seamless

System for Environmental and Agricultural Modelling; Linking European Science and Society

Specification of databases of SEAMLESS-IF

Andersen, E., van Diepen, C.A., Hazeu, G.W, Elbersen, B.S., Terluin, I.J., Verhoog, A.D., Godeschalk, F.E., Genovese, G., Confalonieri, R., Kyhn, M.

Partners involved: FLD, Alterra, LEI, JRC

Report no.: 20
November 2006
Ref: PD1.3.3
ISBN no.: 90-8585-047-9

Skov & Landskab

ALTERRA WAGENINGEN UR

Logo's main partners involved in this publication

Sixth Framework Programme

SEAMLESS integrated project aims at developing an integrated framework that allows ex-ante assessment of agricultural and environmental policies and technological innovations. The framework will have multi-scale capabilities ranging from field and farm to the EU25 and globe; it will be generic, modular and open and using state-of-the art software. The project is carried out by a consortium of 30 partners, led by Wageningen University (NL).

Email: seamless.office@wur.nl
Internet: www.seamless-ip.org

Authors of this report and contact details

This report has been written by WP4 partners of SEAMLESS

Contacts:

Erling Andersen E-mail: Eran@kvl.dk	FLD
Kees van Diepen E-mail: Kees.vandiepen@wur.nl	Alterra
Gerard Hazeu E-mail: Gerard.hazeu@wur.nl	Alterra
Berien Elbersen E-mail: Berien.elbersen@wur.nl	Alterra
Ida Terluin E-mail: Ida.terluin@wur.nl	LEI
David Verhoog E-mail: David.verhoog@wur.nl	LEI
Frans Godeschalk E-mail: Frans.godeschalk@wur.nl	LEI
Giampiero Genovese E-mail: giampiero.genovese@jrs.it	JRC
Roberto Confalonieri E-mail: roberto.confalonieri@jrc.it	JRC
Martin Kyhn E-mail: kyhn@kvl.dk	FLD

Disclaimer 1:

"This publication has been funded under the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2). Its content does not represent the official position of the European Commission and is entirely under the responsibility of the authors."

Disclaimer 2:

Within the SEAMLESS project many reports are published. Some of these reports are intended for public use, others are confidential and intended for use within the SEAMLESS consortium only. As a consequence references in the public reports may refer to internal project deliverables that cannot be made public outside the consortium.

When citing this SEAMLESS report, please do so as:

Andersen, E., van Diepen, C.A., Hazeu, G.W., Elbersen, B.S., Terluin, I.J. et al. 2006. Specification of databases of SEAMLESS-IF, SEAMLESS Report No.20, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 40 pp, ISBN no. 90-8585-047-9.

Table of contents

General information

Task(s) and Activity code(s):	A1.3.3.
Input from (Task and Activity codes):	
Output to (Task and Activity codes):	A1.3.8
Related milestones:	M1.3.1

Executive summary

In this report the first selection of data sets for SEAMLESS-IF is described, including the survey of available data and the first consultation of data needs. It is stressed that the selection of data sets are not final and that other data sets will be added based on a new consultation and direct communications with other SEAMLESS partners. Furthermore, it is briefly described how the data will be made available to the partners.

Specific part

1 Introduction

The objective of this report is to specify a selection of data sets, which is used as input for the analyses especially in WP3 but also in WP2 and WP6. This has been done based on a survey on available data sets and a first consultation of data needs targeting WP leaders and selected key partners.

It was planned that this report should include a final selection of databases of SEAMLESS-IF. However, it can be envisaged that new data needs will be identified as the work in other WPs progresses. Furthermore, a few data sets need further exploration before a decision to include them can be taken.

Apart from the specific selection of databases, the PD also describes other important issues dealt with in the work of WP4.

1.1 General introduction to WP4 work

The diversity of models used in the integrated modelling of Seamless, the inclusion of both quantitative and qualitative indicators and the hierarchical system approach result in a very complex demand for data (see figure 1). The data have to provide information on the biophysical, social, economic and institutional aspects of a system. Furthermore, the data need to be organised to support modelling and assessments at different spatial scales.

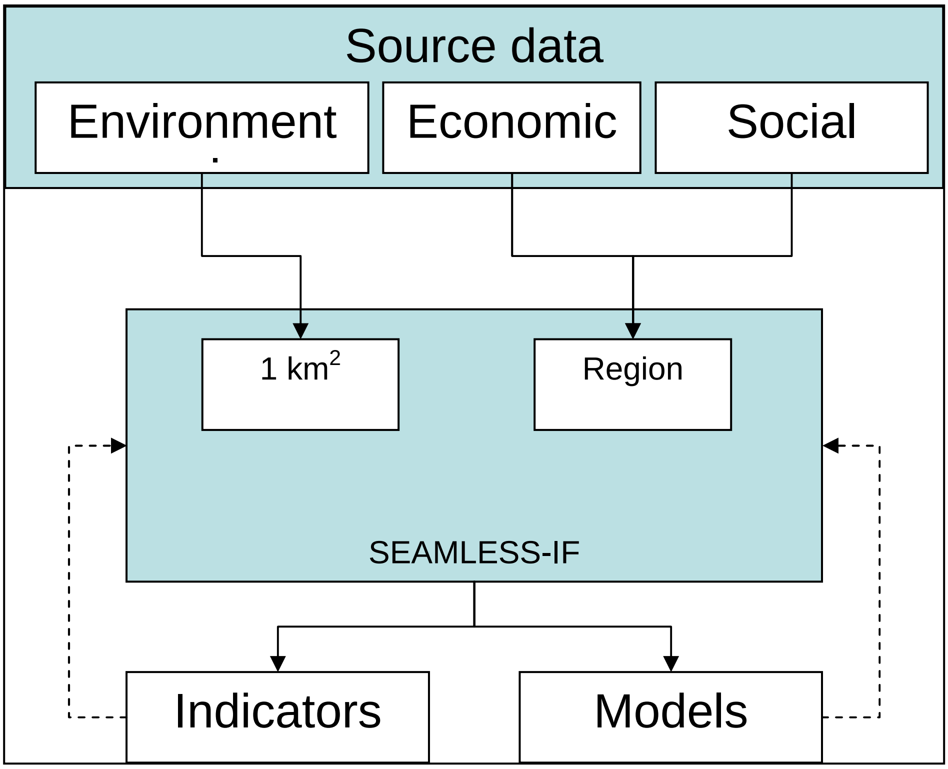


Figure 1: Data need and integration in SEAMLESS-IF

To provide the necessary data the challenge is to bring together existing datasets that have been generated for very different purposes and with very different methods. For example data on agriculture have been generated with the purpose of monitoring the agricultural sector mainly from an economic point of view. On the other hand data have been gathered on environmental issues mainly to provide information on the state of the environment and linked to assessment of environmental issues. Furthermore, the data provides information at very different spatial scales and levels of organisation. Typically, economic and social data are linked to administrative regions whereas environmental data are linked to landscape units or grids (see figure 1). For the knowledge base of Seamless the different data sets will be pre-processed and linked to a common spatial framework enabling linkages between the different data sources. The finest scale in the spatial framework will be 1km grid cells. The most challenging part of the work in relation to the spatial integration of data will be the spatial allocation of farm types.

The knowledge base will primarily be based on existing Pan-European data sets. Only in very specific cases will other data sources, such as national data sets, be included and it is not planned to do any new data collection. The scope of the work with the data sets in Seamless is thus to bring together different existing data sets in a common framework.

It is the aim that all the data that are integrated in the SEAMLESS-IF should be available to the public. Thereby, the scope of using open source software developing the integrated framework will not be limited by restrictions on data use. This means that data sets where the original data might not be available to all must be pre-processed before being included in the knowledge base of SEAMLESS-IF. This is for example the case for single farm data from FADN, which have very strict confidentiality rules in relation to the identification of specific farms. This will be solved by providing information on farm types rather than single farms in the data to be integrated in the knowledge base. Another example is CORINE data, where processing to another spatial resolution, enables inclusion in the knowledge base without limitations on use.

2 SCOPE OF SEAMLESS WP4

The original version of this chapter was finalized in June 2005 and presented at WP1 meeting in Montpellier in the same month.

