MICROECONOMIC AND GEO-PHYSICAL DATA INTEGRATION FOR AGRI-ENVIRONMENTAL ANALYSIS, GEOREFERENCING FADN DATA: A CASE STUDY IN ITALY

Authors

Andrea Fais, Pasquale Nino, Antonio Giampaolo

National Institute for Agricultural Economics

Rome - Italy

andrea.fais@inea.it; pasquale.nino@inea.it;

Paper prepared for presentation at the XI[th] seminar of the EAAE

(European Association of Agricultural Economists),

"The Future of Rural Europe in the Global Agri-Food System"

Copenhagen, Denmark, August 24-27, 2005

Abstract

The authors, inside a TAPAS[1] action, have developed a methodology - based on the integration between FADN[2] data with statistical, administrative and cartographic information - to improve the production of statistical data related to agricultural policies impact on land and environment. This methodology allows - using Geographic Information System (GIS) technology - to produce and organise data at geographical level. Data spatially referenced respond to the specific needs of agri-environmental analysis and problems, mostly related to specific areas (environmental vulnerability areas), inside defined boundaries (e.g. river basin). The GIS is implemented with several cartographic layers (Topographic and cadastral maps, land use, soils, water sources, climate, Digital Terrain Model etc.) and is related to the regional FADN database.

Keywords: Agricultural statistics, Agri-environment, FADN, GIS, TAPAS, DPSIR
Jel: C81, O13, Q21

Introduction

The need of reliable, meaningful and accurate statistics to monitor agri-environmental policies and their application is continuously growing in EU since almost the last ten years. Thanks to the TAPAS (Technical Action Plans for Agricultural Statistics) actions, the EU Commission is strongly supporting the improvement of agri-environment statistics.

DPSIR model (the EEA modified version of the OECD[3] conceptual scheme) is the basis, commonly accepted at international level, to analysis the complexity of the relationships between agriculture activities and environment. The central point of the model is the **State**, represented by quality and quantity of natural resources. **Pressures,** caused by human and natural activities (**Drivers**), can alter this state (Natural resources degradation; Water pollution; Air pollution; Climate change). This alteration has effects (**Impact**) on the whole ecosystem. **Responses** are elaborated (from Public and Private Bodies) to prevent the risk derived from the Impact e/o to restore appropriate environmental condition. The Commission [doc. COM(2000) 20 final], fix priorities and topics in the development of agri-environmental indicators, and focuses the attention on two critical elements:

1. Make better use of existing data, when they have not yet been fully exploited;
2. Other sources of information, such as geographical databases, should be better exploited.

At OECD and EEA level (DPSIR model) and in the IRENE's[4] project some methodological priorities and recommendations have been fixed, the following being absolutely relevant:

- needs of the integration between statistical, socio-economics, administrative and spatial data, towards the introduction of geo-referencing into existing data sets. Precise spatial referencing of all relevant data sets (a high spatial resolution and/or stratification is required), allowing their integration, is considered the key element to improve and regionalise environmental analysis;
- FADN combines data on farm structure, input use and economic variables, and provide information on structural change and other farm trends. The combination of such different variables in one data set is a key factor for linking different issues in agri-environmental analysis;
- possibilities for crossing-over and validation between different data sets are very important and such options should be re-inforced.

[1] TAPAS action represented one of the Commission instruments (Decision 96/411/EC) for the improvement of agricultural statistics and monitoring the integration between agriculture and environment.

[2] The **F**arm **A**ccountancy **D**ata **N**etwork (FADN) is an instrument for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy. The concept of the FADN was launched in 1965, when Council Regulation 79/65 established the legal basis for the organisation of the network.

[3] Organisation of Economic Cooperation and Development - European Environment Agency

[4] The IRENA project (Indicator Reporting on the integration of ENvironmental concerns into Agricultural policy), carried out by the European Environment Agency in Co-operation with the EUROSTAT "Agriculture and Environment" working group, has defined the general and specific rules to utilise agricultural statistics in the agri-environment analysis.

A key element to respond to these needs and requirements is the definition of a methodology to link socio-economical and structural agricultural data to physical/territorial data. The baseline to develop such methodology is the experimentation of methods for georeferencing farm data. This allows to produce and organise socioeconomic data at geographical level (normally these data are organised at administrative – NUTS II or III – level), and to utilise agricultural statistics in agri-environmental analysis (environmental problems are related to specific areas, inside defined geographic boundaries, usually never corresponding to administrative units). This responds also to the needs of establishing quantitative relationship between farm and pertinent land.

At the beginning of the INEA's TAPAS experiences, and still now, it was quite impossible to find any work on agricultural socio-economics georeferencing in literature (Liu, J. 2001). The authors main experience in this field concerned the use of GIS (Geographic Information System) technology for the integration of different data typologies and sources (cartographic layers, administrative and statistic data, manuals and data on diseases distribution and pesticides requirements of each crop) to develop indirect statistics and analysis in water use in agriculture (Fais, A. 1997 - Fais, A. and Nino, P. 2000, 2001).

The general goal of the methodology developed is to establish the possibility of utilising the already existing agricultural statistics as a tool for analysing agriculture's impact on environment. In particular, the full utilisation of FADN data (one of the permanent EU agricultural statistics network, with annual information on structural change and other farm trends), could be a key element to develop the use of agricultural statistical data on environment monitoring and analysis.

The present paper refers to the Italian TAPAS actions "**Use of pesticides**" and "**Spatial referencing of FADN data**", on the possibility of utilising agricultural statistics as a tool for analysing agriculture's impact on environment. Key issues of the methodologies developed are:

- Exploit if inside the FADN database is possible retrieve data to environmental analysis;
- Spatial referencing/GIS compatibility, in the integration of microeconomics (FADN) and statistical, administrative and physical (topographic and cadastral maps, land use, soils, water use, climate, Digital Terrain Model etc.) data;
- Spatial resolution requirement for the integration of FADN with other data sets;
- Cost/effectiveness of different FADN georeferencing methods.

Methods and activities

Area of interest

The area of the province of Chieti (Total surface: 258.648,42 hectares) in the Abruzzo Region, has been selected for the realisation of both the TAPAS action. In the 145.627 hectares of agricultural land-use (from CASI 3[5] - see table 1), Vineyards, Olive grows and orchards (76.307 hectares) have a significant role in agri-environmental analysis, being cultures and farms of high capital investment.

