

COMMISSION OF THE EUROPEAN COMMUNITIES

SEC(92) 2084 final

Brussels, 12 November 1992

REPORT FROM THE COMMISSION TO THE COUNCIL

**on the operation of the EUROFARM project
(according to Article 13 of Regulation 571/88/EEC)**

Council Regulation (EEC) No 571/88 of 29 February 1988 is on the organization of Community surveys on the structure of agricultural holdings between 1988 and 1997. The EUROFARM project was designed and set up in close cooperation with the national statistical services to process the data submitted by the Member States. Under Article 13 of Regulation No 571/88, the Commission has to submit to the Council a report on the operation of the project by 31 December 1992.

1. Objectives

In order to assess the Community agricultural situation and monitor structural trends, regular statistical surveys are needed covering holdings which have a certain size of agricultural area, some produce for sale or output in excess of certain physical thresholds.

The Member States may select the most appropriate form of survey, either a full survey or a sample survey, taking into account their statistical organization and the need to obtain reliable data at reasonable cost at the various levels of aggregation.

For the purposes of national and Community agricultural policies, a more flexible and more rapid data analysis and dissemination system is needed, in order to lighten the Member States' workload.

One of the aims has been to harmonize and synchronize, in order to obtain Community-level data which are comparable for all Member States. This work has resulted in the publication of the following legal acts:

- 89/651/EEC
Commission Decision of 26 October 1989 relating to the definitions of the characteristics and to the list of agricultural products for the surveys on the structure of agricultural holdings during the period 1988 to 1997

- 89/652/EEC
Commission Decision of 26 October 1989 establishing, for the surveys on the structure of agricultural holdings between 1988 and 1997, deadlines for the communication of survey results to the Statistical Office of the European Communities

- 89/653/EEC
Commission Decision of 26 October 1989 establishing, for the surveys on the structure of agricultural holdings between 1988 and 1997, a Community programme of tables to be stored in the Tabular Data Bank (BDT) of the Eurofarm system.

The following Decision was in force until the results of the 1989/90 survey were processed.

- 85/377/EEC
Commission Decision of 7 June 1985 establishing a Community typology for agricultural holdings.

This Decision has been adapted to the new list of characteristics, and from the 1993 survey onwards the typology will be established according to the adapted Decision.

2. The EUROFARM project

The EUROFARM project was designed to process individual data, compile the planned standard tables and produce analyses in response to ad hoc requests.

2.1 The components of the EUROFARM project

There are two important aspects to the project:

- the processing software, and
- the system of databases.

2.1.1 The processing software

EUROFARM is designed to process individual data sent by the Member States and respond to requests for statistical analysis by producing tables classifying farms according to a set of characteristics and size categories. These tables have to comply with rules regarding the consistency of results and with rules to safeguard confidentiality.

The processing software is subdivided into processing units which each correspond to a stage in the process of compiling and disseminating tables based on individual data. Annex I contains a detailed description of the processing units.

2.1.2 The system of databases

Depending on the nature of the data and confidentiality requirements, two sets of databases are available with the data on the agricultural holdings and their structures.

2.1.2.1 The individual database

The individual database (BDI) contains, by country, all the data transmitted on either the total number of holdings or the representative sample of holdings covered by the survey. The individual data are depersonalized: no information is sent on the name or address of the holding.

2.1.2.2 The tabular database

The tabular database (BDT) contains all the tabular data compiled according to the programme of tables referred to in Decision 89/653/EEC. Because confidentiality has to be guaranteed, the BDT is in two parts:

- The private BDT, with protected access, which contains all the raw data from the tabulation programmes and the results which have not yet been approved by the Member States. Each Member State has a special domain to which it alone has access.
- The public BDT, which may be consulted and contains the results approved by the Member States after processing to ensure confidentiality and reliability. The public base also includes the results of the surveys for the period 1975 to 1987.

2.2 Progress report on development work

A software package has been developed to carry out the tasks in the processing units described in Annex I. This package has three components:

- COBOL for special programs,
- SAS for compiling tables,
- ACUMEN for storing and disseminating tables.

The latest version of the package can be used to:

- obtain validated individual results,
- calculate tabular results,
- carry out processing to ensure confidentiality and reliability,
- disseminate authorized tables.

The pioneering nature of the EUROFARM project, particularly with regard to safeguarding confidentiality, has meant that certain processing operations have been defined gradually by the Working Party, with an increasingly rigorous and precise approach to confidentiality requirements; moreover, an examination of the results from the initial individual data has led to a revision of certain strategies.

