Inputs for AgrIS - Checking the consistency of the agricultural sector database

J.A. Boone
C.J.A.M. de Bont
K.J. Poppe (eds.)

Project code 63538

July 2002

Report 8.02.04

Agricultural Economics Research Institute (LEI), The Hague
(In close co-operation with INEA, INRA, MTT)

The Agricultural Economics Research Institute (LEI) is active in a wide array of research which can be classified into various domains. This report reflects research within the following domain:

- ☐ Statutory and service tasks
- ☐ Business development and competitive position
- ☐ Natural resources and the environment
- ☐ Land and economics
- ☐ Chains
- ☐ Policy
- ☐ Institutions, people and perceptions
- ☑ Models and data

Inputs for AgrIS - Checking the consistency of the agricultural sector database
Boone, J.A., C.J.A.M. de Bont and K.J. Poppe (eds.)
The Hague, Agricultural Economics Research Institute (LEI), 2002
Report 8.02.04; ISBN 90-5242-744-5; Price € 19.00 (including 6% VAT)
115 p., fig., tab.

This report presents a consistent set of data for Eurostat on inputs (intermediate consumption) related to agricultural activities and proposals to produce this also in future. The set of these data is an important element in the development of AgrIS, a database of Eurostat for statistics, modelling and research. Appended to the report a preliminary database on inputs related to activities is presented to Eurostat. This database can be improved and enlarged for implementation in future. The suggestions in the report are based on analysis of the existing data sources on EU level as well as on the information and views of experts and interested persons of all 15 Member States in the EU and within the European Commission (Eurostat and DG-AGRI). The project has been financed by Eurostat, and carried out by four institutes in co-operation with a number of experts.

Orders:
Phone: 31.70.3358330
Fax: 31.70.3615624
E-mail: publicatie@lei.wag-ur.nl

Information:
Phone: 31.70.3358330
Fax: 31.70.3615624
E-mail: informatie@lei.wag-ur.nl

Content

Project team

LEI	Koen Boone	INEA	Antonella De Cicco
	Kees de Bont		Carla Abitabile
	Frans Godeschalk		Franco Mari
	Boudewijn Koole	INRA	Ahmed Barkaoui
	Krijn Poppe		Jean-Pierre Butault
	Geerte Cotteleer	MTT (former MTTL)	Heikki Lehtonen

Preface

Since more than thirty years the European Union collects economic data on agriculture. For a part these data are based on the information of individual farms in the Member States (mainly in the context of the Farm Accountancy Data Network, FADN or RICA, of DG-AGRI). For another part the data have a macro-economic nature; the Economic Account for Agriculture, EAA, of Eurostat is an important example of this.

These data were used on EU and national level for many purposes, for instance development and evaluation of farm policy, economic research and information. For Eurostat as well as other users of these data on EU level and in the Member States, it is important to have available a common and consistent database. Eurostat is developing this database (AgrIS). AgrIS could help to get a clear view on the consistency between all data sources.

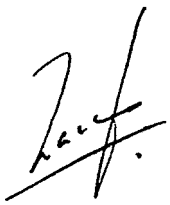
In the context of this aim, LEI is asked to develop a database on the relation between intermediate costs (inputs) and the returns of specific agricultural activities.

Given the different nature of the mentioned and other important European data sources (as Standard Gross Margins, Prices and Farm Structure Survey), it was important to examine the definitions of them in detail. Besides that it was important to have a view on the use of data per Member State; not only on the data delivered to Eurostat and DG-AGRI, but also on the availability of other useful data in the context of the project. To reach a balanced result around 50 experts involved on FADN and EAA in the 15 Member States were interviewed by the project team.

The study has been carried out by four institutes from different Member States. They also interviewed the national experts. The interviewed experts supplied very useful information.

We are grateful to all persons for their kind co-operation. We thank the staff of Eurostat, and especially Marcel Ernens and David Verhoog, for their guidance. Mrs. Brigitte van Oord has to be thanked for her secretarial support.

The managing director,



Prof. Dr. L.C. Zachariasse

Summary

Aim of the project

Eurostat invited LEI to develop - in co-operation with other institutes in the EU - a database on inputs related to the returns of specific agricultural activities for the current 15 Member States of the EU. The database is essential in the process of the development of a common agricultural database (*AgrIS*) by Eurostat. While Eurostat itself carries out the work to develop a main part of the database, as for instance on returns and output of activities, this report deals with questions on costs of production (intermediate consumption) per activity (production of crops and animal production).

Procedure

In the context of the project it was important to start with an examination of desires of Eurostat and DG AGRI as well of the existing data sources at EU level. Main sources are FADN (Farm Accountancy Data Network), SGM (Standard Gross Margins) and EAA (Economic Account for Agriculture). These data sources are developed during the last thirty years by DG-AGRI and Eurostat in co-operation with representatives of the Member States. The examination of the definitions used resulted in an overview of the differences between them. During a discussion at the start (workshop, February 2001) of the project, a number of criteria were fixed to formulate a provisional list on the agricultural products or activities and inputs to be taken into account in the database (EAA+, appendix 2). For a large part this list is equal to the list used for the EAA (EAA being the minimum level of detail on inputs). Some extensions and specifications are made in the context of the desirability of data, for instance related to environmental policy, the development of the production (especially in horticulture) and the characteristics of agriculture in some Member States (Mediterranean products).

It was decided that inputs allocated to activities (e.g. having dairy cows) are more important for users, and probably easier to obtain than when allocated to output (e.g. milk, beef).

FADN/INRA model

Parallel to this examination the project team could made use of the study and results of a project of INRA to develop a model (in this report identified as the FADN/INRA model) on the relation between aggregated costs of production and returns. The (initial) model was produced on request of DG-AGRI. This model resulted in a (draft) comparison with the data in EAA, as they are produced by the Member States. Because the results and the definitions of EAA themselves were not subject of examination, it was necessary to find out what are the reasons for differences (deviations) in the comparisons (where FADN results

do not match with EAA) and what kind of solutions could be used. This was done for a part in contacts (interviews) with the experts in Member States and for a part by discussions in the frame of the project team, in which the INRA expertise was extremely important.

To use the FADN/INRA model for the purpose of the project it was necessary to make some adjustments in (the definitions of) the model. To start with this it was important to analyse the reasons of the less satisfactory results of the FADN/INRA model, as described in chapter 2, compared to EAA data. It was concluded that some main reasons for this are:

- scope of EAA and FADN; EAA for all farms, FADN represents, depending on the threshold (minimum size in ESU), less farms;
- representation: depending on the structure of the sector in a member country the FADN sample is not representing adequately specific types of production (specific crops or categories of animals); besides that in some countries forestry is a part of the farm, for which costs, as machinery costs, can not be separated easily;
- definitions and differences in interpretation of definitions between EAA and FADN (inputs); as is shown in chapter 2 the main differences in definitions are on veterinary costs and other goods and services (insurance and other direct costs animals). Smaller problems exist for internal use of (mainly) feed (grass and silage) and seed, as well as on contract work and maintenance of materials. As far as possible corrections have been made to harmonise these items;
- definitions and differences in interpretation of definitions between EAA and FADN (categories of output); the differences in definitons are mentioned in chapter 2; they mainly exist on inseparable non agricultural activities and processing of agricultural products, other crops, other animal products and other animals. Smaller problems exist for vegetables, flowers and plants and fruits;
- time period: EAA by definition per calendar year; FADN in some member countries not on calendar year, sometimes different per type of farm;
- contract farming; when some inputs, as feed and piglets, are not paid by the farmer it results in a diverging distribution of costs; at the same time the farmer receives a reward for his labour and investments. Because of these characteristics these farms are excluded in the analysis.

We expect that the differences in definitions and especially the differences in interpretation of definitions are the main causes for the differences between EAA and FADN. These adjustments gave an important improvement in the results, but because still deviations came up, the project team had to decide it was necessary to make use also of other data sources: SGM and national.

Detailed SGM

Detailed information on SGM - only on 1996 - came available during the project for 9 Member States. In fact this information is in principle very useful because it provides information on the value and costs of production per activity. However the SGM are calculated in most case per region (only Belgium, Luxembourg and the Netherlands have

unique SGM for the whole country) and are not based on one calendar year, but mostly on an average of 3 years (as is requested by Eurostat). For these reasons the project team made some adjustments (see chapter 2) on the available SGM data to use them in the context of the database. It is clear that for Member States without detailed SGM there is still a lack of information to be filled. Besides this, it would be very helpful if all member countries deliver from their calculations the detailed SGM for each year in addition to the present SGM for a three year period (the present SGM maintain their function for the typology of farms).

Contributions of Member States

This examination made also clear that it was necessary to find out in detail what kind of experience exists in the Member States on the use of the different data sources and if additional data could be provided. Experts in each Member State delivered information on this, in reaction on a common questionnaire and during interviews. The result of this process is a picture ('state of the art') of the data on inputs in the EU-15 (see chapter 3, also figure 3.1). It makes clear that the situation per Member State is quite different; for instance SGM are in some countries produced mainly on the base of FADN, while in other countries they are calculated and fixed by specialised institutes using other information sources. A conclusion is that in a number of Member Countries attractive additional information and data are available, that can be used to improve and develop the database. As far as possible this information is imputed during the project in the database. On the other hand some countries do not have additional information in this context and or they are, at least currently, not in the position to deliver this on a regular base.

Conclusions and recommendations

The project results in an (initial) database on inputs allocated to agricultural activities for all 15 Member Countries. So far the database is developed only for one year (1997). For the production of this database use is made of the (adjusted) FADN/INRA model, detailed SGM and additional national data. Given the lacks in information and the differences in definitions it was not possible to arrive at a database in which all desired specifications of activities and inputs (as mentioned in EAA+ lists) are included. Chapter 2 provides a description of the process followed to develop the desired database, including the encountered problems, in first instance. This is followed by some remarks and additions on the use of national data in chapter 4. This Chapter also concludes that this database is a starting point for further consultations with the Member States.

It is clear that this (initial) database is not fully corresponding to what is aimed by Eurostat. The procedures followed during the project make clear that a lot of problems still have to be overcome. Many solutions imply additional contributions of Member Countries.

The project provides a number of recommendations for this, for instance on making available more information to Eurostat or DG-AGRI (see chapter 5). To stimulate this, it is important that Eurostat and DG-AGRI make some extra efforts, for instance on the use of (detailed) SGM in the context of AgrIS and to check the (aggregated) FADN data in relation to EAA results. This implies a more intensive co-operation with the Member States (in task force meetings), for instance to analyse the reasons for differences in results and how

to cope with these issues (definitions and interpretation of definitions of outputs and inputs in databases, accounting and calendar year, non-represented farms). On some issues a more intensive co-operation between working parties of Eurostat and DG-AGRI is needed, for instance on the harmonisation of the definitions used in the different agricultural statistics. Related to this, the specification of variable costs (especially other variable costs) in SGM could be better matched with EAA definitions.

This means that Eurostat and DG-AGRI are advised to concentrate their efforts in the coming years on matching and making consistent the (links between) definitions and variables (on output as well as on inputs) used in EAA, SGM and FADN. This leads automatically to the need of more co-operation - in each Member State - between the institutes responsible for the data used in the EU. The co-operation in the Member States might reveal the differences in the interpretation of definitions and makes it possible to correct them.

For some problems indicated in the project it is advisable to make on request of Eurostat and or DG-AGRI more detailed studies in the coming years, for instance on:

- the calculation of results FADN/INRA model 1990-1999;
- the calculation and comparison of (detailed) SGM- results for all Member States;
- the organisation of task force discussions (or seminars) on the results of the two previous activities and make transparent the use of EAA, FADN and SGM by publications (dealing with the consistency of data);
- the use of cost price calculations in member countries in the context of the development of AgrIS (What are the definitions and data used for these calculations);
- the differences in interpretation of definitions between the databases;
- the use of next generation models (for example maximum entropy) to combine data of different sources;
- the application of a more fine-tuned, tailored list of outputs and activities and inputs, linked to the desired level of information and the capacity of member countries to deliver the information on a regular base;
- the application of the approach followed in this project in and with the Candidate Countries; in fact the results here have to be matched with the possibilities of the future Member States, which are subject of discussion in another project on request of Eurostat.

1. Introduction

1.1 Problem description

The Statistical Office of the European Union (Eurostat) has asked a consortium of institutes (LEI, INEA, INRA and MTTL) to construct a consistent database on agricultural inputs in such a way that these data are linked to agricultural activities. Eurostat collects a lot of statistical data of the agricultural sector, mainly on the value, prices and volume of products, on the number of farms, animals, areas with different crops (land use), as well as on the income of the sector, trade in agricultural products etc. These statistics are used on Community and on national level for several purposes, for example for information, research and economic analysis. The statistics are also essential for the preparation and the evaluation of the Common Agricultural Policy. Models can provide decision relevant information for DG-AGRI of the European Commission on the basis of these statistics. For all these purposes it is important that the data are consistent.

The statistical data are coming directly from the Member States. Within Eurostat different units are responsible for the harmonised collection and presentation of these data. In the past a number of these statistics were brought into one database (SPEL), together with results from calculations with models in order to fill gaps in the data or to smooth inconsistencies. This practice has been abandoned, as it lacks transparency and blurs responsibilities between statisticians in Member States, statisticians in Eurostat and researchers that develop models for Eurostat or DG-AGRI. In stead it has been decided to develop a database called AgrIS to bring a major part of data together (Verhoog, 2000). In this approach it is all the more necessary to have consistent data. In the concept of AgrIS data are not calculated or smoothed in the database, but facts are stored to spot inconsistencies and improve them. The data have to be based on a set of procedures (or mechanisms) to manage the quality of the data.

The central problem for Eurostat is that there are no specific data on inputs (intermediate consumption) linked to farm activities. EU-15's Economic Account for Agriculture provides only aggregated data on this related to the total value of production per Member State.

1.2 Aim of study

The main aim of the project was to produce a consistent set of data for Eurostat on inputs (intermediate consumption) related to agricultural activities. By this it will be possible to eliminate the number of internal calculations and calculated variables produced by Eurostat.

In addition to the database itself, the methodology (including a 'protocol' to ensure continuity for the future) for the data set is one of the deliverables of the project. In addi-

tion reports per category of inputs and per country on the feasibility of this methodology should become available.

The base for all calculations is formed by the EAA. One important issue in this is to formulate the level of detail of the inputs (and of the agricultural activities) for the database. In this project the problem concerns only the data of the actual Member States (EU-15). A separate project is carried out on Candidate Countries.

1.3 Data sources

Current data sources available at the European level are the point of departure for the construction and improvement of AgrIS regarding inputs per activity. Main sources of data available at European level are:

a. Economic Accounts for Agriculture (EAA)

The EAA per Member Country per year provide essential data on the value of agricultural products and of inputs (intermediate consumption). These data are aggregated values (in millions of Euro), and without providing information on the link between the value of specific agricultural products and the related costs of production.

The EAA based on the recently reviewed methodology itself was not subject for a review during the project, although some attention has been paid (also on request of Eurostat) on the level of detail in the EAA concerning Mediterranean and ornamental products (flowers, pot plants). Indirectly, because of the consistency with other data, the result of the project can help to improve the results of the EAA in future. In this project the relation between the different agricultural activities and the inputs used for them was a central issue.

b. Farm Accountancy Data Network

FADN (or in French RICA) is database of DG-AGRI with financial and technical data of 60,000 farms in the European Union. The basis of the results is formed by the Balance Sheet and Profit and Loss Account of the farms. Because all costs and outputs of the farms are included in this database, it could be useful for allocating inputs to outputs.

c. Standard Gross Margins (SGM)

Eurostat collects every two year data of the standard gross margins per product (crop or animal). The SGM are used for the calculation of the economic size of the farms and the typology of farms. The SGM are calculated for all products that are included in the Farm Survey. In larger countries SGM are calculated for several regions. Officially SGM are based on an average of three years [1]. Until recently Eurostat only collected the total (the margin) and not the underlying data (output, variable costs etc). For 1996 however Eurostat also gathered information on specific inputs as underlying data up of the calculations. The SGM might be a good source, since the SGM are also available per agricultural activity.

[1] Some countries use 5 years instead of 3 years or make some corrections to the calculations to smooth the differences over the years.

d. Prices

Eurostat collects prices of inputs (and products) on a monthly basis. It was not foreseen that the project would concentrate itself on the completeness of the data and the consistency between these prices for the sector account and other purposes. However, it was of interest to analyse prices on inputs as a background for the development of the costs of production. The main work on outputs and related to that on prices of products however is (already) done in a separate activity of Eurostat that checks the consistency of outputs in the databases ZPA, PRAG, FAO and EAA (see figure 1.1).

e. Farm structure

Data on the number of farms, animals, and the acreage of different crops etc from the Farm Structure Survey (FSS) are relevant to get information on the development of production of the agricultural sector and might be of interest as a declaration for the development of inputs (costs of production). Moreover farm structure data could be useful to calculate (weighted) national SGM on the basis of regional data.

The data on prices and farm structure (d. and e.) are used less intensively in this project. The main information used is based on EAA, FADN and SGM. The data on inputs in the EAA are compared with the data of DG-AGRI's FADN (Farm Accountancy Data Network) and the Standard Gross Margins (SGM).

Besides the data sources available on European level, the availability of data sources in the Member States is important. In the concept of AgrIS, the Member States should feel themselves responsible for the quality of the data and the inconsistencies. Differences between national data sets and comparable data available at European level should be transparent and explained in meta-information.

The results of this project are a starting point for discussion with Member States. They can agree with the results or deliver, at a certain time, better data.

1.4 Method of work and procedure followed

A team of researchers from four different institutes in the European Union carried out the project. The project was split in four stages:

1. inception: identification of needs;
2. data analysis;
3. country-specific research;
4. finalising.

Figures 1.1, 1.2 and 1.3 give an impression of the work method. At the moment Eurostat itself is working on the consistency of output data, by comparing data from the EAA (COSA), Price statistics (PRAG) and farm structure (ZPA1). This is carried out with the help of a consulting procedure with the Member States as shown in figure 1.1: Eurostat fills the database AgrIS, writes a discussion note, asks Member States for clarifications (meta data) and improved data, and then updates AgrIS.

For input data per activity, data can be added in AgrIS from the SGM data set, from the FADN and from Member States (figure 1.2). This can be done in the project on past years, as well as after the project. The consulting procedure with the Member States, as currently available on outputs, can then also be used for inputs and inputs per activity.

The aim of this project can be reformulated as testing the feasibility of this protocol by installing it. Therefor this figure 1.2 has been the base of the activities for the project team (figure 1.3). These activities are described below more in detail grouped per stage of the project.

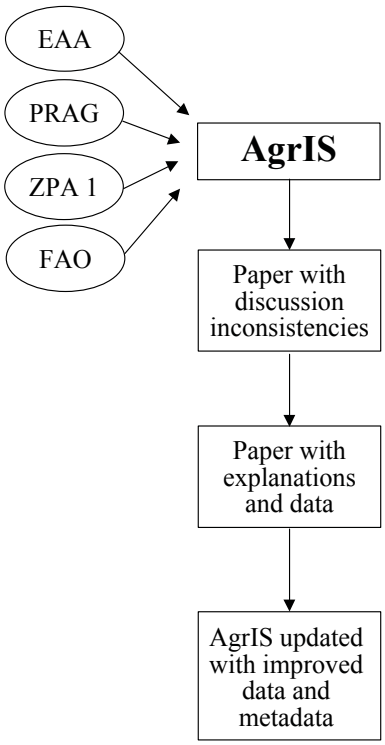


Figure 1.1 Information flow on data for AgrIS concerning output per activity, current situation

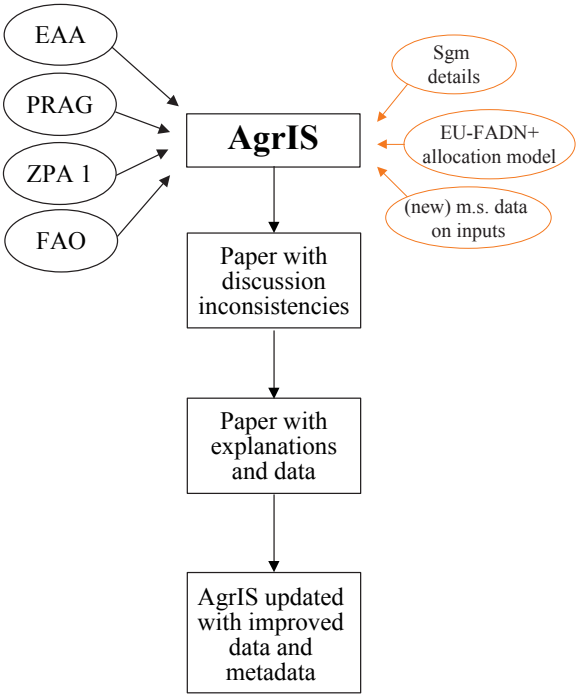


Figure 1.2 Information flow on data for AgrIS concerning input per activity, after the project

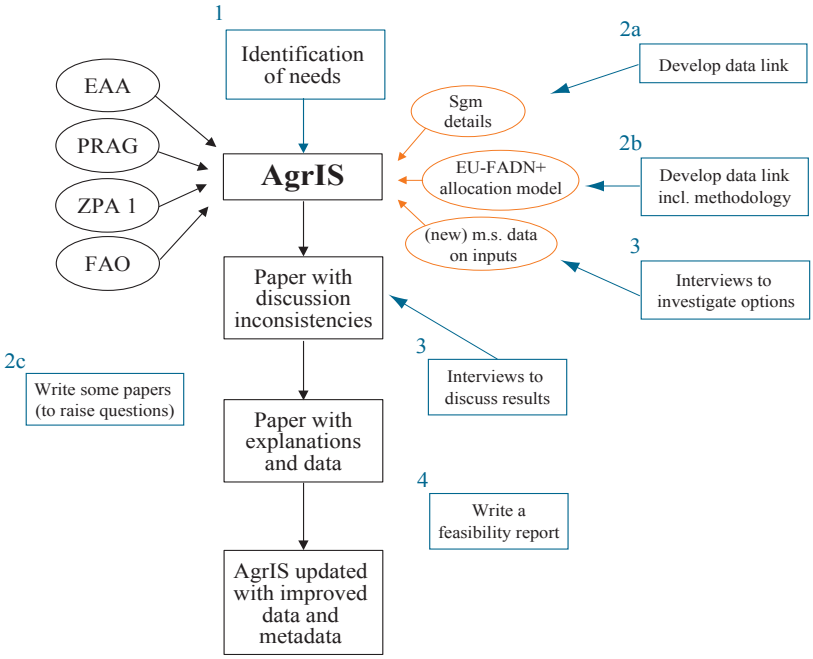


Figure 1.3 Main activities to support the set-up of the information flow on inputs

Part 1. Inception: Identification of needs

To identify the needs and current problems related to (the lack of) data on inputs in the database of Eurostat (AgrIS), the project team started the work with a workshop with database managers, modelling experts and others in collaboration with Eurostat. During the workshop criteria were formulated to elaborate a list of agricultural activities and inputs, which could be used during the project to elaborate the database. Based on the criteria a (provisional) list (EAA+) was fixed. The main results are reported in the review in chapter 2.

Part 2. Data analysis

The main outcome of part 1 of the project, the (provisional) list of activities and inputs, has been used to analyse the problems in more detail. This is done per category of products and activities as well of inputs e.g. by comparing data sets and the relevant meta data (including the handbooks and guidelines) from Eurostat, as well as making comparisons with FADN and SGM. As in the FADN costs are not allocated to outputs, a regression model was used to allocate costs to farm products. This regression model was developed by the INRA for France and has been extensively used there. Just before the start of the project, a first experiment was made to extent the model to European level based on EU-FADN (Butault et al., 2001 and Pollet et al., 2001).

This delivered descriptions of each identified problem, as well as some possible explanations. This information was used to formulate questions to the Member State(s) involved.

During the project it became clear that a lot of problems are related to the definitions used in the different data sources (EAA, FADN and SGM). Although Eurostat was helpful in collecting SGM data from Member States, for not all the Member States the detailed data were available (only for 9). Therefor the work carried out had to focus on describing the methodological pittfals in connecting the data sets and in trying to estimate coefficients, for all countries, on the relation between the value of activities (outputs) and inputs concerned.

The work in this part of the project was split in:

A. develop data link SGM-EAA in AgrIS. This includes issues like differences in definition and data availability;
B. develop data link FADN-EAA in AgrIS. This work was carried out in co-operation with the EU FADN (DG-AGRI);
C. write paper that discusses inconsistencies between SGM, FADN and EAA;
D. improve the INRA model so that it can be used on EU scale and that it can be used to allocate inputs to outputs in EAA;
E. compare the totals of the INRA model (FADN) and SGM, with the EAA totals;
F. write questionnaire for interviews in Member States discussing, among others, the differences between EAA on the one side and INRA model on the other side.

The results of this analysis are mainly reported in chapter 2.

Part 3. Country-specific research

The AgrIS concept asks for an involvement of the Member States in solving inconsistencies and in taking responsibility for the data quality at EU level. That alone already implies an involvement of the Member States in this project. It also became clear that solving the problems on inputs related to activities, which came up in part 2, requires direct additional information of the countries concerned. So members of the working party on the EAA and experts on cost of production data (most of the time the FADN representatives) were invited to co-operate in the project. They were interviewed by members of the project team. They were asked to check the quality of the INRA model and to explain why EAA data might deviate from INRA model and SGM results. Besides they were asked to provide additional data and to explain why some data are not available. In addition they were asked to answer questions on how they implement EAA guidelines.

The national experts (see list in appendix 14) were interviewed in their own country on the base of a questionnaire (see appendix 12), including general as well as a country specific questions. Before the start of the interviews, both the working party of Eurostat on the Agricultural Economic Accounts and the management committee of the FADN (with the representatives of all Member States) in Brussels were informed on the project and the need for information. In addition extra information was gathered directly by e-mail and/or by telephone and by using written reports. The results of this stage are reported in chapter 3.

Part 4. Finalising

The analyses (in part 2) and the gathered information with the analyses per country (part 3) were brought together to make a picture of the whole situation ('state of the art') on the data on inputs in the EU-15.

Based on that state of the art and the reports per country a first database was made. The starting point of this database are the results of the INRA/FADN model. If however better quality SGM data or national data were available, these data were used.

A list of conclusions and recommendations for the maintenance and improvement of the database were made. Chapters 4 and 5 report on this. The (initial and provisional) database has been handed over by the project team to Eurostat.