2.1 Aim of this section

This paper aims at providing a better understanding of the scope of SEAMLESS WP4, the approach to be followed and the way WP4 tasks should interact with the tasks in the other 5 WP. More specifically this paper outlines the understanding of the WP4-partners on what is needed to support the activities in the other work packages WP2, WP3, WP6 taking into account the concepts agreed in WP1. Such decisions have a direct impact on both WP4 and WP5.

In the DOW of Seamless the WP-4 tasks are described as follows:

> *In WP4 the knowledge base will be populated with data and with the information needed to access and combine the data in different formats and at different spatial levels. The data includes model inputs, source data for queries and statistics, metadata and SEAMLESS-IF analysis outputs. The relevant and available farming system, environmental, economic and social data sets – with coverage for EU-25 and at the global level – will be collected and adapted. The adaptation includes developing typologies of farming systems and of regions to be used for organising the data in the knowledge base. Specific routines, procedures, protocols and knowledge rules will be developed to facilitate access and inclusion in the domain editor developed in WP5. This includes developing protocols for combining spatial and statistical data.*

From this description it becomes clear that the activities of WP 4 can be sub-divided into 4 sub-activities:

1) To collect and standardize model input data
2) To collect and integrate model output data
3) Design specific routines, protocols and knowledge rules for facilitating access and inclusion of input and output data in the domain editor of Seamframe (to be developed in WP5), including the specifications for the development of metadata.
4) The adaptation of input (and output) data for inclusion in the knowledge base. The adaptation of data involves the production of typologies but also procedures for aggregation and disaggregation of input and output data.

Sub-activities 1-3 mainly involve the collection and standardization of data and their distribution to other WPs. However, there are several reasons why we cannot limit our activities to only collecting, standardizing and re-distributing data. This is why we also need to do activity 4. This activity is of special importance because in Seamless input and output data have several different sources. Each have their own coding, classification, spatial resolution and selection of parameters which need to come together in one knowledge base. The heterogeneity in data sets will require data integration through harmonization, categorization and sometimes aggregation and disaggregation. The integration of data in WP4 will support the development of an innovative SEAMLESS-IF with the following characteristics:

- SEAMFRAME will enable integrated assessments linking farming, environmental and socio-economic issues. In order to allow these assessments to be integrated, protocols are available that link data sources related to different issues (farming, environment, socio-economics) together. The best way to do this is by making in- and output data spatially explicit.

- SEAMLESS will provide assessments at different spatial levels (e.g. global, market, national, region, farm, field levels). This is enabled by the inclusion of protocols for aggregation and disaggregation of in- and output data which ensures consistency and enhances the exchange of information. This requires a hierarchical and typology based approach.

- SEAMLESS will cover EU-25 and include assessments based on cross regional comparisons. To enhance these comparisons territorial typologies are needed to provide contextual information on environmental and socio-economic issues.

- SEAMLESS aims to be a (set of) tool that offers transparency in the assessments. This requires that detailed descriptions of the data (metadata) are elaborated.

This paper is structured as follows:

2. An overview of our understanding of the division of tasks in other WPs which are data related and therefore connected to the tasks in WP4.
3. The first selection of datasets to be collected, processed and standardized in WP4.
4. Selection of spatial units and procedures for making data spatially explicit
5. Agricultural systems and farm data
6. Providing contextual information – territorial typologies
7. Metadata issues
8. Other data processing issues

2.2 Data related tasks in SEAMLESS

The basic subdivision of tasks between the SEAMLESS WP's related to data can be summarized as follows:

WP1 will define basic concepts

WP2 and WP3 and WP6 will define data requirements

WP5 will design data base and all related software, and build the data base

WP4 will collect the data, process the data if necessary, and fill the data base with base data and processed (classified) data.

WP4 will provide a data base server, and operate and update the data base thematically

WP5 will provide technical maintenance on the data base

The Seamless partners working in WP2 (indicators), WP3 (models) and WP6 (case studies) need access to a data base to download data for model input and for analysis, and to upload model output data and results of their analysis. WP4 should populate the database from which the necessary input data can be downloaded and should support the up-loading of the output data in this same database. In most cases this will first require some standardizing, processing, adding the standardized metadata and providing for integration with other data if needed.

WP5 will deal with the database design taking into account the use of data, database query and search facilities and data transfer protocols. The design of the data base requires the availability of metadata (description of the data) which will be provided by WP4 in a standardized format for every in- or output data to be included in the SEAMFRAME database.

2.3 First selection of datasets to be collected and processed by WP4

WP4 will deal primarily with European data sets covering EU25 completely. A first selection of data to be collected and processed has been made based on a consultation with other WPs. These databases will be targeted in the work until October 2005, where the aim is to provide access for all partners to the selected databases adapted for Seamless. This does not mean that all doors are closed for all other datasets in this period. We will at the same time continue exploring other data sets of which most have already been mentioned in the inventory of European datasets already presented in the consultation with other WPs. A specific issue in the work in the coming period is detailed data on farm management. As the Pan-European data sets hold only a limited information on this, we need to explore whether it is feasible to use national data sets.

In the coming period we will target to combine and standardize the selected datasets into four core sets:

1) Farming data: A database containing a large range of farm variables (from FSS, FADN et., see Table 1). Since these farm data come from different databases the data can only be linked together in one Seamless database by using farm type keys. These keys ensure that e.g. the FSS variables and the FADN variables refer to a common cluster of farms.
2) Environmental data: A database with environmental data linked to a grid map (1x1 km.).
3) Socio-economic data: One database with information linked to a map of Nuts regions (Combination of Nuts 2 and 3, see annex 6).
4) Global data: Adaptation of the GTAP database to Seamless.

The first selection of data is shown in table 1

Table 1: First selection of databases. Priority in relation to work until October 2005.

Issues	First selection of data sets
Farming data	FADN, COCO, CAPREG, Eurostat: Eurofarm, Eurostat: AGRICULTURE, FAOSTAT, AMAD

Environmental data	MARS soil and climate data, Digital elevation model Pan Europe, , European rivers and catchments ERICA, Water catchments Pan-Europe, European Environmental classification, Corine Land Cover 2000.
Socio-economic data	Eurostat: Region, Eurostat: Eurofarm, Eurostat: Agriculture, Eurostat: Environmental statistics, Eurostat: Sirene
Global data	GTAP

At a later stage 1) will be linked to the grid map of 2) by disaggregating the data to 1x1 km using cropping zones information. Also at a later stage 3) will be linked to the grid map, but the data will not be disaggregated but only translated into grid cells which implies that all grids cells in a given Nuts region will have the same value for a given variable). This is further described in the following section on spatial issues.

We do not expect that all variables included in the selected databases will be needed in Seamless. We do therefore in the coming months require input from other partners regarding the specific variables to be chosen from the databases covered by the first selection

The selection of the data sets described above does not imply that all doors are closed for other datasets. WP4 will continue exploring other data sets of which most have already been mentioned in the inventory of European datasets already presented in the consultation with other WPs.

Specific emphasis of the work in the coming period is on detailed farm management data. As the Pan-European data sets hold only limited information on this, we need to explore whether it is feasible to use national data sets or to collaborate with other projects to get access to data.

Other sources of socio-economic data will also be further explored based on requests from WP2 partners. These include the database on Political Institutions DPI Compiled by the Development research Group of the World Bank, The Eurobarometer, Eurostat database Compilation of Sustainable Development Indicators, European Values Study (www.gesis.ord) plus possible databases on the publics perception and evaluation. Finally, it will probably be necessary to include other data sets that will be identified as the work in the other WPs progress.

2.4 Selection of spatial units and procedures for making data spatially explicit

In the DOW of SEAMLESS it is proposed to make an integrated and operational framework (SEAMLESS-IF) to support ex-ante analysis of policies that enables an analysis at the full range of scales, and an analysis of the environmental, economic and social contributions of a multifunctional agriculture towards sustainable rural development and rural viability. Because of the regional variation in climate, natural resources (soils, vegetation etc) and social structures and the increasing move towards de-centralisation of policy implementation, there is an increasing need to appraise the multi-functional agriculture at a range of scales from global to field level. Therefore the requirement is for spatially explicit appraisal ie. it is not sufficient for policy makers to know what the impact will be, but it is important to know where the impacts will be and how they will vary in different regions.

