> <p>Potentially relevant information for the INBALUD project: site spatial delineation, IUCN category.</p> <p>Potential support to target 2, and 3.</p>

Article 17	<p>Reporting on distribution, conservation status and other information of species and habitat types of Community interest under Article 17 of the Habitats Directive (based on surveillance under Article 11 of the Habitats Directive)</p> <p>Potentially relevant information for the INBALUD project: distribution of species/habitat types, range trend, trend, main pressures, threats, overall assessment of CS.</p> <p>Support to target 1 of biodiversity 2020 strategy.</p>
HNV farmlands	<p>High Nature Value Farmland</p> <p>Potentially relevant information for the INBALUD project: spatial distribution of HNMF.</p> <p>Support to target 2 (ecosystem services) and target 3 (contribution of agriculture to biodiversity).</p>
LUCAS	<p>Very detailed in-situ survey on land-cover and land-use.</p> <p>Potentially relevant information for the INBALUD project: information on occurrence of biodiversity related land-cover types at fine scale level.</p> <p>Support to target 2 (ecosystem services) and target 3 (contribution of agriculture to biodiversity).</p>

The following data sets have been considered, but were not integrated in the project database.

Data set	Reason for not being selected
EUNIS	Not a data set, but a clearinghouse, compilation from other sources (e.g. N2000). No spatially explicit information.
WISE	Not a data set, but a portal leading to other data sets concerning water.
CAPRI	<p>EU-wide quantitative agricultural sector modelling system, based on well-documented, official and harmonised data sources, especially data from EUROSTAT, FAOSTAT, OECD, FADN.</p> <p>Livestock and nitrogen data were examined but have too much uncertainty to add value to planned improvement of HNV farmlands.</p>
Soil sealing layer	Derived data set, does not support any of the finally selected project indicators.
JRC forest type map	Very detailed (25m) forest map, but does not differentiate between native and introduced forest species. Very relevant data set, but does not support any of the finally selected project indicators.
Urban Atlas	No specific indicator, does not support any of the finally selected project indicators.
Soil database	Important scale difference to selected data sets.

LANMAP	Very complex data set (built from different sources), better suited as reporting unit than as analytical input data
Landscape fragmentation (effective mesh size)	Derived data set, “negative information” – shows where already problems exist,

Knowledge about the existence of the following data sets was gained during the course of the project, which were suggested by the Steering Committee to be included in further assessments, like on Green Infrastructure.

Data set	Reason for not being selected / Relevance for 2020 biodiversity target
JRC Landscape mosaic pattern	Added value to be tested for the project indicators selected for improvement Support to target 2 (ecosystems and their services)
JRC Mosaic pattern of natural & semi-natural lands	Added value to be tested for the project indicators selected for improvement Support to target 2 (ecosystems and their services)
JRC Forest mosaic pattern	Added value to be tested for the project indicators selected for improvement Support to target 3 (contribution of forest to biodiversity)
JRC Ecosystems services map	Data not available to the project, still in research and validation phase Support to target 2 (ecosystems and their services)

“Green Infrastructure” is an umbrella term that describes all elements of an interconnected network of green spaces that conserves natural ecosystem values and functions and provides associated benefits to human populations. It ensures efficient and sustainable use of land by integrating interacting functions or activities on the same piece of land. By giving back space to ecosystems, green infrastructure can maintain and create landscape features which guarantee that ecosystems continue to deliver services such as clean water, productive soils and attractive recreational areas.

Given the broad definition of Green infrastructure (GI) many European data sets are available that could be considered as information sources for and description of Green Infrastructure. Pattern and connectivity models available at JRC were recently used in a JRC research study on the pattern components of green infrastructure and on the impact of greening measures in the CAP reform policy options (Lavalle et al 2011⁴). The following European data sets / data sources that are available were examined:

⁴ Lavalle, C. et al (2011): Implementation of the CAP policy options with the Land Use Modelling Platform – A first indicator-based analysis.- JRC Scientific and Technical Papers.- Luxembourg, EUR 24909

Theme	Data set	Acronym
Land Cover	Corine Land Cover	CLC
Protected Sites	Common Database on Designated Areas	CDDA
	Natura 2000 data - the European network of protected sites	NATURA 2000
Internationally Significant Natural Areas	Wetlands of International Importance	RAMSAR
	Biosphere Reserves	MaB
Internationally Acknowledge Areas	Important Bird Areas	IBA
	Prime Butterfly Areas	PBA
	Important Plant Areas	IPA
Ecological Network	Pan-European Ecological Network	PEEN
Ecotones	Urban natural, semi natural, wetlands transitions (Ecotones)	
	Urban large crops transitions (Ecotones)	
	Urban forest transitions (Ecotones)	
Forest connectivity and change	Forest Pattern, Fragmentation & Connectivity maps and indices year 2006. Change 1990-2000-2006	JRC_Forest
Forest pattern and changes		
Management	High Nature Value Farmland in Europe	HNVF
	Distribution of land take 2000-2006	Land Take
Landscape	Landscapes of Europe (LANMAP2)	LANMAP2
	EuroRegionalMap	ERM
	Landscape pattern	JRC_Forest
Biodiversity	Net Landscape Ecological Potential of Europe and change 1990-2000	NLEP
	Wilderness Quality Index	WQI
Water	Large rivers and large lakes	
	River basin districts	RBD
	European catchments and Rivers network System	ECRINS
	WISE Lakes	
	WISE Rivers	
Pollution	WISE Water Quantity	
	European Pollutant Release and Transfer Register	E-PRTR

For more information see annex on Green Infrastructure to the Interim Report.

Within the ETC SIA implementation plan 2011 GeoVille has reviewed a number of data sets as potential contribution to GI. A dedicated technical paper is available at EEA.

4.1.2 Integration of LUCAS data

The integration of LUCAS sample points with wall-to-wall spatial data is not straightforward. While LUCAS as well as CLC provide information about the land cover / use of a particular area, their spatial reference units are different.

The field of observation for a LUCAS point in the field is a circle of 3 m respectively 20 m diameter. The sample points themselves are defined with a sampling distance between them large enough to avoid spatial autocorrelation, the avoidance of which in fact is the goal of a good statistical sampling design⁵. For Corine Land Cover the minimum mapping unit is 25 hectares, i.e. a surface which is more or less homogeneous and is represented by its major constituent (e.g. buildings, trees, agricultural use). In a simplified way, one may say that LUCAS provides exact information at the site level, while CLC provides generalised information on a broader area.

The other difficulty in integration of these two data sets is in fact, that the concept of the two surveys is fully different. CLC aims at a full coverage of the territory, while LUCAS is a statistical survey, drawing its information from a number of samples following a pre-defined sampling design and serving a specific problem (population) of concern.

LUCAS was designed to be statistically representative at NUTS level 2 for the 7 main land cover classes, therefore LUCAS land cover sample points do not provide a sufficiently large amount of information for large scale (i.e. detailed resolution, small areas) analysis. This means that the density of LUCAS points is not high enough to use LUCAS data to characterise single CLC polygons with respect to their composition. This can be seen by a simple comparison of the number of CLC polygons and LUCAS points for a given area (see Table 1).

Table 1: Number of CLC polygons and LUCAS sample points in selected countries

<i>Number</i>	AT	BE	DE	ES
CLC polygons	30,942	24,826	146,672	151,053
LUCAS points	4,966	1,809	21,138	29,913

But as LUCAS is a statistical survey also any attempt to spatially extrapolate the point data is not correct. Apart from that, most of the LUCAS survey variables are nominal data which cannot be interpolated.

In order to correctly use LUCAS together with other information, the reference area of the second data set needs to be large enough in order for LUCAS points to be representative within. To know these statistical limits of the LUCAS database an assessment of the minimum sample size (number of points) was carried out.

⁵ Spatial dependency is the co-variation of properties within geographic space: characteristics at proximal locations appear to be correlated, either positively or negatively. Spatial dependency leads to the spatial autocorrelation problem in statistics as this violates standard statistical techniques that assume independence among observations. (Wikipedia)

Estimating the sample size meeting the required accuracy

The following information is usually required to be able to calculate minimum sample size:

1. The **Statistical significance criterion** is a statement how unlikely a result must be to be considered significant. Commonly used criterion is 0.05 (i.e., 5 % or a chance of 1 to 20). Other commonly used values are 0.1, 0.01 etc. The statistical significance criterion is also referred to as α -level.
2. **The pivotal attribute** is a selected attribute from the dataset for which you would like to obtain the desired accuracy. If you are concerned with more than one attribute, the calculation of a minimum sample size is performed for all of them and the maximum obtained sample size is selected to ensure that the results will indeed reach the desired accuracy.
3. **Error limit** is the value of acceptable error in percentage from the mean (total, etc.) of the pivotal attribute.
4. **Estimated proportion of classes within population** using post-hoc analysis is an estimate of the class proportions calculated from your database for the region of interest.

Some of these parameters are set by the user (e.g., α -level, pivot/evaluated attribute), while others depend on the conditions of the specific locality (e.g., proportion of classes: for instance proportion of forest differs among Scandinavian, central European or Mediterranean countries). Unfortunately this means that there is no possibility to calculate a single size of "reporting unit" for all statistics based on the LUCAS data. In this case, the so called "worst case scenario" should be used, which is a conservative estimate that ensures reaching the defined accuracy threshold with respect to the selected statistical probability.

The LUCAS project uses a two-phase sampling scheme. The first phase uses a systematic sampling grid point classification to estimate proportions of land cover/use classes within the area of interest (e.g., the areas of NUTS 2 regions). The second phase is a field data collection on subsets of the sample points.

Once the desired precision " d " and the confidence level α are chosen, the respective sample size " n_0 " can be estimated using standard statistical procedures for calculating the sample size for the estimation of proportions and percentages.

There are two different situations. In the first possible scenario all categories are considered to be equally important for the LUCAS project and the user is concerned with overall accuracy. In such a situation the procedure for estimating several proportions simultaneously is used.

Another method of determining the sample size is used in a situation when single proportions are considered to be of interest. It is based on the specification of the margins of error for the item that is regarded as most vital to the survey. It means that we would like to estimate a selected land cover/use class proportion (e.g., "Woodland")

with a defined precision, while the precision of the other classes is not of concern. In such a situation the task simplifies to the binomial distribution, i.e. the sample point either belongs to the class of interest or not.

If the estimated proportions of all classes are known, one may use alternative approaches to estimate the appropriate sample size for such particular case, which commonly, but not necessarily, may lead to less conservative estimates.

Sample size when estimating several proportions simultaneously

All proportions must sum to one $\sum p_i = 1$, where p_i is true population proportion.

Typically, the worst case in terms of the required sample size is a situation in which virtually the whole population is equally distributed among few categories m and some numbers of the categories $k-m$ has $p_i = 0$. The exact case which is worst depends on the α -level chosen, but does not depend on the number k of categories in the population, provided $k \geq m$ (Thompson S.K. 1992). The sample size needed may be obtained from Table 2.

Table 2: Sample size n_0 for simultaneously estimating several proportions within distance d of the true values at confidence level $1-\alpha$ (Thompson S.K. 1987). Other values of d than shown here may be derived from the member $d^2 n_0$.

α	$d^2 n_0$	n_0 with $d=0.05$	n_0 with $d=0.1$	n_0 with $d=0.2$	n_0 with $d=0.3$
0.50	0.44129	177	45	12	5
0.40	0.50729	203	51	13	6
0.30	0.60123	241	61	16	7
0.20	0.74739	299	75	19	9
0.10	1.00635	403	101	26	12
0.05	1.27359	510	128	32	15
0.025	1.55963	624	156	39	18
0.02	1.65872	664	166	42	19
0.01	1.96986	788	197	50	22
0.005	2.28514	915	229	58	26
0.001	3.02892	1212	303	76	34
0.0005	3.33530	1335	334	84	38
0.0001	4.11209	1645	412	103	46

Example: We wish to estimate the proportion of all land cover/use classes at a confidence level $\alpha=0.05$ and precision within 5% from true population parameters. Therefore we would need roughly 510 sample points to reach the required precision.

Estimating sample size based on known proportions

If we take into consideration the estimated proportions of each class using the LUCAS dataset, then another approach to estimate appropriate sample size n is possible. Let p_i be a proportion of class i , $i = 1, \dots, m$.

The maximum likelihood estimator of p_i is:

$$\hat{p}_i = N_i / N$$

where N_i is number of units belonging to the class of interest in population, and N is total number of units in population.

Asymptotically, the vector

$$\sqrt{N}(\hat{\mathbf{p}} - \mathbf{p}),$$

where $\mathbf{p} = (p_1, \dots, p_m)$ and $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_m)$ has m -variate normal distribution with a zero mean vector and a variance-covariance matrix with the elements (see, e.g., Genz 2001)

$$\sigma_{ii} = p_i(1 - p_i)$$

$$\sigma_{ij} = -p_i p_j$$

for $i \neq j$.