Table 1. Agricultural distribution of the Chieti Province area (INEA - Casi 3 Land use map)

Culture	Hectares	Culture	Hectares
211 – Cereals (not corn)	62.786	221 - Vineyards	30.215
2121 - Industrial Cultures	2.739	223 – Olive groves	42.285
2122 – Autumn vegetables	1.979	222 - Orchards	3.807
2123 – Summer vegetables	1.816		76.307

[5] Land use cartography comparable to CORINE Land Cover (CLC), obtained from satellite images interpretation (Landsat TM), digital restitution based on digital orto photo (grey tones, 1 meter resolution). The classification system derived from CLC, with a IV[th] level for the irrigated classes. The acquisition scale was 1:50.000 for the irrigated classes, while for the other classes (urban, woods, waters) the CORINE scale of 1:100.000 was maintained. The geometric accuracy is 1:25.000 scale equivalent. Other important crop are derived from short term ISTAT data National Institute of Statistic

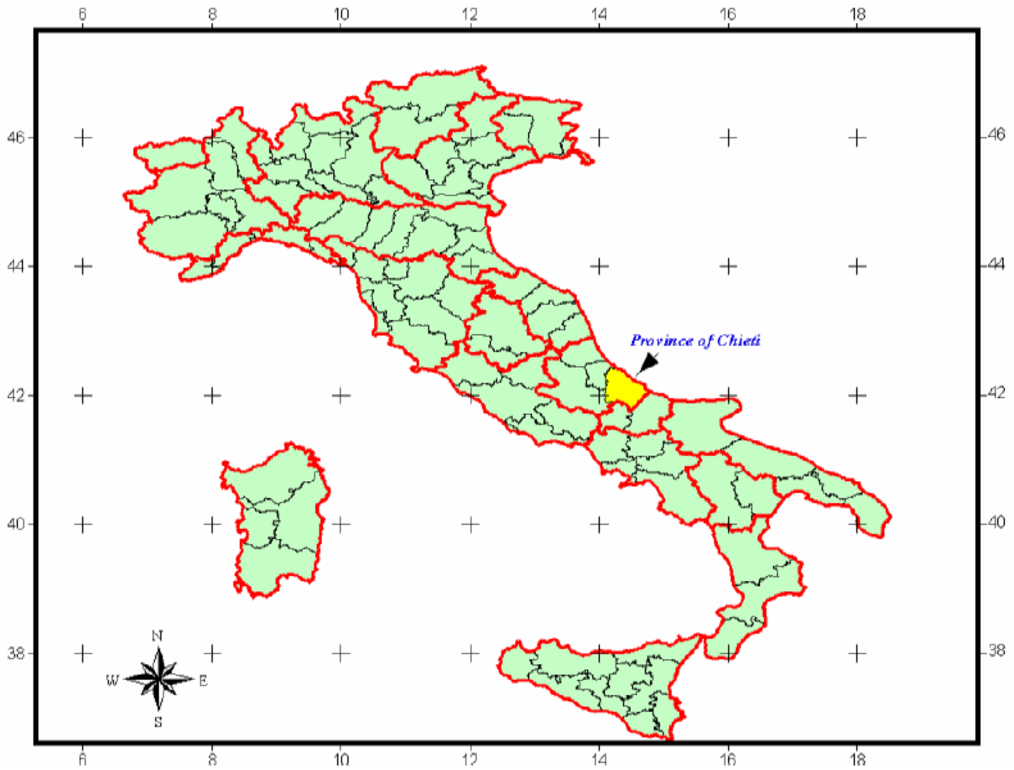


Figure 1 - Study area location

Within the class 211 - not irrigated arable land - the most represented culture is wheat 24.650 he (80,56 % of the whole class).

Within the class 2121 - Irrigated industrial crops at spring/summer cycle - the most represented culture is tobacco with 1.100 he (38,04 % of the whole class).

Within class 2122 - irrigated Horticulture at summer/autumn/spring cycle - the most represented culture is artichoke with 315 he (28,10 % of the whole class).

Within the class 2123 - irrigated Horticulture at spring/summer cycle - the most represented culture is tomato with 978 he (63,38 % of the whole class).

Within the class 2221 - irrigated fruit and berries plantation - the most represented culture is peach with 1.846 he (64,86 % of the whole class).

FADN mini data bank

FADN is a EU network that provides useful information on agricultural microeconomics, obtained from systematic surveys by permanent networks of regional surveyors. The evaluation of data collected from the individual farms may support the formulation of agricultural policies and their evaluation of the impact on the environment. Furthermore, FADN, combining data on farm structure, input use and economic variables could be a key factor for linking different issues in agri-environmental analysis. **The mini data bank** derives, after several formal and contents controls, from the FADN regional network data collection. The regional data bank (RDB) is structured in eight files plus a farm register file (due to the privacy legislation). The mBDR includes all the farm elementary data coming from the regional FADN surveys network. Data are organised to respond to final user query requirements. The physical implementation of the RDB is managed by INEA's regional researchers. The batch command creates 8 output files in DBF format. Each table include farm data on general information, crops, livestock, subsidies and labour. The table name is characterised by a suffix on the information typologies, followed by the UE NUTS code and the accountancy year. The single tables contain several variables that can be usefully utilised in the agri-environmental indicators

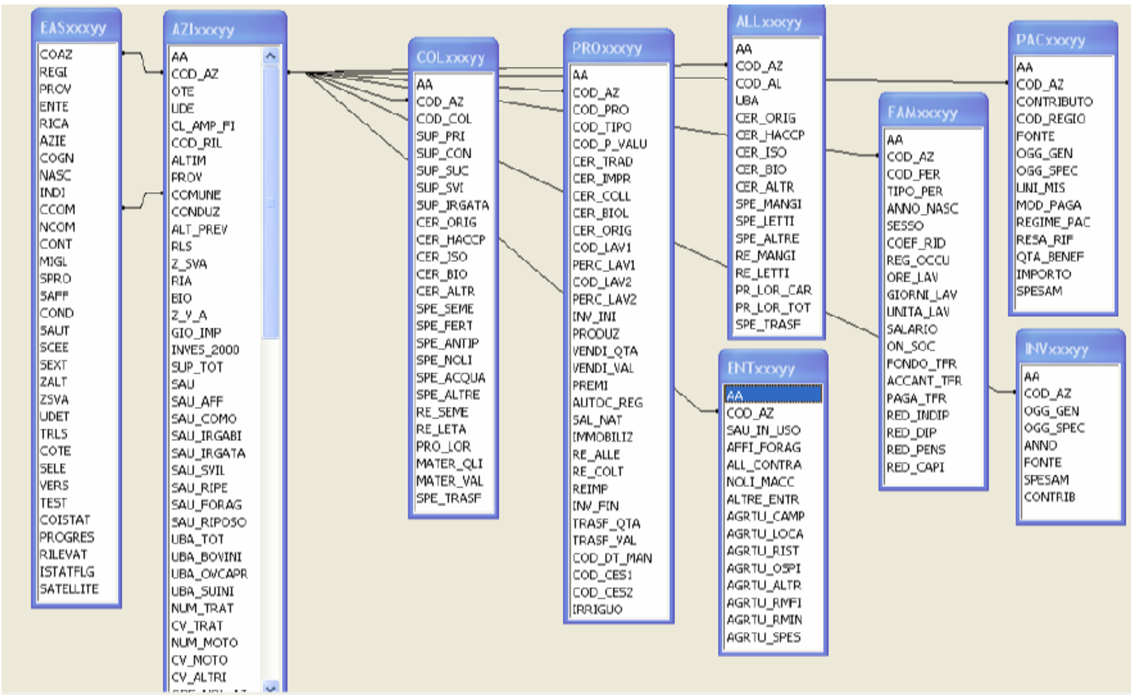


Figure 2. FADN mini data bank architecture

implementation, as well in the definition of impact indicators lines to evaluate farmers attitudes towards environment and land conservation.