EUROFARM is to pay particular attention in the future to the dissemination of results to users. This will involve publication of results, interfacing for interrogation and data extraction so that data can be transferred to the user's particular location for analysis. Satisfactory operation will depend on close cooperation between Eurostat and the users involved.

2.3 The supply of individual data

In order to achieve the set objectives, EUROFARM has to receive correct individual data from the Member States.

Eurostat has therefore developed validation programs, which it has sent to the Member States. The Member States were to implement these programs in their computer systems, with the assistance of Eurostat.

3. Member States' participation

The following operations in the EUROFARM system have been defined and implemented in cooperation with representatives of the Member States:

- definition of the system,
- supply of individual data,
- system for subregional data from exhaustive surveys
- approval of the tabular results.

3.1 Definition of the EUROFARM system

For certain aspects, particularly data validation, ensuring confidentiality and assessing reliability, the experience and current practices of the Member States proved invaluable in defining a system to meet their individual requirements.

3.2 Supply of individual data

The data have to be collected, coded and recorded by the Member States, which then have to validate the results using the checking programs provided by Eurostat.

Those Member States which have equipment similar to the Commission's send in individual data containing very few errors.

Some Member States have not been able to carry out the individual data checks since they have not been able to put the Eurostat programs into operation, and compilation of the final results is therefore delayed because Eurostat has to carry out a series of checks and corrections.

3.3 Approval of tabular results

The Member States wanted to carry out final checks and give their approval before any of the tabular results were disseminated.

Not knowing how great a workload would be involved, the Member States have been reluctant to commit themselves to a binding approval procedure. To facilitate this work, Eurostat is developing computerized procedures to check the consistency of the results and highlight elements requiring manual intervention.

Since Eurostat, as well as the Member States, lacked experience in setting up such a procedure, the first exercise had to be conducted empirically to define a more complex and comprehensive procedure based on data-processing and data-transmission facilities.

4. Proposed adaptations

In the light of the experience gained, the analysis of initial results or the lessons learned from other projects, various adaptations are proposed for the EUROFARM project.

4.1 Statistical confidentiality

4.1.1 EUROFARM and statistical confidentiality

The statistical confidentiality project is moving towards the processing of confidential data, which include individual data on farms, in a protected area managed by Eurostat, and on Eurostat's responsibility. As the EUROFARM project had been established before the statistical confidentiality project, the former has already put in place measures in agreement with the Member States to protect individual data and safeguard confidentiality.

Bringing EUROFARM into line with the rules and procedures applied in the statistical confidentiality project requires:

- sufficient computer calculation and storage capacity,
- an infrastructure which can cope with the extra capacity,
- adaptation of the EUROFARM project to comply with the new rules.

4.1.2 EUROFARM and the special situation of the FRG

As no Community regulations on statistical confidentiality had existed, the FRG had been exempted from the obligations to supply individual data when the EUROFARM project was established. With the publication of Regulation 1588/90 on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, and following implementation of measures to guarantee confidentiality under the EUROFARM project (measures or developments to which the Federal Republic of Germany has contributed extensively), it may now be proposed that the Federal Republic of Germany participate fully in the EUROFARM system by supplying individual data.

4.2 Validation of individual data

The Member States must validate individual data on their own equipment using the programs developed by Eurostat. Some of them are finding it difficult to set up the Eurostat checking routines on their computers, for the following reasons:

- computer equipment varies from one country to another. Programs therefore have to be adapted, with no guarantee that the results will be satisfactory;
- the programs are designed for high-powered equipment. Lower-powered equipment may not do the checking quickly enough.

It is proposed that:

- the Member States be consulted, so that the problems each of them has run up against can be examined and the peculiarities of their computer systems studied;
- the checking routine be revised to take account of national peculiarities, viz.:
 - technical constraints due to the computer equipment;
 - tolerance thresholds;
 - checks to be carried out.

4.3 Systematic approach to standard gross margins

The gross margin of a farm is the monetary value of its gross output, with certain specific costs deducted. The standard gross margin is the gross margin corresponding to the average situation in a given region for each type of farming. The details of how to calculate standard gross margins are set out in Commission Decision 85/377/EEC of 7 June 1985 establishing a Community typology for agricultural holdings.

Hitherto, the processing of SGMs was justified only for the Farm Accountancy Data Network (FADN). By processing individual data and allowing itself sufficient flexibility to apply SGM coefficients for a period closer to the survey period, EUROFARM in turn is using the SGM coefficients collected for the FADN project to set up a classification of agricultural holdings by typology and economic size.

Currently, the processing of standard gross margins is not incorporated into the EUROFARM project.