If the sample size n is large enough, we can apply the multivariate central limit theorem and the confidence interval for p_i will be (Correa 2001):

$$\hat{p}_i \pm z_{\alpha/2m} \sqrt{\frac{\hat{p}_i(1 - \hat{p}_i)}{n}}$$

Consequently, if we want to provide the same precision d for all categories, then under given α and

$$d_i = z_{\alpha/2m} \sqrt{\frac{\hat{p}_i(1 - \hat{p}_i)}{n}} \leq d$$

we obtain

$$n = \max \left[\left(\frac{z_{\alpha/2m}}{d} \right)^2 \hat{p}_i(1 - \hat{p}_i) \right], i = 1, \dots, m.$$

Note: $z_{\alpha/2m}$ is $1 - \alpha/2m$ quantile of normal distribution.

The number of sample points within a particular class n_i is calculated for all classes and the maximum value is chosen to secure that required precision d is reached.

Example: We wish to know the proportion of the land cover/use class "Woodland and shrubland" with desired precision within 5% from its true population parameters, under α level of 5%. For the selected region we consider the known proportion of the land

cover/use classes (obtained from LUCAS 2009 pre-stratification) and estimate the required sample size n for the example region as shown in Table 3.

Table 3: The estimates of the minimum sample size (n) for given land cover/use class and its representation (Proportion, %) within a given region.

Class	Proportion (%)	Sample size (n)
Arable land	25.4	548
Permanent crops	2.9	82
Grassland	16.7	403
Woodland and shrubland	45.8	719
Bare land	2.1	60
Artificial land	4.0	111
Water	3.1	87

Finally, $n=719$ sample points are selected, which represent the most conservative estimate among the classes with known proportions within the example region.

Conclusion

The results of the statistical assessment is that at least 510 LUCAS sample points are needed if a statement is to be made for the whole database. If only one land cover class is of interest, then this class should be represented at least by 719 sample points within the area of interest.

Considering the average representativity of a LUCAS sample point of 13 km² (see table 4) any entity to be further characterised by LUCAS points should have a minimum area of **6 630 km²** (510x13 km²) respectively **9 347 km²** (719x13 km²).

Table 4: LUCAS 2006 sample density and point representativity by LC class (source: Statistics in Focus 33/2008)

Land cover	Estimated Area in km ²	Sample size in Phase 2	Density in Phase 2	Representativity of sample point in km ²
Artificial Land	122,863	5,592	4.55	22
Cropland	667,369	78,467	11.76	9
Woodland	635,851	19,247	3.03	33
Shrubland	112,934	5,029	4.45	22
Grassland	521,555	50,914	9.76	10
Bare land	78,801	7,702	9.77	10
Water	36,745	1,453	3.95	25
Total	2,176,117	168,404	7.74	13

The entities in which these minimum number of sample points are located can be administrative regions, but also natural regions such as drainage basins, landscapes, biogeographic regions or High Nature Value Farmlands.

Integration of information into the LUCAS database

Since the combined assessment of LUCAS point information with spatially explicit information is limited (i.e. mainly to an area covering at least some 6 600 km² or the respective number of sample points), it was decided to reverse the combination of information and to join additional land cover (CLC) and biodiversity information to the LUCAS point (database).

This will allow the user to put LUCAS data in the context of other information occurring at the same place. The following information was added to each point:

- CLC1990 land cover class
- CLC2000 land cover class
- CLC2006 land cover class
- N2000 site ID
- CDDA ID
- Biogeographic region
- Article 17 Habitats
- Article 17 Species
- HNV farmlands probability value
- Grid cell (1 km²) index (just as a reference, not to be integrated into the OLAP Cube)

Thus for each LUCAS point it is now possible to say in which CLC polygon it falls (i.e. what is the wider land cover context), if it lies in a Natura 2000 site or what is the probability of falling into an HNV farmland. However, also in this approach the statistical rules of the LUCAS survey need to be respected (i.e. representativity).

4.1.3 Indicator proposal

A series of biodiversity related indicators were reviewed against the potential contribution of land cover / use information for improving their message. A common problem to most indicators is that limited information on species and habitats exist and, if it does, time series are not available, while land cover, where time series are available, does not provide "biodiversity" information, but only a proxy for it.

In addition to that, existing data sets and especially LUCAS data were reviewed for their potential to provide supplementary biodiversity relevant information. As LUCAS was originally conceived as an agricultural survey most field survey variables do not provide directly biodiversity relevant information. Even in case they do, the surveyors in the field

often have not been able to differentiate between two sub-classes with different biodiversity values (e.g. forest subclasses).

Potentially useful information, as the survey transects, is limited in practical applications, due to survey of the "wrong" parameters (e.g. woodlands are only mapped with their CX0 woodland type, not with their biotope code (i.e. forest type: CX1 – CXE) and the "statistical" (sampling) character of the survey which impedes the extrapolation of the results to entities of "biodiversity relevant scale". Biodiversity relevant scales, in our understanding, are areas of a given size which provide habitats to species, these habitat areas are mostly much smaller than the $\pm 10.000 \text{ km}^2$ that are needed to encompass the minimum number of LUCAS samples to achieve statistical representativity.

The Eurostat assessment of landscape structure in the EU Member States⁶ uses the number of different land cover types encountered along the 250 m long transect for the calculation of a land cover richness indicator. Unfortunately the results are only presented (and valid) at national level – which has little relevance for real biodiversity assessments needed for local action.

Similar, but geometrically superior, structural information of landscapes is or will shortly be available from other, more detailed sources like the JRC landscape mosaic pattern maps and indices (fragmented/unfragmented pattern of natural & semi-natural lands and patch sizes, interface/ecotone zones, and components of green infrastructure (Estreguil and Caudullo, 2011, Lavalle et al, 2011 and <http://forest.ec.europa.eu/forest-pattern>), the JRC forest pattern maps and indices related to forest morphology, mosaic/interface and connectivity in year 2006 and trends in 1990-2000-2006 period (Estreguil and Caudullo, 2011 in Forest Europe 2011 report and on line at forest.ec.europa.eu/forest-pattern and efdac.jrc.ec.europa.eu/pattern/map), the EEA study on ecotones or the ESA project on small linear features. Except for the EEA ecotones study, which is based on Corine Land Cover all other approaches are based on earth observation data with a pixel resolution of 100m, 25m or better; thus providing a spatially exhaustive coverage of the landscape structures in a continuous form.

Based on some of the aforementioned considerations (e.g. the lack of time series for habitat and species data, the limited added value of LUCAS (transect) information for biodiversity purposes), we identified a selection of biodiversity indicators, which offered the highest potential for improvement. The following indicators were suggested to the project Steering Committee for evaluation and selection of the most relevant ones.

- SEBI No 3 - Species of European interest
- SEBI No 5 - Habitats of European interest
- SEBI No 7 - nationally designated areas, sub-indicator on percentage of PA
- SEBI No 7 - nationally designated areas, sub-indicator on proportion of PAs primarily designated for biodiversity protection

⁶ Statistics in Focus, 21/2011

- SEBI No 8 - Sites designated under the EU Habitats and Birds Directives
- SEBI No 20 - Area under management practices potentially supporting biodiversity (i.e. HNV farmlands)

The results of the implementation of the selected indicators are described under part 3.

4.2 Part 2: Produce as useful first result a processed database

The objective of part 2 of the INBALUD project was to set up a database which serves as the basis for part 3 – the work on selected biodiversity parameters. The part 2 report submitted as separate document presents to full detail on the underlying GIS fundamentals, the storage of the data and tools for accessing and querying the integrated databases.

4.2.1 Data storage

The different datasets (i.e. CLC, LUCAS, CDDA, N2000, HNV farmlands and Article 17) and their metadata are stored as an ArcGIS Geodatabase with the creation of Feature Datasets, Feature Classes and Raster Datasets. This ensures a coherent, consistent database, which can be easily stored and imported to different formats. The data are stored within a DataBase Management System (DBMS), such as MS SQL Server, using ESRI ArcSDE technology, this ensures a permanent remote access to the database.

The SDE Geodatabase can be remotely connected to any ArcGIS system by establishing a new database connection using the following connection parameters:

Server: 158.109.174.72	Username: INBALUD_U
Service: 5156	Password: \$inbalud_user\$

In order to access the project database (including the project outputs) a Geonetwork⁷ web interface, has been set up, to discover the metadata, view and eventually download the indicators. This ensures a visibility and accessibility of the INBALUD results, in line with the INSPIRE specifications and the current European developments in terms of spatial data sharing. For more details, please see the report on part 5, entitled “Feed in the future strategy on data storing and handling within European Institutions”.

4.2.2 Data integration and query

In order to combine the selected datasets and being able to query them together by reporting units, such as NUTS regions or catchments, it was decided to implement OLAP technology⁸, which is based on a multidimensional data model, allowing complex analytical and ad-hoc queries with a rapid execution time.

⁷ Geonetwork is an open-source catalog application to manage spatially referenced resources. More info at: <http://geonetwork-opensource.org/>

⁸ Some OLAP information resources:
http://en.wikipedia.org/wiki/Online_analytical_processing
<http://www.cs.sfu.ca/CC/459/han/papers/chaudhuri97.pdf>

This model integrates all the datasets into the 1 km² Reference Grid⁹ (based on a proposal at the 1st European Workshop on Reference Grids¹⁰) by means of an aggregation or disaggregation according to the nature (resolution) of the data. The transformation into the 1 km² grid is needed for the OLAP cube to be built, and it is not directly linked to the analytical units used to spatialise some of the datasets, e.g. LUCAS data.

Once the data are integrated into the OLAP Cube, they can be queried by one or more reporting units available within the Cube, i.e. by NUTS region, by catchment, by biogeographic region, etc.

In the case of the biodiversity data, the OLAP cube includes the selected datasets and some measures that have been aggregated using a set of dimensions. The dimensions or themes of interest available will be represented by different types: spatial dimensions, e.g. administrative units for Europe (NUTS), biogeographical regions; a number of thematic dimensions: land use data or dominant land cover types; and a third type, which is the temporal dimension, which shows the difference / changes between two reference years.

The scheme in Figure 1 shows the general process of data integration within an OLAP cube. Once an OLAP Cube is created, it can be queried both online or offline (by means of a single file .CUB), which can be connected from common applications such as MS Excel (Pivot Table or Pivot Graph tools) or ArcGIS.

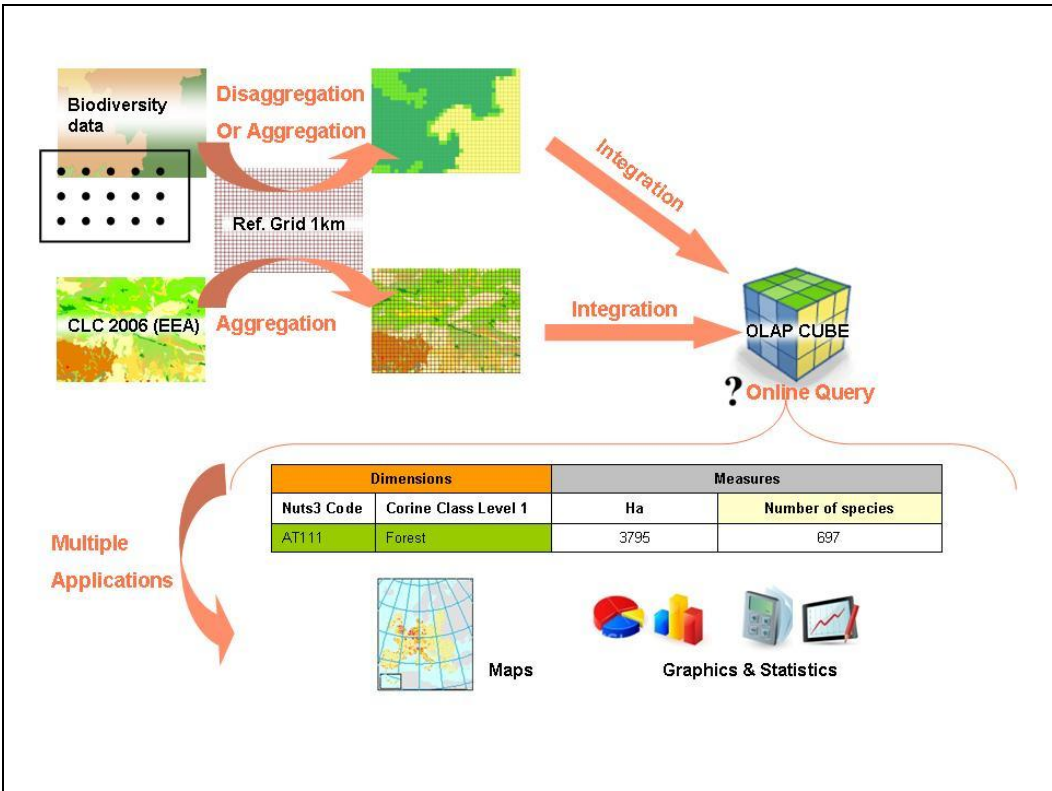


Figure 1: General schema of data integration in an OLAP Cube

⁹ <http://dataservice.eea.europa.eu/dataservice/metadetails.asp?id=760>
¹⁰ http://eusoiils.irc.ec.europa.eu/projects/alpsis/Docs/ref_grid_sh_proc_draft.pdf

Querying the OLAP cube contents is based on the pre-processing of the input data. All data sets (input data, reporting units) are aggregated or disaggregated to the European 1 km reference grid. This grid is the central element which allows the creation of the needed relationships between the different sources and will set up the basic schema of the OLAP Database model.