Table 2: Typical resolutions or characteristic length scales

	Char. Length Of mapping unit	Grid cell	Minimum size spatial unit	Nr of spatial units EU15	Nr of spatial units EU25	Nr of spatial Units Pan-Europe
NUTS0	50-200 km			15	25	
NUTS1				77		
NUTS2				206		
NUTS3				1031		
MARS climate grid		50x50 km	50x50 km			
CORINE (vector/grid)	500 m	100x100 m 250x250 m	25 ha			
PELCOM		1x1km				
Soil map		1x1 km				
Elevation map (DTM)		250x250 m				

This means that the different data types should be balanced in their spatial resolution and if possible should be available at flexible scales. But it also implies that all the data incorporated in SEAMFRAME should at least have one spatial dimension which could be a reference to a polygon and/or a grid. Polygons can be administrative boundaries such as country and regional borders (equivalent to NUTS regions) but can also be pedo-climatic entities such as river basins, altitude zones, soil entities etc.

Grids can vary in size from e.g. 1*1 km to 100*100 km. WP4 prefers and recommends to use a spatial raster of a fixed grid cell size rather than vector systems for mapping. The smallest grid size should be a 1*1 km and the fixed grid will be the same as used in the SENSOR project which is the EEA Reference Grid Fishnet Tool for ArcGIS 9. The lower left coordinates are X = 1500000 and Y=900000 and the coordinate system is ETRS89-LAEA. The coverage of the grid has the extension of the Corine Land Cover 2000 data, which is all EU25 countries.

The great advantage of working with a fixed grid cell is that it makes the integrated assessment and the visualization of output of different data fairly easy. One of the disadvantages of a grid system however is that in the case of relatively large homogeneous zones many identical grid values are stored in the data base. This can be optimized however by data base techniques (WP5), so that data storage redundancy is minimized. Another disadvantage could be that differences in natural conditions exist within a grid when the grid cell size is too large. However, such problems will even be larger if spatial data need to be integrated with varying polygon forms and sizes. Working with a fixed grid and a range of fixed grid sizes (1*1, 5*5, 25*25, 50*50 and 100*100 km) and translating all data to this fixed standard will require intensive data processing before inputting it in SEAMFRAME. But after this it will make data handling and using it by other WPs and in the eventual SEAMLESS-IF system easy and efficient.

As already indicated in section 2.3, WP4 suggests to make all data spatially explicit by linking it to a standard grid. This both applies to spatial databases such as the soil information data from MARS as to the Farm Structural Survey information which provides statistical data at the NUTS level. The decision to link the data to a grid means that a standard selection of Nuts regions is less crucial. However, where appropriate we will use the same standard selection of Nuts regions as used in the Sensor project (see annex 6).

In the former section it was already described what data will be initially collected in WP4. Overall it is also clear that each of the input and output data sets in SEAMLESS will have their own typology, accuracy, spatial and temporal resolution. However for a consistent integration in SEAMFRAME we need to ensure that data are consistent across spatial scales. In order to do this systematically and to ensure the integration of input and output data, a set of standards and procedures, called spatial framework, is designed in WP4. This spatial framework does not only imply the collecting, standardizing, adding metadata and distributing to other WPs, but also the integration of the often heterogeneous data. The 2 main reasons for this are:

1. because within a given category of data there may be several different sources of data, each with its own coding and classification, and selection of parameters. This heterogeneity in data sets will require data harmonization, categorization and sometimes aggregation and disaggregation.

2. because within SEAMLESS we foresee that the aggregation and disaggregation issues play a role when exchanging data between the farm level and regional level and between the farm models and the market models.

Given the former, in WP4 procedures will be defined as part of the spatial framework of SEAMLESS for:

1. spatial aggregation from the finest resolution to larger spatial units, for disaggregation from the larger spatial units to finer spatial units and for combining data of different spatial resolutions.
2. the translation of polygon information (e.g. NUTS region) to grid
3. the production of typologies of farm systems and territorial typologies in order to classify information into meaningful entities which enables further analysis and integration with other data sources.

Ad. 1. . Statistical data on the agricultural sector are generally given per administrative region, which ignores that such regions are spatially heterogeneous. The internal spatial variation within a region is related to the history of human occupation and to the natural conditions of the land. In most European statistical sources the most generalized land use information is on cropped area and crop production per NUTS0, NUTS1, NUTS2 and sometimes NUTS3 region. When we want to know in which locations within a Nuts region the crops are really grown we have to identify the cropping zones (cropping zones are defined as zones with relatively uniform suitability for the production of different crops) and estimate the share of the crop per zone. A parameter derived from a map or from a combination of several maps (as spatial data sets) could serve as basis for the disaggregation of a Nuts region into different cropping zones. Candidate parameters can be selected from spatial data sets on land cover, soil or soil suitability, geology, climate, elevation, slope. The allocation rules can be based on expert knowledge or derived through statistical methods from ancillary data sets. The Dynaspat project (see Annex 1) is an example of an disaggregation exercise in which cropping zones have been identified.

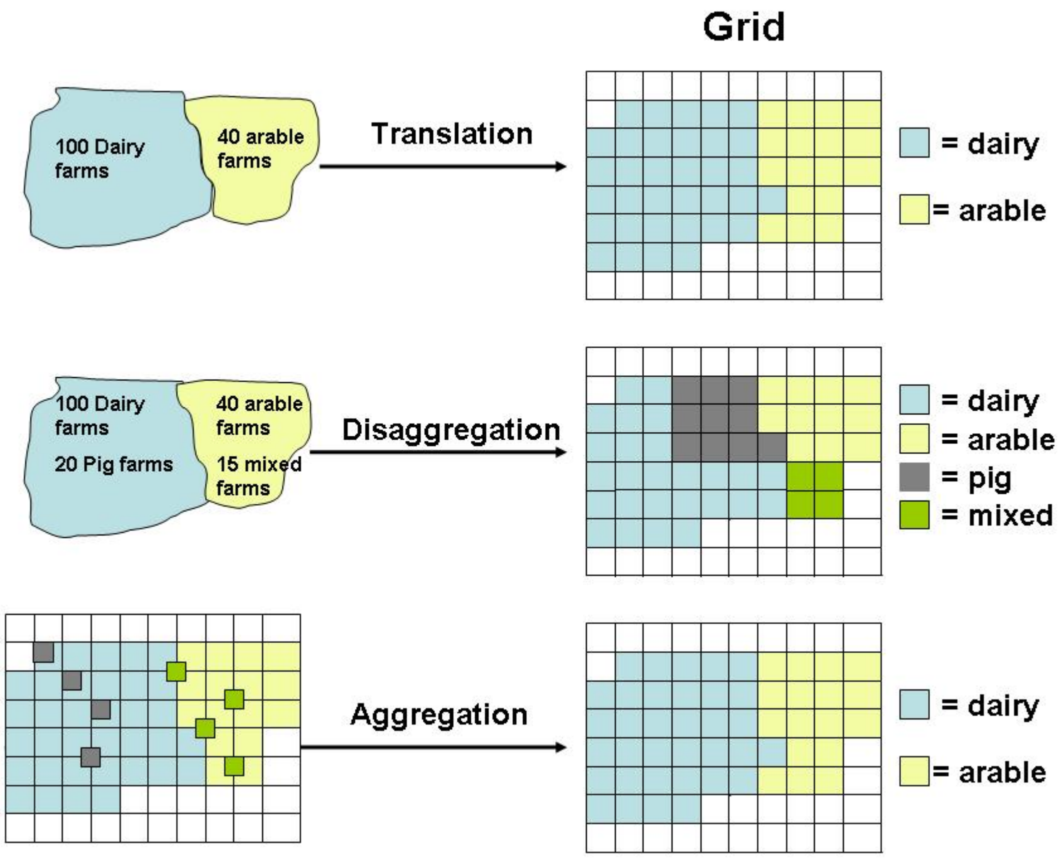
Ad.2. For the aggregation, disaggregation and integration of information the translation of polygon information to a grid is often a necessary step.

For the switch from polygon to grid there are 3 possibilities:

1. Translation
2. Disaggregation
3. Aggregation

The three are explained diagrammatically in Figure 1. The simplest step is the translation of polygon data to grid. In this case no knowledge is needed to differentiate data further since all grids that are located in the same spatial unit at which the statistical data is available (e.g. a NUTS region) get the same (or dominant) value for the specific variable.