Table 2. Description of the mBDR tables

Table	Contents	Example of others variables for agri-environmental indicators
AZI	*Data on farm typologies (ote, ude, rls, comune, dimensione fisica, elementi strutturali, patrimoniali e reddituali);*	**Z_V_A:** *farm parcels on agricultural area with environmental constrains (SIC, ZPS and other sites of natural interest, NATURA2000, etc);* **SAU_IRRIGATA**: *farm areas annually irrigated, in hectares (in the table COLTxxxyy each irrigated crop area is fixed);* **SUP_RIPE**: *annual multiple crops areas, in hectares (e.g., second harvesting maize, industrial horticulture, flowers, green houses horticulture, etc.);* **SUP_RIPOSO:** *farm set-aside surfaces for both EU schemes obligations (arable land OCM CAP) and farmers choices (without subsidies);* **CAP_MACC:** *farm ownership (agricultural structures, machines and equipment) total value.*
COL	*Productive and economic data of each orchard and arable land crop (occupazione, rotazione, certificazione di processo, produzione fisica, reimpieghi aziendali, trasformazioni, costi, ricavi e autoconsumi)*	**SUP_IRRIGATA:** *hectares of each irrigated crop;* **PROD_LOR:** *total value per crop of the gross production.*
ALL	*Productive and economic data of each farm livestock typology (specializzazione produttiva, consistenze medie annuali, certificazione di processo, produzione fisica, reimpieghi aziendali, trasformazioni, costi e ricavi, autoconsumi)*	**UBA:** *number of adult livestock units;* **RE_LETTI:** *total value of fodder and straw farm production and reused for the farm's livestock*
PRO	*Data on vegetal and animal productions (principal and transformed) at farm level (tipologia, certificazione di qualità, produzione fisica, tipo di lavorazione, costi, ricavi, contributi specifici, modalità di commercializzazione);*	**COD_P_VALU:** *codes of crop and animals productions typologies (principal, secondary and transformed) of the bookkeeping year;* **CER_BIOL:** *it points out farm biological certified productions;*
ENT	*Data related to the extra-characteristic/multi-functionality activities of the modern farms*	
CAP	*Data on public subsidies (classified on the basis of source, general measure and single farm sector);*	
INV	*Data on public subsidies for farm investments (classified on the basis of source and investment typology);*	
FAM	*Data on farm's wage and familiar labour composition (worker typology, type of occupation, days and hours of annual job, salaries and other charges);*	
EAS	*Farm register data (denomination, municipality and province, address, codes ote and ude).*	

The eight mBDR tables, with the description of the related contents and some example of variables usable on the agri-environmental indicators, are in the table 2. The system/dB architecture/scheme is described in the figure 2.

The FADN database analysis allowed to establish the variables, according to the georeferencing type requirements and correlated spatial and administrative data, could be usefully used in the agri-environment indicators development. The following table show the selected economic variables and the possible parameter that is possible to derive from them.

Table 3. FADN data and georeferencing type requirements per agri-environmental indicator

FADN Data	Georeferencing type	Spatial/administrative Data linked	Parameter derived	EUROSTAT Area
Expense in pestisides for crop + designed questionnaire	Farm centre	Land use, DTM, river basin	Application rates of three pesticide categories: • insecticides/herbicides/fungicide • disease distribution	Pesticide consumption
Farm capital /ha + intermediate consumption /ha	Farm centre	Land use, DTM, river basin	Intensity distribution of farm management practices (IFP)	Farm management (practices)
Expenses in fertiliser use	Farm centre GPS, Cadastral parcel	Land use, river basin	Intensity distribution of Fertiliser Consumption (FC)	Fertiliser consumption
Pasture data	GPS, Cadastral parcel	Land use, DTM (to derive slope), administrative data on breeding consistence (ISTAT)	Intensity Grazing Rate Indicator, IGR	Soil erosion
ZVA[1]+ BIO	Cadastral parcel	Land use, protected areas boundary, administrative data on bio subside (AGEA) at municipality level	% of agriculture area under eco - compatible techniques	Area under agri-environment support and under organic farming
Agro - environmental measures + RIA[2]	Cadastral parcel	Land use, administrative data on Agro - environmental (AGEA) at municipality level		
N ° tractors + CV + repeated surface	Cadastral parcel	Land use, DTM (to derive slope)	Under development	Soil erosion
Irrigated surface + Expenses in water use	GPS	Land use, Irrigated areas map, irrigation network extension	Under development	Water consumption

[1] Environmental constrain (agricultural area prevalence in protected or constrained area: Dir. EEC .92 /43 and 79/409)
[2] Low environment impact perceived by the farmers on the basis of the agri-techniques

For the application of the methodology 4 indicator has been developed:

1. Used quantities of different pesticide categories, using FADN farm centre georeferencing type
2. Intensity Farm Practices (IFP), using FADN farm centre georeferencing type
3. Fertiliser Consumption (FC), using FADN farm centre and GPS georeferencing type
4. Intensity Grazing Rate Indicator (IGR), using FADN cadastral parcel georeferencing type

The methodology to derive the first indicator has been developed inside the TAPAS actions "**Use of pesticides**" while the other three inside the TAPAS "**Spatial referencing of FADN data**"

Other common task in both TAPAS action is represented that FADN farm centres are point dataset with known value for the variable of interest, the problem is how to predict the value of the variable at unknown location. This is normally solved with interpolation techniques, thus we test different type of these techniques by fixing the following requirements to retain the choose of the method that better achieve our task:

- Integration within GIS modelling environment and easy to use
- Exact interpolator (predict a value identical to the measured value at a sampled location)
- Fast computing and modelling time

Analysis of different interpolating techniques to distribute FADN sample data

There are many interpolation techniques, that can be grouped in two principal categories:

- Deterministic - create surface from measured point, based on either the extent of similarity (e.g. Inverse Distance Weighted) or the degree of smoothing (e.g. radial basis function). Deterministic techniques are dependent on mathematical model.
- Geostatistical - create surface using the statistical properties of the measured points. Many methods are associated with geostatistics, but they belong all at Kriging family (ordinary, simple, universal, probability, indicator, and disjunctive). Geostatistical techniques depend on mathematical and statistical model.