To conclude

The main points and steps in the project are:

1. defining the needs of Eurostat as regards data on inputs in relation to agricultural activities and data in the Economic Account for Agriculture;
2. analysing the available data sources at EU level in relation to what is desired on data on inputs for AgrIS;
3. the use and development of the INRA/FADN model as an method to construct the model;
4. evaluation of the results of this model;
5. analysing the use of other sources (SGM, national data) in this respect;

6. review of the position of individual Member States to the development of the desired database (interviews with selected experts);
7. formulating conclusions and recommendations for further development of the database.

2. Review and analysis

2.1 Need for details on activities and inputs

Improving the data on inputs in the AgrIS database and linking them to activities leads to discussions that can only be solved by looking to the need for such information and the level of detail needed. This project has taken as a starting point the general impression that agricultural, rural and environmental policies are supported by policy analyses that provide a true and fair view of the economics of farm level activities. This view was underlined during a workshop with stakeholders from Eurostat and DG-AGRI organised at the start of the project, in February 2001, by the project team in Luxembourg. Such a picture includes costs or inputs, revenues or outputs as well as the link between them (see also Abitabile et al., 1999 for a similar analysis on the FADN).

In this section this general point of departure is worked out into two topics: the concept of 'activity' and the level of detail of inputs needed.

Outputs and activities

The EAA manual (Economic Accounts for Agriculture Eurostat edition 2000, rev. 1.1, appendix V) prescribes the different outputs and the elements of intermediate consumption (inputs) of the production account. This list has to be seen as the minimum level of detail for the inputs. In a workshop with stakeholders and Eurostat it was concluded that some extension is attractive. This is true for a more detailed level of inputs (see below), but it might also be attractive to discuss some new output items. For historical reasons specified data on some ornamental plants (e.g. flowers, flower bulbs, pot plants) as well as on some Mediterranean products (vegetables, fruit (stone fruits, citrus fruits and nuts fruits), wine and olive oil, tobacco and rice) are lacking [1].

To facilitate the allocation of costs to activities, a distinction between activities and outputs is necessary: outputs and activities are different concepts. Outputs are the products of an agricultural activity. One of the problems is that one activity can produce several outputs and that one output can be produced by several activities (figure 2.1 shows this in the form of a data model). Therefore it seems to be less complicated to allocate inputs to

[1] As regards horticultural products, thanks to the fast development of both technology and agronomy, several of them are now successfully grown and marketed by North European countries (e.g. tomatoes, cabbages, cucumbers). For instance looking at EU-FADN figures for type-2 specialist farms (cf. http://europa.eu.int/comm/agriculture/rica/index_en.cfm) the outstanding results of the Dutch and Danish farms catch the eye. Thus, a more detailed input data for single horticultural products would be recommended to allow analysis of economic efficiency and competitiveness to be carried out, especially taking into account the near joint of candidate countries, several of them already active on the fruit and vegetable markets. Furthermore, recently there has been a growing request of analysis for evaluating both ex ante and ex post the impact of CAP measures. In this view, detailed input data on wine, olive oil, tobacco, rice, tomato for processing should be available considering that for these products specific Common Market Regulations exist.

activities instead of outputs; it is for this reason that for instance a standard gross margin (SGM) is calculated for dairy cows, not for milk, and that the number of dairy cows is collected in the Farm Structure Survey (FSS), not the beef production.

It is also important to note that production methods (e.g. organic farming or production in the open-air/under glass, or irrigated/non-irrigated) can apply to several activities and vice versa. Several of the inconsistency problems between statistics can be traced back to blurring the difference between the concepts of output, activity and production method. Heroic trials to integrate some of the concepts into one hierarchical list in a database (e.g. a list of activities with some of them split to different production methods like irrigated maize and open-air vegetables) lead often to databases that are not very transparent and hard to manage.

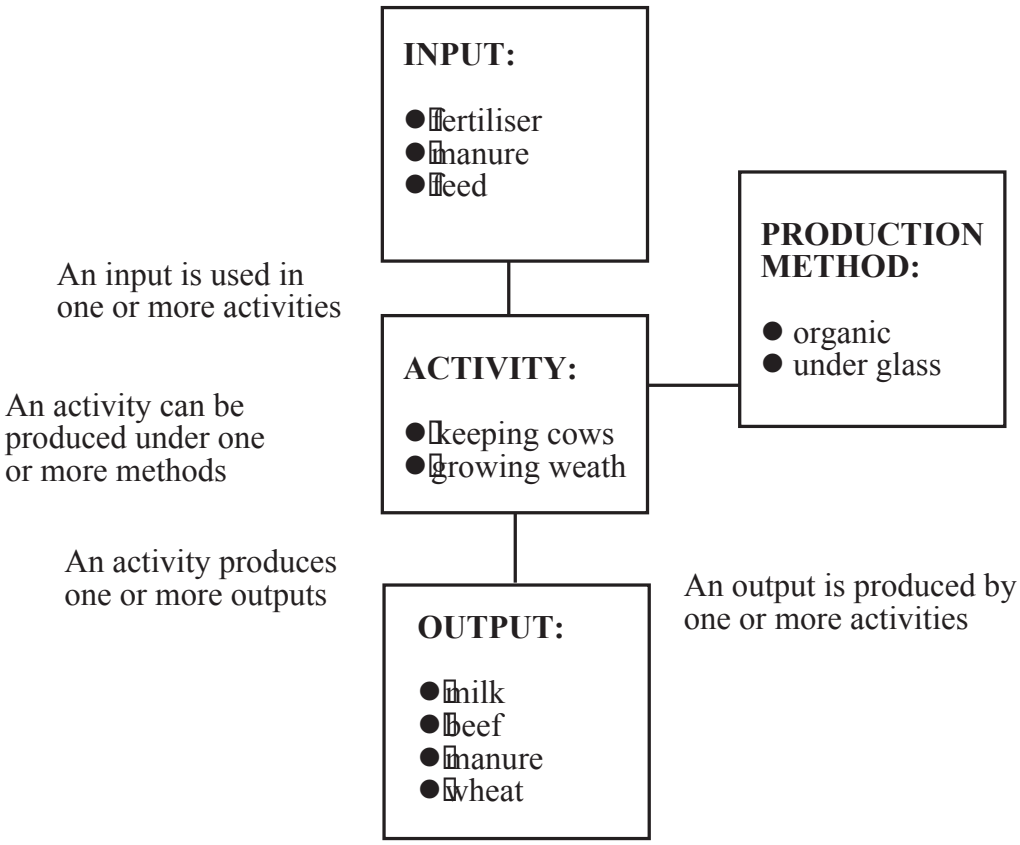


Figure 2.1 Data model to show the relationship between the concepts Input, Activity, Output and Production Method (with examples identified with a •)

The data model described in figure 2.1 has been used to make a list of activities. Based on the current EAA also a list for inputs was made. For both lists it is important to have common criteria. These criteria were formulated during the workshop mentioned and were used to decide on the list of activities and inputs. Of course the definitions in Farm

Accountancy Data Network (FADN) and SGM database were kept in mind. This resulted in the following criteria to define the lists of activities and inputs:

- economic importance;
- environmental importance;
- models and policy relevant;
- consistency with EAA desagregated;
- availability, quality;
- rural development issues.

For the outcome of this decision process it was important to take into account what is desired by the users in future by Eurostat and others as well as what is feasible and available. At the same time it is relevant to be aware of developments in agriculture (for instance the growth of horticultural production, flowers and plants) and the society. For example it is important, if possible, to include information on environmental issues (minerals, pesticides) as well as on food safety. This may have as a consequence that the list will be reviewed in future. The list of activities and inputs defined with the use of the criteria mentioned above has been labelled EAA+ (see appendix 2). This list has been used during the project with the objective to test its feasibility.

The main changes on the input side have to do with environmental issues:

- split of fertilisers into N, P, K and others;
- split of pesticides into:
  - fungicides;
  - insecticides;
  - herbicides;
  - others;
- add water and manure;
- a different split of feedingstuffs.

The main differences on the output side are:

- add triticale, grass and mushrooms;
- split ornamental crops into (like in FADN):
  - live plants;
  - flowers;
  - flower bulbs/tubers;
- split potatoes into:
  - ware potatoes;
  - seed potatoes;
  - potatoes for starch production;
- split of animal categories into activities;
- a more detailed split of cattle.

2.2 Analysis of sets of data

As described in the introduction of this report, the first step was to investigate if other databases that have data available for all Member States of the EU, could be useful to improve the AgrIS database. The two most promising databases are the FADN and the SGM in combination with FSS. The SGM calculations are in most countries based on FADN data but sometimes also other data is used for this calculations. In section 2.2.1 some additional remarks on the databases are presented. While comparing different databases, one should keep in mind the diverging objectives and assumptions behind these databases. Therefor section 2.2.2 starts with a description of some general differences between the databases. In section 2.3 it is described how the databases were used in this project. Among others, the differences in definitions of inputs and outputs between the databases are described. In the last section the conclusions of the coupling of the databases are described.

2.2.1 Closer look to the databases [1]

SGM and FSS

Eurostat assembles every two year standard gross margins per activity. The methodology is described in a working document, Classex 44, and Eurostat is working on a manual for this domain. The SGM are used for the calculations of the economic size of the farms in both the FADN and the FSS. Therefor SGM are calculated for all activities included in the FSS.

To calculate the margins, variable costs have to be allocated to the different activities. As stated before, the categorisation into activities makes the allocation of costs less complicated. Although Eurostat procedures require Member States to make the details of their calculations available, this is often not done, due to a lack of capacity and the fact that the details are not used or published by Eurostat. Therefor normally, Eurostat only assembles the total (the margin) and not the underlying numbers (output, variable costs etcectera). The standard gross margins are published on paper, but not much used in statistics or research except for typology purposes. For 1996 however Eurostat also assembled the split up of the calculations. This means that the margin is split up in returns and costs. The direct costs are split in: fertiliser, crop protection, seedlings, feeding stuffs and other direct costs.

This means that not all inputs as specified in the EAA+ list (appendix 2) are available; energy and veterinary costs are not made explicit in the detailed SGM. Besides this not all Member States have sent details on SGM calculations of 1996 to Eurostat yet or have not given permission to use the detailed SGM in this project.

In the smaller countries one SGM is calculated for the whole country while in the other countries SGM are calculated for several regions within the country. To arrive to the inputs used for the total production for a country (as presented in the EAA), the SGM should be multiplied by the total number of hectare or animals in a country (or its regions). These total numbers can be found in the FSS and national Census. SGM should be based

[1] See section 1.3 for a general introduction.

on an average of 3 years although some countries use even a longer period to smooth changes over the years. This makes it difficult to compare SGM data with e.g. 1997 data from the EAA.

In the context of this project is has to be underlined that SGM are mainly calculated so far for the typology of farms and the economic size of farms (farm structure, in relation with FSS) and not to provide information on costs of production. A new aim on the use of SGM data will automatically lead to extra needs on information.

FADN

FADN (or in French: RICA) is database of DG-AGRI with financial and technical data of 60,000 farms in the European Union. The costs are not allocated to the outputs in the FADN, nor to activities. Therefor the FADN could only be useful if also a method for allocating costs to activities can be found. The advantage in comparison with the EAA is however that data of individual farms are included which might make it easier to allocate costs by econometric methods or based on specialised farms.

2.2.2 General differences between the databases

Some relevant differences in underlying assumptions between the databases to be used in the project (EAA, FADN, SGM and FSS) exist on:

(1) time and availability, (2) representativity and (3) taxes and subsidies. It is important to be aware of these differences in the context of the project; it means that some data can not be compared with data of other databases without correcting for this differences in assumptions.

Time period and availability

The databases have a different time horizon and deliver data for a given period (for instance year x) with a different time gap (figure 2.2).

	Period	Differences between countries in period	Most recent period available during project (summer 2001)
EAA	1 calendar year	No	2000 (forecast) 1997 (final)
FADN	1 year (not necessarily calendar year)	Yes and also within countries (e.g. between horticulture and agriculture)	1998/99
SGM	Average of (mostly) 3 years	Yes, often based on FADN data so also differences within countries. Because SGM are averages of 3 years, differences are small. Some countries smooth data over a longer time frame	'1996'
FSS	One moment	Yes, time differences up to 1,5 year (for intermediate surveys 15 months and for basic surveys 27 months)	'1997'

Figure 2.2 Differences between databases in time period and availability

The table shows that for countries where the FADN year is not equal to a calendar year, the numbers are incomparable to the EAA unless corrections for the differences in period are made. In appendix 1 the bookkeeping years of all Member States are presented. The SGM can only be compared with the other databases when an average of some years is used.

Representativity

Not all statistics have the objective to represent all agricultural activities in a country (figure 2.3).

	Representativity	Definition per country
EAA	All agricultural (incl. hunting and contract work) activity (including agricultural activity by non agricultural units) + non agricultural activity produced on agricultural units for which inputs can not be separated from inputs used for agricultural activity. Farms that are too small to be included in FSS and that produce only for own final consumption are not included.	Unclear if some Member States have more restricted definitions
FADN	All 'commercial' farms, which implies more than 90% of agricultural production.	Between >1 ESU and >16 ESU (different thresholds)
SGM	Officially it should be representative for all farms in FSS. But for ease of calculation some Member States use only large and/or specialised farms.	Unclear, probably between >1 ESU and >40 ESU
FSS	All units with agricultural activities above a minimum size.	Officially survey units are holdings with: UAA > 1 ha a) Less than 1 ha UAA for certain products b) Some countries might have slightly different definitions (for example the Netherlands >3 ESU)

Figure 2.3 Differences between databases in representativity
a) UAA is utilised agricultural area; b) More detailed information in Eurostat publication: Farm Structure - Methodology of Community surveys, Brussels, Luxembourg.

In some countries the SGM calculations are based on the FADN results of only the larger specialised farms or on management systems meant for large farms. This means that they do not have to be representative for the complete production in the EAA.

The EAA represents a larger part of agricultural production than the FADN. There are three important differences:

A. EAA also represents hunting and contract work;
B. EAA also represents agricultural production by non agricultural units;
C. The FADN only represents the agricultural production of commercial farms which means that the smaller farms are not included.

In a lot of countries however the FADN is one of the main databases used to calculate the total input use in the EAA (see chapter 3). Although the production of contract work is not included in FADN, the consumption is (see section 2.2.3).

To compensate for differences B and C (in those Member States where links between FADN and EAA are established), a lot of different methods are used. The method that is used might be important in explaining differences between the total in the EU-FADN and the EAA. In appendix 3 some of these methods are described. It would be a large improvement for the harmonisation between countries of the EAA if more information would be available and if all countries would use the same method.

Taxes and subsidies

Both inputs and outputs are exclusive VAT in all databases. Direct income subsidies are always included in the SGM calculation and can be separately identified in FADN and EAA.

	VAT	(Direct income) Subsidies
EAA	Only non deductible VAT	Producer prices, subsidies on product, taxes on product and basic prices are separately booked.
FADN	Exclusive (total difference between paid and received VAT is recorded but no split to products)	Direct income subsidies can be separated from receipts.
SGM	Exclusive	Direct income subsidies are included in the total margin of a product.
FSS	No data in monetary value, so irrelevant	No data in monetary value, so irrelevant

Figure 2.4 Differences between databases in taxes and subsidies

2.2.3 Using FADN to allocate inputs to outputs/activities

2.2.3.1 FADN/INRA model

As described before, inputs are not allocated to outputs or to activities in the FADN. This means that extra information is needed. This information could be gathered through other sources (for example technical relations) but inputs could also be allocated using econometric models. The INRA has built such a model (COUTPROD) for the French FADN (Appendix 4, Butault et al., 2001). It is for instance used for the calculation of SGM in France.

Just before the start of this project, a first try was made to extend the model to European Union level for DG-AGRI. This version of the model is based on the data of EU-FADN (Butault, J.P. and J.M. Roussel, 2001). Because the model has been used for a long time in France already, it has been fine tuned to the French situation. This means that not only the model itself is improved but also experience with the use and the interpretation of the results of the model has been gathered. These steps are not yet made for the European Union version. The model has been used to make FADN data available for this project to improve AgrIS (box 2.1).

The model is an econometric model based on regression analysis. It uses micro-economic data from the FADN farms and calculates average use of some inputs per 100 euro (or Ecu) output. When multiplied with the total production, macro-economic aggregates on costs of production related to specific products can be calculated. For DG-AGRI, calculations were made for the main agricultural products for the years 1990-1997 (for 12 Member States until 1994; for 15 from 1995 onwards). Because for most products also the price of the products, the output per ha, and the number of animals are known in FADN, the costs of production per unit of product (quintals), per hectare and per animal can be calculated.

DG-AGRI confronted the INRA model calculations with cost of production calculations of milk based on their own calculations. In their own calculations, they used simple allocation rules for allocating costs: like the percentage milk cow livestock units in total livestock units and the percentage milk output in total output. The differences between the INRA and DG-AGRI calculations were rather small for most countries. For some countries (Austria) differences were larger. The DG-AGRI calculations were more stable over the years than the INRA results.

The value of production is defined in the model as the total sales plus the household consumption and changes in stocks (at producer prices). These stock changes are also taken into account for the inputs. Internal consumption of seeds and feeding stuffs however are not included in the costs and the outputs. This was impossible because not all countries include the internal consumption of some products (for example forage crops) in the EU-FADN. For products where a large part of the production is used for internal consumption (for example cereals used as feed for pigs), this might lead to strange results at first sight. The costs of the production of the feed (seed, pesticides etc) are however included in the costs of the pigs that consumed the grain.

Because nearly all forage crops are used on the farm, no cost of production of forage crops is calculated. For other crops which can be used within the farm and for milk, only the part that is sold to others is included in the model.

(Direct income) Subsidies and taxes are not included in the receipts but as a separate item. The subsidies are defined as negative costs.

The first version of the model is made for DG-AGRI and takes into account all variable and fixed costs. The variable costs specified are:

- fertilisers and soil improvers;
- motor fuel, lubricants and heating fuels;
- plant protection products;
- seeds and plants;
- feeding stuffs (compound feedings and others);
- other specific costs for animal production (as for instance veterinary costs).

The fixed costs are split into:

- regular costs of maintenance buildings;
- regular costs of maintenance equipment;
- other goods and services (electricity, contract work, water, insurance, other specific costs of crops and other general costs);
- property costs land;
- depreciation;
- paid wages;
- interest (excluding land);
- taxes.

Only the first three categories of the fixed costs are relevant for this project. The first two (maintenance of buildings and equipment) are also separate categories in the EAA and the third category (other goods and services) is split in three categories in the EAA (energy, agricultural services and other goods and services). For this project, it is important that at least some of these categories of costs should be more specified. The maximum level of detail however is the FADN. This level might still not be enough to get a complete match with the EAA definitions or EEA+ requirements.

In the first version of the EU-model the costs of production are linked to 27 selected agricultural products (outputs). These include five specific categories of cereals (soft wheat, durum wheat, barley, maize and other cereals), as well as milk of cows separated from milk of sheep. For Greece, Italy and Spain specified categories are added on olives, olive oil and citrus fruits.

Because in the end all costs and revenues are included, the model provides as result the profit per product. A constraint is included in the model that the total profit of all products should be equal to the total profit (family farm income) of the farm.

The costs of production are based on the data of all farms (the mixed farms as well as the specialised enterprises). The model does not take into account scale effects which means that it is assumed that larger farms have the same costs per ECU output as smaller ones. A detailed description of the model is given in appendix 4.

Box 2.1 Short description of FADN/ INRA model and its development as contribution to the project

2.2.3.2 Differences in definitions between the FADN/INRA model and EAA

After the analysis of the differences in assumptions and scope between the databases, a detailed analysis of the definitions of the data in the EAA and FADN has been performed. A comparison of the definitions used in the first version of the INRA model and EAA resulted in a number of improvements in the model. Because the model is based on FADN data, the maximum level of detail is given by the FADN definitions [1]. After all the possible improvements in the model were made, there still remained some differences in definitions between FADN and EAA (and EAA+).

Because we do not know yet if all the items in the EAA+ (appendix 2) are available, it would be too ambitious to try to base the analysis on EAA+. Therefor the analysis will concentrate on the items already included in the current EAA. We will check the availability of EAA+ results in the Member States in the next chapter. The analysis deals only with the situation in 1997. It would be too complex to include all the changes in the definitions from 1973 to 1997. We will first discuss the inputs (a.) and then the outputs (b.).

a. Problems with inputs

In the current EAA, the inputs are split into the following items:

- seeds and planting stock;
- energy;
- fertilisers;
- pesticides;
- veterinary Expenses;
- animal feedingstuffs;
- maintenance of materials;
- maintenance of buildings;
- agricultural services;
- other goods and services.

[1] Not only the definitions of inputs and outputs/activities of EAA and FADN have been compared but also the differences with SGM. The definitions of the three databases are linked to each other in a spreadsheet. They are presented in such a way that the differences in definitions are easily visible. In the next section the main differences between EAA and SGM are discussed.

Since the revision in 1995, the use of animal feedingstuffs produced on the own farm, is included, while in FADN it depends on the way of treatment (with differences per Member Country) if this is included and to what extent. For several countries the internal use of feedingstuffs for grazing livestock (like grass and silage maize) is not included. Because in the INRA model all countries must be treated in a harmonised way, the internal use of products is not included at all. Although the amount of internal use of seed and feedingstuffs for pigs and poultry could be added to the model (because it is included in the FADN), the internal use of feedingstuffs for grazing livestock could only be added if information from other sources could be attained. Therefore only part of the total feed costs in EAA can be allocated using the INRA model.

In the improved version of the INRA model, the item 'other goods and services' is split into three separate categories:

- energy, which is added to 'Motor fuel, lubricants and heating fuels'. In this way this item is completely comparable with the EAA item Energy;
- contract work;
- other goods and services.

In the 'other goods and services' insurance is skipped. This is done because in the EAA only the part of the insurance premium that is paid for the services and the profit margin should be included. The part of the premium that is reserved for settling claims, so the part that is expected to be paid back in case of disasters, is not included. In the FADN however the total premium should be included. Because we expect that the premium reserved for settling claims will be a large part of the total premium, insurance is not included in the 'other goods and services' in the INRA model [1].

In this way the categorisation used in the final version of the INRA model is as much as possible the same as in the EAA. It was not always possible however to use exactly the same definitions. In appendix 4 the exact definitions are presented. The largest technical problems that remain in linking FADN data with the INRA model to the EAA are the following:

- cost of inputs corresponding to output period or cost of inputs used in period;
- forestry costs (sometimes partly included in FADN);
- veterinary costs in EAA versus other specific livestock costs in FADN;
- other goods and services (includes other specific livestock costs than veterinary costs);
- maintenance of buildings and materials (investment versus repair);
- maintenance of materials (small purchases);
- seed (use of seed produced on the own farm for next period);
- contract work (machines hired with labour);

[1] This small change to the model has been made after the visit to the Member States. Therefor the final results of the model are slightly different from the results presented in the questionnaire.

- theory and practice (manuals are not automatically followed);
- cost of inputs used by contract workers.

The problems are described more in detail in appendix 5.

b. Problems with output

A general problem in the context of the project is that EAA is based on outputs and not on activities. EAA+ and SGM are based on activities. The FADN is also constructed on outputs, but the micro data of farms provide (detailed) information on the value of production of each product. Most activities lead to one product. Some however, like 'keeping of animals', lead to two products: an animal product (milk, wool, eggs) and meat (cows, sheep, laying hens). It is possible that the level of detail of outputs in FADN is too less, to make a complete allocation to one activity in EAA+.

For some products, like fattening pigs, it is possible to produce them on a contract base. Although a lot of contracts exist, we mean here the type of contract in which another company, for example a slaughterhouse, delivers piglets and feed to the farm. The farmer does not have to pay for this and is not the owner of the piglet and the feed. Per piglet delivered to the slaughterhouse, he receives a compensation based on the cost of housing and labour that he has provided. These farms therefore do not own animals and have no variable costs. This means, that they could strongly influence the regression analysis in the INRA model. Based on this and the results of the first runs of the model we decided to exclude these farms from the FADN sample and therefor from production. Although quite some farms were excluded from the analysis, the results did not change as much as we had expected.

Several corrections have been made to the categorisation of outputs in the first version of the model. Corrections were made for, among others, industrial crops, vegetables, wine and milk.

With the new model two different runs have been made: one with a detailed specification of outputs and one with groups of outputs (see appendix 4). We will concentrate the analysis on the detailed specification because this model leads to slightly better results and more interesting data because data can always be grouped afterwards if needed.

After these corrections the following linkage problems remain:

- vegetables (strawberries, (water)melons);
- flowers and plants (growth of young plantations);
- fruit;
- inseparable non agricultural activities and other output (processing of agricultural products etc).

They are described in detail in appendix 6. For all other outputs definitions could be matched.

2.2.4 Using SGM to allocate inputs to outputs in EAA

2.2.4.1 Coupling of SGM and FSS

Activities

As described before the totals in the SGM 1996 are multiplied with the totals (animals/hectare) in the FSS to arrive at aggregated results that can be compared with the EAA results. In theory there are no problems with the coupling of the SGM and FSS because definitions of activities are perfectly harmonised. The SGM after all are used for the calculation of the economic size and the farm type of the farms in the FSS. In practice however there are a few problems because SGM are sometimes available on a more detailed product level than the FSS data.

Regions

For most countries SGM are only available on a regional scale. There are no national averages. To get a national average, a weighted average of the regional totals should be calculated. The regional SGM are weighted with the number of hectare or animals in that region for each type of crop or animal as a proportion to the total hectare and number of animals in the whole country.

Based on the data of FSS 1997 and SGM 1996 it is not always possible to couple the data sources because of differences in the regional categorisation. When a coupling was not possible on FSS 1997, it was however possible to use the FSS 1990 which has a more detailed regional differentiation than the FSS 1997.

The detailed SGM were available for nine European countries. Belgium, France, Luxembourg and the Netherlands provided information about the country as a whole. The data provided by these countries could be used without further manipulation.