Figure 2 shows the general process of building the OLAP database.

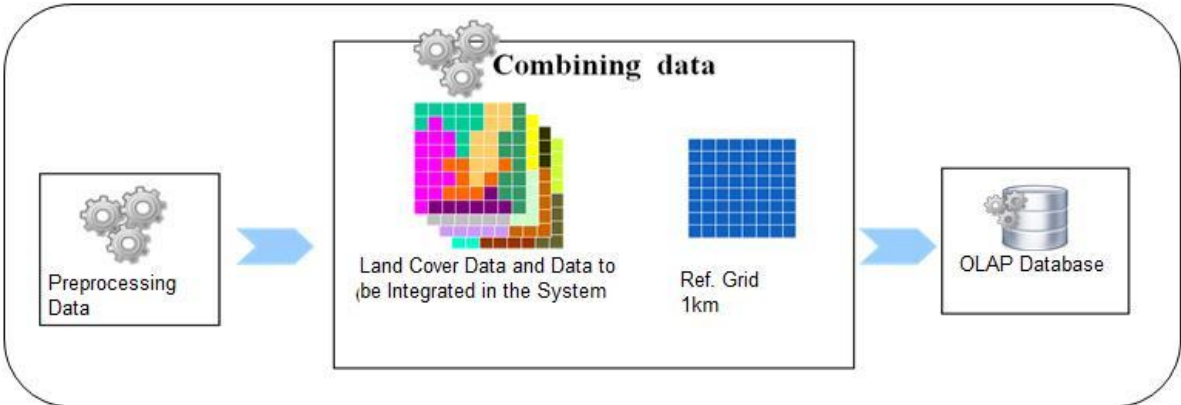


Figure 2: Simplified schema of the OLAP Database built-up

Once the data are integrated into the OLAP Cube, they can be queried by one or more reporting units available within the Cube, i.e. by NUTS region, by catchment, by biogeographic region, etc.

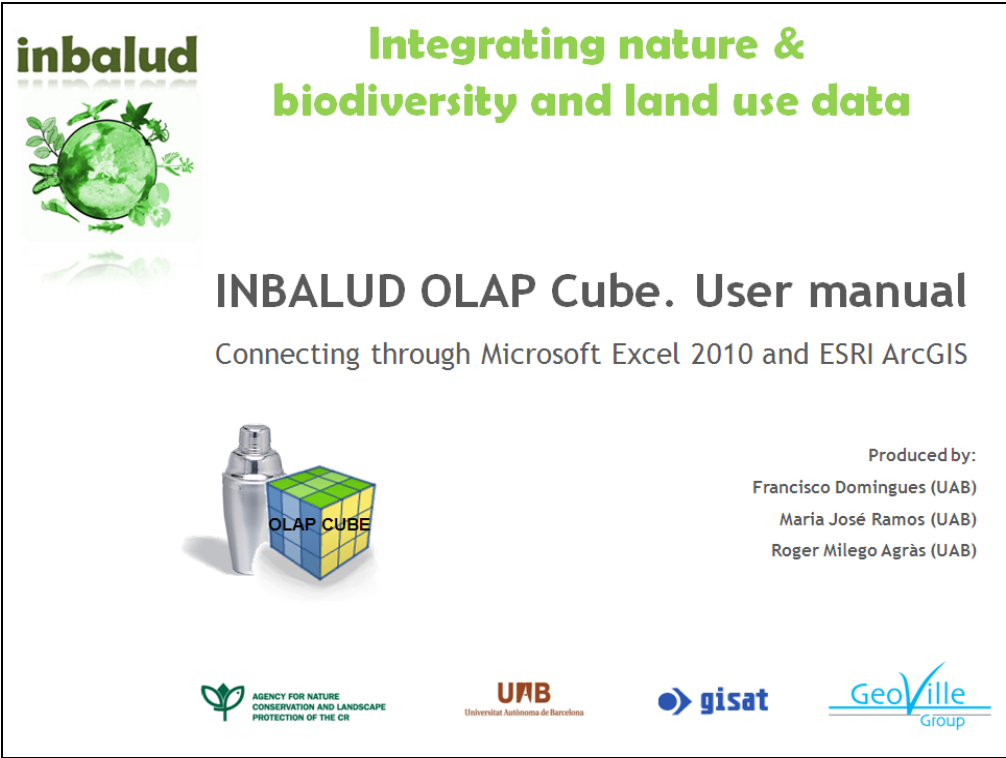


Figure 3: Frontpage of the OLAP manual

The INBALUD OLAP CUBE is delivered in file format (.cub) and it can be queried by using Microsoft Excel. An Excel file with an OLAP connection as well as a manual how to use the OLAP technology is delivered together with this report.

The Steering Committee has confirmed that the enhancement of the OLAP cubes with biodiversity data is fully compatible with the work on-going at EEA and directly applicable by EEA. The data, method and results are in line with the EEA approach (1 km raster) and show a clear link to green infrastructure work done in 2011 by ETC SIA.

4.3 Part 3: Include follow-up work by producing an EU analysis on selected biodiversity parameter

Using the project database included the implementation of new information on selected biodiversity parameters. In part 1 of the project several potential biodiversity related issues (indicators, data sets) were proposed to the Steering Committee which selected the most promising ones (in bold) for actual implementation.

- **SEBI No 3 - Species of European interest**
- **SEBI No 5 - Habitats of European interest**
- SEBI No 7 - nationally designated areas, sub-indicator on percentage of PA
- SEBI No 7 - nationally designated areas, sub-indicator on proportion of PAs primarily designated for biodiversity protection
- **SEBI No 8 - Sites designated under the EU Habitats and Birds Directives**
- **SEBI No 20 - Area under management practices potentially supporting biodiversity (i.e. HNV farmlands)**

4.3.1 Habitat Directive Species and Habitats Index Proposal

The proposal includes an improvement of the indices "SEBI No 3 - Species of European interest" and "SEBI No 5 - Habitats of European interest" by introducing a numerical Habitat Diversity Index (HDI). Similarly to the Red List Index (SEBI No 2), the proposed Habitat Diversity Index shows trends in the overall threat status of species and habitats of European interest. The HDI relates to the proportion of species expected to remain extant in the near future in the absence of additional conservation action.

The Red List Index is a global biodiversity indicator capable of measuring whether the rate of biodiversity loss has been reduced. It has been used as an indicator also at the EU level. It provides simplified and easily communicable figures based on a thorough knowledge of biodiversity change to make informed decisions. The idea is that a single value may summarize the general situation of red-listed species. The Red List index takes a value between 0 (all species "Extinct") and 1 (all species "Least Concern", thus not-red-listed), and it has been devised to indicate the estimated proportion of all species that will survive in a medium-term perspective – all intermediate values. A change in the index means a change in rate of species diversity loss, thus same value at two time

points means a stable, but constant rate of loss. The estimation of the actual absolute value is more complex due to many unknowns, such as the actual number of species.

The proposal of the Habitat Directive Species and Habitats Index (HDI) follows the logic of Red List Index (RLI). Compared to the RLI, using the HDI, even habitat types could be assessed as the methodology/results are uniform to the species assessment.

In both assessment systems, IUCN Red Lists and Habitat Directive (Article 17) Reporting universal categories are present:

RLI: IUCN: LC-NT-VU-EN-CR-EX(RE), DD excluded from RLI;

HDI:Article17: FV-U1-U2; XX excluded form HDI

Therefore their weights are: from 0 to 5 in case of RLI; from 0 to 2 in case of HDI.

The formula of HDI is an adaptation of the RLI formula:

$$HDI_t = 1 - \frac{\sum V_{k(t,a)}}{V_{U2} \cdot N}, \text{ respectively: } HDI_t = 1 - \frac{\sum V_{k(t,a)}}{2N}$$

The resulting index values of HDI lie between 0 (all species "unfavourable-bad") and 1 (all species "favourable").

The HDI is based on the following input data collected under the obligations for monitoring/reporting under

- a) Article 11 of the EU Habitats Directive (92/43/EEC)
- b) Article 17 of the EU Habitats Directive (92/43/EEC)

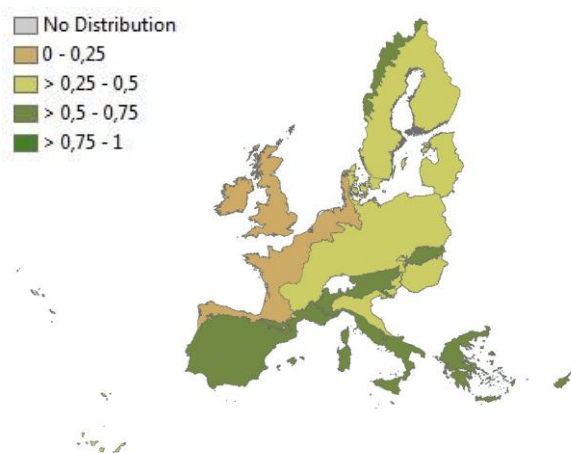
The output of the indicator was calculated for biogeographic regions and member states.

Added value

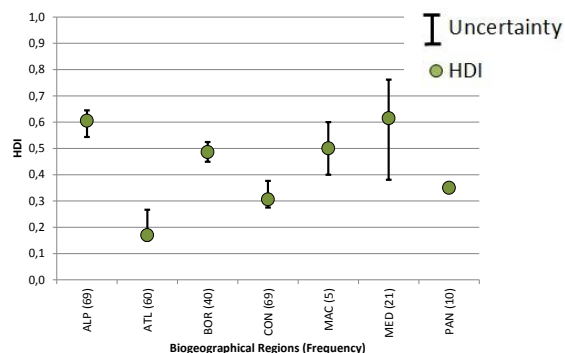
The calculation of the HDI presents a new approach to the assessment of existing information. It is regarded as an "indicator" by the members of the Steering Committee. Nevertheless, in order to avoid confusion, it should be clear that the work done by INBALUD on the SEBI indicators are **proposals**; any final modification of the currently agreed set of SEBI indicators needs to be approved by the SEBI coordination team.

The HDI has a similar added value as the RLI to the general statistics of the red-lists, which exists in cases of IUCN and HD assessments (e.g. in SEBI2010 reports). The resulting index value is a clear and rapid communication means, to be used in policy decision. The index can also be used for trend information through repeated assessments. Methodological consistency in case of Article 17 Reporting ensures the stability of use of the proposed HDI. Such a numerical index was still missing in case of habitats and HDI should solve this lack. An introduction of the index imposes a homogeneity of approaches between indicator No 2, and No 3 & 5.

An Example of the HD-Index together with the display of the associated uncertainty of the assessment is presented below. The report on part 3 contains the full set of maps and graphs for this indicator.



Map 1. HD-Index for "Bogs, Mires and Fens"(Habitats) within Biogeographical Regions.



Graph 1. HD-Index for "Bogs, Mires and Fens"(Habitats) within Biogeographical Regions.

The illustration above suggests an unfavourable status of bogs, mires and fen habitats in the Atlantic region.

The HD Index has been appreciated by the Steering Committee as a new way of presenting information gathered from Habitats Directive Article 17 reporting. There is a possibility to discuss with and present the indicator to DG ENV B3 as well as the SEBI coordination team in the context of the second Article 17 reporting.

4.3.2 Sites designated under the EU Habitats and Birds Directives

The objective of this "measurement" is to address the current state of the Natura 2000 network in terms of quality and quantity of designated sites. The suggested methodological approach is rather general and therefore might be used for scientific evaluations or as background information. The results can be visualised in a map at different scales, which can be adapted to the scope of investigation.