To distribute FADN data we compare the performance of the Inverse Distance Weighted (IDW) and ordinary Kriging method:

1 The Inverse Distance Weighted (IDW) method, is a simple interpolation techniques that can often produce good results. The basic assumption of inverse distance is that data points are weighted by the inverse of their distance to the estimation point. This approach has the effect of giving more influence to nearby data points than those farther away. Additionally, the inverted distance weight can be raised to further reduce the effect of data points located farther away. This approach is mathematically expressed as:

$$v_0 = \frac{\sum_{i=1}^{N(v_0)} \frac{1}{d_i^p} v_i}{\sum_{i=1}^{Nv_0} d_i^p} \qquad (1)$$

where:
v_0 is the estimated value at (x_0, y_0), v_i is a neighbouring data value at (x_i, y_i),
d_i is the distance between (x_0,y_0) and (x_i,y_i), p is the power
$N(v_0)$ is the number of data points in the neighbourhood of v_0.

2 Ordinary Kriging (as all Kriging method) is divided into two distinct steps:

- quantifying the spatial structure of the data (called variography), where a spatial - dependence model of the data is configured throw statistical analysis
- producing a prediction

Table 4.Comparison of interpolation techniques

Method	**IDW**	**Kriging**
GIS Integration	Yes	Yes
Easy to use	Yes	No
Computing time	Fast	Moderately fast
Modelling time	Fast	Slow
Exact interpolator	Yes	Yes, without measurement error No, with measurement error
Advantages	Few parameter decision	Very flexible, allows assessment of spatial autocorrelation
Disadvantages	No assessment of prediction error: produce "bull eyes" around data location	Need to make many decision on transformation, trends, models, parameters and neighbourhood

Use of pesticides

In this part of the paper particular attention is paid on the valorisation of FADN network data to develop geo - spatial distribution map on plant diseases, essential to produce agri-environmental statistics. The work started in September 2001 with the collection of data on land use and on pesticides best practices. The experimentation of spatial distribution methods started georeferencing the farm centre of the FADN sample. The local FADN surveyors network, normally in charge of the on-farm accountancy support and surveys, was involved in farm centre localisation on topographic map (UTM-ED50, scale 1:25.000). The final results were obtained in July 2002.

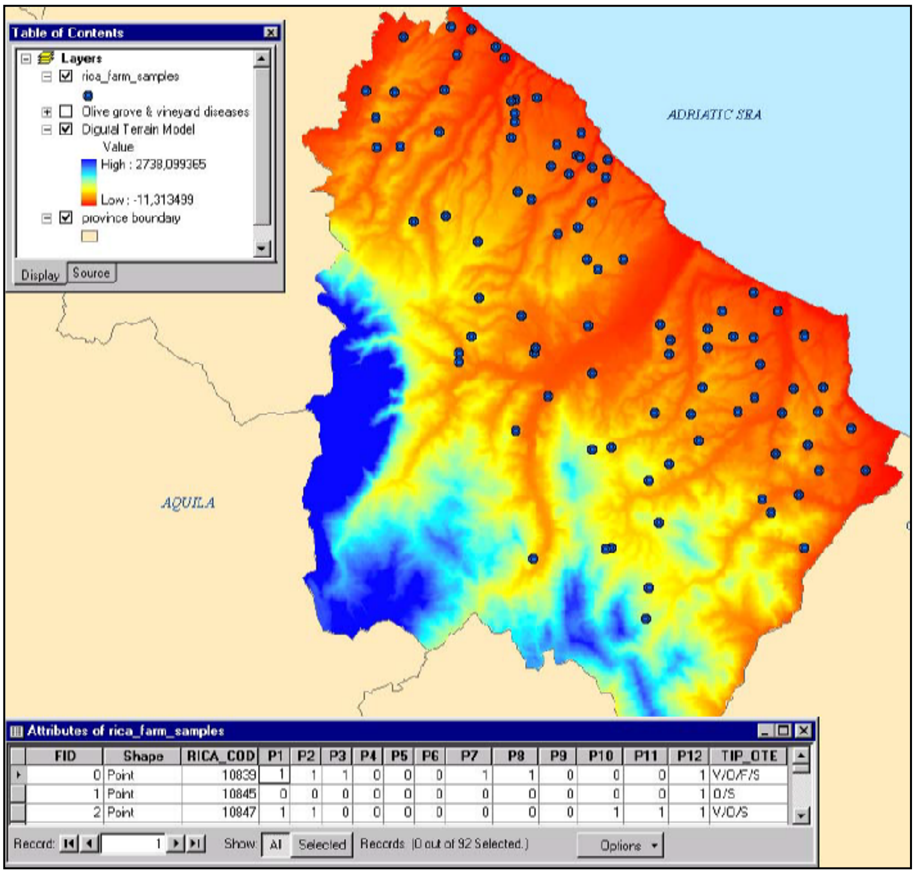


Figure 3. FADN sample distribution inside the study area

The activities carried out to completely meet the methodological approach were:

- realisation of the geo-database on land use and agricultural cultivation distribution in 2001;
- acquisition and elaboration of:
  - national data bank on pesticides: legislation and practical uses, on the base of GAP: Good Agricultural Practice and Pesticides Pathology Institute (ISPaVE) Data Bank. This data bank contains the suggested optimal quantities for each culture, of the different pesticides, to be used depending on the found diseases;
  - research studies has been done with the ARSSA Institute (Plant disease observatory - OMP) especially on the diffusion of the diseases;
  - data on industrial pesticide production and sales (1998-1999, from Agricultural Ministry);
  - data on pesticide uses at farm level (on the basis of a specific questionnaire for FADN sample - 2001);
  - extension services data (2001);
- definition of the methodologies to link the land use database with the pesticide database and to correct the data on the average/optimal quantities of pesticides per disease/crop with the 2001 effective use (based on a farm sample);
- realisation of models and related software;
- calculation of normalised pesticides use per land use polygon;

- realisation of the pesticides use (normalised) map;
- aggregation of data at river basin and administrative (NUTS IV and V) level;
- calculation of annual pesticides use per land use polygon;
- data validation and correlation with administrative data at NUTS IV level.

Methodology used in the spatial distribution of Plant diseas Mapping

1. Using the FADN farms sample, a vector point layer has been created, divided in three categories:
    - Vineyard - Olive grove farms (most important crop in study area)
    - Orchard
    - Arable land

    Each of these farms sample has been coded regarding the disease observed during the farm survey, using the code reported on table 5.