Figure 2: Translating polygon information to grid



However, with disaggregation extra knowledge is needed to make the data more spatially explicit. This can be done by involving additional spatial information on e.g. land use that will enable us to identify the places where it is likely that there is a higher concentration of dairy, pig, arable and mixed farms so that we can differentiate the grids within one polygon (Nuts) in more dominant land use clusters.

Ad.3 Typologies are needed for classification, simplification and for making integration with other datasets possible. A typology is the same as a classification of one of more variables into meaningful classes. This implies that the threshold values between two classes have been chosen from a certain, perspective (environmental, social, economic, etc.). Typologies can be expressed in spatial and in tabular format. Typologies can be one and more dimensional. Examples of typologies and their characteristics are:

- The Community farm typology (FSS/FADN):
  - Classification of farms into production types based on share of income derived from a certain activity
  - Economic bias
  - One dimensional
  - Presentation: tabular format and linked to a Nuts
- ELPEN/IRENA farm typology
  - Classification of farms on basis of intensity and land use
  - Environmental bias
  - More dimensional
  - Presentation: tabular format and linked to a Nuts
  - Hierarchical
- Environmental classification
  - Classification of regions on basis of climate and altitude
  - Bio-physical bias
  - More dimensional
  - Presentation: Spatial
- Urbanisation level
  - Classification of regions on basis of urbanisation level (e.g. population density and/or size of nuclei present, share of active population employed in agriculture)
  - Social-economic bias
  - One or more dimensional
  - Presentation: Spatial

In the following two sections a description is given of typologies to be further developed in WP4.

2.5 Agricultural systems and farm data

It is a basic assumption that the data (in- and output data) on farming in Seamless has to be linked by a common reference to farm types. From a data management point of view this enables re-use of different data sources and enhances the possibilities to link information stemming from different sources. From a result point of view it is necessary to use the same reference farm types as this will facilitate integrated assessment of the modeling results . Finally, and most important, it also allows us to link farming types to environmental data. This is crucial as it is the only way to link economic and environmental resource endowment.

We suggest that the typology of agricultural systems is elaborated as a combination of the following:

1) a specialization (or sectoral) dimension
2) a land use dimension and perhaps an intensity dimension)

These two points have in previous work been used to define farming systems. In Seamless we will add:

3) a locational dimension with a link to cropping zone (see previous section).

By doing so we define agricultural systems. The term agricultural system is thus used when referring to a farming system with a specific environmental resource endowment.

At the moment we do not find it feasible also to include capital and labour resource endowment in the typology. However, it needs to be explored further whether this should be added as a dimension to the typology or as a variable describing the agricultural systems. On the other hand capital and labour endowment could also be added to the analysis at a regional level and included in a territorial typologies.

A first description of the suggested farming systems will be made available at the regional level in month 10 for EU-15. The connection to cropping zones are planned to be finished in month 12. For the new Member States we hope that the necessary data will become available early 2006.

In relation to the farming systems, we have found it feasible to use the already established EU typology as starting point. This typology is already used for market level modeling in Capri (Capri farm types). However, this typology will be adapted to allow for spatial location and to reflect especially environmental issues. Based on previous experience and discussions at the Lund –meeting, January 2005, we suggest that the Capri farm types are further detailed according to land use and intensity of farming, keeping the latter open.

Until now farm typologies have already been developed and tested in several former EU projects (e.g. ELPEN, HNV farmland and IRENA LOT 1). The resulting typology is a hierarchical typology with two dimensions; a land use and an intensity dimension. The two dimensions can be combined, resulting in classes like *intensive permanent grassland system*, but can also be used as separate typologies e.g. *grassland systems* and *intensive systems*. The typology only based on the land use dimension corresponds to the one used in the IRENA operation of EEA. In table 3 the different land use types are presented. In Seamless in WP4 these land use types will be tested further statistically and adjustments might be made at a later stage incorporating the needs of other WPs. These needs might lead to incorporating new dimensions to the typology, but these could also be incorporated as extra variables describing the farming types.

Table 3: Land use types to be used initially

Land use types	Definition

Permanent grassland	>=55% of UAA in grass and <40% of grass in temporary grass
Temporary grassland	>=55% of UAA in grass and >=40% of grass in temporary grass
Cereal systems	>= 55% of UAA in cereals and <12,5% of UAA in fallow
Fallow land systems	< 50% of UAA in permanent crops and <55% grass and >=12,5% of UAA in fallow
Specialised crop systems	< 50% of UAA in permanent crops and <55% grass and <55% of UAA in cereals and >= 25 of arable in either sugar beets, oil seeds, potatoes, seeds for sowing, fodder-maize, cotton, tobacco or vegetables and flowers
Mixed crops systems	< 50% of UAA in permanent crops and none of the above:
Permanent crop systems	>= 50% of UAA in Permanent crops

* (UAA=Utilised Agricultural Area)

In table 4 the intensity dimension of the typology is shown. At the moment these types are based on the expenditure on fertilizers, crop protection and concentrate feedstuff. This corresponds to the approach used in the IRENA operation. However, in WP4 a further assessment has to be made of this dimension and it should be considered whether a nitrogen balance indicator would be a better indicator for the intensity dimension than the expenditure on inputs as used in the IRENA typology. Since the intensity dimension is not crucial for allocating farm types spatially it is less important than the land use dimension and can therefore obtain less priority on the short run. This is also why the intensity factor does not necessarily need to become a dimension of the typology. Instead it could also become an additional variable describing the main land use types.

Table 4: Types of intensity of farming to be used initially

	Expenditure on fertilizers, crop protection and concentrate feedstuff per ha per year
Low-input	< 80 Euro
Medium-input	=> 80 and < 250 Euro
High-input	=> 250 Euro

The end result is that the data bases in Seamless will hold information linked to agricultural systems such as Dairy farms (Capri type) basing their fodder production on arable crops (land use dimension) in a specific cropping zone with homogenous conditions for crop growth (= environmental resource endowment).

NOTE THAT THE FARM TYPOLOGY HAS BEEN DEVELOPED FURTHER SINCE THIS PD WAS WRITTEN: SEE PD 4.4.2 FOR UPDATED VERSION.

2.6 Territorial typologies in the EU

As a part of the SEAMLESS-IF it is envisaged that territorial typologies should be available providing structured information for assessing model results in a spatial context. In the DoW this task is allocated to WP4. It is not the purpose to duplicate the work done in WP2 on indicators, but merely to provide a few carefully selected typologies of regions that can serve as a standard framework for spatial assessments. In some cases the territorial typologies might also serve as model input (for example suitability for crop production) The typologies will as far as possible build on already established typologies for example from OECD or from the ESPON

Typologies of EU regions can be derived from any indicator one wants. For a proper analysis, it is useful that there is some theoretical relationship between the indicator(s) used for the design of the typology and the objective of the study. Further, typologies can have as many clusters as one wants (depending on the level of detail desired). However, the more clusters, the more difficult it becomes to interpret the distribution of the studied indicators. Hence, a typology with a limited number of clusters is preferred to one with many clusters.

In annex 3 a number of typologies of EU regions are presented, which have been used in recent studies of regional performance.

1. OECD territorial typology of predominantly rural regions, significantly rural (or: intermediate) regions and predominantly urban regions. This classification is derived from population density;
2. Extension of the OECD typology with leading and lagging regions, derived from employment growth;
3. Typology of less favoured areas;
4. Typology based on structural handicaps in the scope of the EU Structural Policy;
5. Typology based on the integration of regions in the national/global economy;
6. Typologies on the base of landscape features and bio-geographical characteristics;
7. Typology based on accessibility;
8. Cluster typologies: apart from designing a regional typology based on one indicator, typologies can also be derived from a number of indicators by using cluster analysis;
9. Harmonised urban-rural typology: based on the degree of urban influence and the degree of human intervention;
10. Typology based on natural cultural values and economic development.

In Annex 4 an example is given of a typology of environmental zones of Europe based on climate and altitude information used as a sampling framework for vegetation mapping but also as an analytical framework for assessment of impacts on biodiversity from climate change, suitability for biomass crop production or for the development of indicators of High Nature Value farmland areas.