The current indicator provides

- the area and coverage of Natura 2000 Sites as obtained from the N2000 database and
- the proportion of species/habitats of Community Importance evaluated as sufficient during the biogeographical seminars

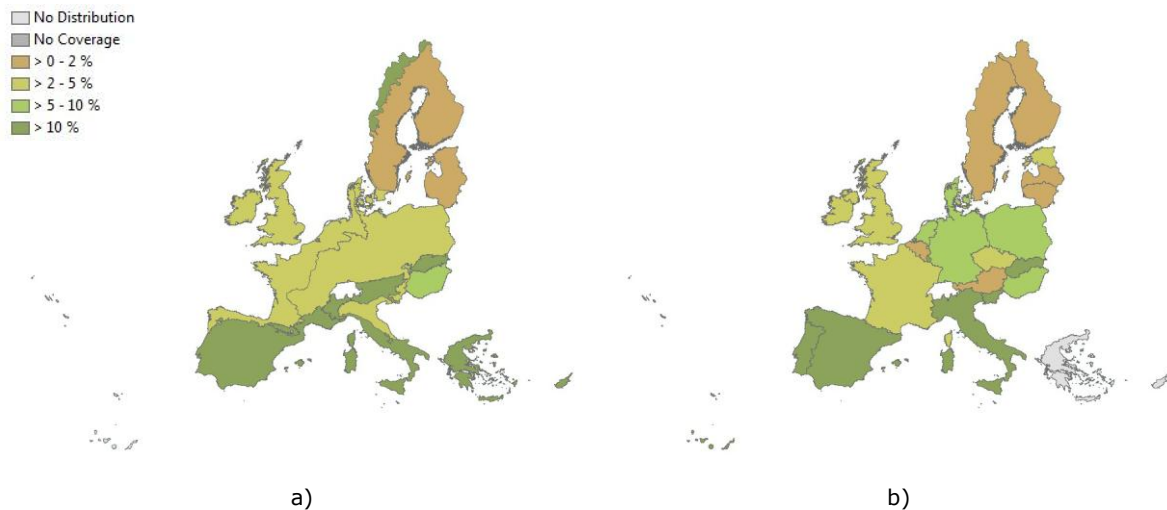
The suggested improvement includes:

- the proportion of total distribution of each species/habitat within the focal area covered by the SCI or SPA and
- the relative quality of the SCIs/SPAs within the focal area counted as weighted mean global quality of SCIs/SPAs within the focal area.

The calculation is based on the following input data collected under the obligations for monitoring/reporting under

- a) Article 17 of the EU Habitats Directive (92/43/EEC)
- b) Natura 2000 sites designated by EU Member States under the Birds Directive (79/409/EEC) and the Habitats Directive (92/43/EEC)

The output is calculated for biogeographic regions and member states, however similar approaches may be used also for other focal areas.



Map 2: Proportion of total distribution habitat type 6210 - Semi-natural dry grasslands and scrubland facies on calcareous substrates (*Festuco-Brometalia*) (*important orchid sites) in Natura 2000 sites for a) biogeographical regions; b) EU 25 Member States. Percentage coverage is visualised.

NOTE: Bulgaria and Romania were not analysed because data on Rep17 are not available.
No distribution - habitat type has no distribution within the area according to Rep17 data, No coverage - no Natura 2000 is designated for this habitat type.

A detailed description of the method and evaluation of examples is provided in the report on part 3. Map 2 provides an example of the results for the proportion of habitat type 6210. Similar results are available for the relative quality of SCI / SPA sites.

Added value

The Natura 2000 network is viewed as a multilayer system consisting of separate networks for each target species/habitats. This view corresponds to the aim of the Habitats Directive because favourable conservation status should be achieved for all the target species and habitats, and thus failure in one case means failure in the conservation aim. Therefore, this approach, when each species or habitat type is evaluated separately, may help to find gaps and strengths of the network more effectively than the overall network evaluation. Moreover, adaptability to different scales makes these characteristics applicable for different political and scientific proposes.

The results of this assessment are regarded by the Steering Committee as important background information, which could be used for instance, as elements of context, for the "new biogeographic process". One of the aims of this process is to discuss how Natura 2000 management can contribute to improving the conservation status of

habitats and species which are considered as a shared concern between the Member states of the given region. This information could also be used in support to discussions within WP3 of the Expert Group on Reporting which aim at assessing the contribution of Natura 2000 to conservation status of species and habitats

The report on part 3 contains the full set of maps and graphs for this assessment.

4.3.3 High Nature Value Farmlands

Europe has significant areas of High Nature Value (HNV) farmland, which provide habitats for a wide range of species. Typical high nature value (HNV) farmland areas are extensively grazed uplands, alpine meadows and pasture, steppic areas in eastern and southern Europe and dehesas and montados in Spain and Portugal. Such areas are under threat from intensification of farming and land abandonment. The mere presence of HNV farmland is of course not proof that current management practices are still sustainable, in fact promoting conservation and sustainable farming practices in these areas is crucial for biodiversity.

One of the short-comings of the current version of the HNV farmland data is the lack of explicit information on intensity of use of grasslands / arable lands.

The CAPRI project¹¹ and its database¹² contain information on crop shares, yields, stocking densities, fertilizer application rates at the level of 150.000 Homogenous Soil Mapping Units (cluster of 1x1 km grid cells). Information on farm management (mineral fertilizer application (kg N/ha), manure nitrogen application (kg N at tail/ha) and nitrogen surplus (kg N after considering volatilization losses/ha) as well as livestock densities were downloaded from the database. These are calculated through a procedure that disaggregates NUTS2 data to Homogeneous Soil Mapping Units.

Added value

By overlaying the HNV farmland data (with a probability > 75% HNV presence) with areas of high livestock densities and/or nitrogen surplus areas with high agricultural intensity could be identified. These areas are potential areas for exclusion from HNV farmlands, due to a high agricultural intensity.

When exploring the feasibility of the actual integration of HNV farmland data and the information on agricultural intensity from the CAPRI project it became apparent that the uncertainty of the CAPRI model output was judged by the Steering Committee to be too high for a reliable integration of the two data sources at the required level of detail. Furthermore, areas identified as HNV farmland include sites, i.e. Important Bird Areas that can be characterised by a relatively high intensity of management and should not be excluded.

¹¹ <http://www.capri-model.org/dokuwiki/doku.php?id=capri:concept:spatialdownscaling>

¹² <http://afoludata.jrc.ec.europa.eu/index.php/dataset>

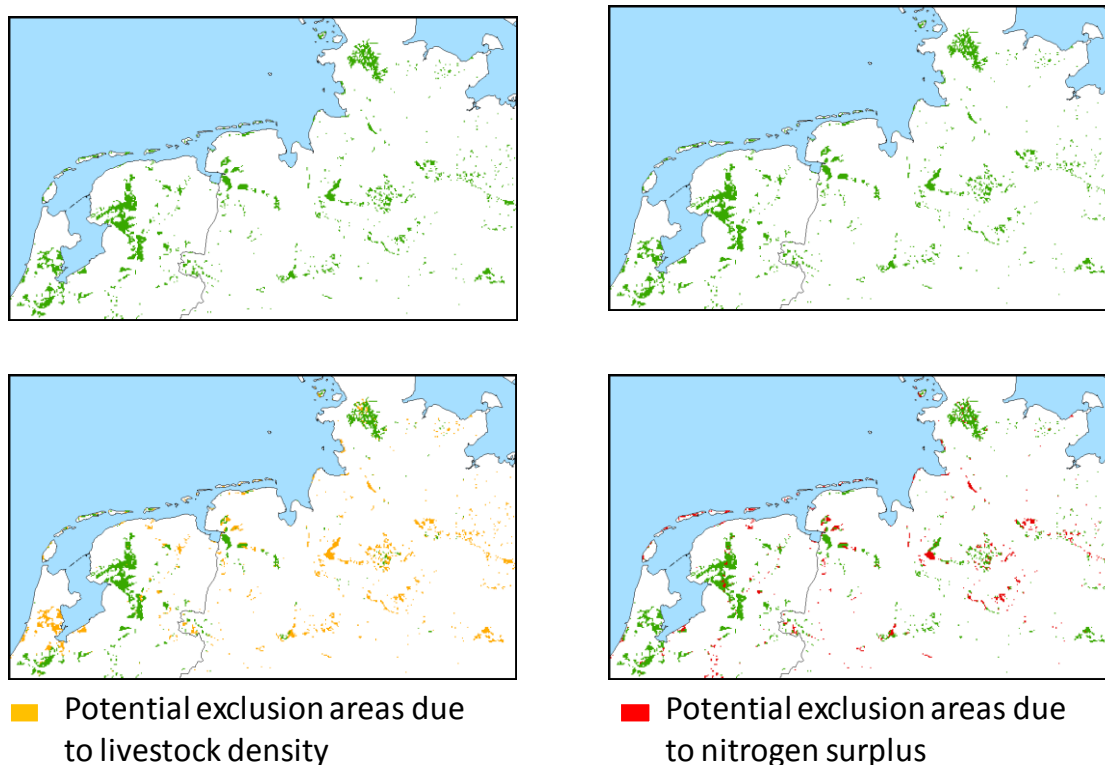


Figure 4: HNV farmlands and agricultural intensity

The further pursuit of this idea was abandoned.

4.3.4 Extracting biodiversity information from LUCAS

In order to assess the potential value of LUCAS information for biodiversity issues, a small interactive script has been developed. The application assigns a “biodiversity value” from 1 - low biodiversity value to 3 - high biodiversity value (based on expert opinion) to selected LUCAS variables (field observations). A first test was carried out for six LUCAS attributes (the INBALUD scores for the LUCAS parameters and their individual class members are provided in the Part 3 – biodiversity elements in in-situ surveys such as LUCAS):

LUCAS field survey parameter	code
land cover 1	bio1
land use 1	bio2
size of the area	bio3
Irrigation	bio4
count of artificial land elements	bio5
count of green land elements	bio6

For each LUCAS point the scores for each of the attributes are collated into a new attribute (CODE) which can be used to map the potential biodiversity value of each point.

This "code" can be summed up (SUM) or the number of particular scores (here 3) can be counted (COUNT SCORES 3)

CODE	SUM	count scores 3
3_3_1_3_2_3	15	4
3_1_1_3_2_3	13	3

Calculating the number of times the maximum score of 3 is reached (min 1, max 6) a map of potential biodiversity was calculated and compared to the map of Wilderness areas produced by the University of Leeds and downloaded from the EEA "maps and data" service (see Figure 5).

Aggregating the results to NUTS 2 regions a map of potential biodiversity of European regions can be created (see Figure 6). When looking at this map the selection of input LUCAS attributes as well as the biodiversity value score need to be kept in mind, both can be easily changed in the application using expert opinion and thus the resulting map adapted to different questions and assumption.

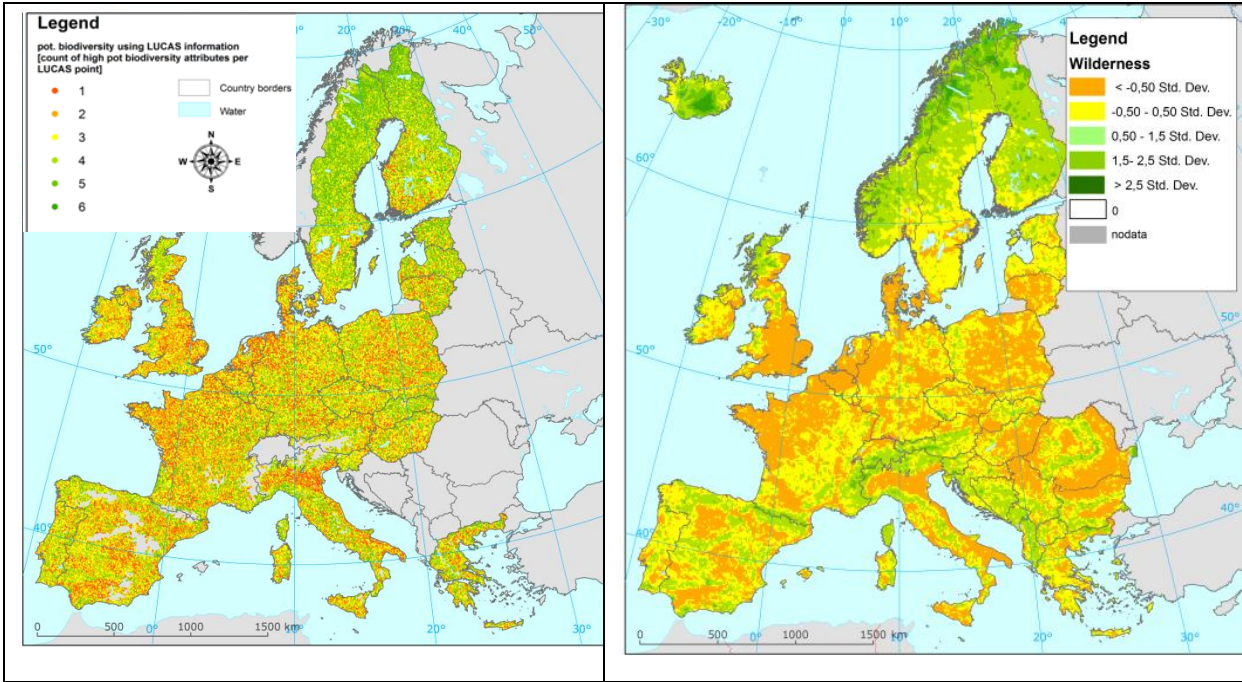


Figure 5: Comparison of LUCAS potential biodiversity (left) and Wilderness Quality Index (right)

In order to interpret the classification of the regions in Figure 6 the individual scores per LUCAS parameter are needed as well as their contribution to the overall result. Table 5 provides for one region how often a score of "3" (from 0 to max. 6 times) occurs for each LUCAS parameter within a given NUTS 2 region (e.g. for the Spanish region Castilla-La Mancha (ES42) the score 3 occurs 2.73 times for the land cover parameter. The overall average of score 3 counts is 2.61, which is the basis for display in Figure 6.