    The structure of the code used (field P1……..P2) is a Boolean type:
    - 0 = absence of disease or infestant
    - 1= presence of disease or infestant

Tab. 5 – Main diseases on FADN sample

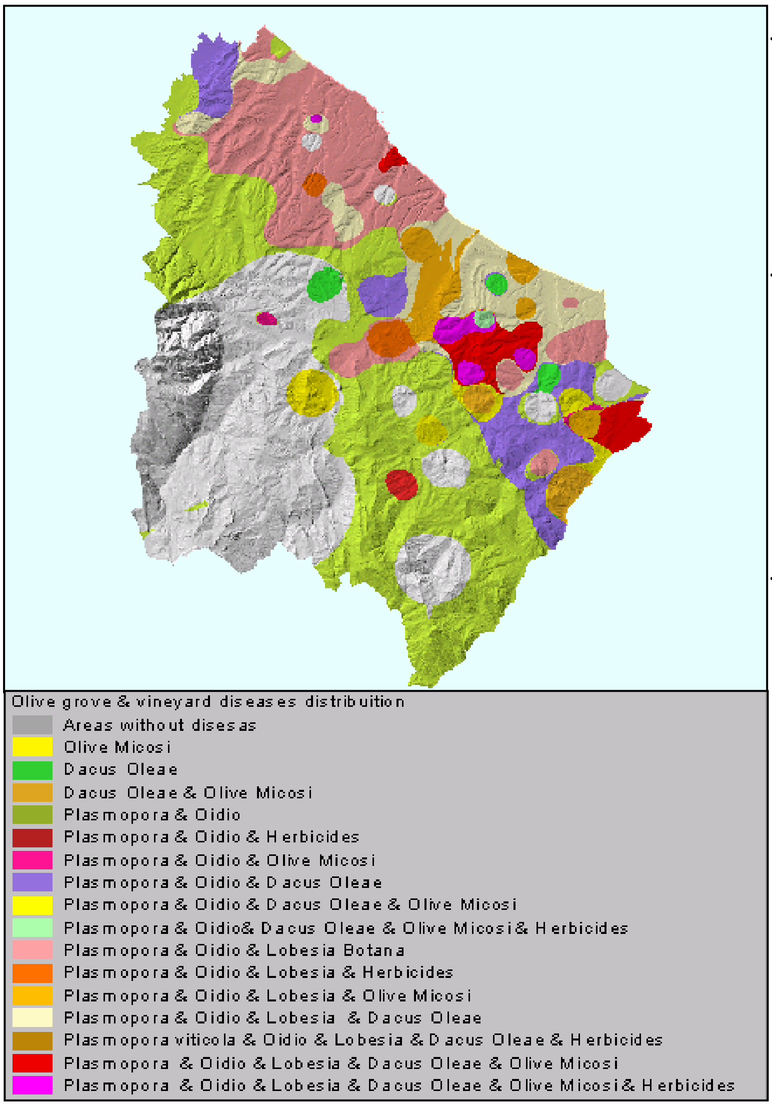
Crop	Observed disease	Code
Vineyard	Plasmopora viticola	P1
	Uncicola Necator	P2
Vineyard	Lobesia Botrana	P3
Olive grove	Spilocaea oleagina	P4
	Dacus oleae	P5
Vineyard + Olive grove	Infesting	P6
Orchard	Mycosis various	P7
	Insect various	P8
	Infesting	P9
Arable land	Mycosis various	P10
	Insect various	P11
	Infesting	P12

2. Next step has been the interpolation between the sample point to obtain a predict value for unknown location, using the Inverse Distance Weighted (IDW) method
3. Therefore trough the interpolation process, 12 different layers have been created (in GRID format with a 100 meter resolution) for each code number;
4. The obtained layers have been summarised into two final grid layers (one for the olive grove and vineyard – code P1, P6 and the other one for the disease distribution for Orchard and Arable land disease distribution code P7, P12); the grid summarising "need to compute an offset that will scale each grid so they are added together without overlapping cell value" Ersts (2001). To compute the offset following map calculator has been used, scaling the different grid by a power of 10:

$$NewGrid = (GridP1 * 10) + (GridP2 * 10) + ...(GridPn) \qquad (2)$$

Table 6. Example of grid summarising

Grid P1					Grid P2					Summarise P1 P2				Legend
1	0	1	1		0	0	1	1		10	00	11	11	00 = no disease
1	1	0	0	+	0	1	0	0	=	10	11	00	00	10 = presence of Plasmopora
0	0	1	0		0	0	1	1		00	00	11	01	01 = presence of Uncicola Necator
1	1	0	1		1	0	1	1		11	10	01	11	11 = presence of both Plasmopora and Uncicola



. 5. The Grid layer has been converted into a vector layer, which has been intersected with the land use cover, thus obtaining for each polygon the information about the different type of disease found.

. Using the surveyed farms data about the pesticides quantities used on the farms and those from best practices data (ISPaVE) a JOIN table has been created which has been linked with Land use-disease diffusion vector layer table, getting in this way the average quantity used for each polygon;

. Calculating and dividing alphanumeric data regarding the Fungicide, Insecticide and Herbicide total quantities in two categories:

- Data obtained from the farms data survey;
- Data obtained from the best practice data.

The difference between the two sets of data has been calculated.

Figure 4. Grid summarising related to vineyard and olive grove diseases distribution.

Spatial referencing of FADN data

The work started in September 2002 with data mining and collection (land use and soil maps, DEM, organisation of the FADN regional mini data bank), and with the selection and analysis of the agri-environment indicator feasible with FADN data. As aforementioned three indicators has been developed inside this TAPAS action.

1. Intensity Farm Practices (IFP), using FADN farm centre georeferencing type
2. Fertiliser Consumption (FC), using FADN farm centre and GPS georeferencing type
3. Intensity Grazing Rate Indicator (IGR), using FADN cadastral parcel georeferencing type

This part of the paper describe the methodologies applied on computing the indicators.

Intensity Farm Practices Indicator

Using the FADN farms sample georeferenced, a vector point layer has been created, divided in three farm type (representing the most important crop in study area):

- Vineyard: 127 farm (Corine Land Cover code 221)
- Olive grove and Orchard: 103 farms (Corine Land Cover code 223 and 222)

- Arable land: 43 farms (including glasshouses) (Corine Land Cover code 211 - 212)

Each of these farms sample has been linked with the FADN database which store value of the farm capital investment (KI) and intermediate consumption (CI), using as common key the FADN code.