Finland provided SGM for four different regions that could be perfectly matched with FSS data of these four regions. Sweden provided SGM for three different regions. Because the FSS data used for the weights concerned eight different regions, this information had to be reduced to three regions to be able to multiply the weights with the variable values. The eight smaller regions perfectly fitted into the three bigger regions, so the hectare and number animals were added over the smaller regions in order to obtain those for the bigger regions. The same method was used for the calculation for the UK (SGM from six regions were weighted with FSS data from eleven sub-regions that perfectly fitted into the six). The national totals for Denmark were calculated in the same way, except that the weights of 1990 were used instead of those of 1996. Denmark provided information about two regions and the weights were available for twelve different regions that could be reduced to the two bigger regions just mentioned.

For Austria another problem existed. Austria provided SGM for nine regions, whilst information about the number of hectare etc was only available for three regions in 1997 and not at all in 1990. To be able to reduce the number of SGM regions to three, the SGM were weighted with data from 1999 about agricultural land acreage's in the smaller regions divided by that in the whole country. In this way the number of regions was reduced to

three before the multiplication with the number of hectare and animals took place to obtain the national totals of all variables.

LFA

Sometimes a total SGM was available while the number of hectare and the number of animals were split into LFA numbers and non-LFA numbers. In this case it is assumed that the total SGM is representative for all hectare and animals. In the case that a separate SGM for LFA regions was available, this SGM was used for the LFA regions and the other SGM for all other regions.

2.2.4.2 Coupling of SGM/FSS and EAA

This paragraph presents information on the problems registered in the process of coupling SGM and EAA data as well as the corrections made to solve, if possible the problem. Some problems have a general nature (a.), others are on specific input (b.) and on products/or activities (c. output).

a. General observations

The (calculated) national averages of the SGM were multiplied with the totals in the FSS 1997 for all activities. Nearly all countries however did not deliver a detailed calculation of the SGM for all activities. When this was the case for activities with a small economic value, it could be assumed that the split of costs is the same for other activities. Also for larger activities (mainly horticultural ones) however sometimes no detailed calculation was available. In this case the total input use in the SGM can not be compared with the input use in the EAA.

Because all countries delivered their data in a different data format, it took a lot of time to find out how the data could be transformed in a harmonised format with the other countries.

Corrections made

Before the comparison started some corrections were made in the SGM calculations [1]:

- some subsidies (like direct payments, Mac Sharry) are included in the SGM calculations. The INRA model calculated cost of production coefficients at producer prices (without subsidies). To make SGM and FADN/INRA model calculations comparable, subsidies were subtracted from output of SGM. For some countries however, the subsidies could not be separated in the SGM calculation. In this case the subsidies included in the EAA were subtracted from the output total in the SGM to end up with an output value exclusive subsidies;
- for the animal SGM, the purchases of animals were subtracted from both variable costs and output. This is needed because in the EAA, the trading of living animals is

[1] If the detailed SGM's are included in AgrIS these changes could also be made within AgrIS which makes it possible to see the changes to the orginal data.

not included. Output of these purchased animals is already included in the SGM of other animal categories so including them would lead to double counting;
- the SGM of wood were excluded because they are part of the EAF (Economic Account for Forestry) instead of the EAA;
- the SGM of piglets were excluded. In the SGM of breeding sows, the outputs and inputs of the piglets are already included;
- in holdings with grazing livestock, fodder is mainly consumed by the grazing livestock on the holding. Fodder is therefor not included in intermediate consumption and does not need to be valued in the calculation of the total SGM of a holding. Only in the case where a fodder deficit or surplus exists, the fodder should be given a separate SGM. On national level however, one can assume that there is no fodder deficit or surplus. In that case no SGM for fodder should be calculated. In the EAA however the consumption of feed that is bought from other farms and the feed that is produced and consumed on the same farm should be included as a separate item. This means that fodder is both included as output of fodder crops and feed costs of grazing livestock. Therefore also the SGM of both fodder crops and grazing livestock should be included. The margins of fodder areas that are produced and consumed on the farm might however be included in both the fodder areas SGM and the SGM of the grazing livestock which eat the feed. Only when in the SGM of animals, the feed is valued at market price instead of variable cost of production, it can be concluded that the margin on the production of feed is not included in the SGM of animals but only in the SGM of fodder crops. In this case both the SGM of animals and fodder should be included in the comparison. It is however not always clear how countries have calculated their SGM and how the fodder is valued in the EAA. Some countries do not include the fodder that is bought from other farms and the fodder that is produced and consumed on the same farm in the EAA yet. In the Netherlands, the fodder produced and consumed on the own farm is valued at variable costs (in the EAA) in both the inputs and the outputs. In this case the same solution should be used for the SGM. So for the Netherlands the best solution would be to correct for the part of fodder crops that is used on the own farm. Instead of the market price, the variable costs are included as output of fodder crops and input of grazing livestock. The discussion above shows that it really depends on the situation in the Member States what should be the best solution. Therefore for other countries no correction is made;
- in the FSS the number of poultry is shown in thousands. The SGM however are calculated per 100 animals.

Some remarks

For two reasons one would expect the SGM totals to be a slightly higher than the EAA output totals. First, the SGM calculation assumes that all hectare are harvested and that there are no large outbreaks of diseases. In practice part of the crops can not be harvested and sometimes large diseases do occur. In the Netherlands for example an outbreak of classical swine fever took place in 1997. Therefor the production based on SGM was much higher than in EAA.

Second, the SGM calculation is most of the time based on (larger) commercial farms. These farms have on average a slightly higher output than non commercial farms [1]. The difference in output between SGM and EAA is for example very large for horses. In the Netherlands, most horses are kept for hobby (horse riding) and do have no margin (or a negative margin). The SGM however is calculated for farms that have horses for commercial reasons (riding schools, breeding). The multiplication of the SGM with all horses in FSS leads to a much higher estimated output and input values. In general however, both reason will only lead to small differences in results between SGM and EAA.

b. Inputs

In the detailed SGM data (only available for SGM 1996), the direct costs are split in: fertiliser, crop protection, seedlings, feeding stuffs and other direct costs. The SGM are often based on FADN data. This means that general problems between the definitions in the FADN and EAA are also a problem in the comparison of SGM and EAA. So veterinary costs and energy (heating costs) can not separately be identified in the SGM calculations. The veterinary costs and part of the energy costs (heating) are included in the other direct costs. Also part of the costs which are included in the other intermediate costs in the EAA, are included in these other direct costs in the SGM: water, variable marketing costs, variable insurance costs, costs of artificial insemination and covering, costs of milk checking and animal selection. The definition of 'other direct costs' can therefor not be harmonised with items in the EAA.

In the EAA all feed used is included but a split can be made between bought feed and feed produced on the farm. In the FADN, only the bought feed is included for grazing livestock in some countries. Therefor we only used the bought feed in the comparison between EAA and FADN. In SGM calculations all feed is included. It is not possible to make a split between bought feed and feed produced on the farm. Therefor we can only compare the total feed use with EAA. This has the disadvantage that the cost coefficient for feed based on SGM can not be compared to the cost coefficient for feed based on FADN. The advantage however is that all feed costs in EAA can be allocated using SGM.

Cost of maintenance of equipment and buildings and the cost of contract work are not treated as variable costs and are therefor not included in the SGM calculation. Thus the SGM can not be used for the allocation of these costs.

c. Outputs

SGM and FADN are more or less equal concerning outputs [2]. So the same coupling problems appear. Besides this the SGM output is inclusive subsidies and based on a three-year average. In appendix 7 the exact definitions used in the coupling, are presented.

For most crop activities, definitions can be harmonised. SGM of secondary crops and kitchen gardens are not included. It is possible that some differences appear in the category 'other crops'. Rough grazing (F/02) is included in this category for example but it might

[1] The influence on inputs is less clear. The larger farms will have higher output but also a higher input use.
[2] Sometimes the FADN is bit more detailed.

also be included in the fodder crops. It may be expected however that both inputs and outputs will be very low for this uncultivated area. For some countries however specific coupling problems occurred because the SGM were not available on the most detailed level.

Problems on output per specific product and activity
A specific problem occurs for the animals and crops with two or more different products as outputs in the comparison with the EAA, based on outputs instead of activities. The costs of these activities (milk cows, laying hens etc) should be split to the products in the EAA. In theory this is impossible. Based on some assumptions, estimates could however be made. This is however only possible if the value of the different outputs can be identified in the detailed SGM data calculation. If this was the case the procedure described in box 2.2 was used (milk cows is used as example).

2.2.5 Analysis of results of the INRA model and SGM

In this chapter, we discuss two methods for using additional data sources at EU level to calculate input data and link it to activities or output data in AgrIS. In theory it is attractive to introduce into AgrIS data from the FADN and SGM/FSS data set on inputs (aggregated to national totals) and linked to activities as well as outputs, independent from EAA data. FADN and SGM however can only be used to link input to activities in EAA if definitions are harmonised and when the reasons for differences in the values of output and total input are known. Differences in definitions are discussed before but the influence on the values is not yet known. Therefor the total output value in FADN and SGM/FSS are compared in this section.

This analysis can also be used to check the quality of the current EAA data and to improve AgrIS and/or FADN and SGM/FSS data. If definitions are the same, values should be the same. The quality of the allocation of inputs to outputs itself can not be checked with the EAA data. It can be checked however if the use of the input/output ratios from the FADN for the allocation of EAA inputs to EAA outputs, would lead to the input totals in the EAA.

In the following section, the first results of the comparison of the aggregated totals in the databases are presented. The results have also been presented to the Member States. Their comments will be discussed in the next chapter. First the coupling of the INRA model (FADN) with the EAA is discussed and afterwards the SGM and EAA.

If the SGM of milk cows includes also a positive gain in value of the animals itself (calves and value of cull cows for slaughtering) > value of heifers), the following costs are allocated to the milk:

(output of meat) * (costs all other cattle categories/output all other cattle categories).

It is however more complicated if a loss in the value of the animals results. As a start all costs of milk cows are allocated to milk and all costs of other SGM cattle categories are allocated to meat. Part of the costs of cattle categories that are also on the dairy farms should however also be allocated to milk.
This part is calculated by the following formula for both feed costs and other intermediate costs:

Neg. cattle output milk cows * (costs all other cattle categories/output all other cattle categories).

Part of the value of the heifer was in fact the potential to give milk. The farmers accept this decrease in value of the heifer because the value of the milk is higher than the decrease in value of the cow. The decrease in value of the milk cow has to be subtracted from the total meat production to harmonise it with the EAA output definition (sales of living animals are not included). But in fact this decrease in value is a cost of the milk production. Assuming that the value of meat of a heifer is equal to an old milk cow, the extra value because of the milk is the difference in value between a heifer and an old milk cow. The costs of creating this extra value, could be estimated by allocating total costs of keeping of non dairy cattle proportionally to the value of the different outputs. So if 50% of the value of the heifer is caused by the potential to give milk, 50% of costs are allocated to the milk.

The above mentioned formula could also be defended by the following reasoning. The keeping of young cows for milk is not different from the keeping of young cows for meat. Both cows need the same treatment. One would expect that in this case, the variable costs as a percentage of output would be the same for both categories of cows. Therefore one may assume that the costs as a percentage of total output of meat is the same on dairy farms as on farms with cows for meat.

The same reasoning could be made for other animals. So if SGM are available for both a category that only produces meat and a category that produces mainly an animal product, this method is also used (for poultry, goats and sheeps). If however only one category exist for that animal, the costs are split assuming that the relation between output and costs is the same for both outputs.

The following formula is used:

Cost milk = (total costs)* (output of milk/total output)

For some animals in some countries the value of the individual outputs could not be identified from the SGM calculation. If it can be assumed that one of the products represents a very large part of the production, the SGM are allocated to this output.

The following decisions have been made in this case:

- all costs of milk cows are allocated to milk;
- all costs and output of goats are allocated to milk;
- all costs and output of sheep are allocated to sheep and goats;
- all costs and output of laying hens are allocated to eggs;
- all costs and output of rabbits (J17) are allocated to other animals;
- all costs and output of J18 (bees) and J19 (other livestock) are allocated to other animal products.

Box 2.2 Procedure to allocate input of one activity to individual outputs

2.2.5.1 INRA model (FADN) versus EAA

In previous paragraphs a number of difficulties have been described related to the coupling of databases. Although not all difficulties have been solved, a first coupling of the data in

EAA and FADN has been performed with the use of the definitions on inputs and outputs described in section 2.2.4. The methodology for the coupling is described in detail in Appendix 9.

In this paragraph a short description and some explanation of the results is given on output values of products (a.) as well as on input data (b.).

a. Output

First the total production of the outputs in the FADN is calculated by multiplying the total production per farm with the weighting factor per farm. No corrections have been made for the 5 to 10% of the production that FADN does not represent. The FADN total is compared with the total value in the EAA. Both the production in EAA and FADN is inclusive the production that will be consumed on the farm itself. Because some countries do not include the production of forage crops used on the farm in FADN, the total production of forage crops is not for all countries comparable between EAA and FADN.

We made this comparison to check the similarity of definitions and the representativity of the FADN for all activities. In figure 2.4 the totals in EAA and FADN are compared for soft wheat and milk. In appendix 11 the totals are compared for all outputs. The match for soft wheat is rather good. For some other products however, the match is less favourable.

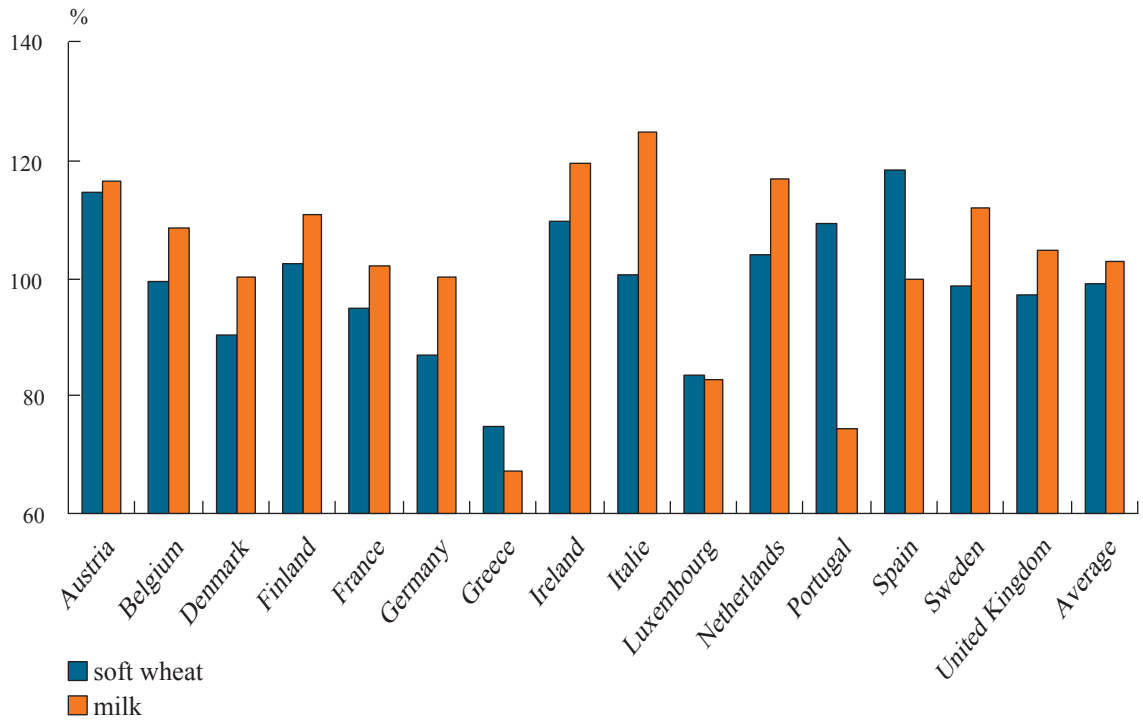


Figure 2.4 Total output of soft wheat and milk in FADN as a percentage of the output in EAA, 1997

It is expected that FADN results in a slightly lower production than EAA because FADN does not cover all production. For some activities (e.g. Protein crops, Other crops and other animal products) the FADN results in higher (aggregated) values than the EAA. As written before, the definitions of other crops and other animal products could not be matched very well. The very high difference between the production of 'other animal products' in FADN and EAA for Germany is however rather strange. The production in FADN includes manure for sale, stud fees received and eggs other than hens etc. One would not expect these items to have a high value.

For other products FADN results in lower values: vegetables, flowers and plants, fruit, equines, poultry, eggs and other animals.

The main reasons for causing the differences are probably [1]:

A. differences in definitions and differences in interpretation of definitions between EAA and FADN;
B. fields of observation;
C. weighting in EU-FADN;
D. use of bookkeeping years in FADN and calendar years in EAA.

Ad. A) The outputs for which differences in definitions exist are described in previous sections. Although the national FADN total is known by the person who makes the EAA, he/she might decide to use another source because this source has a higher reliability. This other source might have a different definition or a different interpretation of the definition than the FADN. In this case the EAA has probably better data than the (aggregated) FADN.

Although definitions might be harmonised, interpretations of definitions might differ (see section 2.2.3). For example in the Dutch FADN, the production of hatching eggs is part of egg production. In the Dutch FADN, farms with hatching hens that breed the eggs (hatcheries), are not included. This means that the sales of 1-day chickens is not included while the sales of hatching eggs is. In the EAA however the sales of these 1-day chickens are included in the poultry production, while the production of hatching eggs is not included in the egg production, because it is treated in the same way as the trade of other living animals. This means that the egg production in FADN is larger than in EAA and the poultry production is lower.

In Italy strawberries and (water) melons are rather important products. As described before, these products are included as vegetables in FADN and as fruit in EAA.

Based on the more detailed analysis of the differences in some countries, we expect that the differences in definitions and especially the interpretation of definitions is the main reason for differences between EAA and FADN.

Ad. B) On average FADN should represent about 90% of agricultural production (that is 90% of total European Size Units) [2]. For some products however, the FADN represents a much lower percentage. Both eggs and poultry are badly represented in a lot of countries.

[1] Because the differences between FADN and EAA were studied in detail in the Netherlands, most examples are from this country.

[2] In practice this percentage is lower however.

These products are produced on very large farms that are not really interested to participate in the FADN. The poultry output (both eggs and poultry meat) is also poorly represented in Swedish FADN. Besides the reason mentioned before, this is caused by the fact that the Swedish FADN sample was not complete before 1998 and still there are very few poultry farms in the sample. Dairy farms, on the other hand, are well represented in the Swedish FADN. However, the total cattle output (beef), calculated using the Swedish FADN sample, is 20% less than the actual. This may be partly explained by the relatively smaller sample of specialised beef farms.

Furthermore, in Italy for instance a large number of both rabbits and poultry are reared for own farm consumption on very small farms which fall outside the FADN field of observation. The commercial production of both rabbit and poultry is done mostly in the very large units, which are either not willing to participate to FADN or if no UAA is present, they do not belong to FADN field of observation.

Italian FADN apparently undervalues both wine and olive oil value of production. The reason is to be found in the different field of observation. According to the new EAA manual, (par. 1.33.1, 2001 ed.), wine and olive oil production to be included in EAA is that made by farms, from grapes and olives grown in the same farm, and by groups of agricultural producers (e.g. co-operatives). The latter are not part of FADN field of observation. In Italy, processing of grapes and olives into wine and oil is frequently made by co-operatives (especially for wine) and other groups of farmers and this leads to the higher output value accounted for in EAA with respect to FADN.

In a lot of countries, producers of horticultural products are not well represented. Some products are not represented at all. This is sometimes the case for important products, for example the producers of seed for horticultural products. But most of the time these products are rather small.

In some countries however there are also some important products that are not very well represented. In Spain for example, the FADN represents less than 60% of total ESU and in Finland about 65%. It is regretted that the FADN itself does not periodically publish information on the representativeness of its sample.

Ad C) The weighting system of the EU-FADN is not the same as for the national FADN. For DG-AGRI the harmonisation of the weighting system between countries is very important. Therefore it is based on the most recent FSS and SGM available for all countries. This is most of the time a FSS of some years before. Besides this their typology system might be different from the typology system used for the selection of farms. In practice this means that large differences exist between the totals in national FADNs (that is in a number of countries used for the EAA) and the totals in the EU-FADN. For the Netherlands, for example milk production is always overestimated in EU-FADN. In 1997, pig-meat production in the Netherlands was highly overestimated. In 1997, there were large problems with classical swine fever in the Netherlands. Therefore less pig farms were included in the national FADN with an under representation of farms that had problems because of the disease (pigs destroyed, not possible to transport pigs to slaughterhouse etc). For this reason the national FADN overestimated pig-meat production already. The EU-FADN however over represented pig production with about 70% in comparison with EAA. This was also partly caused by the difference in bookkeeping year and calendar year (see D).

Besides this it is possible that farms which are included in a national FADN, are not included in EU-FADN. Some farms that do not contain all data needed in the FADN or which data is not available on time for the EU-FADN or is not reliable enough for the EU-FADN, are not included in the EU-FADN but can be used for national purposes. Some countries have higher national samples than is needed for DG-AGRI (Abitabile et al., 1999).

Ad. D) For countries with a difference between accounting year and calendar year, differences between EAA and FADN might appear (appendix 1). This is especially the case for products with high variability of prices (for example potatoes, pork and poultry).

Remark

Values in appendix 11 that differ significantly from 100 do not mean that the input coefficients are necessarily wrong. If the difference between the databases is caused by the fact that the FADN does not represent all production of a product, it is still possible that the farms which are included in the FADN are representative for the input/output relationship on the average farm that produces that product. In that case the information from the FADN/INRA model could be very well used to link EAA input values to outputs or activities. In this case reason B (field of observation) mentioned above, should not be problem. However there is no information available to test this assumption.

Input

In addition to the analysis on output values, an analysis on the inputs has been performed. Given the large differences on output values, it was decided not to aggregate the inputs on the basis of FADN and compare them with EAA values. This would lead to the same type of conclusions as on output. In stead the analysis focussed on an estimation of inputs, given the level of output in the EAA.

The methodology is described in appendix 9. The value in the EAA was for each type of output multiplied by the coefficients of the INRA model. This results in total cost of inputs per output. After this, the total of one type of input (for example feed costs) over all products was calculated. This results in the total input for a country. In appendix 10 this calculated input total is shown as a percentage of the input total in the EAA. The closer the figures are to 100, the higher the consistency between the two data sources (as far as the inputs are concerned).

Because the base for the calculation is formed by the output totals in the EAA, it is not necessary that the output in the FADN should be equal to the EAA as long as the cost of the farms included in the FADN are representative for the complete production in EAA.

As described before, the definitions of 'veterinary costs' and 'other intermediate costs' are not completely comparable. The calculated number often resulted in higher figures than in EAA is presented. The total intermediate costs are about 8% higher than in EAA. Only in Italy, Greece and Portugal a large difference in the total costs between FADN and EAA occurs (figure 2.5).

On average the cost of seed per unit of output [1] is higher in FADN than in EAA. This is rather strange because part of the seed that is produced and used on the farm is included in EAA (when it is produced in the first year and used in the second year) but it is not in FADN, so one would expect that seed use in EAA would be higher. In 7 countries however, the seed use in EAA is higher. In most countries also the fertilisers use and the crop protection is higher in FADN. So variable crop costs are higher in FADN than the totals in EAA. The average difference between EAA and FADN is rather small for energy, feed costs and other intermediate costs. In 10 of the 15 countries the costs of feeding stuffs in FADN are more or less equal to that presented by EAA. Feed costs are clearly underestimated (by 25%) with this model in Sweden. This may be partly explained by the small samples of poultry and specialised beef farms in the Swedish FADN.

For the services, and maintenance of buildings and equipment the differences between countries are very large. Some countries have much higher FADN costs (services in Portugal, Sweden, Belgium and Greece) while others have much lower (services in Austria, maintenance for equipment in Belgium and Spain).

It is difficult to explain why the maintenance costs of equipment and buildings are much higher in FADN than in the EAA for Finland. The repair costs are carefully documented in the FADN system, hence they can be considered somewhat realistic.

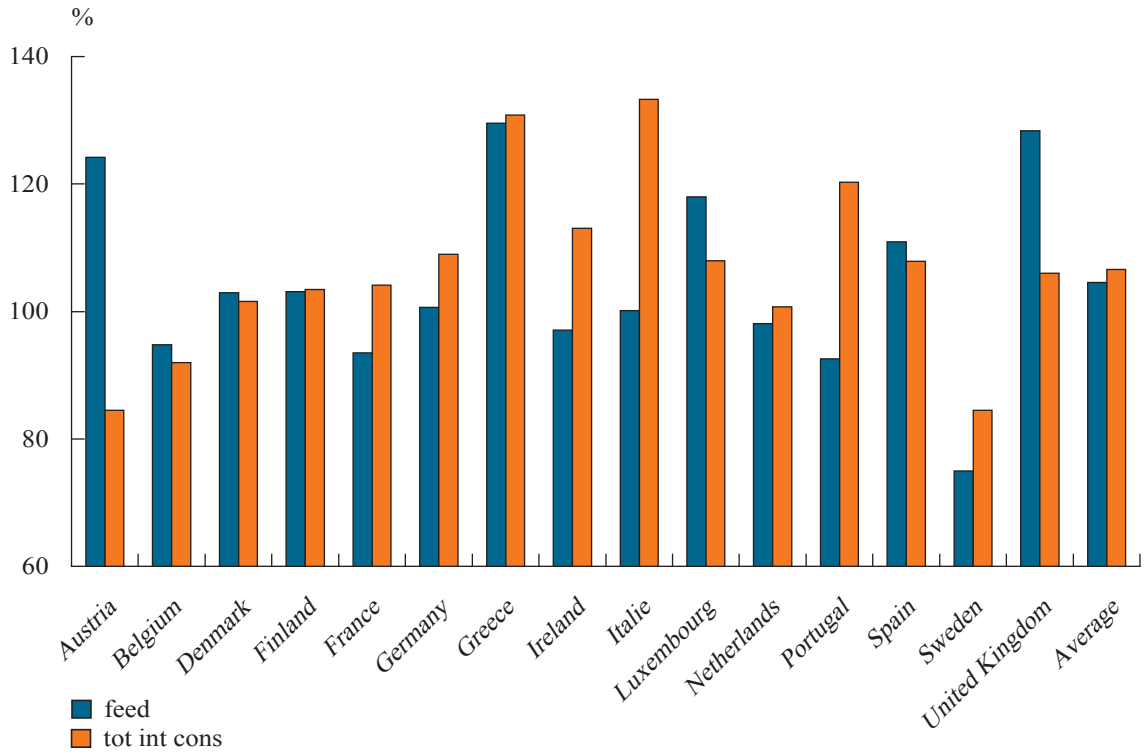


Figure 2.5 Total intermediate consumption and feed costs in FADN/INRA model as a percentage to EAA total, 1997

[1] In the text we will simply label this as a comparison between FADN and EAA.