From a first quick scan of regional typologies it appears that:

1. the design of regional typologies depends on the objective of the analysis;
2. it is rather difficult to interpret typologies that are based on a large number of indicators;
3. simple typologies are preferred to complicated typologies.

We are planning in the first phase to elaborate typologies related to the following variables:

1. Suitability for crop production (a. cereals, b. grass) (linked to identification of cropping zones)
2. Degree of rurality
3. Structural handicaps (linked to regional level labour and capital endowment)
4. High Nature value farmland

2.7 Metadata issues

Metadata is a summary of information about a dataset (data about data). It contains for example information on the content, quality, type, creation, distribution, temporal and spatial information of a datset. The main reason for describing the data used in assessments or delivered by models is to improve data usability. The description of the datasets (metadata) needs to be standardised so the data can be unambiguously interpreted by others. It makes data discovery easier and reduces the risk of duplication.

A commonly accepted metadata standard is the ISO 19115 standard for metadata. This standard describes digital geographical data. The INSPIRE (Infrastructure for Spatial InfoRmation in Europe) initiative supports a metadata profile which follows the guidelines in ISO 19115. The SEAMLESS project has to gear to this European initiative Another reason to select ISO 19115 metadata standard is the use of this standard in the SENSOR project.

Metadata catalogues are selections of metadata of different kinds of datasets covering one or more themes (environmental data, socio-economic data, etc.). A metadata catalogue facilitates the search and selection of data needed for a certain assessment. Also it will help the potential user to obtain the data needed. An important advantage of a metadata catalogue is that the user has to deal with a limited number of information what is made available in a standardised way.

The metadata profile which will be used in the SENSOR project is a useful profile for the SEAMLESS project. The profile is coming from the EEA (European Environmental Agency) and only slightly adapted. In this case the tuning with the INSPIRE initiative, EEA and SENSOR regarding metadata is guaranteed. Annex 5 presents the metadata profile as used in the EEA and the SENSOR project. Selection of metadata describing environmental and socio-economic data can be found on the SEAMLESS website. Both metadata catalogues follow the INSPIRE initiative and ISO 19115 standard.

The ISO 19115 standard is compatible with the ontology-based approach to be taken in Seamless IF. The ISO 19115 standard has already been formalized as an ontology by the Seamless partner University of Vermont (see:
http://esd.uvm.edu/cgi-bin/showl.pl?OWL=http://www.integratedmodelling.org/iso-metadata.owl&CC=MD_Metadata).

2.8 Data server

The selected and adapted datasets will be made available to the Seamless partners via a web server at FLD. Tools to explore the data before downloading will also be available. Basically the partners will have the option to explore the data on maps and, via a link from legend, to access metadata. From the metadata a direct link will allow the partners to download the data in ascii format. The first data will be ready for access before the end of September 2005 at http://gis.slnet.dk/seamless/ .

The decision to locate the data server at FLD now is primarily to facilitate the exchange of data for the duration of data. The decision on where to locate the data server after the duration of the project should be taken later in the light of the progress on software development and data processing and taken into consideration the interactions with the models and more practical issues such as data updating and maintenance of the server.

2.9 Other data processing issues

WP4 will follow the standards set by INSPIRE for reference data and meta-data, projections and European reference grid system. In addition, WP4 will define data specific procedures for data preprocessing, including quality checking, maintenance (updating), data gap filling, and interpolation.

WP4 assumes that the Seamless data base is capable to provide spatial input data to a model and to store spatial output data from the model. However, any dynamic process involving continuous mutual influences between spatial units over time, such as in models for stream flow, erosion-sedimentation, and in cellular automata, should be handled by the model itself. The WP4 data base will not be designed for monitoring events with impact on neighboring spatial units

2.10 Conclusion and discussion

The present position paper outlines the views of the WP4 participants on what could be the contribution of WP4 to SEAMLESS with special attention to the interactions with the other WPs. The views of WP4 are based on the DOW and communication with many SEAMLESS partners since the start of the project. This WP4 position paper could not cover all possible options to process data from many sources for different purposes over a range of scales, but we believe that the most relevant European data sets have been mentioned, and the ways they can be combined thematically and spatiallly. As such the present paper can be considered as a minimum work programme of no-regret activities, with a priority setting. WP4 would like to receive confirmation from the other WPs, specifically from WP2, WP3, and WP6 that the proposed WP4 approach and priorities indeed mirror their requirements in data content and access. If this is not the case please provide alternative or additional specifications. Once agreement on data requirements has been reached, the technical realization will be achieved in cooperation with WP5.

3 Annexes

Annex 1: Examples of spatial disaggregation of regional statistical data in several projects

The CAPRI-Dynaspat project

The CAPRI-Dynaspat project (Kempen et al., 2005) distinguishes HSMU Homogeneous Spatial Mapping Unit, which is formed by overlaying a biogeographical map, the NUTS2 administrative map, the CORINE land cover map, a soil surface texture map, a slope map, and the MARS climate grid. The bio-geographical map and the NUTS0 (country level) map were used for stratification of the territory. The remaining natural location factors in the HSMU are used to estimate the share of a single crop in the HSMU area. This requires the application of advanced statistical methods. For the sake of explanation let us first simplify the HSMU to the combination of NUTS2 and CORINE (Regional Land Cover Units, RLCU). Within each stratum these RLCU's were crossed with the LUCAS crop survey results which allows to estimate the share of each single crop in each RLCU as a result of a regression applied to each single CORINE class. The estimation of actual crop shares per RLCU can be refined considerably by taking into account the natural location factors soil texture, slope, elevation, rainfall and duration and temperature of the growing season. These factors were introduced as explanatory variables in the regression equations, which allows the estimation of crop share for the complete HSMU. This is followed by the application of advanced statistical methods to make the estimated crop shares as consistent as possible with the original NUTS2 cropping area statistics.

The EEA Hydrosol data base

JRC-IES and EEA have carried out a study into nitrogen surpluses from agricultural lands per river basin in Europe (Crouzet and Steenmans, 2001). The study used the Hydrosol data base, formed by overlaying the administrative map, river basins and CORINE land cover. The spatial distribution of crops was estimated from the Corine land cover classes. A given crop is grown in the land cover units of the Corine classes consisting largely or partly of arable crop land. It is estimated which area fraction of those land units is potentially available for the crop. For each subregion the potentially available land area can be calculated, which will be usually much larger than the real area under a specific crop. The potentially available land areas per subregion are used as weighing factors for disaggregating the regional statistics on cropped area over the subregions.

The JRC MARS data base

JRC-IPSC has created the MARS data base for European MARS Yield Forecasting System. The distribution of individual crops is derived from the soil map by application of soil suitability rules. De elementary mapping unit is made by overlaying the NUTS2 administrative map, the soil map, and the MARS climate grid. The soil suitability rules excludes soil units which are obviously unsuitable for a given crop (e.g because of rocks, very steep slope, very shallow rooting depth, saline, toxic, very wet conditions), the remaining soils are considered potentially suitable, and are further clustered on the basis of water holding capacity into simulation units for application of the crop growth model. Model output can be aggregated to climate grids, or to soil units, or to NUTS regions.

The ELPEN Project

The ELPEN project developed methods for a European Livestock Policy Evaluation Network (ELPEN). The spatial analysis followed a two step approach: first the FSS-NUTS-data and CORINE land cover map were used to estimate the location of different sector-based FSS farm types, and in the second step FADN data were used to add to these farm types some selected characteristics on intensity and farm size.

The ELPEN project had created a GIS data base by overlaying maps of administrative regions, agricultural districts formed by regrouping NUTS regions, land cover, soil, climate, elevation, less favoured areas as a tool for aggregation and disaggregation of economic, statistical and environmental data at the level of regions and farms. To the basic maps derived maps have been added, e.g maps of soil suitability, nitrogen leaching sensitivity, farm typology. For the estimation of the location of livestock systems rule-based allocation algorithms were developed, that indicate the probability that a specific livestock activity takes place in a grid cell or other spatial entity. Applying these algorithms to the NUTS1-level FSS data base, maps can be produced of the most probable location of a given livestock system within the EU at the scale of 5x5 km grid cells.