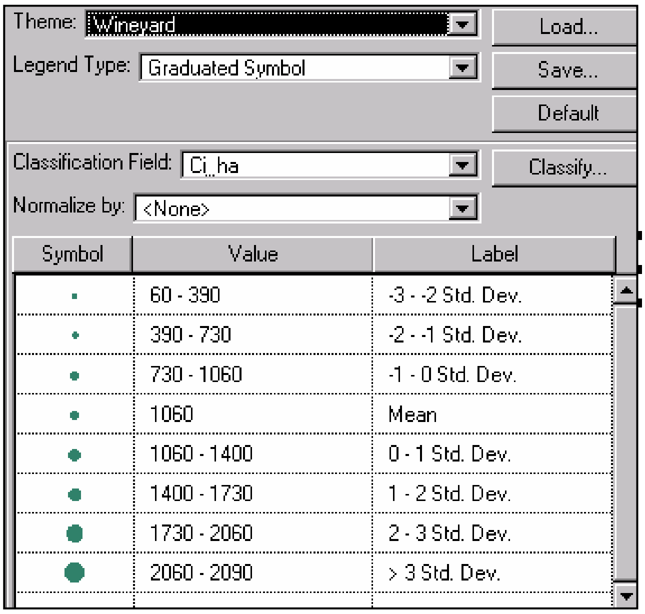
Next step has been the interpolation between the sample point to obtain a predict value for unknown location:

Trough the interpolation process, 6 different layers have been created (in GRID format with a 100 meter resolution);

KI/ha vineyard	KI/ha arable land	KI/ha Olive grove and Orchard
CI/ha vineyard	CI/ha arable land	CI/ha Olive grove and Orchard

Data statistical analysis

The interpolation process create a surface with continuos value of the variable considered. The problem is to decide in which range of this value is possible to identify low, medium and high intensity. Due to the lack on standard reference, we propose the use of the standard deviation as discriminant statistical parameter to define the range classes of the indicators. The values of the IFP (Intensity Farm Practices) indicator have been calculates combining two economic variables (KI - CI).



In the left side figure the distribution of the CI variable values of the vineyards farms are shown.

Thus, Grid's data coming from the interpolation process have been reclassified, utilising the following reference values:

-3 ; -1 Std. Dev. = 1 - low intensity
-1 ; 2 Std. Dev. = 2 - medium intensity
> 2 Std. Dev. = 3 - high intensity

The obtained reclassified layers have been summarised into three final grid layers (one for the vineyard, one for olive grove and orchard and one for arable land); in the grid summarising is necessary to compute an offset that will scale each grid so they are added together without overlapping cell value. In fact without an off-set is not more possible to retrieve the single contribution of the variables.

In attributing the code more weight was given to the CI value in comparison with the KI value.

Table 8. Possible combination of KI and CI variables and the related IFP value.

KI Grid	CI Grid	Summarising	Intensity value
1	1	11	1 - Low
1	2	12	2 - Medium
1	3	13	3 - High
2	1	21	1 - Low
2	2	22	2 - Medium
2	3	23	3 - High
3	1	31	2 - Medium
3	2	32	2 - Medium
3	3	33	3 - High

The interpolation process estimates the indicators values even when the corresponding farm typology doesn't exist in the sample (e.g.: KI e CI values of vineyards in a arable land area). To traduce these estimated values into real land use, the following procedure was developed:

1. the land use cartography (1:25.000 scale) has been converted in Grid format with the same geometric resolution (pixel of 100 meters) of the interpolated data. The affected land uses have been extracted implementing the following three grid layers:

- GRID 221 – vineyard; - GRID 223 - permanent crop (olive & orchard); - Grid 211 - arable land

2. Utilising the COMBINE[6] command in ArcInfo environment, three 3 grid layers, with the KI and CI value, and related combination, referring to the equivalent land use, have bee realised:

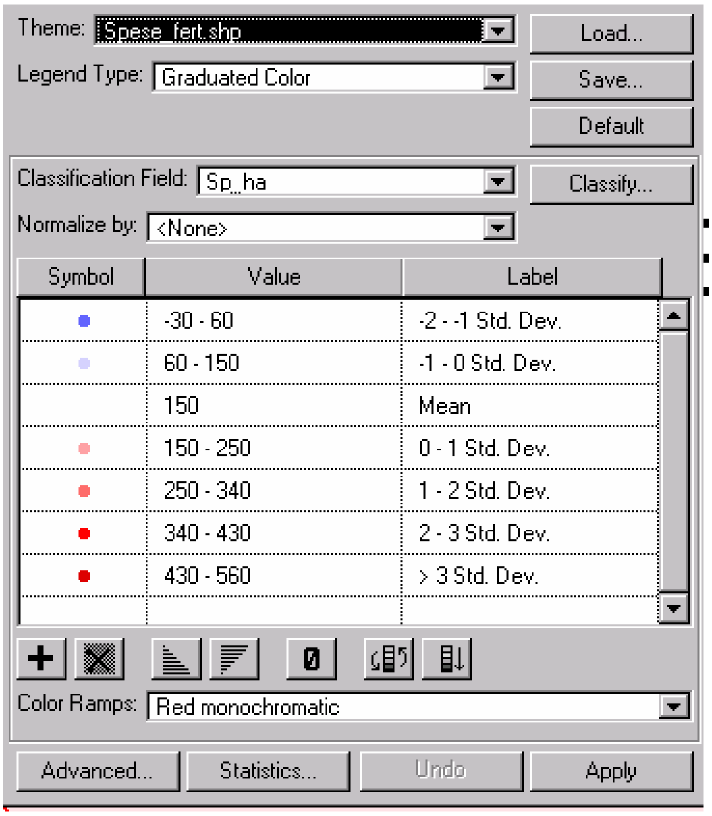
- GRID IFPV = Intensity farm practices - vineyard
- GRID IFPP = Intensity farm practices - permanent crop (olive & orchard)
- GRID IFPS = Intensity farm practices - arable land

The final step consists in the mosaic of the three grid layers, thanks to ArcInfo Merge7 command, and the realisation of a final synthesis layer.

Fertiliser Consumption (FC)

Using the FADN farms sample georeferenced, a vector point layer has been created, that store in the attribute table the fertiliser consumption per hectare. This value has been direct interpolated (using also in this case the IDW method) to obtain a prediction surface on Fertiliser Consumption at the location where the value is unknown.

Data statistical analysis



The same consideration on the IFP are valuable for this indicator. The prediction surface obtained from the interpolation has been reclassified utilising the following reference values (see figure on the left side):

- -2 ; -1 Std. Dev. = 1 - low intensity
- -1 ; 2 Std. Dev. = 2 - medium intensity
- > 2 Std. Dev. = 3 - high intensity

Also in this case the interpolation process estimates the indicators values even when the land use is not an agricultural area (e.g.: FC on forest areas).

To traduce these estimated values into real land use, the COMBINE command in ArcInfo environment has been utilised, using the FC surface and the Land use grid, so each cell of the combined grid contains the value of FC and Land use, this allow, trough spatial query, to have information about the FC of different land use.