The reasons for differences can be split in two groups:

A. differences in definitions;
B. the quality of the INRA model.

Ad. A) Differences in definitions between databases

The inputs for which differences in definitions exist, have been described in the previous sections.

Ad B) INRA model

This reason can be split in:

B1. reasons why models in general will not lead to perfect results;
B2. quality of the current model.

Ad B1) Problems with models in general

Econometric models are of limited use when multicollinarity occurs. This means that there is not much variation between the farms. In this case if the share (%) of the individual products in the total output is equal for most of the farms, there is no implicit information to allocate costs over the products. In the Netherlands for example a lot of farms have potatoes on 25% of their land and 20% sugar beets.

Also within homogenous output categories the production methods might differ significantly. Some farms will use a lot of machines while others use more labour. On some farms workers perform most work and machines from outside the farm (contract work) while others do not use these services at all. The model tries to find one coefficient for all farms independent of their production method. This problem could be partly solved if a variable would be included in the FADN that describes the production method. This variable could be included as one of the independent variables in the model.

A third reason that econometric models have problems to solve the allocation problem, might be the fact that some output categories are very heterogeneous. The costs of the individual vegetables for example differ significantly. The model tries to find one coefficient for all vegetables. In heterogeneous groups, this will never lead to good results.

Ad B2) Quality of the current model

One important point to underline is that the model was built on the structure of the 1985's Farm Return of RICA. This Farm Return was changed several times to introduce new variables and improve the quality of data. For instance more variables are now collected on direct payments (table J) where as subsidies by output are generated in the current model by considering it as negative cost. So it's possible to introduce new information but it is necessary to modify the structure of the files.

The INRA model was used is in several studies to analyse costs of production in EU Member States and some Candidate Countries. Results were discussed with experts in these countries that enabled improvements to be made to the model. Some other possible improvements were discussed in this current study, as shown in box 2.3.

1. On farm use of products could be included in the model
2. Non-agricultural activities
3. Introduction of expert knowledge or exogenous coefficients
4. Scale effects
5. Physical based allocation of inputs
6. Direct subsidies

Box 2.3 Possible improvements on the COUTPROD model

These improvements are explained here:

On farm use of products could be included in the model

Products and charges are estimated without on farm use production (seed and feedingstuffs). The costs relating to crops consumed on the farm as animal feed do not therefore appear as such, but indirectly in animal production costs. The reason of this option is the lack of valuation of fodder in RICA.

In the current model, one can expect that results could be upset for some crops, as barley, a lot of which is consumed within the unit as well as being sold. In fact, when intra-unit consumption is important a bias is reported on coefficients crops (intra-unit consumed) rather than on animal production.

The proposed improvement is to induce the on farm use production in the model. It's then necessary to estimate fodder value. Tests could be done for countries where this valuation exists. When it's not the case (most countries), we can introduce a fixed valuation on fodder areas. The change in the econometric model is quite simple: we restrict to zero the coefficients of animal productions in crop cost equations and introduce on farm use fodder as feed cost to be distributed among animal productions.

Non-agricultural activities

In the current version of the model the value of non-agricultural activities (other receipts, forestry) is not included but some costs that can not be separated from other costs, are included. The problem was not so important when the model was built in 1985, but now the part of these activities is quite high in some Member States. Table 2.1 presents percentages of 'other receipts' and 'forestry product' in total production.

It might be possible that results would be improved by adding an output category of non-agricultural activities. Coefficients of these categories have to be put to zero for crops and animal proportional costs.

Introduction of expert knowledge or exogenous coefficients

In the application of INRA model some results could be very weak and in few cases coefficients of production could be negative. This problem may occur for different reasons:

- one activity is not representative in the sample. For instance, eggs and pigs are often produced in specialised farms. These farms produce also a small amount of cereals;
- multi-collinearity of regressors. For instance beef and milk;
- specification problems.

Table 2.1 Percentage of output from other receipts and forestry products in the EU, 1997

	Other receipts (%)	Forestry products (%)
Belgium	1.2	0.0
Denmark	3.0	0.0
Germany	10.2	0.7
Greece	0.5	0.0
Spain	0.8	0.0
France	4.0	0.0
Ireland	3.1	0.0
Italy	1.2	0.2
Luxembourg	8.9	0.3
The Netherlands	2.7	0.0
Austria	14.9	6.6
Portugal	3.1	2.7
Finland	2.9	0.0
Sweden	7.7	0.0
United Kingdom	7.0	0.0

One proposed solution consists to introduce exogenous coefficients in the model when estimated costs are non reliable. These coefficients could be given by experts or estimated by other methods (for instance by applying the INRA model to a reduced field of the sample). A study of the coherence of this solution has to be made.

Scale effects

Currently the model assumes the absence of economies of scale and economies of dimension, the output coefficients being the same for all farms regardless of their size. Such an assumption could be realistic for intermediate consumption, but it does pose the problem of such more or less fixed factors as land, capital and family work. In the case of land, some extensification of production is observed as the area increases. Given its virtual fixity, however, family work is not at its long-term optimum level for many farms. The lack of a real link between family work and the physical size of farms is reflected in economies of dimension.

The easiest way to take account of economies of scale is to estimate costs for different class size of farms and recalculate average costs.

Physical based allocation of inputs

In the current model, inputs are allocated based on the value of the output. We can assume that the difference in the prices of outputs explains the difference in product quality. One proposition of improvement consists of an allocation of inputs based on physical output (hectare, number of kilogram produced, animals). This could lead to better results, but we

have to apply the model to small homogeneous areas such as regions. The different units of production will also complicate the structure of the model.

Direct subsidies

Subsidies by product are generated in the current model. Subsidies are treated as negative costs and the coefficients permit to allocate total subsidies among outputs. According to the improvements in collecting direct payments in RICA (table J), it will be possible to introduce exact direct payments by product.

Most of the above mentioned improvements however take a lot of time. It was not possible to realise them within the current project. The most advanced way to judge the results of the INRA model, would be the comparison of the cost of production data of the model with other sources. Because no other data sources with cost of production data exist on EU level, this could only be realised by interviewing experts from the Member States. The results of the interviews are presented in the next chapter.

2.2.5.2 EAA versus SGM

In this section we deal with EAA versus SGM results. Some more general remarks (a.) are followed by some remarks on individual countries (b.).

a. General remarks

Calculations on the value of SGM compared with EAA were made for nine countries. For the following reasons, the analysis could not be as complete as we had originally in mind:

- for some products, no split of costs was made in the detailed SGM calculations. So a comparison of total inputs based on SGM and based on EAA, was not possible;
- for some products, the composition of the outputs was not clear. Sometimes the split between subsidies and output could not be made. For activities that deliver output of different products, the value of the individual outputs was not always clear. Besides that problem some countries, like the Netherlands, also include other output in the SGM calculations. This is output that can not be allocated to an animal or hectare but is realised on the farms. On dairy farms for example, the income from leasing of milk quota to other farms is one of these outputs. This output can not be separated from other outputs;
- sometimes, the SGM or FSS data was not available at a level of detail that is needed to make a good match with the EAA output definitions;
- detailed information about the way the SGM are calculated (see section 2.2.4.2) is sometimes needed to make a good match between EAA and SGM.

Most of these problems could however be solved in co-operation with the representatives of the Member States. We did not have the time (and the authority) however to contact all these representatives. In some countries we knew the persons involved and sometimes persons responsible for SGM calculations were involved in the interviews (see chapter 3). In these cases, we asked some questions but in most cases they did not have the

time to answer us. The analysis of differences between databases is a very time consuming job and our questions did not have the highest priority.

Our analysis also proved to be a very time consuming effort. Beside the reasons already mentioned (data were missing, Member States were not always able to co-operate), this was caused by the following reasons:

- the countries delivered data in a different format;
- some mistakes were made in the spreadsheets with the data;
- a lot of calculations were needed to calculate national SGM;
- problems in coupling EAA and SGM (definitions, allocating costs to different outputs of one activity).

In appendix 8, the total output based on SGM/FSS calculations and EAA are compared for the nine countries. For some products, results are missing because either the products are not produced in that country or data is missing. Large differences do exist for a lot of products. For some products this can be explained by volatile prices, like potatoes and pork. The SGM are based on averages of three years and they might differ a lot with the results of the year 1997. For some others, output definitions could not be completely matched. For most products however, one should expect that output totals would be more or less the same.

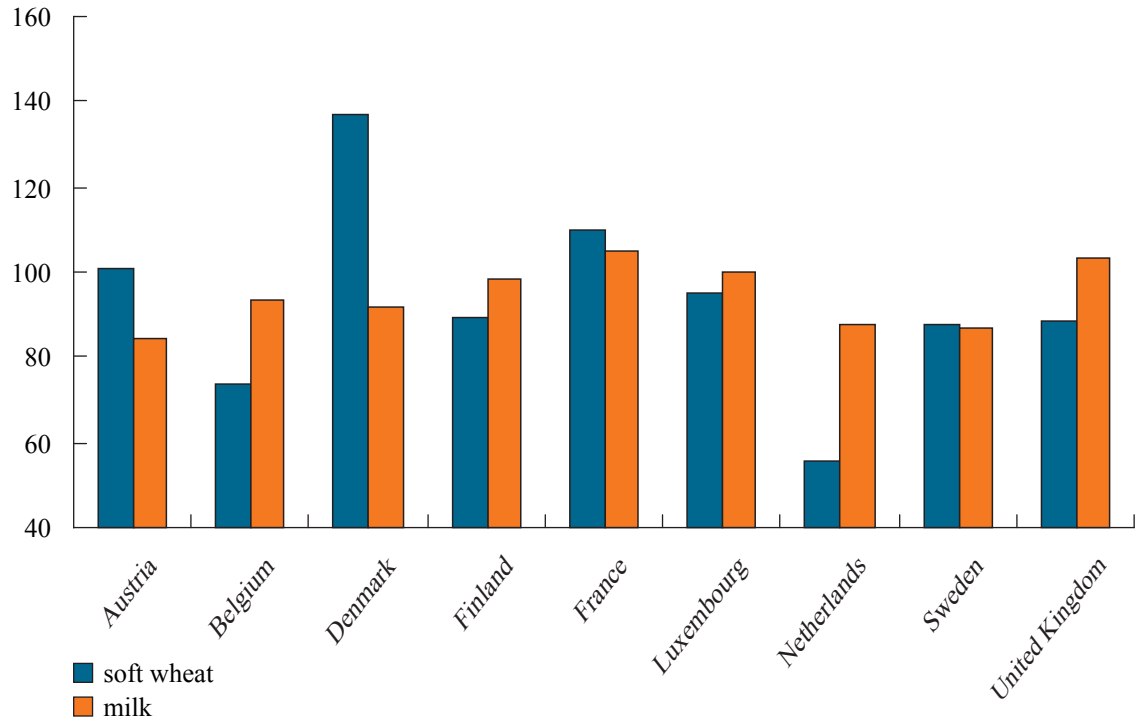


Figure 2.6 Outputs of milk and soft wheat in EAA as a percentage of output based on SGM/FSS calculations in 9 EU countries

In most countries output of soft wheat and milk is smaller in EAA than in SGM/FSS. For wheat this could be explained by the decreasing price. SGM output is based on the average price in the 3 years before 1997. At this time, price were higher than in 1997. In Denmark however wheat output is much higher in EAA.

The high milk production based on SGM/FSS could be explained by the fact that SGM are most of time based on the more advanced specialised farms. The production per cow might be higher on these farms than on the average farm.

For some countries, it was possible to allocate costs to nearly all outputs in EAA. In this case, total inputs in EAA and total outputs based on SGM/FSS could be compared (figure 2.7 and appendix 8). In most countries seed use in EAA is lower than with SGM/FSS. Feed use however is larger in EAA for most countries.

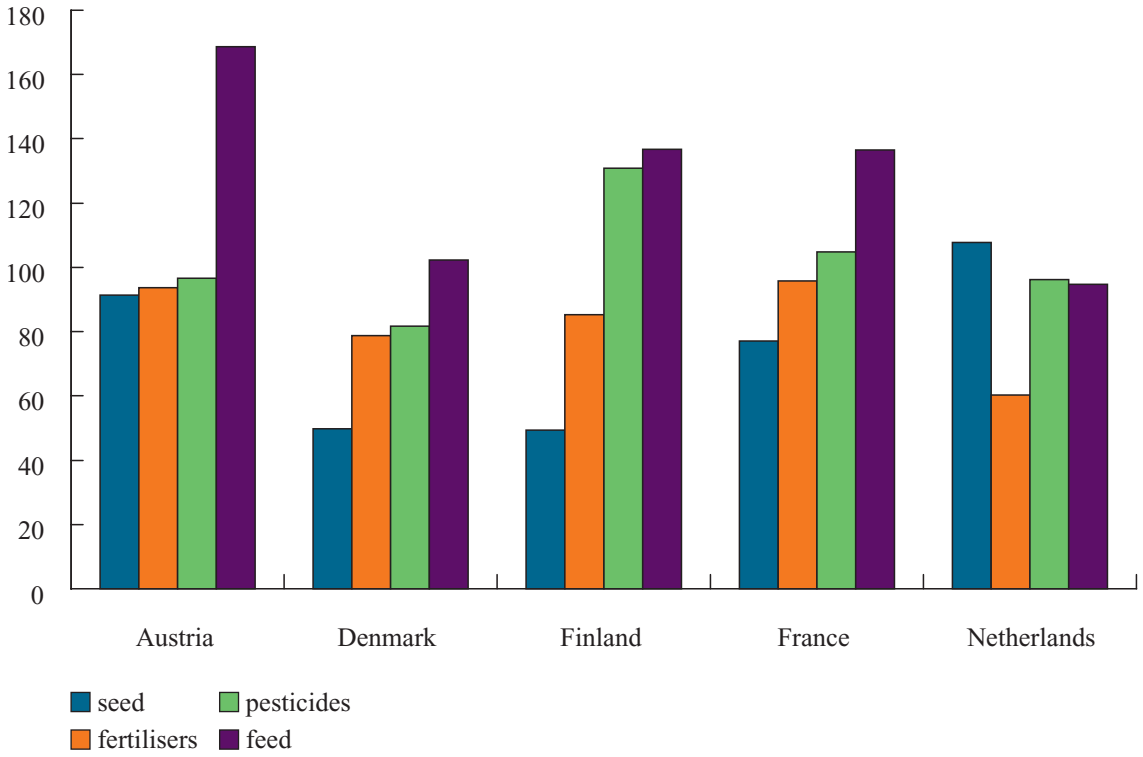


Figure 2.7 Input based on EAA as a percentage of input based on SGM/FSS

b. Remarks per country

It would take too much time to find the reasons for differences for all countries. For some countries, like the Netherlands and Finland, missing data could be assembled. Also the methodology used in SGM calculations and the reasons for differences with EAA could be discussed with the experts involved. The analysis for these countries is presented in box 2.4.

The Netherlands

Output
The total production based on SGM is about 4% higher than in EAA. For individual products however, differences are much higher.

To illustrate the reasons for differences for individual products between EAA and SGM/FSS we studied the reasons for the difference in soft wheat in detail. The total difference is about 80 million euro. Some reasons have to do with the fact that SGM is an average of 3 years:

- price in SGM calculation is higher than price in 1997 (6 million euro);
- yields per ha in SGM calculation are higher than yields in 1997 (20 million euro);
- soft wheat is a combination of winter wheat and spring wheat. In the SGM calculation a higher percentage of winter wheat is assumed than the real percentage. Winter wheat has a higher yield per hectare than spring wheat (1,5 million).

Other reasons have to do with the way the SGM is calculated:

- multiplying the SGM with the number of hectare assumes that on every hectare subsidies are realised. In practice, on a large part, no subsidies are realised. Because the subsidies could not be separated in the SGM calculation, the total subsidies in EAA were subtracted. In fact, more subsidies were included in the SGM calculation (14 million);
- in the SGM calculation also some other output is included (straw and other products): (27 million euro);
- as stated before, the SGM calculation includes VAT 6% (12 million).

A last important reason has to do with the quantity of soft wheat in EAA. EAA estimates that 875 million kg has been produced, while other sources (like FADN and ZPA1) estimate a much higher production (about 1,060 million kg). This is partly caused by the fact that by mistake also the farm use of cereals for feed was subtracted in EAA. Another reason is that sometimes the supply and use of a product are not in balance in the EAA. In this case one of them has to be corrected.

The large difference for potatoes is caused by the low prices in 1997. The strange results for pigs (classical swine fever), equines (most horses are for hobby), poultry/eggs (hatchery eggs are included in eggs instead of poultry) have been described above. The high milk output is caused by a too high milk production per cow in SGM and because the milk that is used at the farm is included in SGM but not in EAA a). The large difference for flowers and plants was also noticed in the comparison with FADN. The reason is not quite clear at this moment.

Input
The use of fertilisers is higher than in EAA. This might be caused by the fact that the SGM of fodder crops (grass for example) is based on large specialised dairy farms. A large area of grass however is used for other purposes and might be used more extensively. It might also caused by the fact that organic fertilisers might not be included in EAA.

The total use of pesticides and seeds is very close to the total in the EAA. The total feeding costs in SGM is also very close to the EAA total. So the large difference in output for some products between EAA and SGM did not result in large differences in input. Most of the differences in output are probably caused by differences in output prices.

Finland

Outputs
When calculating SGM for Finland actual average yields of last three years have been used. Hence the SGM output is very close to the actual output. Since the Union of Rural Centres presents activity based cost calculations at different yield levels, the input specification which results in yield which is closest to the actual yield has been adopted in the SGM calculation. Hence the output in SGM very close to the actual and equal to EAA. There may be small differences between the SGM output and the actual output, however, especially in the case of small volume products.

Inputs

SGM input use specifications are based on data from specialised farms (from recording samples of the Union of the Rural Advisory Centres). Hence one should not expect the input use of each and every input to be the same as the average use of input in the whole country.

In the analysis made it turned out, for example, that the total use of fertilisers is overvalued by more than 15% in SGM compared to EAA. The value of seeds is overvalued as much as 100% in SGM compared to EAA. This is no surprise since the use of purchased seed in the calculation of the Union of the Rural Advisory Centres is based on simple assumptions, not on actual seed use. Pesticide use based on SGM is less than 80% of the EAA total.

The value of 'other variable inputs' in SGM can not be compared to EAA. Part of these costs are included in energy costs in EAA but also some costs might be included in veterinary costs and 'other intermediate costs'. Value of feed is undervalued by nearly 30% in SGM compared to the EAA. This would be not expected since the feed use composition (the relative shares of the different feed stuffs used in feeding) in the data of the Union of the Rural Advisory Centres is based on large farm samples and can be considered rather representative.

One may conclude that there is a lot to be improved in the harmonisation of SGM calculation and EAA in Finland. In fact, the use of each and every input should be close to the value of EAA. However, one should find procedures how to adjust the input use in SGM in order to find the correct total value for each and every input, without abstracting too much from the actual data and agricultural production conditions. If only data from specialised farms is available, arbitrary adjustments in the use of inputs of each product are risky and may result in unrealistic input use specifications. One should find keys how to find the products whose input specifications are adjusted in an appropriate range without risking the realism and connection to the original sample data.

Box 2.4 Analysis of SGM and EAA for the Netherlands and Finland

a) A large difference for 'other crops' existed. Based on this, it was detected that the SGM of horticulture seeds was calculated in the wrong way.

Conclusion

The detailed analysis for some countries and the rougher calculations for other countries made clear that an exercise like this can be done and is useful, for AgrIS as well as to improve the quality and the usefulness of the SGM.

Besides this, agreement has to be reached about some methodological issues of the SGM calculation:

- should all outputs of a farm be included in the SGM calculation?
  Besides the main products, farms will have some other outputs. For cereals for example, also straw may result. By products like straw should be included in the calculation but it is not clear if all countries act like this. Besides should also other outputs that do not have a relation with any output (letting of machines, trade activities, tourism) be included in the calculations?
- average for all hectare/animals or average for commercial farms?
  Some countries state that SGM are not representative for all farms but only for commercial farms. Should SGM be calculated for all farms or only for the commercial farms?
- all products in a category or main products?
  Should the SGM calculation be based on all products included in that category or is it also possible that it is based on (some of) the main product in the category. In the last

case, the SGM might not be representative for the whole category and therefor less useful for improving AgrIS.

- cattle calculations: split into categories?
  For cattle a split should be made into age-classes. It might be difficult to allocate costs between those categories. Sometimes prices of animals at the age that they are changing category, are missing because they are not traded at that moment. Some countries use therefor cost of production calculations while others try to estimate a market price. A common methodology for these problems would help.
- processing of agricultural products
  In some countries a large part of the agricultural product is processed on the farm. In Austria for example most goat milk is used for cheese. Should the output and costs of cheese be included in the SGM calculation or the output of milk?

2.3 Some final remarks

Ambitions to data sets

The data sets available at the European level are an interesting jig-saw. Even for relative insiders like this project team it is an enormous challenge to get information out of a combination of the different data sets. A number of conclusions and recommendation will be presented in chapter 4 and 5, after describing the Member States' view on the usefulness of these data sets compared to their own in AgrIS.

The work on the data sets however also reveals that with common efforts the current data sets could be made even more:

- useful. An example is the SGM that could easily be published, with details, in a database. This would benefit policy (researchers) and would indirectly improve their quality also in their current use;
- transparent. An example is the FADN. If the Member States and DG-AGRI would put a bit more effort in investigating (and publishing) the representativity, the quality and use would be improved;
- consistent. Harmonising data definitions and making differences clear would improve the data quality. The analysis of definitions used in FADN and EAA shows that a lot of differences exist. It would be advisable if Eurostat and DG-AGRI would try to harmonise these definitions. As described before FADN is used a lot for the calculation of the inputs in EAA.

Not a unique issue, other studies

The objectives of AgrIS are useful. We learned during our project that others are working on related issues, concering cost of production on European level.

Eoghan Garvey is working with an international team on a maximum entropy model that allocates inputs to outputs in EU-FADN. In this model, that is built for DG-ENVIR, all kind of a priori data can be included. They included the SGM results as a priori information in their model and included all kind of others (logical) restrictions in the model. All

the way the project, we contacted each other about the progress of our projects. Short before our project was ended, some preliminary results for some products for the Netherlands were sent to us. Because they are still working on the improvement of the model, the results can not be evaluated at this moment.

Graham Brookes from the UK assembled variable costs data for the most important crops in several European countries. Short before the end of our project, he published his book titled European arable crop profit margins 2001/2002.

His method of assembling data deviates in some extent to this project:

- definitions of some costs were not harmonised between the countries;
- results are not always claimed to be representative for a whole country;
- results have been extrapolated to 2001 and the method is not quite clear.

Besides that the results of Brookes were based on 2001 instead of 1997 and are only available for crops and not for all EU countries, while output definitions can sometimes not be harmonised with EAA.

Such studies show that AgrIS fulfils a need. Co-operation with some of these initiatives in future could be useful.

Interests and involvement of Member States

Using the available European data sets, as it has been described in this chapter, seems to be promising, but also has a big disadvantage in terms of transparency to the Member States. Different institutes (or at best different departments in one institute) provide the European data sets in the Member States. They might have difficulty to understand the results that become available from the coupling of databases at European level. SGM are perhaps most easy to understand. The FADN/INRA model needs more explanation as is done in this report. Given the attention paid to that, results of it might be useful to be discussed with experts in the Member States. The sooner the Member States could be involved in these processes on development of data sources and solving the problems of coupling data, the higher their support for the final database.

This makes it attractive to have a look on the position of the Member States, the availability of adequate additional national data sets and the support for the approach followed in this project in the context of the construction of AgrIS. These are some of the topics of the next chapter.

3. Review of Member States

3.1 Organisation and preparation

Mainly to get more information than was available through the data sets discussed in the previous chapter, in each of the Member States interviews were organised with national experts. The invitations to the national experts were in first instance sent by Eurostat to the members of the Working Party on EAA. By this the project team got a lot of positive reactions. The non-responding countries were asked in a later stage by direct contacts (telephone and or email) with the experts. In our contacts with the EAA representatives we emphasised that it would be useful if persons with knowledge of FADN or other micro databases would attend the interview.

The list of experts in the Member States (appendix 14) shows that many of them are active in statistical institutes, economic research institutes on agriculture (and or food) or in the Ministry of Agriculture. For a part the interviewed persons are experts on FADN, others on the EAA and or price statistics.

Members of the project team carried out the interviews. They visited the experts in their office. In the allocation of the interviews to the members of the project team the knowledge of the regional agricultural circumstances (for instance in the Mediterranean and Nordic countries) and the language were important criteria. The interviews were organised in the months October, November and December 2001, in a stage of the project that the first results of data processing, mainly with the FADN/INRA model, were available for discussion. This was important to show more in detail what could be expected from the project and what were the problems, at least the lacks in the desired data as well as the deviations in results with EAA statistics, to settle.

3.2 Questions to the Member States

For the interviews the project team prepared a common questionnaire (appendix 12). This was sent to the experts some weeks before the interview, so they could prepare their reaction. The questions raised in the questionnaire and the interviews concern mainly:

General information:

- a short description on the organisation of the work on agricultural statistics in the Member States (responsibilities, collection of data by the institutes or by using data from others);
- the attitude on the project: do the national experts find it of interest to allocate inputs to output or to activities? (for research, modelling, etc);
- the possibilities to provide to Eurostat additional (national) data in the context of the desired database, on a regular base.

Availability of relevant data and statistics:

- detailed Standard Gross Margins (SGM), at least for '1996', and what is the quality of SGM in the context of the project;
- national data sources (cost price calculations etc), suggestions for co-operation with national institutes to receive results to be used in the context of the project and AgrIS;
- detailed EAA data on output and inputs available (as is presented in the EAA+ list): what kinds of problems arise and what are the restrictions to assemble specific data? Related to this also it is useful to know for which data FADN is the source.

Discussion on (first) results of FADN/INRA model compared with EAA:

- are the results (in general) satisfying for the member country?
- on which data exist large deviations?
- possible reasons for these deviations?
- can regrouping data solve it, for instance?

So Member States were visited for different reasons; in a nutshell, as far as data are concerned:

a. to check the availability of additional data for the EAA (EAA+);
b. to check the quality of the results based on the two databases and to get information to have an explanation for strange, deviating results;
c. to check the availability of additional data on cost of production in the Member States.