Additional examples can be found in the Part 3 report.

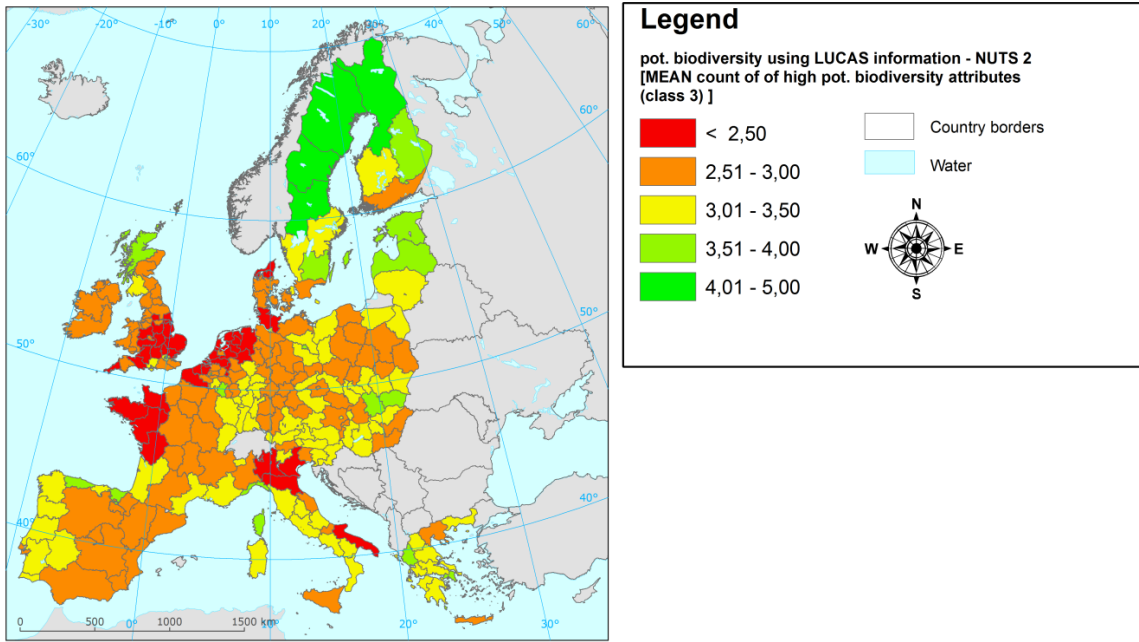


Figure 6: Potential biodiversity of NUTS 2 regions

Spain – orange vs. green						
ES42 (mean: 2,61)						
count of score 3	land cover	land use	size	irrigation	red EL	green EL
0	0,00	0,00	0,00	0,00	0,00	0,00
1	0,02	0,00	0,01	2,56	0,16	0,17
2	1,15	0,51	0,20	3,94	1,77	1,08
3	2,73	0,74	0,34	4,92	2,63	3,47
4	2,89	0,94	0,24	3,20	2,65	2,92
5	0,83	0,79	0,10	0,84	0,83	0,82
6	0,01	0,01	0,01	0,01	0,01	0,01
Total number of samples:	4817					

ES61 (mean: 2,7)						
count of score 3	land cover	land use	size	irrigation	red EL	green EL
0	0,00	0,00	0,00	0,00	0,00	0,00
1	0,11	0,01	0,10	1,82	0,22	0,32
2	1,55	0,08	0,40	3,63	1,42	1,31
3	2,99	0,51	0,52	4,53	2,18	3,00
4	3,08	1,33	0,28	3,31	2,32	2,94
5	1,40	1,36	0,07	1,40	1,38	1,39
6	0,01	0,01	0,01	0,01	0,01	0,01
Total number of samples:	5310					

ES12 (mean: 3,67)						
count of score 3	land cover	land use	size	irrigation	red EL	green EL
0	0,00	0,00	0,03	0,29	0,00	0,00
1	1,68	0,00	0,26	2,26	0,03	0,29
2	3,34	0,83	1,43	3,99	1,27	1,11
3	6,19	3,53	1,61	6,27	3,14	4,35
4	3,24	2,90	1,01	3,24	2,72	3,08
5	0,60	0,60	0,60	0,60	0,60	0,60
6	0,00	0,00	0,00	0,00	0,00	0,00
Total number of samples:	643					

Table 5: Individual score “3” realisations per NUTS 2 region

An advantage of using LUCAS information is the availability of time series to derive trend information, independent from Corine Land Cover.

We are fully aware of the sensitivity of the results to the “biodiversity values”, but using the script these values can be freely chosen and adapted to the topic of interest. Despite of – or maybe even because of – this interactivity the Steering Committee has judged this LUCAS biodiversity exploratory exercise to be interesting which should be followed up with some more concrete objectives. Especially the fact to see that two completely different data types (i.e. wilderness map and LUCAS) exhibit such a high degree of similarity was striking. In terms of usefulness, this the map is probably an additional layer that could be used in contribution to the Green Infrastructure debate, knowing that it will be easily updated in the future and thus enable to show trends which are often lacking from other data sets.

4.3.5 Characterising other data sets with LUCAS

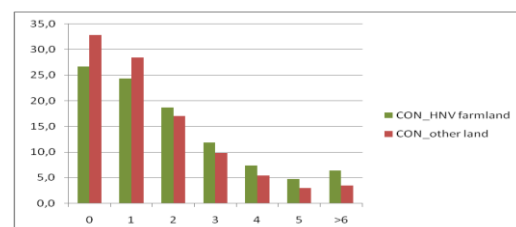
One of the findings of the assessment of statistical validity of LUCAS points was the need to have at least 510 sample points within one object / theme under investigation (see chap. 4.1.2). This requires significantly large enough entities (not single land cover polygons).

In order to test the added value of LUCAS information for biodiversity a characterisation and comparison of HNV farmlands and their surroundings based on LUCAS points was undertaken. The following LUCAS information was used for the characterisation:

- Transect information (number of artificial linear elements)
- Presence of grazing
- Grasslands with / without trees
- Crop types

The assessments were done for LUCAS points within HNV farmlands and outside HNV farmlands within a given biogeographic region. The different realisations were compared. The results show that – with respect to the selected parameters – there is little difference inside and outside HNV farmlands, except for the presence of grazing in HNV farmlands.

- The number of artificial elements per LUCAS transect shows no significant difference inside and outside HNV areas.



- The presence of grazing at LUCAS points within HNV farmlands is significantly higher than outside.

Biogeo. regions	HNV farmland (No. of points)	HNV farmland percent of total	Non-HNV land (tot. points)	Non-HNV land percent of total
ALP	960	10,6	4307	6,5
ATL	1753	36,3	30785	25,4
BOR	604	7,5	29121	3,6
CON	5404	19,1	44444	6,6
MED	9300	27,3	25211	12,1
PAN	758	17,0	4568	3,9

- There is little difference in the presence of trees on grasslands inside and outside HNV farmlands

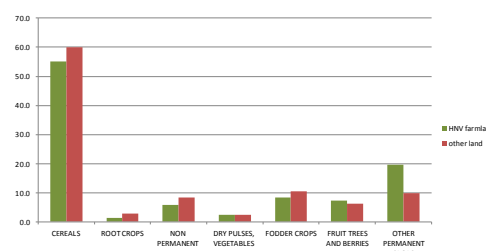
Some differences exist for the alpine and the continental region.

For the boreal and pannonian region not enough sample points are available.

Biogeo. Regions	HNV farmland (No. of points)	Presence of trees (E10)	Non-HNV land (No. of points)	Presence of trees (E10)
ALP	519	12,5	991	20,1
ATL	1269	9,3	2007	13,4
BOR	258	21,3	1885	23,2
CON	2902	11,0	5590	17,2
MED	2509	41,8	1742	41,4
PAN	344	20,9	504	32,5
TOTAL	7801		12724	

- With regard to crops there is little difference inside and outside HNV farmlands.

A significant difference exists only for "other permanent crops".



A similar exercise was carried out for a combination of Corine Land Cover and LUCAS. The identification of "high nature value" forests by differentiating natural and semi-natural forests vs. non-native (LUCAS code CXE) failed due to the fact that almost no non-native (or introduced) forest types had been recorded. Almost 45% of all forest sample points in the LUCAS database had no information on forest types. IN DE, AT and CZ only 4 sites out of 8000 were classified as non-native, which is unrealistic if we compare these results with information published by Ministerial Conference on the Protection of Forests in Europe¹³, stating that on average 13% of the European forests area are dominated by introduced species (with figures for Germany, Czech Republic and Austria ranging below 5%).

A possible reason for the high amount of points without any information on forest type (ca. 37,000) might be the lack of expertise of the surveyors to identify the three non-

¹³ MCPFE, 2007. *State of Europe's Forests 2007*, the MCPFE Report on Sustainable Forest Management in Europe, jointly prepared by the MCPFE Liaison Unit Warsaw, UNECE and FAO Ministerial Conference on the Protection of Forests in Europe, Warsaw.
http://www.foresteurope.org/filestore/foresteurope/Publications/pdf/state_of_europes_forests_2007.pdf

native forest species of interest: except Eucalyptus, the other two coniferous species might be difficult to identify for non-experts. In combination with the mapping criteria (i.e. "all the following criteria have to be met: one or two species at plantation, even aged, regular spacing, systematic thinning regimes") it might be difficult to correctly identify this class.

An exercise similar to the Eurostat landscape structure (land cover richness) assessment¹⁴ was carried out to assess the heterogeneity of CLC classes (grasslands, forests). By determining the number of "intersections" (changes between recorded landscape elements) along the transects falling into the particular land cover class, the heterogeneity of the land cover classes in Germany, Austria and the Czech Republic were compared. Interestingly enough the heterogeneity is rather similar in all three countries, but again the results are only valid at country level and regional differences might apply. Heterogeneity is one of the most important and widely used concepts in ecology. It is highly related to ecosystem conditions (e.g. health) and features wildlife habitats.

The method described here is a first attempt to assess the structural diversity of landscape elements (i.e. forests, grasslands) below the resolution threshold of CLC. So far, the link between the heterogeneity value and whether a landscape is highly valuable from the biodiversity point of view has not been fully established. This would need further research.

However, such basic information about the general character is needed when studies on broader areas are carried out in order to compare the areas among each other. Similar general indices for mainly comparing large areas are also used by SEBI.

These indices are also used as background information for other studies and for general overview of the situation at the broader level as well as for comparison of data from different dates and for the evaluation of changes in landscape structure over time.

Table 6 and Table 7 provide a first overview on the heterogeneity (number of "intersections") in European forests and the share between red (environmentally less favourable) and green intersections.

The lowest proportion of red intersections was found in boreal countries (Estonia, Finland), which are less modified by human management, while the highest proportions

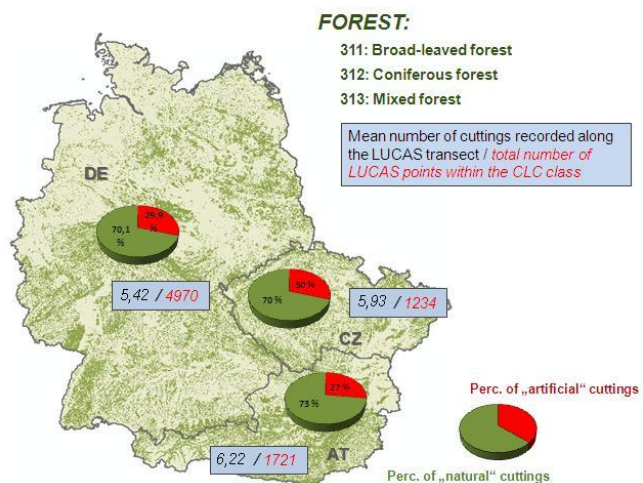


Figure 7: Heterogeneity of CLC forest areas

¹⁴ Statistics in Focus, 21/2011

of red intersections are found in densely populated and thus highly human influenced countries (Netherlands, Belgium).

According to this first investigation, it seems, that suggested the method corresponds well with the general assumption and therefore might be used also for other land cover classes for the evaluation of human influence at broader scale.