[6] Combines multiple grids on a cell-by-cell basis within the analysis window, such that a unique output value is assigned to each unique combination of input values (ESRI on line Help for ArcInfo 8.1.2 Workstation)

[7] Merges multiple input grids into a single grid based upon order of input. (ESRI on line Help for ArcInfo 8.1.2 Workstation)

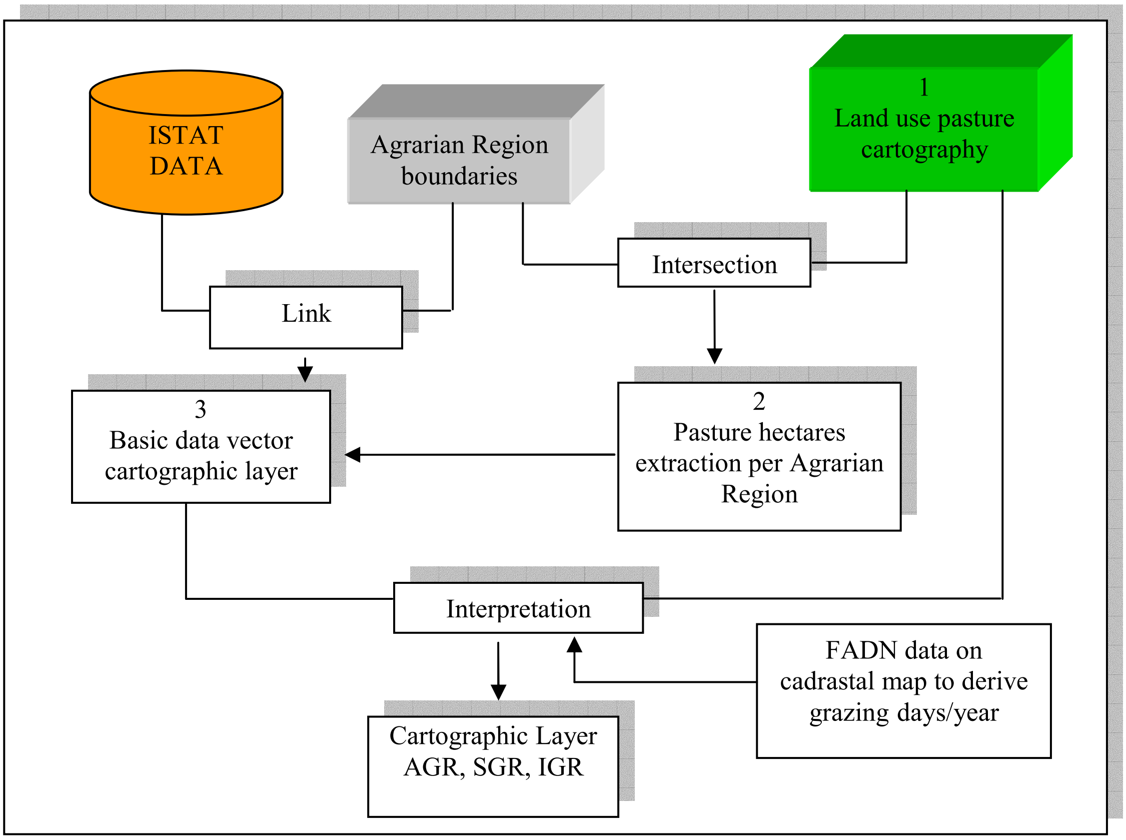
Intensity Grazing Rate Indicator

This indicator - combined with slope and soil depth - is very important to assess the soil erosion risk.

The indicator take into account the number of bovine and ovine/caprine per grazing surface unit and the duration of grazing day, and is derived from the comparison of the following value:

- Actual Grazing Rate (AGR) - define the actual livestock rate - in LU[8] - on the grazing surface
- Sustainable Grazing Rate (SGR) - define the sustainable livestock rate - in LU - on the grazing surface

The Intensity Grazing Rate (IGR) calculation has been carried out through the following steps:



1. extraction from land use cartography of the pasture polygons;
2. intersection between the previous layer and the Agrarian Region[9] boundaries to extract the pasture hectares per Agrarian Region (this administrative/territorial aggregation takes more into account the animal movement from the residential municipality – from where the total number of animals is extracted - to other pastures of the neighbouring municipalities);
3. link between the ISTAT General Census data with the previous cartographic layer to derive the "basic data" cartographic layer, with the livestock number per area in vector format (geographically referenced with the project standard) and related attributes data base containing all the needed parameters for indicators and indexes calculation (AGR, SGR, IGR).

[8] Livestock Unit (LU) - calculate with a conversion index : 1 bovine of advanced age to 2 years = 1 LU, 1 ovine/caprine of advanced age to 1 years 0,15 UBA.

[9]Group of contiguous Municipality with homogeneous natural (geology, climate, orography, etc) and agricultural condition, defined by ISTAT.

4. Actual Grazing Rate (AGR) calculation:
   - calculation of the Livestock Unit (LU): total number of LU per Agrarian Region, according to the total number of cattle, sheep and goats (= Total Actual Grazing Rate)

     Total LU (AGR) = n° bovine + n° buffalo + n° (ovine/caprine) x 0.15

   - calculation AGR per hectar

$$\mathrm{AGR}_{\mathrm{ha}} = \frac{AGR}{Hectares}$$

5. Sustainable Grazing Rate (SGR) calculation:
   - Fodder Units (UF) calculation (average yield from pasture surfaces):

     U.F. = 841*Hectares*0.80

     841 = average U.F from 1 pasture hectare
     Hectares = Pasture surface from Agricultural General Census
     0.80 = pasture utilisation coefficient
   - Feed cattle need (FCN) calculation

     FCN = (5xN)*(LU_bovine)

     5 = average daily feed need of 1 LU in U.F.
     N = n° of gazing days
   - Feed sheep and goat need (FSN) calculation

     FSN = (5xN)*(LU_ovine/caprine)

   5 = average daily feed need of 1 LU in U.F.
   N = n° of gazing days

FADN data are crucial in the calculation of the value of N. In fact, normally constant value are applied, as the following:

To calculate the real N per pasture area, the following methodology has been developed:

- livestock farms extraction from FADN sample;
- list of the livestock FADN sample cadastral parcels;
- link [10] between cadastral parcel and the FADN's crops and products tables, where it's possible to find information about forages surfaces and yields, and for feed farm, from which it is possible to derive the feed needs, in terms of grazing days per geographic area and per farm livestock typology;
- grazing days per farm typology calculation;
- grazing days values interpolation, to produce a GRID with the variable (grazing days) values of the whole area.