Related to this the questionnaire and the interview had mainly the aim to draw conclusions on the following points:

- which data should be used for this country at this moment and which improvements could be made in the future;
- description of the national data source (if used).

A potential by-product of the visit in the Member States was to increase the support for AgrIS, and the fact that it makes sense to check consistency of databases at the European level.

3.3 Results from interviews

Experts in all Member States gave the opportunity to be interviewed. In some cases however they hesitated in first instance, mainly because other activities or projects had more priority in their work in that period. For some experts the last months of the year are very busy to prepare the (provisional) results of the sector account (EAA), prices etc on request of national authorities and Eurostat. In general however the interviewed experts showed an interest in the project and took at least some hours or half a day to respond the questions and for the discussion. They were open to deal with the questions raised and most of them are inclined to co-operate in future to develop the AgrIS database.

From each interview a separate report is made available to Eurostat and the interviewed experts. Some of them reacted in the weeks after the interview to adjust the draft report and or to provide additional information on the situation in their country, data available etc.

Figure 3.1 shows the results of the interviews per member country in qualitative terms, on the following issues:

1. Attitude of the experts on the aim to link input data to outputs or activities.
2. The way of organisation of collecting FADN data; is it done by one of the institutes interviewed and how many organisations are responsible for FADN, SGM and EAA.
3. Is EAA assembled with FADN data?
4. Is SGM assembled with FADN data?
5. Are SGM detailed data available?
6. Are the desired data on output and input values (EAA+) available?
7. Are extra data available, now or in future?
8. Are the FADN/INRA model results satisfying?

On these issues the interviews and the information received from the Member States in a later stage, result in the following remarks and clarifications:

1. Attitude

Most of the Member States have a positive or at least a rather positive attitude on the aim of the project and AgrIS. So this gives, at least in principle, opportunities to go ahead in future. Some countries however (Austria, Germany, Greece) are at least hesitating. In fact they are for several reasons (budget, organisation or lack of adequate data sources) not in the position to deliver more statistical information than what is required and requested in handbooks by Eurostat and DG Agri or they are questioning the aim of the project. Also the UK is hesitating for budget reasons. Some countries (Ireland for instance) remark that EAA is not the right source for the allocation of inputs; it could be calculated from FADN. Some countries, as Belgium and Denmark, remark the project is useful in improving the quality of data, but the allocation of inputs to outputs or activities should lead to correct results, which, however, is difficult in the case of small products. In Italy the person in charge of EAA believes that for the goal on National Accounts there is no need for further desegregated macro data on agriculture and Spanish experts are of the opinion that FADN nor SGM data are adequate to allocate input costs to activities. In some cases the interviewers had the impression that the AgrIS' aim of 'checking the consistency' of different data sources is perceived as not so welcome. This might be caused by three reasons. First of all, they might have the impression that their authority and expertise is questioned. Why is it needed to check our data? Secondly they might not like it that 'outsiders' tell them that they should (better) co-operate with other institutions in their country. There might be institutional barriers to co-operate with other institutions in the country and besides that, they might have the opinion that the co-operation between different national data providers is a national business. Thirdly, the checking of the consistency between data sources is a very time consuming and difficult task. Some institutes do not have the resources for this kind of tasks.

2. Organisation

Most Member States have a central organisation to collect the (micro) FADN data on farms. In some it is decentralised, for instance in the UK: England and Wales, Scotland and Northern Ireland. In some countries however (for instance: France, Germany, Spain) data are collected by private organisations. In most Member States different institutes are responsible for the data on FADN and EAA; some times the responsibility on EAA is shared by the agricultural research institute and the central office for statistics (for instance in the Netherlands, Sweden). In some countries a split is made between input data (research institute) and output data (statistical office). These kinds of division of responsibility can handicap the desired connection between input and output data, depending on the communication and co-operation between institutes involved. On the other hand it can improve data quality, if more experts with different knowledge are involved and discuss the (draft) results.

3. FADN and EAA

About half of the Member States have EAA data derived, at least for a part, from FADN data (mainly on inputs) and or they evaluate or verify the EAA with FADN results. In practice because of the time lag between FADN (provisional and definitive results), at least the first results on EAA - as they are requested by Eurostat at the end of the calendar year - have to be produced on the base of macro-information (development of prices and volumes, estimates on harvest, values of slaughtered animals etc in that year). Besides that, in a number of countries the FADN is not produced on a calendar year base (with differences per sector), which means that (difficult) corrections have to be made for this purpose. In Spain not FADN but SGM data are used to construct EAA statistics. In Ireland, the FADN is hardly used for the EAA. The institute responsible for FADN is also located in another part of the country. In this situation, the harmonisation of FADN and EAA is less likely.

4. FADN and SGM

SGM results in about a half of the Member States are derived from FADN data, at least for the most important products. In most cases for minor products, for which FADN is not representative, in these countries additional technical information of other sources is used. In some of these countries, like Ireland, Denmark, the Netherlands and others, the variable costs are allocated in the accounting process to the products on the farm itself. So the farmer tells the FADN for which products for example the pesticides are used.

In other countries - Austria, Germany, Finland, Portugal and Spain for instance - however (all) SGM data are produced by specialised institutes using economic and technical information and expert opinions, so from other sources than FADN. In fact then, as in Portugal, two data networks exist, both with some thousands of farms and besides that (for crops) an additional survey (MBM) was organised. Spain has a comparable situation.

5. Detailed SGM

Nine countries send detailed SGM data to Eurostat and gave permission for the use of the data in our project.

In some other countries that did not deliver detailed SGM, like Germany however very specified SGM exist, mainly as a tool to advise and inform farmers (management).

They are available per region and related to different production circumstances, but they are not, at least not without treatment, fit to be used in the project. In some countries (like the Netherlands and the UK) farm management pocket books with detailed standard gross margin data exist, in which however the methodology is not equal to that of Eurostat (Classex 44). Such data is used by researchers for policy research and by statisticians for SGM on minor activities.

Some of the (S)GM data sets are available on the internet (e.g. KTBL in Germany) and recently a commercial publisher started to publish a European data set (Graham Brookes, 2001), which suggests that there is also an interest in the market for SGM data.

6. EAA+

Most countries have comments on the EAA+ list. The main problem, practically for all countries, is to make a distinction into activities for the animals. For cattle for example it is very difficult or impossible to make a split of the total beef into beef from dairy cows and beef from other type of animals. The total beef production in the EAA is based on the number of slaughterings and this data is not split into activities.

Another major problem is the value of forage (grass a.o.), for instance in France, Ireland and UK, while Denmark and Finland remark the production of grass seed should be added. Other problems arise on ornamental products; in many countries the value is not split in sub-categories (flowers, pot plants etc), as well in some countries on industrial crops (oil seeds etc), potatoes (no split in ware potatoes and seed potatoes, but in Portugal a split between irrigated and non-irrigated potatoes) and specific cereals, as triticale. In some countries no data for wine exist, but on grapes (Portugal) or most (Germany). In the case of Portugal data on wine and olive oil also represent the value of production outside farms. On vegetables and fruits Italy has some suggestions to add a number of products on the one hand and to delete some others. Italy and Portugal make a split between vegetables (tomatoes a.o.) for fresh consumption and for processing. Some countries have specific data on mushrooms. Several inputs in EAA (as a.o. pesticides) are not split in more detailed subposts (insecticides etc) in most countries. The same problem arises on fertiliser (nitrate, phosphate etc), while most countries have no (financial) information on the use of manure. On feed costs some Member States (France, Ireland, UK for instance) have problems to detail the costs of feed, often forage, produced on the same farm. On energy besides the fact that not in all countries a split is made between the energy for motor fuelling and for heating, the use for household and farm use is sometimes difficult to make.

On other inputs concerned (veterinary expenses, materials, buildings, services) information is rather global; on water some countries have only some data on the volume of water used, but not on what is paid.

More generally the desired information on input costs allocated per activity is scarce, mostly it is not available on a regular base (see also point 7). As far as information is available in a country, mostly it is only for some products, as for instance in the Netherlands on pig meat and milk. In some countries it can, for a part, be derived from (detailed) SGM data and or additional information. For instance in Portugal MBM (Model de base micro-economics) data provide for crops additional information on this, while in Spain a comparable situation exists.

7. Extra data

Most Member States reacted in a positive way on the questions to deliver more data. This means that most countries are willing to invest in collecting more data or at least processing the existing data in such a manner that they can be used by Eurostat to compose the database. On the other hand some countries (Austria, Germany) are not inclined to do more in this direction, mainly due to budget restrictions. But also in these countries as well as in most other countries some institutes or organisations (f.i. producer associations in Portugal) have additional data on cost prices, however only for some products. For instance in Italy, yearly (based on specialised farms) data on cost prices of milk are available. In other countries however not always on a regular base. Sweden has 'synthetic' calculations of Commodity Boards on the costs of meat, milk, cereals, but on (own) other procedures, while in Finland Rural Advisory Centres have (based on activity based cost calculation) additional data on the use of some inputs (including buildings and machinery) per activity, it has also production cost calculations for milk, cereals and pig meat.

8. Results of the FADN/INRA model

The results of the FADN/INRA model as they were presented at the time of the interviews in 2001, were not satisfying for the experts in many countries. Experts in Denmark make the remark that they frequently compare FADN, EAA and SGM results and that the differences are small, while what is presented (the initial results of comparisons) now shows remarkable differences. For some products, often the most important for the country concerned (f.i. milk), however the relation between inputs and output values seems reasonable. Most problems occur on 'smaller' products, for which it is difficult to have a reliable and representative FADN (not enough farms with the specific crop). In some countries however, as the Netherlands, the model produced deviations with EAA also for important products (pigs, poultry, eggs, potatoes, flowers, fruit etc). For Belgium also on some animal products problems arise due to contract farming. A more general problem in the view of the experts is, at least in the Southern countries, Germany and Austria, that many farms have a size under the threshold for FADN, even if it is 2 ESU. In for instance Portugal the remark is made that SGM data would better serve to the INRA model than FADN data. On the other hand SGM provide no information on fixed costs (for example intermediate costs and maintenance of buildings and machinery).

3.4 Conclusions

Each member country has its own characteristics on availability of data, depending on the organisation of agricultural statistics in the country. The overall impression is that in fact in all countries many micro- and macro-data are available, but often not in a (standard) way appropriate for the project and for AgrIS. It is obvious that on the base of the available data (FADN, SGM, EAA as well as others) in all countries improvements can be made in the consistency of agricultural data, and that important contributions to AgrIS could be made if resources were available; but that there is a need to arrive at a model of organisation or a method of work that is adequate for this goal. Sensitivities on checking the work of Member States by Eurostat should be taken into account.

Many interviewed persons in the Member States have a positive attitude towards an improvement in working methods; they are - for instance - aware of the fact that more can be done on the allocation of inputs per product or activity. In some countries data are already available on this, but in most cases only for some (major) products and not for each year and or only for specialised farms.

The response of the countries on the question to deliver additional data for the project and the AgrIS database was rather various on the point what could be delivered. The risk of these various answers is that it results in a large variety of additional data, based on different definitions and samples of farms, for a part on a regular base and for other data ad hoc etc.

On the activities (or outputs) to be taken into account in a database, it is difficult to arrive at a common list, given the wide diversity in production circumstances and structures in the farm sector. Answers on these indicate on the one hand the necessity to simplify the suggested EAA+ on some points (bovine animals, forage) while on the other hand Member States want to have room to implement specific data on for instance some specific fruits and vegetables. Perhaps one could start with a flexible list on this point.

On the inputs to be included in the database a more or less comparable situation comes up on the base of the interviews. Many countries have problems to make a (more) detailed split on pesticides, fertilisers, feed and or energy. One could start by asking a minimum total input data of each of these categories of inputs per product or activity. Member States with more detailed information can provide this additionally on a voluntary base. It will be difficult to get precise data on costs of machinery, buildings and services per activity or product. It seems important to discuss the procedure to allocate these costs more in detail with the Member States.

Nevertheless there is a positive approach of most of the national experts. This offers a good base to find common solutions on several questions. Many experts gave the impression that some data on cost of production are available in their country. An important condition for the deliverance of such data is that it is clear by instructions for them what is desired.

In some countries that are at the moment hesitating on the aim of the project and on the AgrIS database, it is important to find out what can help them to overcome the doubts. It might be important to show them what can be done with the data model working adequate at least for some (other comparable) countries and what was necessary to arrive at that situation. It is necessary to invest some time in discussing the objectives of AgrIS with the Member States more in detail; questions to be responded are for instance: what are the reasons, what will be done with it, how can we benefit in the Member State, what is already available and what has to be added minimally. Assembling of data and harmonisation with EAA will take a lot of time for the Member States. It is very important that the Member States are convinced that they could also profit from this extra work. Eurostat could motivate members by emphasising that the data could also be very useful for national purposes and that this data could improve model calculations and decisions.

Country	Attitude on project and AgrIS	Do interviewed institutes collect alone ?	Is FADN used to make EAA?	Is FADN used to calculate SGM?	Are detailed SGM available in project a)	Reaction to EAA+?	Is Extra national data for AgrIS available?	Quality of FADN/INRA results
Austria	0/-	Collecting; different org.	Yes, partly (inputs)	No +	0/-	0/+	+, -, hesitating	
Belgium	+,0	Collecting, one org. b)	Yes	Yes+	0/-	+(+)	+/-	
Denmark	++	Collecting, different org.	Yes, partly	Yes+	+/+	++	-, negative	
Finland	+	Collecting, different org.	Yes	No +	0/-	+(+)	+,- (on small products)	
France	+/0	No collecting, different org.	Yes, at least for most inputs	Yes+	0/-	+(+)	+	
Germany	0/-	No collecting, in one org.	No	No No	0/-	0/+	+, -, hesitating	
Greece	0/-	No collecting, different org.	No	No No	No comment	0	-, hesitating	
Italy	+/-	Collecting, different	No	Yes No	+, -	+, for some products	=/-, hesitating	
Ireland	+/-	Collecting, different org.	No, only for specific costs	Yes, for most products	No	+/-	+(+)	+/-
Luxembourg	+	Collecting, different org.	Yes, except for feedingstuffs and energy	Yes+	0/-	0	No comment	
The Netherlands	+	Collecting, different	Yes, partly	Yes, on most products	+	+	++	+,-
Portugal	+	Collecting, different org.	No	No?	No	+/-	++	Some comments
Spain	+	No collecting, different org.	No?	No	No	Yes per activity	+(+, hesitating) not available	-, not satisfied
Sweden	+	Collecting, different org.		Yes	No	+	Yes, some remarks	+ +, hesitating
UK	+/-	Collecting, decentra-lised organisation		Yes, for a part	Yes, for a part	+	0/-	+ +/-

Figure 3.1 Information per EU member Country on Agricultural Statistics and availability of extra data
a) For some countries (like Italy) they have been delivered to Eurostat however; b) In future split between Flanders and Wallonie.

4. Database for AgrIS

4.1 Conclusions

Make use of FADN/INRA and SGM

The objective of this project was to develop the AgrIS database on inputs and their allocation to activities, keeping the concept of AgrIS in tact. During the project, a number of hard decisions was needed, to make these tasks feasible. As discussed in chapter 2, linking different data sets and using an econometric model has a certain potential, but easily leads to a data set that contains estimation errors and that is filled with data that is hard to understand by statisticians from the Member States. That could degrade the AgrIS concept and would certainly not support the process of checking the consistency of different data sets in consultation with the Member States (as described in chapter 1).

The interviews in the Member States (chapter 3) revealed that in a number of cases attractive alternative data sets (and know how) are available. Such databases could be used in AgrIS to improve input data and to link these to activities or output. However in a lot of countries such data do not exist, or resources are lacking to co-operate with Eurostat on these priorities.

1997 as a starting point

Confronted with these realities, we had to give up our intention to construct a database from 1973 onwards. In stead we concentrated on 1997, the most recent year available in all data sets. Together with Eurostat we decided to use the FADN/INRA approach as a basis. This is the only data set available for all Member States. In addition to that data set, the data from the SGM-1996 (if available) or national data were added. After a comparison we then created the most likely matrix of inputs per category of output. This data set can be a bases for further discussion in the AgrIS framework to check the consistency of inputs linked to output in the EAA. More details on how the database has been constructed as well as recommendations for further development are given in the next section. If the Member States agree to fill AgrIS in this way with FADN/INRA data, comparable data can be generated for the period 1973 onwards (for newer Member States starting at a later date).

It should be realised that this database is a starting point for further consultations with the Member States, as currently carried out on outputs (see figures 1.1 to 1.3 in chapter 1), and not a database to make publicly available at the moment.

EAA+ still to be made adaptable

This result implies that we have to conclude that the realisation of an EAA+ is not feasible at the moment. Although in some Member States there is support to make some changes in the EAA list of inputs and outputs it has be concluded that the overall support is too weak. The interviews also revealed that EAA data on outputs are often collected in such a way (e.g. questionnaires on slaughterings) that information on activities is not available, and that also inputs cannot be linked to activities. A switch from output to activities, although attractive from the perspective of the user and from a micropoint of view in data collection, is not feasible.

Role of Member States in the process

In the AgrIS concept it would be most attractive if the Member States improved the data set and its consistency. There are a number of reasons for this:

- self control is more motivating and often cheaper than auditing by a third party;
- a check carried out at the end of the pipeline causes more delays and costs then a check at the beginning or somewhere halfway the pipeline;
- it helps Member States to improve quality in their own statistics and provides opportunities for cost savings (e.g. using FADN for the EAA and national accounts);
- resources in Eurostat can be used to concentrate on common methodology in stead of checking.

Besides this specific task, Member States are important to support the development of the AgrIS database. An important element in this is that experts in the Member States have the notion to be co-responsible for the data in AgrIS. With this in mind in chapter 5 some suggestions are made to improve their input in AgrIS.

4.2 Construction of the database for AgrIS

The interviews showed that the first application of the FADN/INRA model did not always lead to satisfactory results. The SGM deliver only details on variable inputs and are not available for all countries and products yet. The interviews also showed that most Member States have some additional data. Additional data however in most cases are not available in an adequate form for the purpose of the project. In nearly all cases it will take a serious effort from the Member States to harmonise these data with EAA (differences in definitions etc). Therefore we decided to use all three sources. For each country, we made an Excel spreadsheet with 4 sheets. In the first sheet, the results of the model are presented. In the second sheet, the results of the detailed SGM are included. The SGM are available for nine countries (Austria, Belgium, Denmark, Finland, France, Ireland, Luxembourg, the Netherlands, UK). In the third sheet, data from national sources are included. Some national data is available for Ireland, UK, the Netherlands, Denmark and Italy. In the fourth sheet, the best available data from all 3 sources is included. If national data is available, this is used in the fourth sheet. In absence of national data, SGM data is used. If both na-

tional data and SGM are lacking, the FADN/INRA model results are used in the fourth sheet.

For the construction of the second (SGM) and third sheet (national data) some assumptions had to be made. These assumptions and the exact calculations have been documented in the (formulas of the) spreadsheet however. Besides a separate document describing the SGM calculations, is handed over to Eurostat. If Eurostat does not agree with the assumptions in the SGM calculations, they could easily correct it because basic data have also been made available. The assumptions could also be checked in the member states and sometimes they might be replaced by extra data available on member state level.

In the following sections we explain what kind of decisions are taken to arrive at the data in the database, as it is constructed now.

4.2.1 FADN/INRA model

In the first sheet we included the results of the FADN/INRA model. It was decided to concentrate on the link between outputs and inputs, by estimating the (aggregate) inputs, given the level of outputs in the EAA. This means that the FADN/INRA model was used to estimate cost coefficients per € 1,000 outputs. The FADN data could also have been used to generate an aggregated level of outputs or inputs as an independent estimation but Eurostat preferred coefficients per € 1,000 EAA output in the database. Besides FADN represents only a small percentage of total output for some products (for example poultry).

One of the problems with the results of the INRA model was in first instance the negative value of some input costs. It could be corrected in the model itself by introducing a constraint but this influenced the other coefficients in the model very badly. Therefor we made the correction outside the model. To keep the methodology transparent and simple all negative values were set to zero and the negative values were distributed proportionally over all other costs.

For example:	Fertilisers	0,1
	Pesticides	-0,01
	Seed	0,2
After correction:		
	Fertilisers	0,1 –0,01* (0,1/(0,1+0,2)
	Pesticides	0
	Seed	0,2 – 0,01*(0,2/0,1+0,2)

In this way the total cost of production remains the same, only the distribution of costs changes. This is important because in this way the model assumption that the output and costs of all products of a farm result in the income, remains valid.

At the end of this procedure in which some corrections have been made, we have to conclude that the results of the confrontation of the model are much better than in first instance, but at the same time we have to admit that further improvements have to be made (see section 5.2.2).

4.2.2 SGM

Nine Member States have managed so far to produce the requested detailed information on SGM for '1996' (or have not given permission to use it for our project). Chapter 2 showed that large differences might appear in the output totals based on SGM and the output totals in EAA. In the end, we should find cost coefficients per 1,000 euro of EAA output. So if a difference in output exists, we should decide if the costs in the SGM calculation should be divided by the output total of EAA or of SGM/FSS. There might be three reasons for a difference in output:

- difference in price;
- difference in output per hectare/animal;
- differences in definitions.

If the difference in output is caused by a difference in price of the output or the output per hectare/animal, the total costs based on SGM should be divided by the total output in EAA. One can expect that a difference in output prices or output per hectare does not influence the use of inputs. If the largest part of the difference is caused by a difference in scope/definition, it might be better to divide the SGM input total by the SGM output total. Without knowledge of the reason for the difference, it is very hard to decide what to do.

As a standard, we divided by the EAA total because our goal is to divide EAA inputs to EAA outputs, but if large differences occurred, we assumed that a difference in definition existed and we divided by the SGM total. Because this decision was rather subjective, we added in the spreadsheet a column which indicates through which total the SGM costs were divided. Besides, we will also deliver the basic data to Eurostat. Eurostat could always correct the data.

As already indicated in chapter 2, the feed cost coefficient based on SGM includes both purchased feed as feed that is produced on the farm. The feed cost coefficient based on FADN includes only purchased feed. Therefor those two coefficient can not be compared.

In chapter 5 some suggestions to improve the results are presented. Because of time limits within the project and the need of co-operation of the Member States, these improvements have not yet been made.

4.2.3 National data

For five countries national sources are included in the database. Because the treatment of this sources is very different from country to country, we will discuss them by country. For some countries we did not manage to get national data although interesting data exist in this country. In this case, we will describe the available database and how the data could be added in the future (appendix 13).

Denmark

Use of inputs and production costs per activity are calculated using FADN data (http://www.sjfi.dk/ go to 'Data og Statistik' and Serie B). These activity based calculations

are calculated for a number of main production activities. Since almost all output of major agricultural products come from large specialised farms in Denmark the activity based cost data is likely to be rather representative. However, calculations for a number of small volume products (horticulture etc) are less representative which could make calculations somewhat inconsistent with EAA. The data are based on accounts of full time farms, i.e. farms on which the annual standard labour input is at least 1,665 man hours.

Production costs of barley, for example, have been calculated not only at specialised grain farms, but also at pig farms, farms of different sizes, and farms at different regions. The activity based cost calculations are also available for hens, broilers and fur animals. The calculations, which are available from years 1991-1999, are documented in SJFI publication nr. 84, Serie B. One can say that Denmark is rather advanced in terms of availability of activity based cost calculations (at least in the case of main activities). The Danish activity based cost calculations are relatively detailed and are already at an easily accessible form via internet. In some other countries, a lot of effort is needed before such calculations at the same level of detail can be made available.

The input definitions in the Danish data are, in most cases, equal to the definitions of the INRA model in the database (this is no surprise since FADN data is used in both). There might be some small differences in the definition of 'other goods and services'.

Ireland

For Ireland, we used data from the National FADN. The data was sent to us by the person responsible for FADN in Ireland, Liam Connoly (Teagasc). We got only data for the most important products in Ireland but based on activities (dairy farming, keeping of sheeps) instead of outputs. Besides the data was based on 2000 instead of 1997. The definitions of inputs were harmonised but the split of costs was not as detailed as we wanted you (although this split is available in Irisch FADN). Energy, maintenance of buildings and materials and other goods and services are therefor integrated in one category and a separate category 'other variable costs' for animals is created that has the same definition as the other variable costs in SGM. We did not have time to ask for new corrected data within the project.

The data has been transformed to the needed format for the database as good as possible. More details about the exact assumptions are described in a separate document. If Eurostat wants to use this data in AgrIS, we would advise to ask for new more detailed data for 1997. The harmonisation with EAA should be realised in close co-operation with the experts from Teagasc.

Italy

Data offered for Agris is coming directly from the National FADN database of INEA. Data from 16,279 farms, surveyed in 1997, was used to compute intermediate consumption per all the 29 outputs chosen for the project.

We first operated at single farm level. Several cost items were already available per output, typically crop cultivation costs (seeds, fertilisers, plant protection), while some others had to be allocated like energy, maintenance of both building and machinery. The latter

were allocated to single outputs via the percentage of specific output value of production (net of subsidies) on total farm value of production (net of subsidies).

In case of intra-unit consumption of crops for animal feeding, appropriate manipulation was required to make data comparable to the estimated results generated by INRA model. According to the INRA model specification costs for cultivating crops used as animal feeding appear indirectly among the costs of production of animal. This required an extra manipulation of data since part of the crop can be intra-unit consumed and part sold outside the farm. Thus, in order to distribute the cultivation costs to the part of the crop sold and that used as feed within the farm, the criterion of percentage of value of production (i.e. value of production of intra-unit used crop/value of production of crop) was used.

Once intermediate consumption was computed per each output at single farm level, the aggregate costs had to be calculated. To do this, each single farm was weighted according to the farm type distribution of 1997 FSS universe. This means that the weight used was referred to the farm as a whole irrespectively of the different outputs one single farm produced. In other words, if a farm had produced two outputs, to weight the intermediate consumption of those two products the same weighting factor was used.

Some caution is needed in using that data given the amount of manipulation and the weighting system, still it is a valuable source of information, thanks to the desegregated nature of original farm data.