It is proposed:

- to tackle the collection and storage of standard gross margin coefficients as an integral part of the EUROFARM project;
- to standardize the methods used to establish the agricultural typology and those of the Farm Accountancy Data Network (FADN);
- if some Member States adopt characteristics defined in greater detail, to plan ways of storing appropriate information which will enable an accurate classification to be produced on the basis of more recent SGM coefficients;
- or, in this same case, to work towards applying the common list of characteristics.

Some Member States expressed concern over the prospect of the EUROFARM project integrating the SGM processing which are already vital components on a technical viewpoint.

4.4 Validating the sample

Obtaining reliable results does not mean carrying out an exhaustive survey on each occasion. The sample method is an extremely useful way of carrying out intermediate surveys and producing limited results at certain geographical levels and for certain size categories in the base surveys.

The base survey is aimed in principle at compiling and validating a new representative sample. Some Member States have already had help from Eurostat with the compilation of a representative sample.

It is proposed to give the same assistance to other Member States which would like help with fixing and validating a sample drawn from an exhaustive survey.

4.5 Revision of the programme of standard tables

The programme of standard tables includes a number of detailed analyses by: geographical level (country, region, district), area status (all agriculture, less-favoured areas and mountain areas), size category, type of farming and selected characteristics.

By a variety of means, these analyses produce a distribution of farms according to the criteria applied to the table cells. The more detailed the criteria applied, the wider the distribution and the greater is the tendency for the values obtained to dwindle.

This phenomenon reduces the reliability of analyses based on sample surveys or samples. Moreover, and this also applies to exhaustive surveys, the rules which are applied to safeguard confidentiality involve altering the data: when values are too low, they have to be eliminated.

When analysis is carried out at the most detailed level, sample surveys produce results based on a non-representative population and all the surveys in general tend to yield results which the procedures for guaranteeing confidentiality are liable to disguise by zero values.

The programme of standard tables could be revised in a number of respects:

- reducing the number of size categories in line with the population of the geographical area observed,
- eliminating certain rows in the tables, particularly the labour force characteristics for distributions by type of farming,
- making more use of the ad hoc table procedure for specific analyses,

4.6 Maintaining the EUROFARM system

The EUROFARM system is scheduled to remain in operation at least until the next base survey. Changes in the operating systems and programs used, changes in the list of characteristics and adaptations to meet requirements and the users' working methods mean that the whole information system needs constant maintenance. EUROFARM's requirements are estimated at 2.5 man-years per annum, not counting new developments.

4.7 Promotion of the EUROFARM system

Direct access to the EUROFARM information system is reserved for preferential users: Commission departments, national statistical institutes and Ministries of Agriculture.

Eurostat handles requests for information from other persons or organizations. There is a charge for the information supplied, and, in order to avoid competition with equivalent services offered by the Member States, these charges should be drawn up in agreement with the Member States.

In order to attract a wider circle of potential customers, there should be a promotion campaign for the EUROFARM system.

Special training schemes would give preferential users a greater knowledge and experience of the use of the information and its access system.

All the Commission's information channels should be used to publicize the EUROFARM system.

4.8 Organization of the 1999/2000 base survey

The 1989/90 base survey was organized over the period from 1988 to 1991 to fit in with national timetables. This staggered arrangement can only delay the availability of the final results, and the aggregates compiled at Community level reflect national situations at different points in time.

In order to avoid these differences in the future and reduce the observation period and the time taken to produce the results, it is proposed that, before the Member States finally set the timetable for their future surveys, the 1999/2000 Community base survey be organized over a period restricted to one year.

4.9 Setting up a system for subregional data

Areas or subregions which meet the eligibility criteria laid down are eligible for Community financing under objectives 2 or 5b (rehabilitation of industrial regions in decline or development of rural areas). The regional statistics which are currently available do not provide appropriate information on these areas and subregions.

Data at a more detailed geographical level can be used to compile statistical aggregates relating to the objectives and areas or subregions concerned. Moreover, the 1989/90 base survey can be used to compile aggregates from data relating to all farms. The results of exhaustive surveys provide appropriate information for a common system of subregional data.

The Member States have two options for compiling this information:

- having sent Eurostat a more detailed code on geographical location and individual data on all farms, the Member States then authorize Eurostat to carry out the requisite compilations. The criteria which were established to ensure confidentiality in the EUROFARM project also apply to this information;
- the Member States perform the required analyses using individual data on all farms.

The Member States have said that they are prepared to provide Eurostat with the information requested according to whichever option is suitable for them.