The interpolation process estimates the indicators values even in non pasture land uses (e.g.: Grazing day on arable land). To apply these estimated values only to pasture areas, the COMBINE (grazing day per hectare and the land use grid are merged) command has been utilised; each cell of the combined grid contains the value of Grazing day and only pastures land use polygons.

6. The value of SGR is calculate with the following formula:

$$Total\_LU(SGR) = \frac{U.F.}{(FCN + FSN)}$$

[10] The link between cadastral parcel is allowed thanks a common key composed of: FADN farm code + province code + municipality code + cadastral sheet number + cadastral parcel number

Calculation SGR per hectar

$$SGR_{ha} = \frac{SGR}{Hectares}$$

7. Intensity Grazing Rate (IGR) calculation
   The following threshold has been defined to evaluate the impact on soil erosion risk

Impact	Grazing Rate	IGR
Vulnerable	AGR ≤ SGR	1
Sensible	AGR > SGR	2

Results and conclusion

The work carried out in these three year has allowed to examine several methods of FADN farm georeferencing. The three different methodologies has been tested on the following farm samples.

- 273 FADN farms with the farm centre georeferenced (point geometry);
- 60 FADN farms with 2287 cadastral parcels (resulting from the total number of parcels of each farm, with an average number of parcel per farm of 38,12) georeferenced (polygon geometry);
- 10 FADN farms with 128 way points (resulting from the total number of field plot centres of each farm, with an average number of field per farm of 12,8) georeferenced (point geometry).

The first methods allows to work with a huge number of farm with very low costs (total costs: per farm = 62,27 €), but it's not performing with all those indicators related to the impact of agri-techniques on the land conservation. The second could respond to many data needs related to agri-environmental indicators implementation. Having the parcels georeferenced allows to a simpler integration with other data sources. But it's not always simple, and often very expensive (total costs per farm = 633,33 €; per parcel 16,62) due to the revenue taxes, to collect digital data on cadastre cartography (absolutely necessary). The third is the more expensive (total costs per farm = 800 €; per field/way point 75 €), but with the distribution of GPS to the FADN surveyors could be realised during the FADN surveys normally carried out. The problem is that it doesn't give any idea of the parcel surfaces and fields geometry.

A method on the pesticides use calculation and geographic distribution allocation is now available, together with the following data, both cartographic and statistic:

- plant diseases distribution (on the basis of FADN surveys);
- pesticides quantities calculated at NUTS III and river basin level and for each cartographic polygon, based on suggested quantity, and on FADN data (used on farm), with the related differences;
- definition of correction factor between agricultural best practices quantities and pesticides real use at farm level (main problem for indirect statistics)

Table 7. Used quantities (Quintals) of pesticides - Year 2001- NUTS III (Provincia di Chieti)

Pesticide typologies	**Recommended/optimal quantities**	**effective quantities used by FADN farms**
Insecticide	276.26	437.66
Fungicide	3.262.65	4.872.39
Herbicide	229.48	230.17

A model, with the related procedures, based on GIS technologies, for the utilisation of FADN georeferenced data to calculate agri-environmental indicators has been realised.

The analysis of the indicators that, in the framework of DPSIR model, can be calculated using FADN data, has produced a matrix with the possible combination between FADN variables/parameters and indicators. The analysis of the FADN data bank level (European, National, Regional) needed to calculate the selected indicators, allows to affirm that the Mini Regional data bank is the minimum level required.

A specific geographic unit level (farm centre, GPS way points, cadastral parcels) requirements per agri-environmental indicator (Intensification/extensification, Pesticide consumption, area under agri-environment support, etc.), and per FADN data typology (total capital invested, general expenses, pesticides expenses, agri-environmental subsides, etc.) has been fixed.

Those results allows to establish the relationships between some of the most important indicators and FADN data, related to the georeferencing requirement, and furnish some key element for the application of the model in other situation.

The analysis of different interpolating techniques to distribute FADN sample data (mini regional data bank) at geographic level demonstrate that the IDW (the Inverse Distance Weighted) produces good results, especially having only the farm centre georeferenced (first method).

For the application of the methodologies, particularly attention has been paid on the development of specific indicators to analyse the Intensification/extensification driving force, fertiliser consumption and soil erosion risk through the Intensity Grazing Rate (IGR).

The Intensity Farm Practices (IFP) and Fertiliser consumption (FC) indicators have been calculated, and indicators are spatialised at land use parcel level. This allows:

- to link/intersect several physical layers (land use/cover, DEM, Soil map, etc.) to this (and any other) agri-environmental indicator and/or with other FADN data typologies;
- a posteriors stratification, with representative sample based on territorial aggregates functional to agri-environmental problems.

The soil erosion risk indicator has been calculated utilising the Intensity Grazing Rate (IGR) combined with slope and soil depth. FADN data have been used to calculate the grazing pressure on the soils, and take into account the number of bovine and ovine/caprine per grazing surface unit and the duration of grazing day, and is derived from the Actual Grazing Rate (AGR) and the Sustainable Grazing Rate (SGR). An IGR threshold has defined the soil erosion vulnerability/sensibility.

In conclusion, this work demonstrate that FADN data can be georeferenced with limited costs, and with a good results in terms of their utilisation for agri-environmental analysis and indicators. More efforts should be spent on to the experimentation of methods and procedure for establishing quantitative relationship between farm and related land. In this sense, all the thematic specialist interested in the agri-environmental analysis should produce a cultural effort for a real interdisciplinary approach to data integration.

References

Ersts, Peter J. (2001). Summarise information between grid. *ArcUser Magazine* Vol. 4 N. 4.

Fais, A. and Nino, P. (2004) Mapping the Spatial Distribution of Plant Diseases. 24th Annual ESRI International User Conference Proceedings: pap1820. San Diego, California USA.

Fais, A. and Nino, P. (2003). Field Water Requirements for Regional Water and Land Use Planning in Italy. 6th Inter-Regional EnviWater Conference Proceedings. Albacete Spain.

Fais, A. (2003). FADN data to produce indirect statistics on the use of pesticides and to define the spatial distribution of plant diseases. PACIOLI 10 Proceeding: 38-50. LEI Wegeningen.

Fais, A. and Nino, P. (2001). Indirect statistics on water requirements in Southern Italian region irrigated areas - A method based on integrated use of different data sources. ETK-NTTS 2001 Proceedings. Hersonissos (Crete) Greece.

Fais, A. (1997). Sulla possibile integrazione di diverse fonti di dati nelle indagini congiunturali in agricoltura. Commissione ISTAT per la riforma delle statistiche agrarie. Roma.

Liu, J. (2001). Integrating ecology with human demography, behaviour, and socioeconomics: Needs and approaches. *Ecological Modelling* 140:1-8.

Petersen, J.E. (2004). Agricultural statistics and environmental analysis - an evaluation of future needs from a user perspective. Executive summary. European Environment Agency.