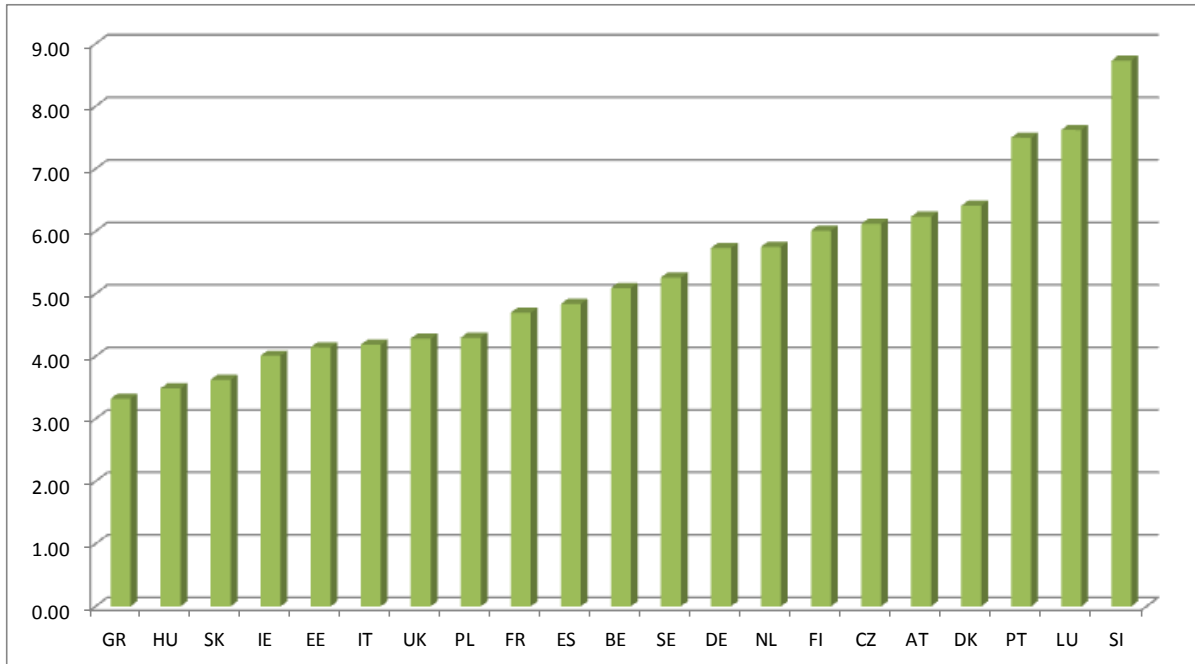


Table 6: Number of "intersections" in European forests

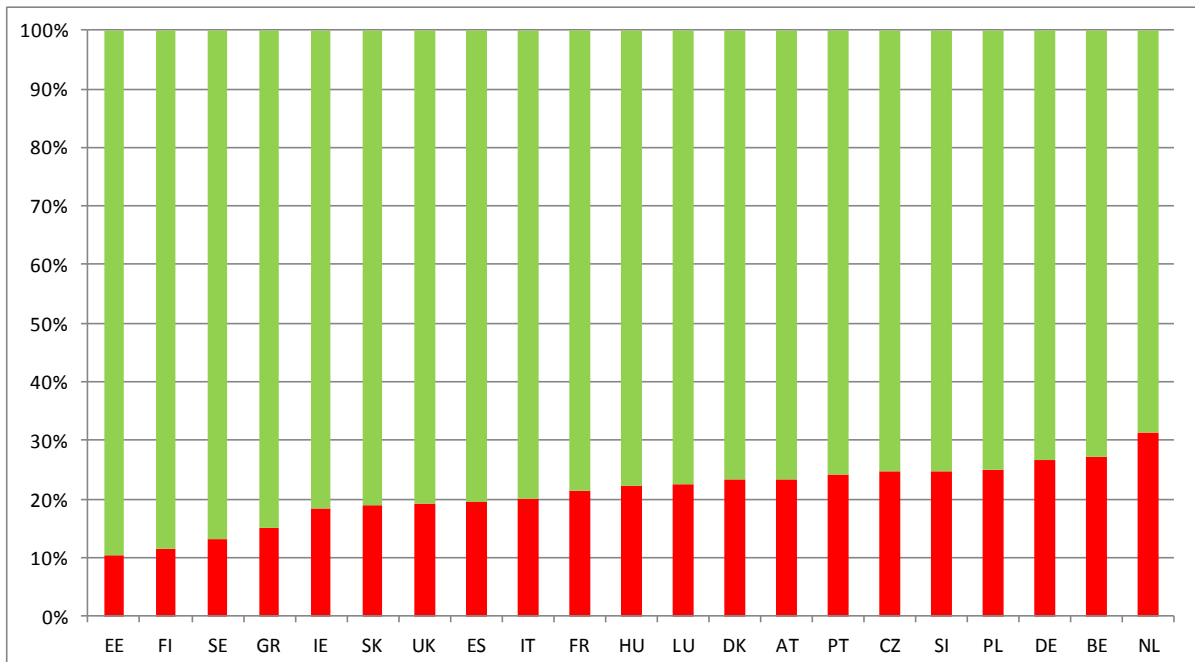


Table 7: Share of red and green "intersections" in European forests

4.3.6 Conclusions

The results of part 3 of the INBALUD project have been successful with respect to improving selected information on biodiversity, i.e. SEBI indicators 3 & 5 via the Habitat Diversity Index (HDI). The development HDI was recognised by the Steering Committee as a new indicator, the characteristics on: i) the proportion of total distribution of each species/habitat within the focal area covered by the SCI or SPA; and ii) the relative quality of the SCIs/SPAs within the focal area, were judged as valuable background information to the existing indicator.

With respect to the integration of LUCAS information with biodiversity data the results are mixed. Due to the "scale" of LUCAS data (i.e. the representativity at NUTS level 2 only) the added value LUCAS is limited. In the context, the transect information has probably the biggest potential (see heterogeneity assessment), if both data types (statistical and spatial) data are brought together in a meaningful manner.

Due to the "scale" differences and the "statistical" nature of the survey, LUCAS cannot profit from the very detailed and localised information. This might change with the introduction of the high resolution layers from GIO – the GMES Initial Operations. These layers will be based on a pixel resolution of 20m which is in the range of the LUCAS in-situ observation. Using the full resolution of these high resolution layers it might become possible to add a spatial dimension to LUCAS.

Joining information from biodiversity and land cover / use data sets to the LUCAS sample points was seen by the Eurostat LUCAS team as an added value for the database:

1. Any assessment can be done in a database environment, no GIS required;
2. The statistical properties of LUCAS remain unchanged;
3. Assessments of the LUCAS points can be widened, i.e. general statements about e.g. protected areas (not for individual sites);
4. In combination with CLC the LUCAS points can be viewed in their larger context

The Eurostat LUCAS team sees a significant potential for further development and testing of the data (i.e. integration of other data sources with LUCAS points, biodiversity exploratory exercise) using categorical analyses. These analyses could be the base for new publications addressing specific questions which should be defined together DG ENV, JRC and EEA.

Overall the Steering Committee has expressed its satisfaction with the project achievements with this central part of the project, including a critical overview and comparison of existing datasets on biodiversity and land cover / use. In addition to that, the analysis of LUCAS data integration and its limitations (i.e. minimum sample size) was highly appreciated.

4.4 Part 4: Make sure that metadata will be available for future data use requests

The objective of part 4 of the INBALUD project was to ensure the availability of a proper documentation of data integrated in the INBALUD database via the standard meta data including input data sets (i.e. CLC, LUCAS, CDDA, N2000, HNV farmlands and Article 17, ...) and the outputs (indicators). Additionally, the processing steps implemented to produce these indicators have been documented using standard Indicator Fact Sheets.

4.4.1 Meta data

Both input and output data integrated in the INBALUD database are accompanied with the standard meta data. The meta data description complies with the INSPIRE Meta data standards - the INSPIRE Meta data Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119 V. 1.2.¹⁵ Meta data are also fully compliant with the European Environment Agency Meta data Standard for Geographic Information (EEA-MSGI), which is a profile of the ISO 19115:2003/19139 meta data standard. Detailed guidelines for EEA standard meta data can be found at ETC - IT infrastructure documentation Wiki¹⁶.

Figure 8: The INBALUD Meta data Editor

¹⁵http://inspire.jrc.ec.europa.eu/documents/Metadata/INSPIRE_MD_IR_and_ISO_v1_2_20100616.pdf

¹⁶https://taskman.eionet.europa.eu/Help/wiki/catalogue/metadata_guidelines

A number of implementations (web or desktop based) of the INSPIRE Meta data standards currently exist, but for the purpose of the INBALUD project and easy meta data filling/validation, an INSPIRE compliant web meta data editor has been launched based on EUOSME (version 1.0.2)¹⁷ as seen on Figure 8 below. The INBALUD Meta data Editor is freely available at <http://metadataeditor.gisat.cz/>.

The meta data contains all optional elements offered by ISO 19115 and INSPIRE, which are regarded as useful. Some of the main aspects covered in that meta data profile are: name of data, description of data, coordinate reference system, source and methodology, responsible party, ownership, and copyrights. Besides that, INBALUD meta data elements compliancy was also validated against the meta data validator included in GeoNetwork v2.5.

4.4.2 Indicator Processing Documentation

The indicator processing is already reflected on general level as part of the lineage element description in the standard metadata. However, for practical purposes this information is often of limited use as it is very general. Therefore for each indicator developed in the INBALUD project a more extensive description including data processing steps has been prepared. This description provides an external user with far more insight into the processing methodology at technical level - information necessary for subsequent indicator reviews or updates. Documentation is based on standard EEA Indicator Fact Sheet template including data flow and data processing flow diagram, links to scripts and functions description as well as additional detailed description of resulting datasets. See example of the Indicator Fact Sheet in Figure 9.

¹⁷ *The European Open Source Metadata Editor (EUOSME) is a web application to create INSPIRE-compliant metadata in any of 22 European languages. It has been developed by the Joint Research Centre as part of the EuroGEOSS project (www.eurogeoss.eu).*



INDICATOR FACT SHEET

Last updated: 30.01.2012

Indicator description

1. Indicator identity

1.1. Title:

SEBI No 3 UPGRADE PROPOSAL - Species of European interest - Habitat Directive Indicator (HDI) (INBALUD project upgrade proposal)

1.2. Code:

SEBI No 3 - HDI (INBALUD proposal)

2. Indicator owner

2.1. Owner:

INBALUD project partners

2.2. Contact information

INBALUD project

Contact: GeoVille Environmental Services

3, Z.I. Bombicht

L-6947 Niederanven

Luxembourg

www.geoville.com

Contact person: Stefan Kleeschulte <kleeschulte@geoville.com>

Phone: +352 26 71 41 35

2.3. Copyrights

INBALUD Project - Indicator Fact Sheet - SEBI No 3 Upgrade - Species of European interest / Habitat Directive Indicator (HDI)

1

Figure 9: The INBALUD Indicator Fact Sheet (example: HDI Habitat Directive – Species, page 1)

For the data and data processing flow diagram (part of the INBALUD Fact Sheet) see an example in Figure 10.

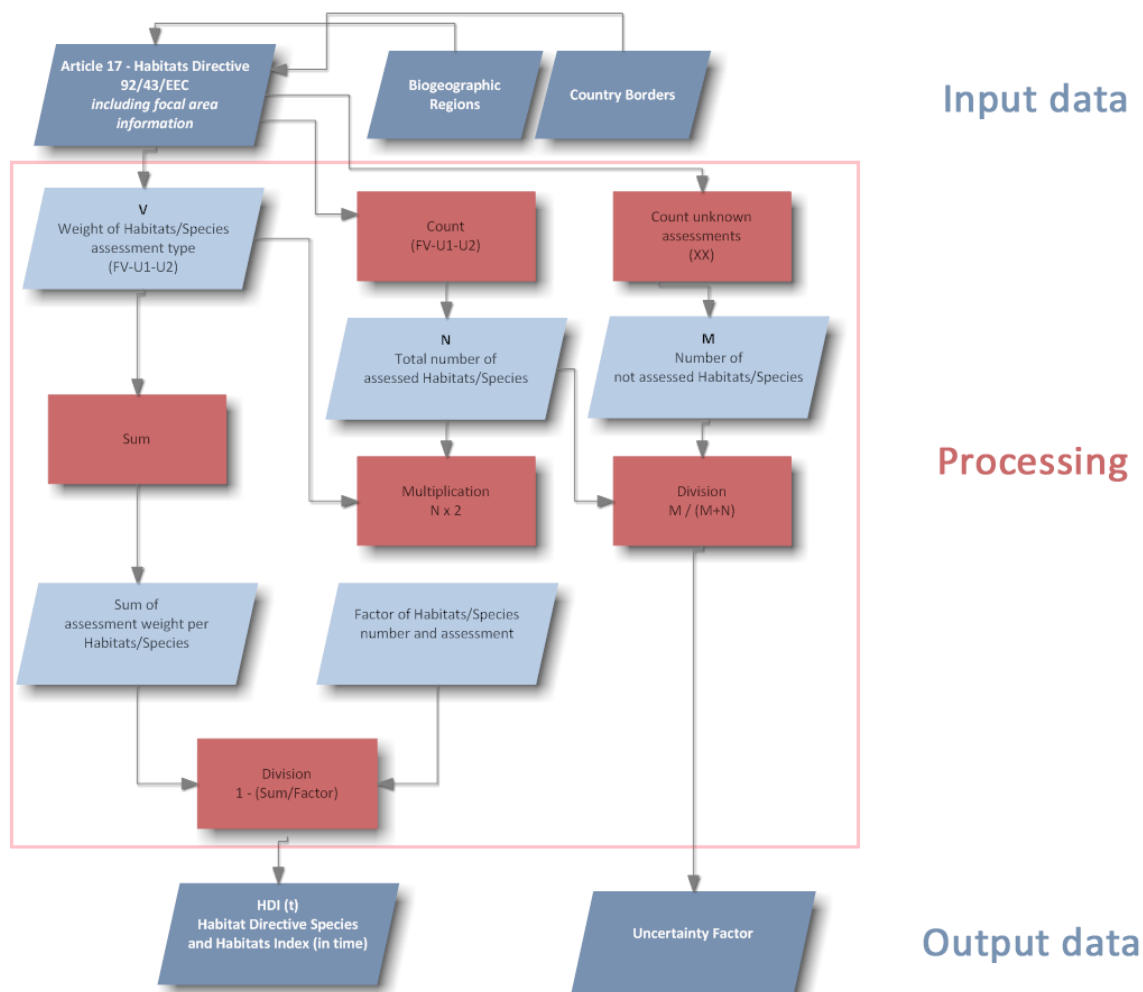


Figure 10: The INBALUD Indicator Fact Sheet – Data and processing flow diagram (example: HDI Habitat Directive – Species)

4.4.3 Conclusions

Data integrated and indicators developed within the INBALUD project are documented using standard INSPIRE compliant Metadata. In addition, produced indicators are further described via the INBALUD Indicator Fact Sheets. Metadata are part of the data package implemented via the GeoNetwork. More about indicators can be found in Part 3 Report. Public availability and accessibility to the project results is ensured by the GeoNetwork service (see Part 5 for details).