Any policy appraisal needs to consider the potential response of different farm types, and their impact in the economic, social and environmental domains. For the distinction of different farm types a farm typology is needed. Traditionally farm typologies have been sector-based and dominated by economic criteria. For the assessment of environmental and social impacts it is necessary to adapt the existing farm typologies by regrouping and refining them, and adding environmental, social and institutional indicators.

In ELPEN a flexible farm typology was proposed, consisting of three dimensions: land use, intensity of inputs and farm size. These three dimensions were chosen because it would allow defining farm types that would behave uniformly in response to policy measures and in their effect on the economic, social, environmental and institutional sustainability indicators. For each dimension of the farm typology, a number of potential parameters (confusingly called indicators) were proposed, to be subdivided into three classes of low, medium, high. This would result in a maximum number of 3x3x3=27 farm types, which number should be reduced by clustering to about 8 within a region, by using farm data from the FADN data base, so that only the most relevant combinations are retained.

The LANMAP project

The landscape map (LANMAP-1) is an database with 2682 landscape mapping units, of which 2600 are larger than 2500 ha. Two hundred and two landscape types are identified in the European Landscape Classification legend, each type having a unique code. The code is based on the dominant altitude, parent material class and land use.

A large advantage of this European Landscape Classification is that its selection of boundaries is consistent, crisp and transparent, based on the underlying layers: topography, parent material and land use. However, if there are any erroneous classifications in any of the three underlying layers, this is reflected in the European Landscape Classification. The European Landscape Classification still lacks information on land use history; although this is a limiting factor, it has proved to be difficult to collect this data at the European scale.

The current landscape classification is now being distributed and revised by a limited number of landscape experts. Improvements on the landscape map will be based on their comments. It is expected that the next version (LANMAP – 2) will be released in 2005 with a pan-European coverage and integrated with climate. Databases like CORINE land cover are being integrated with other land cover data sources such as PELCOM and GLC2000 to obtain a pan-European coverage.

Annex 2: Questionnaire on data needs

WP4 5th of April 2005

PLEASE SEND YOUR ANSWER TO ERLING (eran@kvl.dk) BEFORE THE 25th OF APRIL.

Supporting documents can be found in the folder "consultation" under WP4 shared documents on the Intranet. You will find:

Excel file "overview" with a simple list of the databases we find most relevant. You can use this file if you are familiar with the databases and know exactly which ones to ask for. In this file we have also added some columns that you can use to indicate which of the databases you need to use and some further details.

If you need more detailed information on the selected databases you will find references in the overview file to the 2 reports "Environment" and "soceco_farm" which also can be found in the consultation folder.

In the two reports you will also find descriptions of "close to all" other European databases that could be relevant.

Those of you requiring detailed data on farm management will not find descriptions in the reports. We will include a discussion on this issue in the typology workshop at the Bologna meeting later in April. You are of course welcome to include this issue in the answer to the questions under point B below.

A. Questions in relation to the databases described in the files on the intranet:

(1) Can you indicate which of the databases that are of use as input data to your WP work?

(2) For the databases chosen in (1), please specify briefly per database for what purpose it will be used?

(3) Do you have special requirements in relation to the databases selected in (1) – for example in relation to spatial or temporal scale.

(4) Please specify when you think you will need to use the database chosen in (1).

B. Questions in relation to data not described in the files on the intranet:

(5) Please describe briefly what data you need. (the type of data needed, the purpose for which they need it, preferred temporal coverage, preferred scale). If possible use the same format as in the file "format" in the consultation folder.

(6) Can you direct us to databases that can provide the data described in (5)

Annex 3: Examples of territorial typologies in the eu

In this Annex a number of typologies of EU regions are presented, which have been used in recent studies of regional performance.

OECD territorial typology

In the scope of the Project on Rural Indicators, the OECD has made a typology of rural regions, which covers its whole territory (OECD, 1994). The typology consists of three types of regions, derived on population density:

1. predominantly rural regions;
2. significantly rural (or: intermediate) regions;
3. predominantly urban regions.

The typology is based on a territorial scheme of two hierarchical levels: the local community level and the regional level. Local communities are basic administrative units with a very detailed grid, like cantons in France, districts in the UK and municipalities in the Netherlands. Regions are larger administrative units or functional zones with a less detailed grid, like aemter in Denmark, provincias in Spain and provinces in Belgium and the Netherlands. When population density in local communities is less than 150 inhabitants per square kilometre, the community is classified as 'rural'; when population exceeds 150 inhabitants per square kilometre as 'urban'[1]. As a second step, regions are divided into three groups (Fig. 1):

- when more than 50% of the population of the region lives in rural local communities, the region is classified as 'predominantly rural';
- when between 15 and 50% of the population of the region lives in rural local communities, the region is classified as 'significantly rural' or 'intermediate'[2];
- and when less than 15% of the population of the region lives in rural local communities, the region is classified as 'predominantly urban'.

Moreover, when regions include a city of 200,000 inhabitants or more, the region is classified as intermediate; when regions include a city of 500,000 inhabitants or more, the region is classified as predominantly urban.

[1] For Japan the threshold is 500 inhabitants per square kilometre.

[2] Originally, the term 'significantly rural' was used; as this was difficult to interpret for many users, later the term 'intermediate' was introduced.

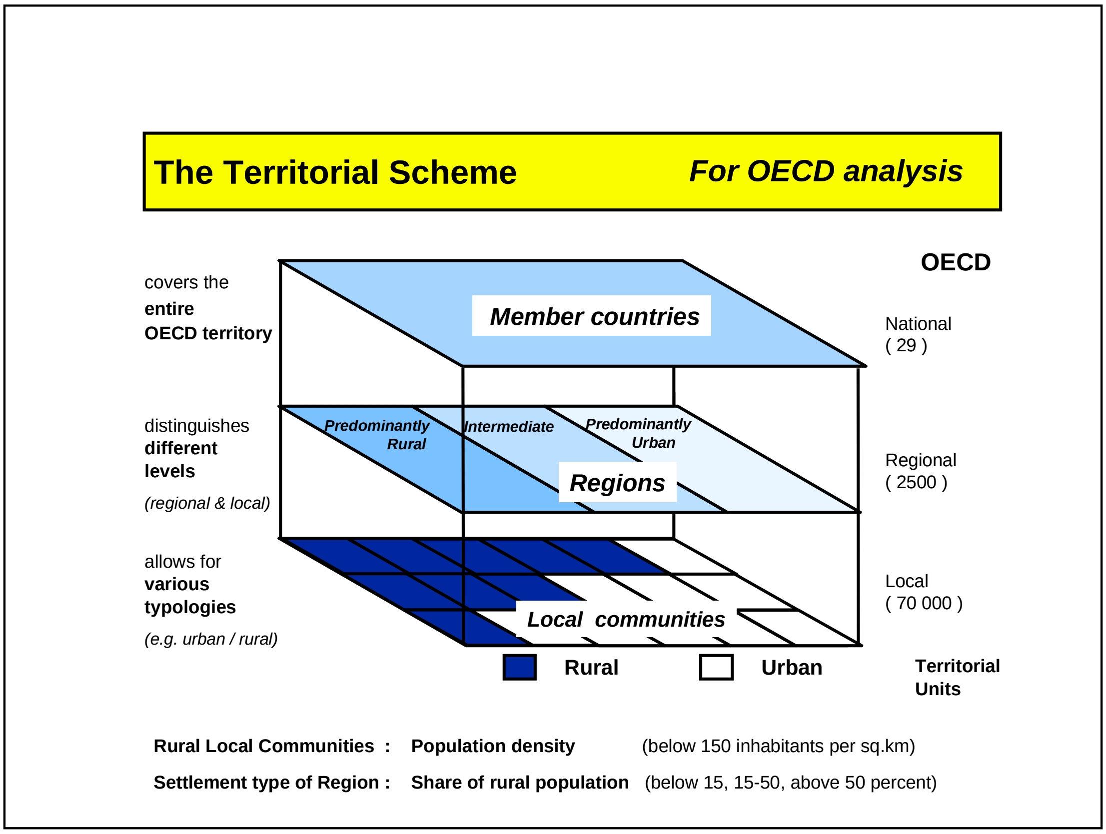


Figure 1 The territorial scheme for OECD analysis

Source: OECD, 1996.