The Netherlands

In the Netherlands, we used the standard added value ('sbe') calculations. In the Netherlands calculations are made by the LEI of average added value in the same way as the SGM calculations. They are based on the national FADN. The split into outputs is exactly the same as for SGM so only a few problems occur. It is possible to get exactly the input definitions used in this project but for most products, this will take a considerable amount of time because this split is not yet made. Besides it has recently been decided not to continue these calculations. The last calculation is SBE 1997 based on the FADN results in the years before 1997. Because the analysis would be very time consuming and the SBE calculations will not be continued in the future, calculations have been made for only a few products. We included some important product for which we knew that INRA model calculations were not very accurate. If by-products resulted (for example straw), we assumed that (costs/output) of both outputs is the same.

United Kingdom

For the UK, we used the Farm Management Pocket book 1998 (28th edition). This book is published every year, recently the 32nd edition is published. This source is however not without problems. First of all output definitions differ. Most of the time, the pocket book had data available on a more detailed level. Because not always data were available to weight the detailed outputs, a weighted average could not be calculated. Therefor we only could make calculations for a few products (soft wheat, barley, protein crops, sugar beet and milk). Only variable costs are allocated to output in this book. A lot of information is also included about the cost of machines, services and labour but these costs are not allo-

cated to products. It might be possible that the author could also make this kind of calculations. If Eurostat would like to use this source, it should be realised in close co-operation with the author of the book and the people involved in EAA in the UK.

4.2.4 Maintenance and quality

It is clear from the analysis that the main sources for the AgrIS database in future are FADN and SGM data. Nevertheless additional national data might be helpful to analyse differences in results and to arrive at improvements of the database and EAA. In first instance however it is important to arrive at a situation that the existing databases (EAA, SGM and FADN) are made fit to serve the AgrIS database as much as possible. For this purpose the next chapter includes some suggestions for improvements.

It would be preferable first to improve the match between the FADN and SGM/FSS on the one side and EAA on the other side before data of other years is assembled. If Eurostat would like to have data of more recent years on the short term, the INRA model could also be used however. At this moment (june 2002), the FADN data of 1998 and 1999 is also available and could be used to fill the database for these years. Besides the SGM of 1998 is available. It would be recommended to ask for a detailed calculation of the SGM 1998 too. The SGM 1996 was very useful for this project and the detailed calculation of 1998 could be even more useful if the recommendations that are described in chapter 2 and 5 of this report, are followed and all 15 countries deliver the detailed calculations. The member states that have delivered national data could also be asked to deliver data for more recent years. For Denmark and the UK publicly available data is used so it would be simple to use this data for more recent years. For Italy and the Netherlands, the project members Antonella De Cicco and Koen Boone could be asked again. For Ireland, the FADN representative Liam Connoly (lconnolly@athenry.teagasc.ie) could be asked again. As described before other countries have also valuable national data but were not able to deliver these data within the time limits of this project. It would be useful to contact them again.

5. Recommendations

Introduction

In section 5.1.1 some recommendations for the EAA are presented. The SGM are a data source which seems to be under-used. The data set is available at a relatively close distance from the EAA within Eurostat. It is a traditional 'bridge' between Eurostat and the FADN and - at least in theory - harmonised. We therefor make some recommendations to improve this data set and link it in AgrIS to the EAA in section 5.1.2.

Recommendations on the use of the FADN in AgrIS, and the way the FADN and econometric models should improve to make it more useful for AgrIS is the topic of section 5.2. In section 5.3 some recommendations for the use of national data are described and 5.4 summarises the recommendations.

5.1 To improve Eurostat statistics

5.1.1 EAA

To improve Eurostat statistics in the desired direction for the AgrIS database, more attention should be paid to build up support for the AgrIS database and especially the objective to check the consistency of Member States' data. This could release useful co-operation by and more resources from the Member States. In some cases persons, interviewed experts and others, in Member States are now not too enthusiastic on this, as it questions their authority and quality and involves extra work.

A potential contribution to this is to focus more on knowledge management than on checking the consistency. Activities in which knowledge on the use of additional data sources (like FADN or administrative data sets) in the EAA are exchanged can be attractive for statisticians in the Member States. The current inventory project of Eurostat that asks member states to make an inventory of the way their data EAA is assembled, is a good start for this. Adding documentation and a common definition of 'best practices' in handbooks are also attractive activities.

A special point could be the exchange in experiences between those Member States that use at a national level the FADN as an input to the EAA or are interested in this. It would be interesting to discuss how Member States cope with common issues like differences in definitions, VAT, accounting and calendar years, valuation methods, aggregation for non-represented farms etc.

An important point in general is the need to have a clear documentation (definitions) of the data on products (activities) as well as on inputs (costs), for instance on the level of prices (farm or market level) and the quantities.

This way of working with the Member States includes the necessity to discuss the EAA+ list in the coming period. This list might be used as an option to collect additional data from Member States. Countries which are in the position to deliver, on a voluntary base, the desired additional, more detailed data could be rewarded.

In this context the differences between the concepts of 'output' and 'activity' (and 'production method') have to be defined well (see figure 2.1). It has to be made clear that they can not be used interchangeably in databases unless clear links between all outputs and all activities are assembled for every year.

Chapter 2 concludes that for some products (output of activities) as well as for some inputs it is desirable to make some changes in the definitions and allocation to be in line with the EAA+ lists (forage, vegetables, fruits, wine, olive oil, cattle, ornamental crops, forage, fertilisers, pesticides, water). It is clear that this requires the co-operation of Member States. Improvements in the data collection and or presentation (if data are already available but not presented to Eurostat) can be made, but need some efforts of the Member States in the coming years.

To explore the data on inputs and activities (combinations of them) in following years, it is useful to analyse the development of prices and volumes separately, with the existing indices or the indices of prices of products and inputs (in the coming years for most countries in euro). It is useful to compare the data of countries.

5.1.2 SGM

This report shows that the use of national data for Eurostat is not without problems for all countries. There might be large differences in scope and definition and the harmonisation with EU definitions sometimes takes a lot of (and sometimes too much) time for the member states. Because the SGM calculations are already harmonised and assembled in all member states, this is a very promising database. Besides SGM is based on (and used for), both FADN and FSS, so it generates already a link between different databases and institutes (on national level and between DG-Agri and Eurostat). Up till now, SGM is only used for typology. Therefor it is possible that the quality level should be improved for some products. If we succeed in this however the SGM could be very useful for:

- the linking of different databases (and the checking of the consistency between the databases);
- a database of detailed inputs used per activity and/or outputs (modelling purposes, comparison of profitability of activities between member states and regions);
- the link between activities and outputs;
- typology and economic size of farms.

It is clear that the reasons of differences between EAA and SGM results have to be discussed in the coming years with representatives of the Member States. This will improve the quality of the SGM and its usefulness in AgrIS.

Available SGM per region (including reliable details like output, total variable costs) might be published in a Geographic Information System (GIS) to improve quality of the data.

The existing detailed SGM of most countries include information on prices and volumes. This could be used to connect these results to EAA, under the condition that the information on prices and volumes used in EAA are available too.

A better use of SGM in AgrIS can be obtained with a number of adjustments on the methodology or the procedure with which the SGM are reported by the Member States to Eurostat. First of all, the way in which detailed SGM are delivered to Eurostat should be improved:

- make a split between different outputs of one activity for all products (for instance milk and meat in the case of dairy cows or cereals and straw in the case of wheat), as well as subsidies;
- produce detailed SGM for all products. So also for products which detailed information is not available at this moment (as crops in horticulture);
- take care that information of SGM is available at the same level of detail (crops, animals) as on FSS. This is important to make a construction in AgrIS to check aggregated data of different sources (EAA, SGM, FADN). At this moment adequate information in New Cronos or detailed SGM is not available for this.

Besides this, some methodological issues of the calculation of SGM should be harmonised between countries. These issues are described in detail in section 2.2.5.2. Also the harmonisation of input and output definitions between SGM and EAA could be improved.

Because SGM are based on averges of 3 years, it is useful to have the detailed information of each year available at Eurostat (as Austria did voluntary). Current capacity in ICT systems make this easier possible than in the past. This makes a comparison with the EAA data for that year easier.

5.2 To use the FADN data in EAA

5.2.1 FADN data set

In some respects it is not clear what basic information and definitions are used for EAA and FADN in the Member States. It is important to have this back ground information to see in which extent data can be used and what are the handicaps to convert data from one data source to the other(s).

To restrict the inconsistencies between EAA and FADN and take away the necessity to estimate data on EU level there is a need to intensify the co-operation and consultation between different groups of experts (Working Parties) of Eurostat and DG Agri (FADN Committee). This could lead to more common definitions and e.g. the same methods of valuations of biological assets.

These different groups should - on the start of the process - be confronted with the (diverging or different) results of their work. The problems behind these deviations have to be identified clearly to start the process to solve the problems of different definitions for the different data sets.

Besides the different definitions on returns (output) and costs (input) have to be made clear. Improvements could also be achieved in the weighting (or the sample) of farms in

FADN. In several cases ('small crops' and specific animal products) however, an adjustment of the FADN sample of farms can not solve the existing problems on insufficient reliable data.

The EU's FADN as well as the Member States should be triggered to provide yearly information on the sample of farms in the FADN network and methods to (and results from) the aggregation of data of individual farms.

FADN is considering a new farm return (Abitabile et al., 1999), once the new IT environment is ready. Eurostat has clearly to gain from such an innovation and should actively support it, and participate in the renewal.

If the FADN is renewed, and has a better weighting system with transparency on its aggregates, it could be a very useful source for EAA and AgrIS data. In several Member States the FADN represents more than 90% of production, and is already used to estimate and or to check EAA data.

Several points mentioned above could be discussed more in detail with EAA and FADN-representatives of the Member States in special meetings (task force meetings or seminars). Informal workshops, as they are organised during last years in the frame of PACIOLI [1], could also be useful (Beers et al., 2001).

5.2.2 Econometric models

To keep the AgrIS' concept as much as possible in tact, any modelling effort to check the consistency of Eurostats databases or to link inputs in the EAA to outputs, should be undertaken with heavy involvement of a task force group and the EAA committee. They should support the outcome of the study, in order to make it a useful exercise.

If Member States have better national data then included now in an EU wide model, model results should be abandoned in favour of national data. High quality data that has the support of the Member State are in principle preferable above a harmonised methodology. It must be clear however what is the base for the national information and the data set has also to be available in the future. As far as possible available agro-technical information can be used to improve data, for instance for 'small crops'.

It is recommended to invest in next generation models. For this it seems attractive to co-operate with the experiences now gathered in projects for DG-Environment. A combination of a priori information of SGM and national sources and the use of a maximum entropy model can serve the construction of the AgrIS database in future. This may also result in more detailed information on specific inputs (pesticides, fertiliser) in line with the EAA+ concept.

[1] During PACIOLI meetings accountancy experts from all over Europe discuss in an informal way problems on accountancy and in particular managing FADN. Besides the presentation and discussion of papers, participants are expected to *work* on common problems during the workshops.

5.3 To use national data

Interviews with experts in the Member States resulted in a number of suggestions on data sources possibly to be used in the context of the project. These can be categorised as follows:

- cost price calculations (several countries - as Denmark, Finland, France, Germany, Italy, the Netherlands -, in most cases not regularly, for one or some products, mainly or exclusively on specialised farms in or outside FADN);
- management data on returns, costs and margins, as well as technical information on yields, volumes of input at national level (UK, the Netherlands) or per region (Germany);
- data on costs per product (Portugal, Spain, Sweden), not based on FADN, not sent to Eurostat or DG Agri;
- specific SGM calculations, not based on FADN, but sent to Eurostat (Germany, Austria);
- data of agro-industries on the volume and value of specific inputs used (f.i. on pesticides, fertilisers, feed) so far or not used in EAA.

It is fairly hard to make a selection on these suggestions to determine what data sources can or has to be used and what is not adequate for the purpose in this project.

In general experts on FADN and EAA in the Member States have by their own the responsibility to select the data sources for EAA as well as to organise the sample of FADN in line with the instructions in Handbooks and EU Regulations. So far the impression is that depending on the history per country different use is made of (in principle) available information to assemble EAA data.

So it is recommended to organise a working method and procedures with the Member States to get a maximum of data in line with the AgrIS context. This needs an intensive communication with experts - mainly members of working parties - to see what is available and is approved to be used. Communication with the working parties across the Member States can stimulate to provide more adequate data for AgrIS.

The results of interviews (chapter 3) make clear that in several countries it is desirable to do some extra effort to collect extra information on some specific costs (for instance on forage and pesticides) and or to regroup the available data in such a way it can be used for the (EAA+) database in the coming years. The policy issues that lead to the definition of the EAA+ list, are also relevant at a national level, and Member States should consider to make the EAA more policy relevant. It is important to do this without or at least a minimum of (extra) costs. Communication between Member States can stimulate procedures to achieve this common goal.

Related to this it is very useful to see to what extent the data collection for the different data sets can be harmonised. This requires in some countries an intensification of the communication and co-operation between institutes, or at least experts dealing with FADN and EAA. By this the process can be made more efficient and provide in a higher quality of results.

An important item for all countries is that data collected for FADN, from which SGM are often derived, can also be used for the EAA and agricultural prices. A restriction

in this is the time lag between the actual moment and the availability of (even provisional) data of FADN.

In Denmark FADN is used in calculating SGM and parts of EAA. These data are already compared regularly and they are found to be rather consistent, at least in the case of major products and inputs. However, written reports of the comparisons and consistency checking would be interesting and they would be also very useful for other countries which would like to improve the consistency of their agricultural database. Since Denmark has some experience in the efforts of creating consistent agricultural data, and calculating production costs per activity, the Danish experts could offer a useful example to experts in other countries in presenting their methods and results.

5.4 To summarise

Conclusions and recommendations

The project results in a (initial) database on inputs allocated to agricultural activities for all 15 Member Countries. So far the database is developed only for one year (1997). For the production of this database as far as possible use is made of the (adjusted) FADN/INRA-model, detailed SGM and additional national data. Given the lacks in information and the differences in definitions it was not possible to arrive at a database in which all desired specifications of activities and inputs (as mentioned in EAA+ lists) are included. Chapter 2 provides a description of the process followed to develop the desired database, including the encountered problems, in first instance. This is followed by some remarks and additions on the use of national data in chapter 4. This chapter also concludes that this database is a starting point for further consultations with the Member States.

It is clear that this (initial) database is not fully corresponding to what is aimed by Eurostat. The procedures followed during the project make clear that a lot of problems still have to be overcome. Many solutions imply the need of additional contributions of Member Countries. It is important to organise this in an efficient way to make it acceptable for the Member States.

The project provides a number of recommendations for this, for instance on making available more information to Eurostat and or DG-AGRI (see above). To stimulate, this it is important that Eurostat and DG-AGRI make some extra efforts, for instance on the use of (detailed) SGM in the context of AgrIS and to check the (aggregated) FADN data in relation to EAA results. This implies a more intensive co-operation with the Member States (in task force meetings), for instance to analyse the reasons for differences in results and how to cope with these issues (definitions of outputs and inputs and interpretation of these definitions in the member states, accounting and calendar year, VAT, subsidies, non-represented farms). On some issues a more intensive co-operation between working parties of Eurostat and DG-AGRI is needed, for instance on the harmonisation of the definitions used in the different agricultural statistics. Related to this variable costs (inputs) of production have to be made more explicit in the calculation of SGM (on a yearly base).

This means that Eurostat and DG-AGRI are advised to concentrate their efforts in the coming years on matching and making consistent the (links between) definitions and vari-

ables (on output as well as on inputs) used in EAA, SGM and FADN. This leads automatically to the need of more co-operation - in each Member State - between the institutes responsible for the data used in the EU. Member states should concentrate on the reasons for differences between the databases for items with equal definitions. The differences might be caused by interpretation of definitions.

For some problems indicated in the project it is advisable to make on request of Eurostat and or DG-AGRI more detailed studies in the coming years, for instance on:

- calculate results FADN/INRA model 1990-1999;
- calculate and compare (detailed) SGM 1996 results for all Member States and ask for underlying yearly data;
- organise task force discussion (or seminars) on the results of the 2 previous activities and make transparent the use of EAA, FADN and SGM by publications (dealing with the consistency of data);
- the use of cost price calculations in member countries in the context of the development of AgrIS (What are the definitions and data used for these calculations);
- the differences in interpretation of definitions (on Member State level) between the databases;
- the use of data of agro-industries (values, volumes, prices of outputs and inputs);
- set up a research database/environment in which individual data from FADN and FSS could be combined for policy research and to improve EAA output/input levels and ratio's;
- the use of next generation models (like maximum entropy) to combine data of different sources;
- the application of a more fine-tuned, tailored list of outputs and activities and inputs, linked to the desired level of information and the capacity of member countries to deliver the information on a regular base;
- the application of the approach followed in this project in and with the Candidate Countries; in fact the results here have to be matched with the possibilities of the future Member States, which are subject of discussion in another project on request of Eurostat.

References

General

Abitabile, C. et al., *Ricastings, the implications of a new farm return on a renewed FADN*. Report 99.06, LEI, The Hague, 1999.

Butault, J.P. and J.M. Rousselle, *Installation of the model 'Coutpord INRA-INSEEE.* Paris, March 2001.

CEC, *Farm Incomes in the European Community in the 1980's.* Brussels-Luxembourg, 1993.

CEC, *Farm return data definitions RI/CC 1256 rev. 2.* Brussels, 2001.

Eurostat, *Manuel on the economic accounts for agriculture and forestry EAA/EAF 97 (rev. 1.1).* Luxembourg, 2000.

Eurostat, *Manuel on 'Standard gross margins and types of farming' (classex 314)*. Draft, 2001.

Eurostat, *Pilot project Agricultural sector Modelling*. Forthcoming.

Eurostat, *AgrIS tutorial*. 2001.

Graham Brookes, *European arable crop profit margins 2001/2002*. 1st edition, Kent, 2001.

Pollet, P., E. Chantry and J.P. Butault, *The agricultural costs model*. INSEE- INRA-SCEES, 2001.

Verhoog, D., 'Introduction to the agricultural information system of Eurostat, 2000'. In: Pauli, *Proceedings of the 7th IWG-AGRI-Seminar*. Luxembourg, July 2000.

Per Member State

Belgium
CLE, *Evolutie van de land- en tuinbouweconomie in 1999/2000*. Brussels, 2000.

UK
Nix, J., *The Farm Management Pocketbook 28th and 32nd edition*. Why college, London, 1997 and 2001.

Ireland
Teagasc, *Farm records and accounts*. Dublin, 2001.

Teagasc, *Costs of cereal production in Ireland and selected EU Member States*. Projectreport 4509, 2001.

Denmark
'Ökonomien in landbrugets dirftsgrene'. In: *SJFI publication nr. 84*, Serie B, 1999.

Beregningsmetoder og fordelingsprincipper. Serie B 1999 (publicly available at http://www.sjfi.dk/

Sweden
Lönsamhetskalkyler för lantbrukets produktionsgrenar. Rapport 1997:21, Jordbruksverket, Statistikenheten, 1997-11-14.

Productionsgrenskalkyler - Metodbeskrivning. Livesmedelsekonomiska Samarbetsnämdens Jordbruksekonomiska grupp, Maj 1995.

Finland
http://www.maaseutukeskus.fi/inenglis.htm

Appendix 1 Book keeping years FADN

For 1997 the following bookkeeping years have been applied.

Belgium (Specialist horticulture) a)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Belgium (Specialist horticulture) a)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Belgium a)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Denmark (Specialist horticulture) b)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Denmark b)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Denmark b)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Denmark b)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Denmark b)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Germany (Specialist horticulture)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Germany (Specialist horticulture)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Germany	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Greece	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Spain	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
France c)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Ireland	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN

Italy	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Luxembourg	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
The Netherlands	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
The Netherlands (specialist horticulture)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Austria	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Portugal	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Finland	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
Sweden	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
United Kingdom d)	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN
	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN

a) In Belgium, some horticultural holdings have accounting years from 1 January to 31 December whilst others have accounting years running from 1 May to 30 April; b) In Denmark, the accounting year for horticulture begins on 1 January and ends 31 December. For half of the agricultural accounts the accounting year starts on 1 January and ends 31 December. For the remainder, the accounting year starts on either 1 April, 1 May or 1 June and ends 12 months later; c) In France, the beginning of the accounting year for a small number of farms falls between 1 October and 31 December of the preceding year (comptabilités fiscales); d) The accounting year in the United Kingdom runs from 31 December to 30 of April (for an individual farm however the year lasts only 12 months).

Appendix 2 List of (proposed) activities and inputs (EAA+)

(Adaptions to EAA in italics)

1.1	**Activities**	**Crop growing activities**
Code New CRONOS		ACTIVITIES
(0)1000	Cereals [1]	
		Soft wheat
		Durum wheat
		Rye
		Barley
		Oats and summer cerial mixtures
		Grain Maize
		Rice
		Triticale
		Other cereals
2000	Industrial crops	
		Rape
		Sunflower
		Soya
		Other oleaginous products
		Protein crops (incl. peas and beans)
		Raw Tobacco
		Sugar beet
		Other industrial crops (including hops and fibre plants as flax, hemp etc)
3000	Forage plants	
		Root crops (f.i fodder beet)
		Fodder maize (for silage)
		Grass
		Other
4000	Vegetables and horticultural products	
		Fresh vegetables
		Cauliflower
		Tomatoes
		Other fresh vegetables
		Mushrooms

[1] No distinction between cereals planted before winter and in spring.

		Ornamental crops
		Live plants (incl. pot plants, nursery and other plantations)
		Flowers (cut flowers)
		Flower bulbs/tubers
5000	Potatoes	
		Ware potatoes [1]
		Seed potatoes
		Potatoes for starch production
6000	Fruits	
		Fresh fruit
		Dessert apples
		Dessert pears
		Peaches
		Other (incl. strawberries etc)
		Citrus fruit
		Sweet oranges
		Mandarins
		Lemons
		Other
		Tropical fruit (f.i. bananas)
		Grapes
		Table grapes
		Other grapes, fresh
		Olives
		Table Olives
		Other Olives
7000	Wine	
		Table wine
		Quality wine
8000	Olive oil	
9000	Other crop products	
		Seed and planting stock (excl. seed potatoes)
		Other crop products
10000	*Total output (1 to 9) of crop growing activities*	

[1] Seed potatoes are also mentioned as inputs.

11000 Animals/Animal husbandry activities

Bovine cattle (mainly for meat production)

Animals under one year

Calves for slaughtering (f.i.veal product)

Calves for rearing

Animals between one and two years

Animals over two years

Heifers/Cows for slaughtering

Bulls/ Oxes

Cows for reproduction (suckling cows)

Bovine cattle (mainly for production of milk)

Calves (< 1 year)

Animals between one and two years

Dairy cows

Pigs

Breeding pigs (including piglets)

Pigs for fattening

Sheep

Sheep for meat production

Sheep for milk production

Goats

Equidea (horses, mules etc)

Poultry

Laying hens

Poultry for slaughtering

Broilers

Other poultry for meat production (f.i. turkeys)

Other animals — For instance rabbits

13000 Total Animal output of animal husbandry activities

1.2 Inputs (intermediate consumption used for agricultural activities)

Code New CRONOS		
19010	*Seeds and planting stock*	
19020	Energy, lubricants	Electricity
		Gas
		Other fuels for heating
		Motor fuels, lubricants and others [1]
19030	Fertilisers and soil improvers	*Manure* [2]
		Fertiliser
		Nitrate [3]
		Phosphate
		Kalium
		Others
19040	Plant protection products and pesticides	
		Fungicides
		Insecticides
		Herbicides
		Others
19050	Veterinary expenses	
19060	Animal feeding stuffs	
	Feedingstuffs consumed and produced by the same holding	
	Purchased feedingstuffs	
	Forage	
	Other pure/straight feeding products and additives used in feeding [4]	
	Compound feeding stuffs	
19070	Maintenance of materials	
19080	Maintenance of buildings	
19090	Agricultural services	
19900	Other goods and services	
	water (for irrigation, drinking cattle)	
	other	
Total inputs (intermediate consumption)		

[1] Propellants, gasoil/diesel.
[2] As far as figures on costs are available.
[3] The reason to make as possible a distinction is the environmental impact of nitrates, phosphates and kalium.
[4] For examle cereals, products from processing industries.

Appendix 3 Methods to make FADN representative for the complete agricultural sector and to deal with difference in calendar and bookkeeping year

A lot of different methods are used for these purposes. One country might use different methods for different outputs/inputs. First of all the methods that deal with the difference in representativity between EAA and FADN are presented:

A. No correction.
B. the Lineair Programming (LP)-method.
C. Farm type method.
D. Simple method.

Ad. A
Sometimes the FADN total is supposed to be representative for the whole country. For some products this might approximatitely be the case.

Ad. B
In the LP method costs by item (ha per crop or animal) are estimated with a system of equations. The sum of the weighted absolute values of the difference between the estimated value and the real value is minimised to estimate the coefficients.

$$Kv_{b,k} = ha_{b,s} * \overline{KS}_{s,k} + u_{b,k}$$

$Kv_{b,k}$ = value of a cost item for a certain farm
$Ha_{b,s}$ = hectare or number of animals by activity
$KS_{s,k}$ = coefficient to estimate
$u_{b,k}$ = amount of disturbance

The goal function of the LP model is to minimise the sum of the weighted absolute values of the disturbances

$$\min \sum_{b} wf_b * |u|_{b,k}$$

Wf_b = weighting-factor by farm
$U_{b,k}$ = disturbance term by farm and cost-item

Having estimated the coefficients per hectare of animal, total area by crop and total number of animals from the FSS are used to calculate macro totals (total intermediate consumption by cost-item). This makes it also possible to calculate sector-totals. For example, when the coefficients for feed consumption are calculated for dairy cows, yearlings and other kinds of cattle, total feed consumption by cattle is calculated. In this way the sector-

point of view is reached. This method has the advantage that not only sector totals but also the inputs per output are found. The method has some similarities with the INRA model.

Ad C
The farm type method is far less mathematical. It consists of two steps. The first step is to aggregate the micro data by using the weighting-factor in the FADN. This aggregation is type specific. The second step is to calculate by farm type total SGM in FADN and FSS and multiply the weighted aggregate with the ratio of total SGM in FSS and in FADN.

Ad D
This method is the same as method C but in this case all FADN totals are multiplied with the ratio of SGM in FSS and SGM in FADN.

Methods to deal with difference in time period

A. No correction
B. Simple method
C. Advanced method

Ad A
The totals of FADN are used for the EAA.