4.5 Part 5: Feed in the future strategy on data storing and handling within European Institutions

The objective of part 5 was to ensure that the project data are delivered in a format compatible with standards used by the European Commission and the EEA (i.e. INSPIRE, EEA meta data catalogue, SEIS, EFDAC).

The INSPIRE directive defines the core services that should be provided in order to build the Infrastructure for Spatial Information in the European Community. These services

include discovery services, visualization services, download services which have been implemented by using GeoNetwork Open Source software¹⁸. This software supports the creation of a metadata catalogue, implementing all the relevant standards to obtain interoperable services.

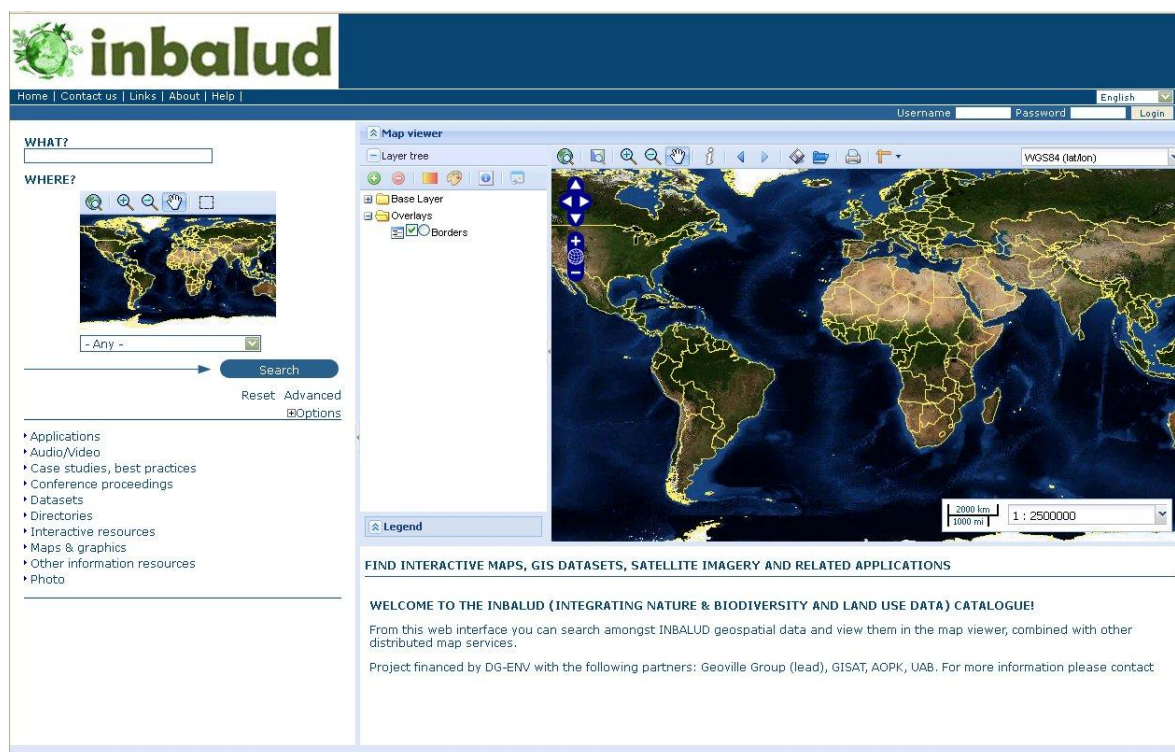


Figure 11: INBALUD geonetwork interface

GeoNetwork has been developed to connect spatial information communities and their data. The software provides a web interface to search geospatial data across multiple catalogues, combine distributed map services in the embedded map viewer and publish geospatial data using the online metadata editing tools.

The INBALUD GeoNetwork service can be accessed via:

<http://etcSDI.uab.es:8080/inbalud>

The part 5 report provides detailed information about the technical solution used for storing the INBALUD data and outputs (indicators).

4.6 Part 6: Formulate input to recommendations on EU biodiversity data policy

The aim of part 6 was to describe main gaps in data availability and data quality on biodiversity at the pan-European level and to recommend how to overcome these insufficiencies. Since biodiversity data and their analysed outcomes – indicators, are

¹⁸ <http://geonetwork-opensource.org/>

important to measure progress of 2020 EU biodiversity targets – links between the data and indicator are provided.

A gap analysis on information needed to support good biodiversity policy making revealed the following issues:

- The number of suitable datasets available at the European level is limited.
- There is a general lack of information on biodiversity (specifically distribution of natural and semi-natural habitats and rare/threatened species) especially outside of protected areas.
- Insufficient spatial resolution and thematic classes limit the value of CLC for more accurate biodiversity related assessment.
- Insufficiency in data and information on current status of phenomena of Community importance within the protected areas (Natura 2000 sites).
- There is no dataset concerning the presence and distribution of alien and/or invasive species with finer resolution than countries.
- Several shortcomings linked to poor harmonization of inventory data over the different countries remain.
- Many other types of data are needed to assess ecosystem services
- Unsatisfactory information on agriculture and forestry supporting biodiversity causes insufficient assessment of High Nature Value Farmland and Forests.

The following recommendations are made by the INBALUD project:

Data sources:

- Improve knowledge about species and habitats outside protected sites by implementing dedicated campaigns for collection of such information, e.g. as proposed by SATCHMO¹⁹ in the context of geoland2 (i.e. area frame sampling).
- Improve spatial and thematic scale of land cover information by fostering the bottom-up exchange of information collected at Member State level through INSPIRE.
- Support data harmonization in the EU
- More detailed information on particular conservation measures set by management plans
- Improvement in understanding of the impacts of climate change, link between biodiversity and ecosystem services is necessary
- Data policy and management
 - Data collected under legislative obligations by Member States (e.g. reporting under article 17 of the Habitats Directive) and data collected at the EU level supported by the EU institutions (e.g. CLC) are important for the EC information support. Therefore, a special attention is to be given to them.
 - Investigate practical solutions to overcome above mentioned data related issues.

¹⁹ <http://www.gmes-geoland.info/service-portfolio/seasonal-change-detection-products.html>

- Continue to support surveillance schemes undertaken by the voluntary and academic sectors.
- Determine target datasets with significant value and arrange a licence from data providers or holders.
- Avoid duplications
- Data generation, quality control and assessment are limited by financial resources. To assure the availability of key datasets the EC should declare financial schemes for their production and maintenance.
- Improve the coverage and consistency of monitoring systems across Europe.
- Establish a process of regular and compulsory updating of the Standard Data Form (SDF).
- Improve data sharing and public access to biodiversity data and associated environmental information.
- Encourage the creation and provision of meta data.
- Completeness of data sets should be encouraged.
- Core set of data sets should be updated periodically to create time series to assess trends.
- Harmonized approaches for ecosystem services, climate change, HNVF and Green infrastructure are needed to gain unified and comparable information across the Member States.
- Better integration of earth observation based data.

Detailed data analyses and concrete suggestions how to overcome the found insufficiencies are described in the report on part 6. The Steering Committee appreciated that the recommendations on EU biodiversity data policy contain a lot of information that could be used in support to a gap analysis in the perspective of the 2020 strategy; only a few recommendations need to be worked out in some more detail before that.

4.7 Part 7: Prepare conceptually the biodiversity element of in-situ surveys such as LUCAS

Based in particular on the work done in part 3 and 6, concrete input to recommendations how to improve the biodiversity element should have been prepared for the next LUCAS campaign.

As the next LUCAS campaign is already planned for 2012, any contributions needed to be made already before summer 2011. Consequently the recommendations made in part 7 could not be fully based on the experiences made in the project.

Nevertheless, the project provided in May 2011 some concrete suggestions for an additional differentiation of grasslands taking into account wetness, presence of shrubs and the intensity of use. The proposal was further discussed with the Eurostat LUCAS team, but finally abandoned as the differentiation of grassland types was judged to be too complex for the non-expert surveyors.

The part 7 report first provides a general overview of different biodiversity monitoring schemes and how the LUCAS survey fits into these. The report moves on to the EU 2020 biodiversity strategy, its targets and how these can be linked to biodiversity data. As the control of meeting the targets (in general meaning target 1 “the protection of biodiversity has to be based on relevant information on its status”) has to be based on relevant biodiversity data, a systematic approach and data collection is urgently needed. The current data are collected unsystematically in various contexts and purposes. The LUCAS-biodiversity element would promote strongly such a systematic approach based on the experience in linking field survey and remote sensing information, both at national (e.g. UK Countryside Survey²⁰ or Biodiversitätsmonitoring Schweiz²¹) and international level (e.g. EBONE project²²), Especially species information in Europe is rather scattered and unequal in quality, the introduction of general EU-wide schemes would be of great importance.

Finally the report presents four different options of a biodiversity survey which contains different levels of complexity. Out of the four options two are more realistic of implementation:

- 1) stage 0 - no change in the methodological approach, but strictly following of the methodology, but with an emphasis shifted to biodiversity data collection;
- 2) stage 2 – use of trained staff (biodiversity specialists) and a new dedicated methodology looking into a selected number of key species.

Stage 2, the real (although reduced) biodiversity component is based on the introduction of specific field-survey point/transect expert activities devoted to monitoring habitat and species groups.

The stage 2 option would compromise the following data collecting activities:

- In case the LUCAS point is a natural area or natural habitat, the point should be extended to a plot of standard size (e.g. 5×5 m; in forest 10×10 m), where standardized phytosociological methods should be applied (preferably phytosociological relevé with cover-abundance assessment (in %) to provide a simple floristic list). This point also enables to estimate the number of points with biodiversity data.
- A transect is fixed on the spot for a birds and butterflies survey using the national methodologies in common the bird monitoring scheme and common butterfly monitoring scheme, if the scheme is existing and operating. In case of missing national monitoring schemes, the methodologies in these two groups will follow the model systems as stated by Butterfly Conservation Europe, with model in UK, see <http://www.ukbms.org/resources.htm> or the Pan-European Common Bird Monitoring Scheme (see <http://www.ebcc.info/pecbm.html>).

²⁰ Regular survey (5 times till now) of the natural resources of UK’s countryside. The field survey covers soils, landscape and habitats as well as vegetation, species are missing. Species in UK are systematically surveyed by targeted schemes, e.g. butterflies, birds, etc. More on: <http://www.countrysidesurvey.org.uk/>

²¹ Switzerland introduced complex monitoring of biodiversity which is very detailed in the targets and surveyed groups; it could serve as a model for an ideal state. More on: <http://www.biodiversitymonitoring.ch>

²² The European Biodiversity Observation Network (EBONE), an FP7 project, addresses the development of a cost effective system of biodiversity data collection at regional, national and European levels that can be used for comparable assessments. More on: <http://www.ebone.wur.nl/UK/>

Although there is a valid indicator on birds and butterflies (SEBI No. 1 Abundance and distribution of selected species) at the EU-level, the spatial coverage is far from being complete and the statistical robustness is doubtful in the view on the whole EU. Bird species are monitored (and therefore covered only by) in 19 countries; butterflies are surveyed only in approximately 5 EU-regions. It is just clear that the recent and existing surveys on birds and butterflies are limited in extent and coverage. The LUCAS in-situ survey could fill these gaps, especially related to the spatial context of the observation: covering of additional states, higher number of plots).

The current survey can be used to provide indicators, but their level of significance could be questioned as the underlying information is not collected in all European Member States and, even if it is, the schemes are based on very limited number of sites / transects / points. Therefore, using LUCAS as a frame could very much contribute to the robustness of the information by increasing the number of plots and therefore the reliability of the results.

The problem that the bird and butterfly survey is sharing with many other existing monitoring schemes is that they are not financed systematically and centrally; they are mostly voluntarily based - and their existence cannot be guaranteed on the long run. By consequence, there is an index for birds and butterflies at EU level, but it is based on a rather small amount of data while on the other hand more data mean a more robust system and therefore more reliable and realistic results and information for the policy.