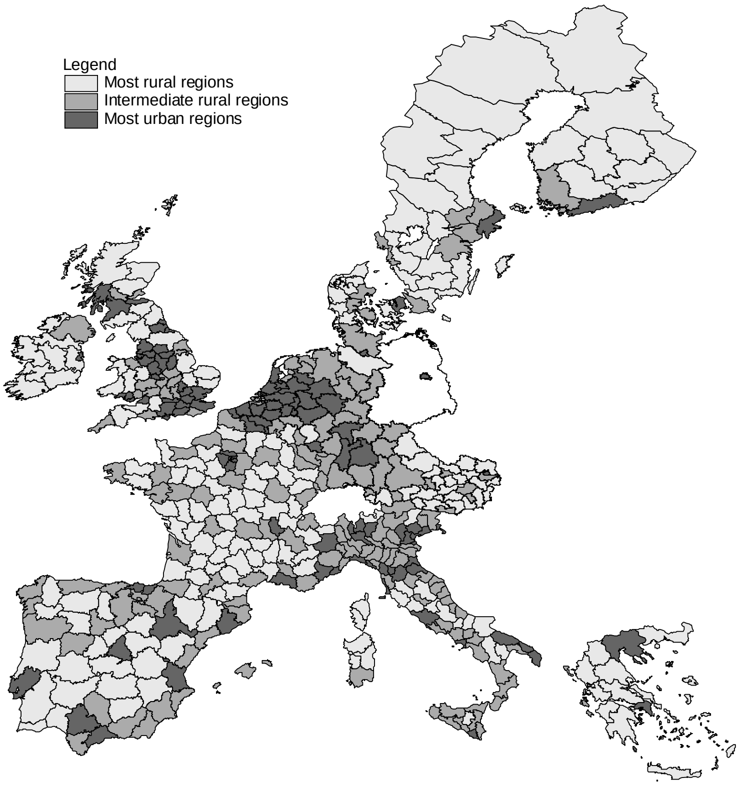


Figure 2 Degree of rurality of the EU15 regions, 1998a)

a) Excluding regions in the former DDR.

Source: RUREMPLO project.

Extension of the OECD typology with leading and lagging

In order to examine differences in employment growth among regions, a distinction of regions into groups with, for example, high, medium and low growth is a useful tool of analysis. In several studies (OECD, 1996, Esposti et al., 1999, Terluin, 2003 and Bollman et al., 2004) extensions are made of OECD typology by distinguishing regions with a high employment growth and a low employment growth during the last decade

LFA typology

In the ESPON 2.1.3 project (2004), a regional typology of less favoured areas has been designed, based on the percentage of LFA in total area. Six clusters are distinguished: 0%, <15%, 15-40%, 40-65%, 65-90%, 90-100%.

Typology based on structural handicaps in the scope of the EU Structural Policy

Since the Reform of the EU Structural Funds (1988), a number of region is eligible for support form the Structural Funds. These are:

- Objective 1 regions: Development and structural adjustment of regions whose development is lagging behind. These regions were defined as those in which GDP/capita was less than 75% of the Community average in the previous three years. They cover the whole of Ireland, Greece and Portugal and large parts of Spain and southern Italy.

- Objective 2 regions: Regions that are suffering from industrial decline.
- Objective 5b regions, aimed at promoting the development of rural areas. These areas were defined according to criteria like a high share of agricultural employment, a low level of agricultural income and a low GDP/capita.

In 1994, to this group were added:

- Objective 6 regions for sparsely populated regions.

In Agenda 2000, these regions a have been reduced to the following two (EC, 2000):

- Objective 1 regions, in which structural development and adjustment in regions lagging in development will be promoted;
- Objective 2 regions, in which support is given for economic and social adjustment in areas with structural difficulties, covering no more than 18% of the EU population

Typology based on the integration of regions in the national/global economy.

In the proposals for 'Agenda 2000' the Commission also uses a classification of rural areas based on the degree of integration into the national economy (EC, 1997a):

1 integrated rural areas: areas with an employment base in the secondary and tertiary sectors, a growing population and potential threats to their environmental, social and cultural heritage;

2 intermediate rural areas: areas relatively distant from urban centres with a varying mix of primary and secondary sectors;

3 remote rural areas: areas with low population densities, heavily dependent on agriculture, isolated by topographic characteristics and providing only the least adequate basic services.

The Organization for Economic Cooperation and Development (OECD) uses also a typology of areas based on the degree of integration into both the national and global economies (OECD, 1996b:4). In this typology, integration in the national/global economy is expressed in terms of distance of a local economy to a major urban centre. Distance does not only refer to physical distance *per se*, but to constraints due to inadequate transportation and telecommunications structure or cultural barriers such as language as well (OECD, 1993:33 a.f.). In this typology three kinds of local economies are distinguished:

1 remote areas: these are peripheral, sparsely populated zones; transportation to major urban centres is inconvenient and time-consuming, which limits commuting or even relatively casual business travel;

2 intermediate areas: local economies that are still heavily dependent on employment in the agricultural sector, due to insufficient economic diversification;

3 economically integrated areas: local economies, often located at the edge of urban centres, with a diversified economic structure. Usually the population in these areas is on the increase due to attractive rural assets.

This typology closely follows the EU classification of rural areas according to their integration into the national economy. In the scope of the Working Party on Territorial Party, this typology based on distances is being linked with economic indicators (OECD (2002), A typology based on an economic approach of the OECD regions; Paris, Territorial Development Service, Working Document DT/TDPC/TI (2002)2).

Territorial typologies on the base of landscape features and biogeographical characteristics

Various territorial typologies on the base of landscape features and biogeographical characteristics can be designed, such as Landscapes, Biogeographical zonations of Europe, Watersheds and Soil Map.

Typology based on accessibility
[to be elaborated]

Cluster typologies
Apart from designing a regional typology based on one indicator, typologies can also be derived from a number of indicators by using cluster analysis. In the ESPON 2.1.3 project (2004) such a cluster typology for the EU 25 is based on population size, population change, population density, GDP/head, change in GDP/head, unemployment rate, change in unemployment, accessibility to markets, number of hotels, change in number of hotels, share of employment in agriculture, forestry and fisheries, average farm size in ESU, share of farmers over 65 years, and land cover of various products. This resulted in a typology with ten clusters for the EU 15:

1. agricultural peripheral regions
2. northern mixed-economy regions
3. vine cultural regions
4. Sweden
5. agricultural tourism (coastal) regions
6. macro-city regions
7. core farming regions
8. southern lagging regions
9. diversified farming regions
10. meso-accessible regions

and seven clusters for the new Member States:

1. Polish cities
2. dynamic remote
3. static remote
4. lagging remote
5. dynamic macro-accesible
6. meso-accessible
7. stable-accessible

Two difficulties arise from such rather complicated clusters. The first one is about their interpretation: what does a specific cluster indicate? The second one refers to the risk that clusters reflect country specific conditions and include only regions from one country.

Harmonised urban-rural typologies
This typology is designed in the ESPON 1.1.2 project (2004) and based on the degree of urban influence and the degree of human intervention. The degree of urban influence is derived from two factors: population density and status of the leading urban center of the region. When a region has a population density above the EU average of 107 inh./km^2, the region is said to have a high urban influence; when its population density is below 107 inh./km^2, the region is said to have a low urban influence. In addition, a region is labeled as high urban influence when its leading urban center is classified as a 'Metropolitan European Growth Area'. The degree of human intervention is determined by the relative share of land cover: artificial surfaces, agricultural areas and residual land cover. The European average of artificial land cover is 3.5 percent of the total land cover. The corresponding figure of agricultural land is 50.4 and of the residual group it is 46.2. High urban intervention corresponds to a situation where the share of artificial surfaces (and possibly one of the two other land cover categories) is above European average. Medium human intervention equals the cases where the share of