Ad B
In this case the total for the calendar year is based on a weighted average of the FADN totals. The weight is based on the number of months that the FADN year and a specific calendar year have in common. If for example the FADN year is from 1st of March untill the 28th of February the EAA total for year 0 is calculated as following:
1/6 FADN (t=-1) + 5/6 * FADN (t=0)

AD C
In this case the total production/input use (based on FADN or other sources) is multiplied by the average price in the calendar year.

Appendix 4 Description INRA model and definitions used in model

Introduction

INRA model has been built in 1985 and improved several times during its use. Costs by activity are approximated by estimating a simultaneous equations model with constraints on coefficients using Seemingly Unrelated Regression procedure. Other more sophisticated procedures was tested (see J-C bureau and M. Cyncynatus, 1991) but improvements was small according to the complication of the proposed procedures. We will not discuss in this note all method issues well as new methods (Maximum Entropy) occur recently, but focus on the possible improvements of the current model.

Specifications and estimation method in the INRA model aim at simplicity and friendly use. Model users could quickly compute costs by choosing a field in the database, a list of outputs and a list of inputs. The econometric approach in the INRA model presupposes a number hypotheses according to the structure of the database and the data availability.

The model

RICA provides information on the total charges paid on each farm according to the type of charge. The charges are not however matched with the various products. It is as though, for each farm, we had only the margins from a table giving the charges borne by each product. This information is inadequate either to tell us the production costs involved in producing a particular good or to identify the income generated by the production of a particular good.

The INRA model is an econometric approach method of breaking down the charges by products. This approach presupposes two strong hypotheses:

- the amount of use made of each factor of production depends only on the product manufactured and not on the farm. All farms are therefore assumed to use the same production technique;
- the value of the input used is proportional to the value of the output.

The model's specification

The database from which the model described here is constructed is none other than the French FADN. The initial statistical unit is therefore the farm. The output of the various goods is X_i $(i = 1, ..., n)$ and C_j represents the total non-allocated costs of the factors of production $(j = 1, ..., m)$. Finally, C_{ij} is the production cost in factor j of the good i.

Assuming proportionality, the cost C_{ij} is a linear function $a_{ij} X_i$ of the output of good j. In all, the total cost observed is the sum of the costs relating to the various products. In practice, the model is only approximate. On every farm, the observed costs differ from the theoretical costs by a random factor u_j:

$$C_j = \sum_i a_{ij} X_i + u_j \quad \text{where} \quad u_j \quad \text{iid}$$

The factors u_j are of zero expectation and independent from one farm to the next, which means that the consumption of input *j* by a given farm is not affected by another farm's consumption of the same input. There is therefore no constraint of supply. Moreover, the link between the residues for two different inputs on the same farm depends on the inputs and not on the farm. The phenomena of size and technology peculiar to each farm are therefore disregarded.

The non-labour income R_i derived from the production of the good *i* is the difference between the output X_i and the sum of the costs occasioned by that output. The model's estimate also assumes that this income (sum of outputs less sum of variable and fixed costs except labour) is a linear function b_iX_i of output. In all, the income generated by all outputs is

$$R = \sum_{i=1}^{n} b_i X_i + v \quad \text{where v iid}$$

If the model is to retain its logical consistency, we have to introduce the constraint that the output of a good is the sum of the income and costs, i.e.:

$$\sum_j a_{ij} + b_i = 1 \quad \forall i$$

Estimation procedure

The model then becomes a simultaneous equations model with linear constraints on the coefficients. It is estimated as such, the constraints being integrated at the time of estimation, from the individual data. We made use of an 'SUR' (Seemingly Unrelated Regression) procedure to estimate the model.

Cost breakdown by imputation

The INRA model procedure provides coefficients that allow an average costs and income structure to be calculated for the farms in the considered sample. However, using it presents problems for two reasons:

- equality between the sum of theoretical costs and the observed costs is not verified at individual level because of the presence of the residues u_j. The same situation arises with income. Statistical exploitation at a more detailed level is therefore impossible once some heterogeneity appears;
- the model is only approximate. In particular, the lack of a constant in the estimates means that, on average, there is nothing to ensure that the sums of the residues are equal to zero. Consequently, even on average, the sum of the theoretical costs for a factor *f* with the observed costs is not guaranteed. In other words, the model does not conserve the masses. This is particularly the case in the recent past.

For these two reasons, a further stage was introduced into the process, consisting of imputing costs per output at individual farm level using the model's results.

A fairly simple way of ensuring identity between the sum of the theoretical costs for a factor *j* and the observed costs, that is of reconstructing the masses, is to recover the residues u_j and distribute them among the different products pro rata to the outputs. Such distribution pro rata to the outputs seems quite natural given the model's specification.

Breakdown of work by product

It is necessary to take account of family work in particular so as to avoid making false comparisons of net income not including family work as between countries or even between products. Because the proportions of paid employment and family work differ according to the production structures, the analysis in fact had to be refined by distinguishing between the two types of work.

We have the valuation of paid work for every farm, but not that of family work. We begin by calculating the total number of family Annual Work Units (AWU). Then, in order to put a value on family work, we apply the average regional wage rate to the number of family AWU. This gives an overall valuation of family work for each farm.

In the model's present version, this breakdown is not made in the same way as for other charges, that is by regression from the work on outputs. That method has already been tried and has given unsatisfactory results. Family work is in fact a 'fixed' factor and its level varies little with the size of the farm. The assumption of a proportional link between the levels of cost and output is not realistic for this charge. Another method was therefore used on the assumption that if the factors of production are 'normally' remunerated, work forms the greater part of the value added. We therefore chose to break down family work and paid work for each farm in proportion to the margins on each product.

Definitions used

- *Inputs*

Name	*FADN code*
Feed	C0606+C0607+C0608+C0609
Veterinary costs	C0613
Seed	C0614
Fertilisers	C0616
Plant protection	C0617
Energy	C0604+C0621+C0622
Maintenance materials	C0603
Maintenance buildings	C0620
Services	C0602
Other goods and services	C605+C0623+C0626+C0618

- *Outputs*

Aggregated model	*Detailed model*	*EAA code*	*FADN code*
Cereals	Soft wheat	01110	p2120
	Durum wheat	01120	p2121
	Barley	01300	p2123
	Maize	01500	p2126
	Rice	01600	p2127
	Other cereals	01900+01200+ 01400	p2128+p2124+p2125+ P2122
Industrial products	Oilseed	02100	p2132
	Protein crops	02200	p2129
	Sugar beet	02400	p2131
	Tabacco	02300	p2134
	Other industrial products	02900	P2133+P2135
Forage	Forage	03000	P2144+P2145+P2147 +P2150
Vegetables and horticulture products	Vegetables	04100	p2136+p2137+p2138+ p2139
	Flowers	04200	p2140+p2141+p2156+ p2157
Potatoes	Potatoes	05000	p2130
Fruits and permanent product	Fruits	06100+06300+ 6410+6510	p2152+P2158+P2159 +p2285+p2291+ (p2154-p2282-p2283)
	Citrus	06200	p2153
	Wine	06490+07000+ 17140	p2155-p2285-p2291
	Olive oil	08000+06590	p2282+p2283
Other crop products	Other crop products	09000+17110+ 17120+17130	p2142+p2143+p2146+ 2148+p2149+p2151+ p2160+2161
Cattle	Cattle	11100	P502
Pig	Pig	11200	P506
Equines	Equines	11300	P501
Sheep and goats	Sheep and goats	11400	P504+P505
Poultry	Poultry	11500	P507
Other animals	Other animals	11900+17150	P508
Milk	Milk	12100+17161	P2162+P2163+P2164 +P2165+P2167+ P2168
Eggs	Eggs	12200	P2169
Other animal products	Other animal products	12900+17162	p2166+p2170+p2172

Appendix 5 Remaining problems in coupling of input definitions in FADN and EAA

Cost of inputs corresponding to output period or cost of inputs used in period

In the FADN two different definitions might be used for the inputs. In the first, and preferred, definition all inputs that are used to produce the output of that period are included. Inputs that are already used for the output of next year (for example seed sowed) are not included.

In the second definition all inputs used in the period are included. It does not matter if the output has already been realised or not. In the EAA, the input use is often based on sales of the industry. This will be the closest to the second definition.

Forestry costs

Some forestry costs may be included in the FADN. Although there is a separate item in FADN called 'variable forestry costs', the fixed costs of the forestry part of agricultural farms are included in the fixed costs. These costs should not be included in the EAA but in the EAF.

Veterinary costs and intermediate costs

In the EAA the veterinary costs are a separate category. In the FADN however, the veterinary expenses are included in other specific livestock costs. Although the veterinary costs might be a large part of the other specific livestock costs, all kind of other costs are included in the FADN item: artificial insemination, castration, milk tests, herd book subscription and registration, products for cleaning livestock equipment, packing and processing materials, costs of storage and preparation for market of livestock products (including short term rent of any buildings). These items should be included in the 'other goods and services' in the EAA.

Therefore the items veterinary costs in EAA and other livestock costs in FADN are not comparable. Only when it is assumed that the distribution of the 'other specific livestock costs' to all outputs is representative for the distribution of the veterinary costs, the INRA model could be used to allocate the total veterinary costs in the EAA over the outputs in the EAA.

Other goods and services

Because of the problems with the definition of veterinary costs, described above, also the 'Other costs of good and services' in FADN is not completely comparable with the 'other costs of good and services' in EAA. Beside this problem, in the EAA a small part of the insurance premium is included, while this item is excluded in the INRA model.

Maintenance of buildings

In the FADN the 'upkeep of land improvements and buildings' includes both maintenance of buildings and the maintenance of land improvements. In the EAA, only the maintenance of buildings should be included. Besides in the FADN, a split should be made between major repairs that increase the value of the building and small repairs. The first one should not be included in the maintenance costs of buildings but should be treated as an investment in buildings. It is not always that clear however which repairs increase the value of the building and which ones do not. In the end all repairs will increase the value of the building because otherwise they were not made. In the EAA, repairs intended to extend their normal service life or increase their productivity, should be treated as investments. This difference in definition and possible interpretations of the definitions might lead to a difference between EAA and FADN.

Maintenance of materials

In the FADN some material costs are included in the category 'current upkeep of machinery and equipment'. In the instruction it is written that small purchases of materials like forcing frames, tyres, protective clothing and detergents for general cleaning should be included. Detergents used for livestock production should be included under the 'other specific livestock costs'. As for the 'maintenance of buildings', large repairs that increase the value of the machines, should not be included.

The EAA includes in the category 'maintenance of materials' beside the goods and services for maintenance, the purchase of crop protection equipment such as detonators, anti hail protection etc.

Seed

In the EAA seed that is produced and consumed by the same holding in the same reference period should not be included. If it is used in the next period however, it is included because it is included in the stocks at the end of the year. In the FADN, the cost of seed should always be included even if it is used in the same period.

Contract work

In the contract work in FADN also machines hired without labour are included. In the EAA only contract work with labour is included in this category.

Theory and practice

Although theoretically definitions of inputs differ between FADN and EAA, this does not always mean that the totals are not comparable in practice. A lot of countries use FADN for the calculation of the total input use in the EAA. Of course the definitions in the national FADN might be more detailed than in the EU-FADN, which makes it possible to use the definitions necessary for the EAA. It might also be possible however that the EU-

FADN definitions are used for the EAA because one is unaware of the differences in definitions between the database or has no better source available. In this case the totals in FADN and EAA are comparable.

It might also be the other way around. Although input definition are described in both databases, these definitions are not that strict and detailed that it is always clear what exactly should be included in the items. The interpretations of definitions might be different in the EAA and FADN which makes the totals in practice incomparable. Besides it might be impossible in some Member States to use the exact definitions because the data might not be available. In this case a slightly different definition might be used.

An example:
In the Netherlands, the maintenance of materials in EU-FADN includes the purchases of small materials. In the FADN data for the EAA, these purchases of small materials are not included.

Cost of inputs used by contract workers

In the EU-FADN, the inputs used by contract workers (pesticides, fertilisers, seeds and energy), should be included in the costs of the inputs and not in the costs of contract work. This might however be rather complicated because this data is not always available on farm level. Sometimes the farmer receives a bill with only the total amount to be paid to the contractor. The FADN instruction prescribes in this case, that the total use should be estimated and subtracted from the contract costs [1]. This is a complicated procedure and it might be expected that not all countries will act like this if data is not available.

In the EAA, also the contract workers themselves are included. This means that the inputs used by the contract workers are included in the input costs and the total amount that has to be paid by the farmer to the contractor is included in both the production of agricultural services and the consumption of agricultural services.

It is not clear how all national FADNs treat this subject but if they are not able to split the input costs from the other costs of contract work the definition of fertilisers, pesticides, energy and seed is stricter than the definition in EAA. If (some) countries follow the FADN instruction, the total of contract work in FADN will have a stricter definition than in EAA.

In the Dutch FADN, a split of the costs of contract workers is made for the fertilisers, pesticides and seed but not for energy (use of fuel etc). So for the Netherlands, the FADN definition is stricter than the EAA definition.

Besides the inputs described above (seed, fertilisers etc), also the other costs of the specialised contract workers (maintenance and other goods and services) are not included in the FADN. So also these intermediate consumption items will have a stricter definition in FADN than in EAA.

[1] At other places in the instruction however, it is written than these costs should only be split from other costs of contract work, if available.

Appendix 6 Remaining problems in coupling of output definitions in EAA and FADN

Vegetables

In the FADN the vegetables are subcategorised in seven categories:
- cabbages, cauliflower, broccoli;
- leaf vegetables;
- tomatoes;
- vegetables grown for their fruits of their flowers;
- vegetables grown for their roots, bulbs or tubers;
- legumious vegetables;
- fruit of non-perrenial plants;

and a division is made into three production methods:
- field vegetables;
- vegetables (market garden);
- vegetables under glass.

In the EAA, vegetables are only subdivided in three categories:
- tomatoes;
- cauliflower;
- other fresh vegetables.

So the FADN definitions can not be coupled on a detailed level to the EAA definitions. The only match are the tomatoes. This is however only an important product for some countries. Therefor we created one category vegetables. In this case only one problem remains: Strawberries and (water)melons are included in the vegetables in FADN and in EAA in Fruit.

Besides a very heterogeneous group with large differences in cost of production between the different vegetables results. Although the definitions match rather well, the results of the model could not be used to compare cost of production between countries because the mix of products might be completely different.

Flowers and plants

The FADN distinguishes between:
- flowers and plants in the open air;
- flowers and plants under glass;
- nurseries;
- (growth of young plantations).

EAA has one category Plants and Flowers that is split into:
- nursery plants;
- ornamental plants and flowers;
- plantations.

The EAA does not make a split into the production methods open air and glass. A detailed match seems not to be possible. If one category 'flowers, plants and nursery' is created however, the total seems to be the same as in the EAA. It is a very heterogenous category.

Only one problem remains which might be specific for the Netherlands. In the Netherlands part of the production of plant and flowers (the gain in value of flower bulbs which will be used within the farm for the production of next year) is included in the FADN code 'growth of young plantations'. In this code are however also the gain in value of fruit trees etc included so it can not completely be included in the output of plants and flowers. Because we expect that this gain in value of fruit trees will be the biggest part of this code for most countries, the growth of young plantations is included in fruits.

Fruit

The FADN distinguishes between:
- fruit and berry orchards;
- citrus fruit orchards;
- olive groves;
- vines (including table grapes and raisins).

A subdivision is made for all these categories. The EAA distinguishes the following categories:
- fresh fruit;
- citrus fruit;
- tropical fruit;
- grapes;
- olives;
- wine;
- olive oil.

The following match has been made in the INRA model:
- fruits (including nuts, tropical fruit, growth of young plantations, table grapes, raisins, table olives and by-products of olives);
- citrus fruit;
- wine (including wine by products);
- olive oil (both olives for olive oil and olive oil).

The category fruit is very heterogenous. Besides the problem of the 'growth of young plantations' and strawberries/melons mentioned above, it seems that the categories match good however.

Inseparable non agricultural activities and other outputs

Other crops products and other animal products. Therefore the definitions of these categories are not completely harmonised with EAA definitions.

Some FADN outputs (like intrest on liquid assets, receipts relating to previous accounting years) are not included in the model. These items are not included in the EAA items 0 to 18000. They can be skipped without problems because no costs are involved for these outputs. For some items it is more complicated because costs are involved and/or these items are also included in EAA:

- processing of products (2160, 2163, 2167 and 2168 in FADN, 17100 in EAA);
- contract work for others (code 2177 in FADN, 15000 in EAA);
- land leased to others (2149 and 2172 in FADN, 28000 in EAA);
- farm tourism (2179 in FADN, 17900 in EAA).

It is very complicated how to treat the FADN outputs which are not included in one of the EAA output categories that are included in codes 0 to 14000 in EAA. On the one hand they should be included because the costs of these activities are also included in FADN. If they are not included, these costs are allocated to other products [1].

On the other hand if they are included in one of the other output categories of the model, they might disturb the results of the regression analysis in the model because they might have a completely different cost of production than the product(s) already included in the category.

They could also be included as a separate output categorie in the model but the output of these activities is rather small and FADN does not represent these activities very well. No perfect solution exist.

For the first group of products (processing activities), we decided to include these in the output of the product needed for the processing activity:

- products of milk (2163, 2167 and 2168) are included to milk in both the model and EAA (17161);
- products processed from crops are included in 'other crops' in the model (2160) and in EAA processing of products from cereals, vegetables and fruits (17110, 17120 and 17130) are included in this group;
- vines is both in model (155) and EAA (17140) included as a separate group;
- processing of animals of EAA (17150) is included in 'other animals' and other animal products (17162) is included in 'other animal products'. There are no such categories in FADN.

The receipts from occasional letting of fodder area and agistment (2172) are included in other animal products. The land ready for sowing leased to others (2149) is included in other crop products. It is not quite clear where these items should be included in EAA.

We decided not to include contract work (2177) and tourism (2179) because these outputs will have a completely different cost structure than other outputs and therefor

[1] This is however not always the case for land leased to others because in this case no costs could be involved.

would disturb the results too much if they are added to another output category. They are too small to be included as a separate category. The similar items in EAA (agricultural services, 15000 and other inseparable secondary activities, 17900) are also not included in the comparison.

Appendix 7 Definitions used in coupling output definitions SGM and EAA

Soft Wheat	d01
Durum Wheat	d02
Barley	d04
Grain Maize	d06
Rice	d07
Other Cereals	d08+d03+d05
Oilseeds	d13di
Protein Crops	d09
Sugar beets	d11
Raw Tobacco	d13a
Other industrial crops	d13b+d13c+d13dii+d13diii
Forage plants	d18+f01
Vegetables	d14a+d14b+d15+i02
Flowers and plants	d16+d17+g05+g06
Potatoes	d10
Fruits	g01+g04c+g04d+g03a+g07
Citrus Fruits	g02
Wine	g04a +g04b
Olive oil	g03b
Other crop products	d19+d20+d21+d22 + F02
Cattle	j02-j06+j08+part J07
Pigs	(j11)+j12+j13
Equines	j01
Sheep and goats	Part j09 + part j10
Poultry	J14+j16+part j15
Other animals	j17
Milk	Part j07 + part j10
Eggs	Part j15
Other animal products	j18+j19

Appendix 8 Comparison of output and input totals based on SGM/FSS and EAA (million euro)

Comparison of SGM/FSS output and EAA output (million euro) for some countries

	Austria		Belgium		Denmark		Finland		France		Luxembourg		Netherlands		Sweden		Uk	
	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM	EAA total	EAA/ SGM
Soft Wheat	138	101	187	74	582	137	64	90	3,863	110	7	95	110	56	230	88	1,961	88
Durum Wheat	7	108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Barley	145	121	40	83	476	66	236	93	1,218	102	7	94	29	56	217	94	954	76
Grain Maize	175	88	3	11	0	0	0	0	1,920	100	0	78	10	63	0	0	0	0
Rice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Other Cereals	55	98	11	76	81	99	144	86	255	118	3	93	6	49	189	89	74	69
Oilseeds	62	183	3	54	71	61	22	119	1,253	83	1	103	4	0	28		372	108
Protein Crops	20	105	3	120	58	91	2	34	441	0	0	98	19	275	14	64	5	4
Sugar beets	140	103	333	127	149	95	61	95	1,177	104	0	0	335	88	134	106	475	103
Raw Tobacco	0	109	1	0	0	0	0	0	94	105	0	0	0	0	0	0	0	0
Other industrial crops	2	44	44	77	0	0	0	0	225	64	0	0	4	45	0	0	300	1273
Forage plants	644	0	671	51	560	96	446	152	4,527	93	15	0	534	51	466	85	135	0
Vegetables	117	134	705	129	137	108	155	134	2,623	53	1	0	2,061	99	131	0	1,372	111
Flowers and plants	184	945	426	97	454	108	110	587	2,085	67	7	0	4,243	133	128		969	134
Potatoes	45	81	255	120	110	112	70	83	1,107	50	3	95	533	70	107	108	563	71
Fruits	245	224	286	148	32	111	41	529	2,186	82	3	0	321	109	37		287	115
Citrus Fruits	0	0	0	0	0	0	0	0	10	100	0	0	0	0	0	0	0	0
Wine	203	72	4	0	0	0	0	0	8,060	99	14	147	0	0	0	0	0	0
Olive oil	0	0	0	0	0	0	0	0	15	103	0	0	0	0	0	0	0	0
Other crop products	0	0	37	4	125	23,995	2	27	285	57	2	8,407	764	598	8	122	60	168
Cattle	495	78	923	70	387	86	241	61	6,400	113	40	78	1,428	93	319	0	2,242	129
Pigs	802	97	1,725	116	2,510	133	275	80	3,391	85	18	158	2,089	73	513	134	1,759	131
Equines	1	2	8	0	5	22	3	6	145	87	0	0	13	4	91	0	164	69
Sheep and goats	16	56	7	0	6	64	1	9	657	88	0	0	70	65	8	23	1,499	93
Poultry	87	343	326	244	176	241	60	208	3,166	14	0	0	747	143	89	1,053	2,163	155
Other animals	0	0	17	0	302		9		399	66	1	0	11	89	92	0	50	0
Milk	761	85	918	93	1,498	92	821	98	7,703	105	81	100	3,491	87	1,132	87	4,587	104
Eggs	103	73	184	99	82	116	41	98	755	107	2	0	415	72	98	167	599	138
Other animal products	20	2,698	1	0	4		220		182	174	1	0	60	0	8	0	75	0
Total	4,464	116	7117	84	7219	104	3,023	107	54,141	89	207	117	17,298	96	3,800	116	20,666	110

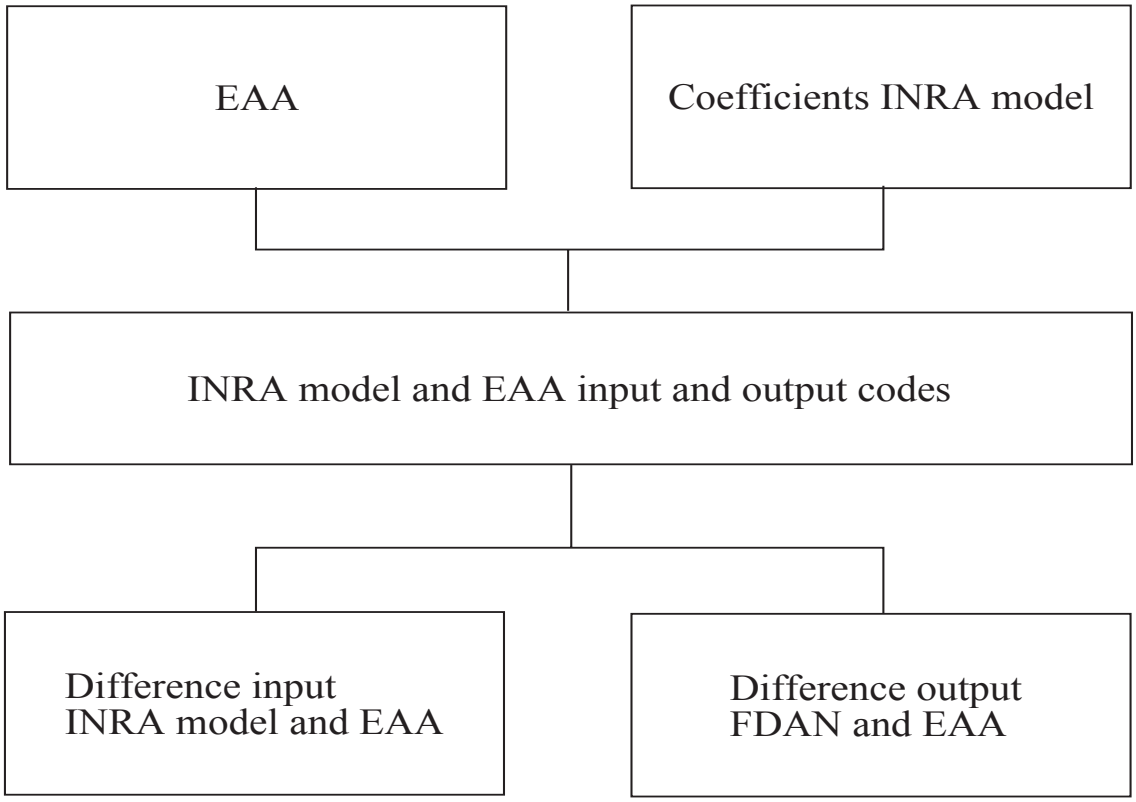
Comparison of inputs based on SGM/FSS and EAA for some countries (million euro)

	Austria			Denmark			Finland		
	EAA	SGM	EAA/ SGM	EAA	SGM	EAA/ SGM	EAA	SGM	EAA/ SGM
Seed	150	165	91	121	243	50	66	134	49
Fertiliser	148	158	94	244	309	79	228	268	85
Pesticides	79	81	97	155	189	82	49	38	131
Feed	1,308	775	169	1,699 a)	1,661 a)	102 a)	1,053	771	137

	France			Netherlands		
	EAA	SGM	EAA/ SGM	EAA	SGM	EAA/ SGM
Seed	1,574	2,042	77	843	783	108
Fertiliser	2,956	3,088	96	298	495	60
Pesticides	2,383	2,274	105	289	300	96
Feed	12,510	9,167	136	3,679	3,882	95

a) Exclusive own produced fodder.

Appendix 9 Comparison of INRA model and EAA



EAA
- EAA 97: values at current prices (producer price) for inputs and outputs in 1997.