Our proposal quotes and builds on the existing experience, therefore, the same methodology enables the inclusion of new data and while ensuring compatibility with existing data. The implementation of the proposal (e.g. the introduction of the butterfly / bird / vegetation component) should be consulted with the national focal points, when and where existing (e.g. BirdLife partners, cooperating subject with Vlindersichting etc.). This consultation process supports the compatibility and use of the results also outside of the scope of LUCAS process.

The financial effort is estimated at approximately 1.4 Mio Euro staff cost for 5 000 sample points (at 280 Euro / point) which partly have to be visited up to 5 times per year (amounting to a total of 171 000 person hours).

Detailed information is provided in the part 7 report.

As additional expenses for the LUCAS in the foreseen order will be difficult to justify, one other path to follow could be the introduction and use of new technologies for recording additional information on site by a surveyor and their web-based dissemination for later analysis by dedicated experts off site. By new technologies we understand going beyond the standard questionnaire and pictures towards "live" data, such as videos of "smart image capture" e.g. www.lytro.com) allowing a targeted post-processing in the office.

The Steering Committee acknowledged the suggestions, but partially doubted whether the scenario provides realistic perspectives. In order to conclude on this, a dedicated workshop should be organised to discuss with all involved stakeholders.

5 Conclusions

The main project conclusions are summarised by the individual “parts” of the INBALUD project.

Part 1: Examine possibilities of how to bring existing data collections together

The different mapping data sets on land use and biodiversity have been integrated in the European reference grid at 1x1 km² resolution. Using some easy to understand analytical tools (based on XLS spreadsheets) the database content can be queried and outputs can be created for different administrative (e.g. NUTS regions) or natural (e.g. watersheds) units.

The integration of the LUCAS database was not as straightforward. The ability to integrate the statistical information of the LUCAS point survey with spatially exhaustive mapping information is limited. By nature the character of a statistical survey and mapping data are different. The underlying rules impede a direct comparison of both data sets without paying attention to issues like statistical representativity and minimum mapping units.

In order to provide statistically sound assessments, the minimum number of LUCAS sample points should at least include 510 points, corresponding to an average area of approximately 6 600 km².

Notable achievements mentioned by the Steering Committee include:

- Analysis of minimum sample size of LUCAS data
- Overview on existing biodiversity and land use / cover data sets

Part 2: Produce as useful first result a processed database

The project input data and its results (indicators) were integrated in an on-line accessible geodatabase. Using current GIS technology and an open-source catalogue application allowed a straight-forward development of an INSPIRE compatible web interface for discovering, viewing and downloading the INBALUD results has be set up.

The processed database as well as the query tool (using OLAP technology) have been designed to be extendable and open for the integration of new, additional layers, which so far have not been considered (e.g. EFDAC - <http://efdac.jrc.ec.europa.eu/>).

Notable achievements mentioned by the Steering Committee include:

- Integration of project data in a framework which allows direct data analysis (OLAP cubes) which is fully in line with the system applied by EEA.
- Willingness of EEA to make project results accessible via the EEA website.
- Open access to the project database

Part 3: Include follow-up work by producing an EU analysis on selected biodiversity parameter

After lengthy testing of different combinations of input data, especially LUCAS data, main contributions were made to two existing SEBI indicators. One offering a new view on existing information and the second providing important background information. The ability to provide the information at different spatial scales (e.g. different NUTS regions) makes these characteristics applicable for different political and scientific purposes.

Both proposals were derived by GIS processing of biodiversity input data sets without considering LUCAS information. The Habitat Diversity Index represents a different way of presenting information that is collected under the Article 17. Its final use will have to be discussed in relevant working groups.

Many options for using LUCAS information for biodiversity issues have been explored. LUCAS transect information was tested for its potential to provide information about landscape structure, information on forests was assessed for its ability to discern semi/natural forests types from non-native ones, finally it was tried to further characterise HNV farmlands by LUCAS variables.

Most of these more detailed or specific approaches failed for different reasons, being the lack of sufficient information collected / differentiated during the field survey or insufficient number of observation points to provide information for detailed assessments.

At coarser spatial scale when sufficient LUCAS points are available to make statistically sound statements often other (e.g. remote sensing based) data sets are available which provide the same information. Nevertheless, the LUCAS survey is able to add an independent contribution to current approaches which sometimes suffer from an overuse of one and the same input data set.

A solution that did work and could provide some benefit was to add value to the LUCAS database by attaching information from other data themes to the LUCAS survey point. This allows to "see" the LUCAS point in its wider context, i.e. as part of a Natura 2000 site or a particular land cover class.

Also attaching scores for specific applications to individual LUCAS parameters has shown good potential. At European level LUCAS provides an independent data source in addition to conventional country based in-situ surveys and earth observation based (modelling) approaches.

Notable achievements mentioned by the Steering Committee include:

- Valuable proposal for assessing Article 17 data via the Habitat Diversity Index
- Suggestion to include HDI into second round of Article 17 reporting
- Improvement" of SEBI n°8 with useful background information
- LUCAS biodiversity exploratory exercise as potential additional layer in Green Infrastructure debate, independent source of information.

Part 4: Make sure that meta data will be available for future data use requests

Meta data is a key element for the exchange of data and to record what has been done to create the information, i.e. to ensure a repeatability of the process. The meta data created in INBALUD complies to the INSPIRE Metadata Implementing Rules and the European Environment Agency Metadata Standard for Geographic Information (EEA-MSGI).

Going beyond the standard "lineage" meta data element INBALUD created for each retained indicator an indicator fact sheet which describes in more detail the different data processing steps. Based on experience, this provides an external user with far more insight into the processing methodology at technical level - information necessary for subsequent indicator reviews or updates.

Part 5: Feed in the future strategy on data storing and handing within European Institution

The use of open source software allows to create cost-efficient software solutions for INSPIRE compliant web services for discovering, viewing and downloading input data and results (indicators). The software even allows to search geospatial data across multiple catalogues (e.g. from different European Data Centres), combine distributed map services (e.g. in EEA and JRC) in the embedded map viewer and publish geospatial data using the online metadata editing tools.

Notable achievements mentioned by the Steering Committee include:

- Developments fully in line with developments in EEA and EC

Part 6: Formulate input to recommendations on EU biodiversity data policy

A gap analysis with respect to measuring progress towards the 2020 EU biodiversity targets has been carried out. Not surprisingly it identified a lack of suitable data sets for monitoring species and habitats at European level, a lack of time series information, a lack of information outside protected areas, inconsistent information across different countries. As a result the few existing data sets are used for the creation of often complex derived data sets sometimes ignoring correlation or interdependencies between the input data sets.

As a consequence of these gaps, practical solutions need to be sought to overcome the data issues, determine target data sets, improve their coverage and consistency across Europe and install financial schemes and quality control mechanisms to ensure a consistent data flow.

Notable achievements mentioned by the Steering Committee include:

- Recommendations on EU biodiversity data policy contain a lot of information that could be used in support to a gap analysis in the perspective of the 2020 strategy.

Part 7: Prepare conceptually the biodiversity element of in-situ surveys such as LUCAS

Biodiversity is a local phenomenon which in theory could be well monitored by in-situ surveys, but to be biodiversity relevant such information must be collected with a denser sampling design and survey variables better suited at biodiversity issues.

The current LUCAS survey approach was reviewed and recommendations were made how to better incorporate biodiversity issues in the survey. Two feasible options were presented a) the stricter implementation of current survey guidelines and b) a biodiversity targeted survey to be carried out by expert surveyors aimed at habitats and selected groups of species.

For the suggested biodiversity survey a scheme was drafted and a first cost estimation was provided.

Comments mentioned by the Steering Committee include:

- Not sure if the part on the "in situ" biodiversity component of LUCAS really brings realistic perspectives.
- Overall, well written discussion of the issue at stake, worth to have a specific in-depth workshop on this topic to discuss pros & cons of the suggestions.

Summary of LUCAS experiences

- The minimum number of LUCAS sampling points needed to make statistically sound statements is about 510; this corresponds approximately to an area of some 6,600 km².

If assessments about individual land cover classes (e.g. agricultural land) are of interest, then a minimum number of 720 points (or approx. 10,000 km²) is required.

As many assessments are addressing individual land cover classes (e.g. share of forests per NUTS or biogeographic region), it is probably safer to assume that reporting units should at least cover an area of 10,000 km² for any assessment based on LUCAS information.

Consequently, individual land cover polygons or other features smaller than 10.000 km² (e.g. *individual* Natura 2000 sites) cannot be characterised by LUCAS data. Larger areas, such as biogeographic regions or *all* Natura 2000 areas in one country (provided that the country is large enough to have at least 510 / 720 sample points within all Natura 2000 sites) LUCAS data can be used to characterise these features.

- The LUCAS field survey is not targeted at collecting biodiversity relevant information. There is little information on species (forest species, crop types) and no information on habitats. LUCAS information therefore can serve only as proxy for biodiversity issues.

A problem of the field survey is the number of empty records at thematically more detailed and potentially more biodiversity relevant survey parameters (i.e. forest

species). The more "expert" the surveyor knowledge needs to be, the more empty fields are in the database. This issue can of course be overcome by a better training or the use of more skilled surveyors, but it will also increase the cost as described in part 7 (chapter 4.7).

Adding more biodiversity relevant information to the field survey will add value to the survey. For the selection of which parameters should be added the added value of in-situ information over other sources (e.g. earth observation, smart image capture) needs to be reviewed.

- It is possible to attach simple biodiversity values ("scores") to individual LUCAS survey parameters. For example, urban areas would score lower than agricultural or forest areas in a simplified assessment of their value for biodiversity. Transects in agricultural areas with many "green" intersections would score higher than those with fewer or many "red" intersections: many green intersections suggesting smaller field sizes and therefore potentially less intensive land use.

Combining several of these parameters and their biodiversity values can be used to assess assets of any particular "reporting unit" (i.e. NUTS region, watershed, biogeographic region, HNV farmlands), provided that it encompasses the minimum number of sample points.

- Such assessments based on LUCAS are of great value as they provide independent information over similar evaluations based on (mostly) Corine Land Cover. The "overuse" of CLC leads to a correlation of information layers which in the end do not provide independent information, but different interpretations of the same basic data sets (i.e. CLC).

Even if LUCAS might not provide new information, it is an important, independent source for validating and cross-checking other information layers.

- Due to the statistical character of the survey, individual LUCAS sample point data cannot be directly crossed with spatially explicit (GIS) data at European scale.

This might change with the arrival of the new spatial data layers of higher geometric resolution, such as the GIO high resolution layers. These layers have a resolution of 20m (at pixel level) and thus correspond much more to the LUCAS observation unit.

The high resolution layers will help to attach a representative area to each LUCAS sample point, which eventually will facilitate the spatial extrapolation of LUCAS information.

- The integration of the LUCAS database with the other project data sets needs to be treated with care. Whatever assessment using LUCAS information needs to ensure that the statistical design of the LUCAS survey is not violated (i.e. representativity).
- The exploratory exercise for assessing the biodiversity value of LUCAS data by attaching scores to different survey parameters has revealed a potential of LUCAS

data, keeping in mind that it will be most useful in terms of visualisation and communication at broad scale than for real implementation on the ground. But this limitation LUCAS shares with most European-level data sets.

- The information on the presence / number of landscape elements / linear features along the LUCAS transect can provide information on the heterogeneity of objects (landscape elements, land cover classes), under the pre-condition that the basic LUCAS sampling design rules are kept (i.e. minimum number of sampling points per entity of interest).

In our view the study confirm the use of LUCAS information for biodiversity purposes under a set of prerequisites:

- Application of a suitable reporting unit, i.e. NUTS 2 or an area equivalent to approximately 10.000 km².
- Respecting the above point, the biodiversity information obtained from LUCAS is mostly relevant for European-level organisations (EG ENV, Eurostat), not for local actors.

To make the data interesting for local actors a denser network of in-situ monitoring sites would be needed.

- Biodiversity information obtained from LUCAS can only be used as proxy information, as very little directly relevant biodiversity data are collected by the survey.

Information on biodiversity provided by LUCAS suffers from the fact that similar information is also provided by other "proxy" sources which are spatially explicit.

To make the data more interesting to the biodiversity community more survey parameters need to be added to the in-situ monitoring which can only be provided through field surveys and which are target at responding to biodiversity issues (e.g. intensity of use, presence of protected / rare species).

- LUCAS information has an important value in the provision of independent information, which makes the data a unique source for validation and cross-checking of information derived from other sources.

6 References

- Official Journal of the European Union, *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*, 2007.
- Official Journal of the European Union, *COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata*, 2008.
- INSPIRE Network Services Drafting Team, *INSPIRE Metadata implementing rules (version 1.2)*, 2007, revised 2010.
- INSPIRE Network Services Drafting Team, *Draft Implementing Rules for INSPIRE Transformation Service*, 2009.