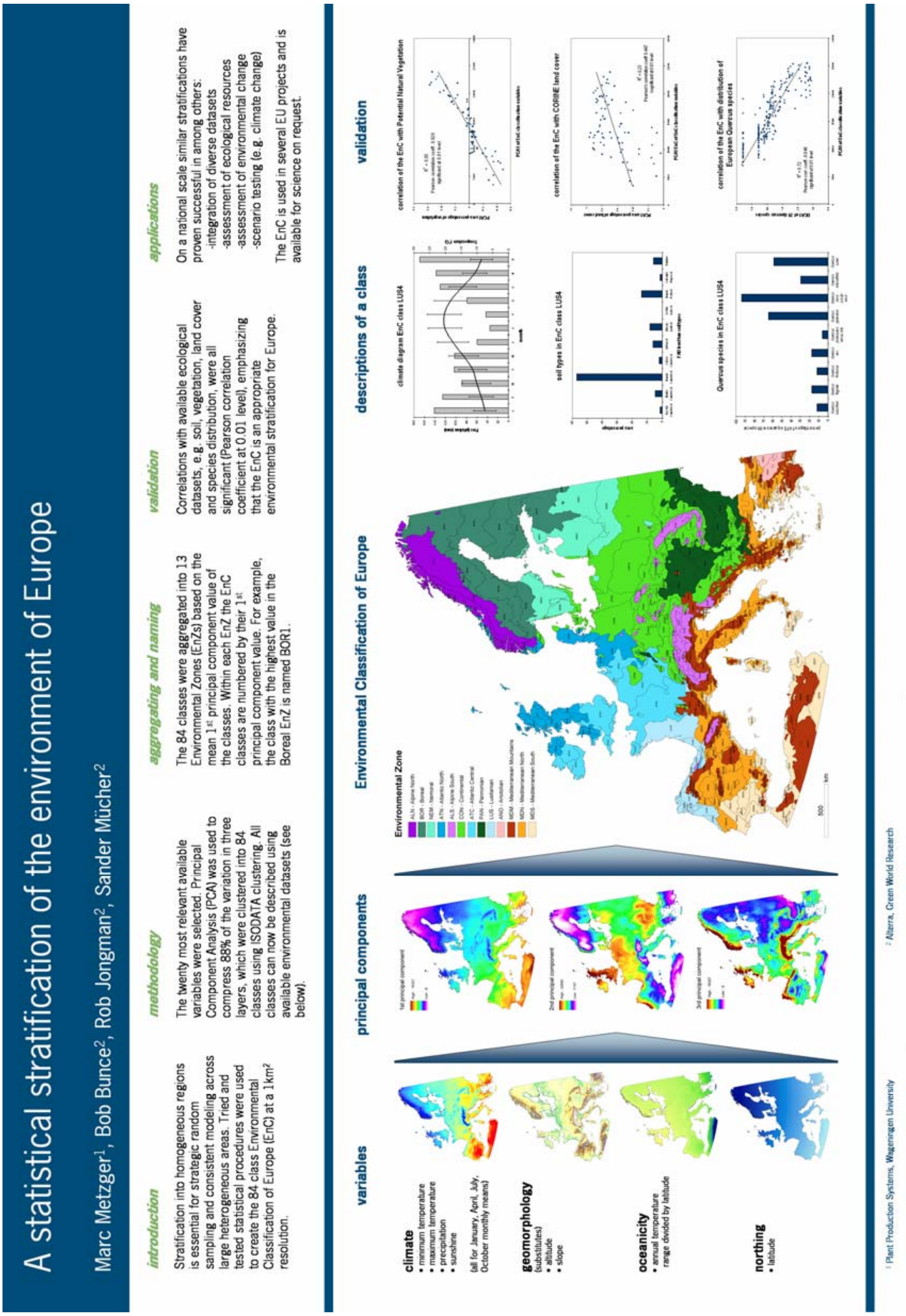
agricultural land (and possibly the share of residual land cover) is above European average. Low human intervention concerns all cases where only the share of residual land cover is above European average.

The harmonised urban-rural typology consists of the following types:

1. High urban influence, high human intervention
2. High urban influence, medium human intervention
3. High urban influence, low human intervention
4. Low urban influence, high human intervention
5. Low urban influence, medium human intervention
6. Low urban influence, low human intervention

Typology based on natural cultural values and economic development
This typology is designed in the ESPON 1.3.2 project (2003) and based on the natural and cultural heritage structure (mountainous regions, smaller islands, coastal zones and other regions) and the socio- economic structure.

Annex 4 : Environmental Zones of Europe



Annex 5: Metadata profile SENSOR

Metadata on metadata **ISO Code**

- Point of contact
    - Name of contact organisation * 8.376
    - Name of contact person * 8.375
    - Position of contact person 8.377
    - Role of organisation 8.379
    - Address: Delivery point * 8.378.389.381
    - Address: City * 8.378.389.382
    - Address: Province, state * 8.378.389.383
    - Address: Postal code * 8.378.389.384
    - Address: Country * 8.378.389.385
    - Address: E-mail * 8.378.389.386
- Last modified * 9
- Name of standard 10
- Version of standard 11

Data set identification

- Title of the data set * 15.24.360
- Alternative title * 15.24.361
- Abstract * 15.25
- Keywords * 15.33.53
- Topic category * 15.41
- Version of data set * 15.24.363
- Date of version * 15.24.362.394

Reference system

- Name of reference system (*) 13.196.207
- Datum name (*) 13.192.207
- Ellipsoid (*)
    - Name of ellipsoid (*) 13.191.207
    - Semi-major axis (*) 13.193.202
    - Axis units (*) 13.193.203
    - Flattering ratio (*) 13.193.204
- Projection
    - Name of projection (*) 13.190.207
    - Standard parallel (*) 13.194.217
    - Longitude of central meridian (*) 13.194.218
    - Latitude of projection origin (*) 13.194.219
    - False easting (*) 13.194.220
    - False northing (*) 13.194.221
    - False easting northing units (*) 13.194.222
    - Scale factor at equator (*) 13.194.223
    - Longitude of projection centre (*) 13.194.224
    - Latitude of projection centre (*) 13.194.225

Distribution information

- Owner
  - Name of owner organisation * 15.29.376
  - Name of contact person 15.29.375
  - Position of contact person 15.29.377
  - Role of owner organisation 15.29.379
  - Address: Delivery point 15.29.378.389.381
  - Address: City 15.29.378.389.382
  - Address: Province, state 15.29.378.389.383
  - Address: Postal code 15.29.378.389.384
  - Address: Country 15.29.378.389.385
  - Address: E-mail 15.29.378.389.386
- Originator
  - Name of originator organisation 15.29.376
  - Name of contact person
  - Position of contact person
  - Role of originator organisation
  - Address: Delivery point
  - Address: City
  - Address: Province, state
  - Address: Postal code
  - Address: Country
  - Address: E-mail
- Processor
  - Name of processor organisation
  - Name of contact person
  - Position of contact person
  - Role of processor organisation
  - Address: Delivery point
  - Address: City
  - Address: Province, state
  - Address: Postal code
  - Address: Country
  - Address: E-mail
- Distributor
  - Name of distributor organisation
  - Name of contact person
  - Position of contact person
  - Role in distributor organisation
  - Address: Delivery point
  - Address: City
  - Address: Province, state
  - Address: Postal code
  - Address: Country
  - Address: E-mail
  - On-line delivery
- Access rights
  - Type of constraint 20.70
  - Description of restriction 20.72

Other information

- Language within the data set * 15.39
- Exchange format

- Name of exchange format   *   15.32.285
- Version of exchange format   *   15.32.286
- Methodology description   18.81.83
  - Link to methodological report
- Changes since last version
- Process steps
  - Description of process steps   18.81.84.87
  - Resource name   18.81.84.91.360
  - Resource date   18.81.84.91.362
- Scale   *   15.38.60.57
- Geographic accuracy   15.38.60.57
- Geographic box   15.38.61
  - West bound longitude   (*)   15.45.336.344
  - East bound longitude   (*)   15.45.336.345
  - South bound latitude   (*)   15.45.336.346
  - North bound latitude   (*)   15.45.336.347
- Geographic coverage by name   *
- *List of attributes*
- *Data type (vector / raster)*

Annex 6: Standard selection of Nuts regions

We aim to use the standard selection of Nuts regions elaborated by Alterra where appropriate. This comprises a selection of regions aiming for a relatively uniform region size across EU and taking into account data availability. A total of 573 Nuts regions are included in EU25+2+2+1 (pus Romania, Bulgaria, Switzerland, Norway and Iceland). This selection will also be used in the Sensor project.