Coefficients INRA model
- Input coefficients for 1997
  (share of input value in output value for different products).

INRA model and EAA input and output codes
- Match of input and output codes of EAA to input and output codes INRA model

Difference input INRA model and EAA:
- aggregation EAA outputs to INRA model output codes;
- calculation of inputs per product/output (= aggregated EAA outputs * INRA input coefficients);
- aggregation of inputs per product to total inputs (for all products);
- aggregation EAA inputs to INRA model input codes;

- comparison of calculated total inputs (based on EAA output values and INRA model input coefficients) and the aggregated EAA input values;
- resulting in calculated input values as percentage of EAA input values (appendix 11).

Difference output FADN and EAA by product:
- aggregation EAA outputs to INRA model codes;
- comparison of total output values of FADN and the aggregated EAA output values;
- resulting in output values of FADN as percentage of EAA output values (appendix 12).

Appendix 10 Calculated total inputs as percentage from EAA inputs

	Seed	Energy	Fertilisers	Pesticides	Veterinary cost	Feed	Maintenance materials	Maintenance buildings	Services	Other IC	Tot IC
Austria	0.80	0.82	0.90	1.15	0.83	1.24	0.85	0.95	0.40	0.93	0.84
Belgium	0.81	0.70	0.97	0.88	1.42	0.95	0.37	0.79	3.57	0.75	0.92
Denmark	1.76	0.91	1.07	1.07	2.15	1.03	0.99	0.73	0.98	0.70	1.02
Finland	1.07	1.05	1.13	1.00	2.09	1.03	1.46	1.41	0.80	0.70	1.03
France	1.29	0.93	1.06	1.08	1.26	0.94	1.03	1.06	1.14	1.05	1.04
Germany	1.52	0.97	0.95	1.28	2.32	1.01	0.82	1.06	1.14	1.10	1.09
Greece	0.83	0.66	1.66	1.44	1.18	1.30	1.72	0.46	2.99	2.29	1.31
Ireland	0.82	0.51	1.05	1.18	3.66	0.97	0.85	2.27	1.17	1.28	1.13
Italy	2.32	1.14	1.44	1.57	17.86	1.00	2.43	1.81	1.32	1.07	1.33
Luxembourg	0.87	0.86	0.99	1.00	4.33	1.18	0.74	1.00	1.83	0.73	1.08
Netherlands	1.07	0.81	1.49	1.19	1.77	0.98	0.90	0.78	0.59	0.95	1.01
Portugal	0.65	1.58	1.29	1.42	4.24	0.93	3.89	1.00	29.35	0.63	1.20
Spain	1.66	0.90	1.44	1.61	1.11	1.11	0.30	0.76	1.53	1.31	1.08
Sweden	0.79	0.76	0.93	0.78	3.28	0.75	1.05	0.57	4.80	0.39	0.84
UK	1.65	0.97	0.95	0.99	3.26	1.28	1.00	0.99	1.08	0.54	1.06
Avg	1.19	0.90	1.16	1.18	3.38	1.05	1.23	1.04	3.51	0.96	1.07

Appendix 11 Output values of FADN as percentage of EAA outputs

	Aus-tria	Bel-gium	Den-mark	Fin-land	France	Ger-many	Greece	Ire-land	Italy	Luxem-bourg	Nether-lands	Portu-gal	Spain	Swe-den	UK	Ave-rage
Soft Wheat	1.15	1.00	0.91	1.03	0.95	0.87	0.75	1.10	1.01	0.84	1.04	1.10	1.18	0.99	0.97	0.99
Durum Wheat	0.89	1.00	1.00	1.00	1.27	0.26	0.87		1.20	1.00	1.00	0.38	0.70	1.00		0.89
Barley	0.84	0.84	0.96	0.97	0.89	0.90	0.63	1.05	1.25	0.79	1.13	0.94	1.35	0.93	0.95	0.96
Grain Maize	0.79	7.24	1.00	1.00	1.01	0.49	1.08		0.93	0.26	0.24	1.17	0.85	1.00		1.31
Rice	1.00	1.00	1.00	1.00	1.22		1.12		1.03	1.00	1.00	0.91	0.12	1.00		0.95
Other Cereals	0.97	0.70	0.65	0.79	1.00	0.90	1.63	0.90	1.21	0.76	0.89	1.36	1.09	0.95	1.04	0.99
Oilseeds	0.89	1.98	0.78	0.82	0.98	0.94	0.96		0.87	0.74	0.98	0.84	0.94	0.92	1.06	0.98
Protein Crops	1.47	0.68	0.73	2.74	1.05	0.78	4.70		0.88	1.04	0.19	0.84	4.52	0.80	26.86	3.38
Sugar beets	1.10	0.86	1.16	1.26	0.86	1.03	1.15	0.95	0.79	1.00	1.11	0.42	1.17	1.03	1.17	1.00
Raw Tobacco	0.00	0.90	1.00	1.00	1.28	2.05	2.26		0.24	1.00	1.00	2.32	1.15	1.00		1.17
Other industrial crops	10.55	0.72	1.00	1.00	0.32	1.53	2.13		1.25	1.00	4.38	0.21	0.33	1.00	0.10	1.82
Forage plants	0.03	0.01	0.05	0.01	0.00	0.04	0.64	0.04	1.18	0.01	0.10	0.91	0.86		1.27	0.34
Vegetables	0.56	0.80	1.20	1.70	0.85	0.70	0.51	0.03	0.49	0.00	1.05	0.44	0.57	0.26	0.65	0.65
Flowers and plants	0.00	0.87	1.00	2.32	0.40	0.77	1.21		0.61	0.00	0.88	0.06	0.06	0.00	1.10	0.66
Potatoes	1.35	0.58	1.13	0.92	0.40	0.88	0.40	0.62	0.66	0.74	1.51	0.95	0.35	0.97	1.13	0.84
Fruits	0.52	0.62	0.65	0.24	0.72	0.22	1.22	0.17	0.95	0.00	0.89	0.60	0.83	0.00	0.75	0.56
Citrus Fruits	1.00	1.00	1.00	1.00	0.70		1.09		0.85	1.00	1.00	0.36	0.38	1.00		0.86
Wine	1.31	0.00	1.00	1.00	0.87	0.79	1.02		0.95	1.15	1.00	0.72	0.59	1.00		0.88
Olive oil	1.00	1.00	1.00	1.00	0.11		0.87		0.91	1.00	1.00	0.73	0.92	1.00		0.88
Other crop products	1.00	1.58	0.65	5.75	1.03	1.45	1.85	0.18	0.35	0.48	0.19	2.98	0.10	3.54	3.31	1.63
Cattle	0.86	0.99	0.91	0.98	0.74	0.75	0.62	1.04	0.57	0.76	0.51	0.85	0.43	0.78	1.01	0.79
Pigs	1.04	0.69	0.88	0.98	0.93	0.86	0.03	0.65	0.45	0.69	1.61	0.47	0.32	1.03	0.86	0.76
Equines	1.94	0.02	0.44	0.53	0.13	0.34	-0.01	0.07	0.76	0.52	0.52	0.10	0.03	0.04	0.01	0.36
Sheep and goats	0.61	0.20	0.82	3.04	0.72	0.17	0.57	1.26	1.80	0.00	1.22	0.38	0.71	0.25	0.97	0.85
Poultry	0.62	0.99	1.30	0.31	0.41	0.31	0.67	0.05	0.05	0.46	0.69	0.07	0.21	-0.01	0.08	0.42
Other animals	1.00	1.00	0.51	0.00	0.38	0.22	0.04		0.60	0.00	0.49	0.08	0.31	0.01	0.38	0.36
Milk	1.17	1.09	1.00	1.11	1.02	1.00	0.67	1.20	1.25	0.83	1.17	0.74	1.00	1.12	1.05	1.03
Eggs	0.89	0.08	1.47	1.25	0.34	0.22	0.20	0.44	0.15	0.04	1.33	0.24	0.11	0.02	0.85	0.51
Other animal products	0.84	15.82	1.75	0.01	0.66	41.58	0.10	1.29	1.30	0.25	0.85	0.36	0.10	0.95	1.58	4.50

Appendix 12 Questionnaire for the interviews with experts of Member States

Introduction

On request of Eurostat (Statistical Office of the EU) a project team consisting of researchers of LEI in the Netherlands, MTTL in Finland, INRA in France and INEA in Italy is conducting a research project to improve the quality of agricultural statistics on inputs.

The main objective of this project is to construct a consistent database on inputs, related to agricultural activities for the actual Member States of the EU. This database can be used for statistics, agricultural economic research, modelling etc.

So far the project team worked with information of Eurostat and the European Commission: mainly the EAA (economic account for agriculture), SGM (standard gross margins) and FADN (farm accountancy data network).

It appears however that this information is not in all respects sufficient to construct the database. So it is necessary that the project team works in close co-operation with experts in the Member States. This means that it would be very helpful if the team gets appropriate information and backgrounds of each country on the existing statistics in this field and the possibilities to use them on a harmonised base in the EU.

In this context it would be welcomed very much if you are willing to discuss the matter with one of the members of the project team and to answer the questions in this paper. Besides that it will be appreciated if you can provide additional statistical information, reports, studies etc related to this subject. This can help the project team to achieve a sharper understanding of the backgrounds and constraints of the statistical data.

The results of the project will be presented to Eurostat in 2002. Of course you will be informed on this. At least you will receive the report with the results of the project.

Questions for the interviews

In this paper a number of questions for the interviews is presented. The questions are grouped in such a way that all relevant points can be discussed in an efficient and logic order. So there are questions raised on the availability of statistics in general (1), on additional data (2), on the specific inputs and activities in the EAA (3), model calculations (4) as well as the possibility to send additional data to Eurostat (5). Besides that it is possible to finish the interview with the provision of suggestions that can be helpful for the project (6).

1. Agricultural statistics in general

The project team has information per country available on EAA (sector account); on EU-FADN/RICA (farm results) and SGM (standard gross margins).

1.1 Do you think it would make sense to improve the inputdata and to allocate the inputs to activities?
1.2 If yes, do you think it is useful to use SGM and or FADN for that?
1.3 Are there other methods to improve the EAA inputdata?
1.4 Are the FADN data used to construct the EAA and SGM?
1.5 Are there comparisons made on national level of the data in EAA, FADN, SGM and perhaps also Price statistics? Can you provide results on such comparisons?

2. Additional data and statistics

2.1 Allocation of inputs to sectors/activities in EAA

2.1 Do you have data on the value of production and intermediate consumption of specific sectors/activities of agricultural production? For example sector accounts for specific parts of the farm sector (f.e. based on input- output studies).
2.2 Are such data produced on a regular base or incidentally (see also question 5)?

2.2 Cost of production calculations

2.3 Are there costs of production calculations of some products or activities? For example based on technical data, (national) FADN, or other databases (Commodity Boards).
2.4 Are these results based on calculations/databases or estimates by experts?
2.5 Are such data produced on a regular base or incidentally (see also question 5)?

3. EAA+: availability of individual activities and inputs

3.1 Activities

In the frame of the project an adapted list of activities and inputs is fixed. This list is mentioned EAA+ (with a provisional status): Appendix 4 [1]. The biggest difference with the old EAA list is the use of activities instead of outputs. An activity (for example keeping of milk cows or growing wheat) might lead to more than one output (respectively milk and beef and straw and wheat). An output (beef) might be produced by different activities (keeping milk cows, keeping bovine cattle for meat production). Because it might be difficult to split the cost of an activity into the resulting outputs, we assumed that it might be easier to calculate the inputs per activity instead of inputs per output.

[1] Appendix 2 in this report.

3.1 Is it less complicated to calculate inputs per activity instead of inputs per output in your country?

The following questions should be answered in figure B12.1:

3.2 Can you agree with the list of activities (annex 4) and if not, which items should be skipped or added and why (colum 2)?

3.3 Are the data yet available (column 3) and if not, could this data be assembled in the future (column 4)?

3.4 Are (in general) data available about the inputs used for this activity (column 5)?

Activity	Items skipped or added and why?	Total value yet available?	If data not yet available, willing and possible to record data in the future?	Inputs per activity available?
Cereals				
Industrial crops		X		
Forage plants				
Vegetables				
Plants and flowers				
Potatoes				
Fruits		X		
Wine		X		
Olive oil		X		
Other crop products				
Bovine animals (meat)				
Bovine animals (milk)				
Pigs and poultry				
Sheep, goats, equidea and other				

Figure B12.1 Availability of detailed info on activities
X = No changes proposed, so already available (EAA = EAA+).

3.1 Inputs and inputs per activity

The following questions should be answered in figure B12.2:

3.5 Can you agree with the list of inputs (annex 4) and if not, which items should be skipped or added and why?

3.6 Is the total value yet available?

3.7 Is the value per activity available and if not could this data be assembled in the future?

3.8 If not, is the INRA model estimation (annex 2, 3 and 4) satisfactory (see also question 4)?

Please fill in figure B12.2.

Name of input	Items skipped or added and why?	Total value yet available?	Available per activity?	If not, willing to assemble?	If not, estimation model satisfactory
Seeds and plants		X			
Energy					
Fertiliser					
Pesticides					
Veter. costs		X			
Feedingstuff (incl. own forage production)					
Maintenance of Material		X			
Maintenance of Buildings		X			
Agricultural services		X			
Other goods and services					

Figure B12.2 Availability of detailed information on inputs
X = No changes proposed so already available (EAA = EAA+).

3.9 At the moment a split is made into changes in price and changes in quantity for the items in the EAA. Do you foresee any problems in the split of activities and inputs into price and quantity for the above mentioned items. If so, for which inputs and activities and why?

4. Model calculations

The project team made already some calculations based on an INRA model. The INRA model calculates the costs of specific inputs in relation to the value of specific outputs. The model uses the individual farm data in the EU-FADN. It allocates costs to outputs by a (advanced) regression with the output as independent and the inputs as dependent. It is extensively used in France (for example for SGM calculations) but it has only recently been adapted for use in the EU. Because the model is based on FADN, the use for the EAA is not without problems (definitions differ, bookkeeping year are not equal to calendar year etc).The model estimations could be used if there is no national data available about (some) inputs per activity.

In annex 1 to 3 some results of the model are presented based on the year 1997:

- Annex 1 Total output FADN as a percentage of output EAA [1]
  This table shows the representativity of the FADN for the EAA. It is calculated by dividing the (weighted) total production in FADN by the output in EAA. Results for

[1] Appendix 9 in this report.

some outputs may be disturbed by differences in definitions between EAA and FADN.
1.22 for example means that for Austria the weighted production of Soft wheat in FADN is 22% higher than in EAA.

- Appendix 2Total costs of inputs of INRA model as a percentage of the input costs in EAA [1]
The INRA model results in percentage of a specific costs per euro output. These percentages have been multiplied by the value of all individual output in the EAA. These totals have been added together to reach the calculated total input used for the production of the outputs in the EAA. In annex 2 this calculated total is presented as a percentage of the total input value in the EAA.
0.77 for example means that given the production values of EAA, the INRA model expects 23% lower use of seed that the actual use of seed in EAA.

Annex 3 Inputs per output, *Country,* 1997

4.1 Do you think that the model could be used for those activities and inputs for which there is no other data available in your country?

4.2 Do you have reasons why use of FADN and the INRA model might lead to unsatisfactory results (for example for the items where percentages are above 125 or below 75 in appendix 1 and 2)?

4.2 Specific questions per Member State

5. Possibilities to send data to Eurostat

If you have data available for Eurostat about the inputs per activity:

5.1 At what time are the data available (time lag) and for which period (1973 untill 2000)?
5.2 On which conditions can it be used by Eurostat (privacy, extra costs, co-operation of other institutes)?
5.3 Would it be possible to send this data to Eurostat before the end of January 2002?

6. Other points/miscellaneous

6.1 To finalise the interview the expert is invited to give suggestions, names and addresses of institutes/organisations and persons, which can provide additional, (more) adequate information etc.

[1] Appendix 8 in this report.

6.2 Perhaps it is also possible to receive reports, statistical overviews (perhaps on diskette or by e-mail) etc.

Example annex 3 of questionnaire (based on an old version of the INRA model) Inputs per output INRA model, Belgium 1997
Inputs per 1,000 euro output

----	SoftW	Barley	Sugarb	Othin	Veget	Flowp	Potat	Fruit	Othcr	CATTL	PIG	Poult	MILK	Eggs
FEED	0	0	0	0	0	0	0	0	0	291	597	723	121	476
Veterinary	0	0	0	0	0	0	0	0	0	111	56	54	39	42
SEED	17	60	92	20	64	127	100	24	86	15	4	1	19	1
FERTIL	167	301	34	133	69	39	29	29	9	57	-1	-1	41	3
CROPS_PROT	138	196	80	136	24	16	125	46	12	7	3	1	11	1
ENERGY	35	40	-4	72	133	107	-18	23	49	23	13	13	31	20
MAINT_MAT	79	46	2	41	11	12	27	25	1	29	5	4	36	8
MAINT_BUIL	16	5	-3	9	19	17	-4	5	4	5	5	6	6	5
SERV	24	205	158	49	21	11	84	31	36	48	14	13	49	22
OTHER_IC	97	22	5	33	109	131	2	150	185	39	11	7	31	15
Tot IC	573	875	363	494	450	461	345	333	382	625	706	821	383	593

Inputs (euro) per hectare, livestock unit or kg

	SoftW ha	Barley ha	Sugab ha	Potat ha	CATTL LU	PIG LU	POULT LU	MILK kg	EGGS LU
FEED	0	0	0	0	201	434	1,089	113	445
veterinary	0	0	0	0	77	41	82	36	40
SEED	15	49	267	308	10	3	1	18	1
FERTIL	156	246	98	88	39	-1	-1	38	3
CROPS_PROT	129	161	231	383	5	3	1	10	1
ENERGY	33	33	-13	-55	16	9	20	29	19
MAINT_MAT	74	38	7	83	20	4	6	34	7
MAINT_BUIL	15	4	-8	-12	3	4	9	5	5
SERV	23	167	459	259	33	10	20	46	21
OTHER_IC	91	18	14	6	27	8	11	29	14
Tot IC	536	715	1,054	1,059	432	514	1,237	359	556

Appendix 13 Description of available national datasources

Austria

Only SGM are available in Austria. These have been used in the second sheet of the database. No other cost of production data is available.

Belgium

In Belgium no other costs of production calculation database exist. The SGM calculations are the only available source. The Belgian LEI (responsible for SGM, FADN and EAA) offered to make a gross margin calculation based on the year 1997. They were however not able to realise this within the time limits of this project.

Finland

Union of Rural Advisory Centres (http://www.maaseutukeskus.fi/inenglis.htm) calculates production costs per activity (per hectare or per head) in different regions in Finland using various data material (like dairy or pig recording data from specialised farms). The variable costs of this data is used in calculating Finnish SGM, but the cost calculations of most important activities also include all other cost items. Production costs per activity are calculated at different yield levels and at different level of variable inputs. The same input use specification is assumed in all agricultural support regions. The purpose of the calculations is to show farmers the profitability of agricultural activities at different yield and subsidy levels. The calculations are for advisory purposes and they are not checked for consistency with EAA or FADN. For this reason the data is not included in the database prepared in this project.

France

In France accounts per type of farm are developed. This accounts use RICA and INRA model. Every year these accounts are built and adjusted on national accounts. Professional agriculture (FADN field) and no professional agriculture are distinguished. In the field of professional agriculture, accounts are made for 13 type of farm:

- cereals, oil seed, protein crop;
- other field crops;
- horticulture and flowers;
- wine with label of origin;
- other wine;
- fruit tree cultivation;
- bovine-milk;

- bovine-meat;
- bovine-mixed;
- sheep and other herbivore
  of which sheep;
- off-area;
- mixed farming;
- other type of farming.

The approach adapted in France is different of the context of our project and it is not possible to harmonise definitions.

Germany

In Germany SGM ('Standarddeckungsbeitraege, StDB') are published each year [1]. These StDB are detailed per activity (crops and animals), in relation to the yield (5 classes) and the region (Bundesland). In relation to that for each Bundesland the yield class per sub-region is given. The variable costs per unit are split to some extent:
- seed and plants, fertiliser, pesticides, machinery and other costs for arable crops;
- seed and plants, fertiliser and pesticides, energy (for heating), electricity and water, small materials, marketing costs, insurance , machinery and other costs for horticultural products (with some other specification for wine);
- compound feed, forage, veterinary, machinery and other costs for animals. So a lot of detailed information is available.

The data are based on 'Testbetriebe', a network of farms that is not equal to the network of farms used for FADN. So it might sometimes be difficult to harmonise data with FADN and EAA. The StDB data are (actually) not the same as the data, which are presented to Eurostat by KTBL. Eurostat receives the SGM data on the base of the EU classification, while the StDB are based on a national classification system. However using the data of yield 'class 3' should give a result comparable to the national result for a crop or animal presented to Eurostat. From 2002 onwards the EU classification will be used also by KTBL.

Each year a publication on StDB is, on request of the Ministry (internet: www.ktbl.de).

Greece

In Greece no cost of production data is available at all.

[1] KTBL, Standard Deckungsbeitraege 1999/2000; Norbert Sauer, Ralf Uhte, Darmstadt, 2001.

Luxembourg

No national data is available in Luxembourg. Experts explain that some details exist in the context of SGM calculations, but this data is not immediately available. They want to co-operate with Eurostat, but they have not resources (human) for this.

Portugal

In Portugal the Ministry of Agriculture via its Cabinet for Planning and Agrofood Politics (GPPAA) between September 1997 and April 1998 conducted a farm survey (Model de Base Microeconomica, MBM) referring to 1996/97 campaign, to obtain crop production costs which would serve, among others, the goal of simulating and evaluating the impacts of CAP. The sample chosen was representative of the mainland agricultural crop production (1997 basis), thus Islands were not surveyed. Data was collected for 20 groups of products and 776 technologies. Some detailed information is already published (a written report and a cd-rom containing Excel data files) for 191 technologies of production and 15 crop products. Data refers to Agrarian Regions, which are different than usual EU region. The list of products does not perfectly match the output list used in the project. As regards items of intermediate consumption data on energy and on maintenance of both materials and buildings is not available.

Spain

In Spain there exists a yearly survey on as much as 7,000 farms devoted to calculate cost of production and technical and economic indexes per crop on a regional basis. For each crop, 20-30 farm in each region are surveyed on average. The data is available for several crops. The average value as well as the minimum and maximum values are presented for all the information, thus providing the range of variability. Unfortunately this data was not made available to the project team. All items of intermediated consumption are collected, while no information is available at the moment on the list of agricultural products covered by the survey. To make them suitable for AGRIS it should be made clear, beside the check on the output list available, the representativity of the sample and consequently the weighting factor to be used.

Sweden

In addition to SGM, activity based 'Synthetic' cost calculations ('Lönsamhetskalkyler för lantbrukets produktionsgrenar' - 'profitability calculations') are available. The calculations are made for main products only for 5 different regions. They are based on specific assumptions and documented methodology ('Productionsgrenskalkyler - Metodbeskrivning'). The data material used in the calculation represent specialised farms. The calculations are available from 1990 until 1997. There is little systematic comparison between EAA or efforts to ensure the consistency. For this reason the data is not included in the database prepared in this project.

Appendix 14 Interviewed persons

Austria
Martin Hellmayr (LBG Wirtschaftstreuhand- und Beratungsgesellschaft m.b.H)
Martin Kniepert (Institut fuer Wirtschaft, Politik und Recht, Universitaet fuer Bodenkultur, Wien)
Christina Mayer(Statistik Austria)

Belgium
Roger de Becker
Dirk van Lierde
Els Demuynck
Els Bernaerts
All from Centrum voor de Landbouweconomie

Denmark
Ole Olsen (Statistics Denmark)
Boerge Nielsen (SJFI, The Danish Research Institute of Food Economics)
Henning Porskrog (SJFI)
Steffen Möllenberg (SJFI)
Mona Kristoffersen (SJFI)

Finland
Martti Kankaanpää (Statistics Finland)
Arto Latukka (Agrifood Research Finland, MTT)
Olli Rantala (MTT)

France
Emmanuel Chantry (Agricultural ministry, SCEES)
Maurice Desriers (Agricultural ministry, SCEES)
Sylvain Moreau (INSEE)
François Rageau (INSEE)

Germany
Volker Appel
Manfred Ehlerding
Sibylle Hilgenstock
Axel Lipinski
All from Ministry of Consumer protection, Food and Agriculture (BMVEL)

Greece
Mr. Anastasios Nikolaidis (National Statistical Service of Greece, NSS)
Mr. Pavlos Tonikidis (NSS)
Mr. Apostolos Vainas (Hellenic Republic Ministry of Agriculture)
Mrs. Elektra Tzanellou (Hellenic Republic Ministry of Agriculture)

Ireland
Mary Smith
Gery Brady
Richard Mahon
Keith McSweeny
All from Central Office of Statistics Ireland

Italy
Mr. Giorgio Seroglia (INEA, National Institute for Agricultural Economics)
Mrs. Antonella Finizia (ISMEA, Institute for Services for the Agrofood Market)
Mr. Domenico Ciaccia (ISTAT, National Institute for Statistics)

Luxembourg
Jean-Paul Hoffmann (Service d'Economie Rurale)
Gerard Conter (Statec Luxembourg)

The Netherlands
Ron van der Wal (Central Statistical Office, CBS)
Boudewijn Koole (Agricultural Economics Research Institute, LEI)

Portugal
Mrs. Maria José Correia (INE, National Statistical Institute)
Mrs. Ana Cristina Ramos (INE)
Mrs. Maria da Luz Correia (Ministry of Agriculture)

Spain
Mr. Fombellida Aragon, Raimundo (Ministry of Agriculture)
Mr. Picòn Alonso, Eugenio (Ministry of Agriculture)
Mr. Rodrìguez de la Cuétara, Manuel (Ministry of Agriculture)
Mr. Concha Leal, Diego de la (Ministry of Agriculture)
Mr. Fuertes Fisher, Antonio (Ministry of Agriculture)
Mr. San Juan, Carlos (University Carlos III of Madrid)
Ms. Postigo Rodriguez, Ma José (Ministry of Agriculture)

Sweden
Gunnar Larsson (Statistics Sweden)
Hans Jönrup (Swedish Board of Agriculture)

UK
Jim Holding (Department for Environment, Food and Rural Affairs, DEFRA)
Barbera Boize (Department for Environment, Food and Rural Affairs, DEFRA)