4.10 ASSESSMENT OF RESOURCE REQUIREMENTS

	1993		1994		1995		1996		1997	
	Manpower Resources	Cost/ECU	Manpower Resources	Cost/ECU	Manpower Resources	Cost/ECU	Manpower Resources	Cost/ECU	Manpower Resources	Cost/ECU
Statistical confidentiality	1 man-year	120 000								
Data validation	0.5 man-year	60 000								
Incorporation of SGMs	1 man-year	120 000	0.5 man-year	60 000						
Sample validation	0.5 man-year	60 000								
Help with sampling	0.5 man-year	60 000								
Revision of programme of standard tables										
- redefining	0.5 man-year	60 000								
- programming			1 man-year	120 000						
System maintenance										
EUROFARM	2.5 man-year	300 000	2.5 man-year	300 000	2.5 man-year	300 000	2.5 man-year	300 000	2.5 man-year	300 000
Promotion	0.25 man-year	30 000	0.25 man-year	30 000	0.25 man-year	30 000	0.25 man-year	30 000	0.25 man-year	30 000
Infrastructure		1 Mio ECU		1 Mio ECU				1 Mio ECU		
Total manpower resources	6.75 man-year	810 000	4.25 man-year	510 000	2.75 man-year	330 000	2.75 man-year	330 000	2.75 man-year	330 000
Total infrastructure		1 Mio ECU		1 Mio ECU		0.5 Mio ECU		1 Mio ECU		0.5 Mio ECU

ANNEX I

The EUROFARM processing units

The EUROFARM project has been subdivided into processing units which each correspond to a stage in the process of compiling the results.

1. Security

A fundamental concern of Eurostat and the Member States is to protect the confidentiality of information.

As EUROFARM receives individual data and compiles tabular results to which access is restricted prior to dissemination, the project's data-processing environment has been organized so that sensitive information is protected against attempted access by unauthorized persons.

2. Validation of individual data

Data on farms are collected, coded and recorded on magnetic media by the Member States. Before any other processing takes place, the individual data undergo syntax and plausibility checks to establish their authenticity and quality. Using error-free data as a basis for tabular analysis minimizes the risk of aberrant results.

3. Compilation of standard tables and checks

On the basis of validated, error-free, individual data, the standard tables are calculated according to the tabulation programme laid down on by Decision 89/653/EEC of 26 October 1989. They then undergo consistency checks within and between tables.

4. Safeguarding confidentiality and assessing reliability

The processing to safeguard confidentiality detects and eliminates or flags the following cases in which individual farms could be identified:

- values or number of farms too low,
- dominance,
- derivation or calculation by difference of values which would have been eliminated under the above two cases.

As far as possible, the programs generally try to eliminate the calculated values in the cases which are detected, in order to limit the work involved in settling individual problem cases of divulgence. A more comprehensive procedure will be introduced following the experience gained from the processing of the initial results.

The processing to assess reliability establishes an indicator of the quality of the results, taking account of the error introduced by sampling and by the procedures to ensure confidentiality.

5. Approval procedure

Prior to publication, the tabular results compiled during the previous stages are sent back to the Member States, which have to comment on any remaining cases where confidentiality is not guaranteed and no manipulation has been carried out by the programs. They are asked to give their authorization for the data to be disseminated, and to make any appropriate comments.

Several methods of approval are proposed:

- off-line approval based on hard copy results;
- off-line approval based on results transmitted on magnetic media using the software adopted by Eurostat;
- on-line approval based on a consultation of results in each Member State's national domain in EUROFARM.

The approval procedure will be implemented in cooperation with each Member State, using a method which is suited to each State.

6. Consultation of the disseminated tabular results

The tabular results authorized for dissemination are entered into the public tabular database (BDT), where all authorized users may consult them.

Appropriate menus enable users to access the results required.

7. Historical data

The FSSRS (Farm Structure Survey Retrieval System) was used to disseminate the results of the 1975, 1979/80, 1983, 1985 and 1987 structural surveys, which it is planned to transfer to the EUROFARM system. As the structure of the FSSRS tables differs from that of the EUROFARM tables, conversion will involve moving from one format to another in order to incorporate the historical data for use in EUROFARM.

The results of the surveys on fruit trees (FRUCTUS) and winegrowing (VITIS), which are likewise contained in the FSSRS, are also being transferred to databases from which they will be disseminated.

8. Compiling ad hoc tables

A procedure for compiling ad hoc tables has been devised to meet specific requests which cannot be met by the standard tables. This procedure always involves producing tabular results from individual data. The tables thus compiled have to comply with the confidentiality rules and be approved by the Member States concerned before they can be sent to the requester.

9. Publications

Only certain preferential users have access to the full results (public BDT). They are the Commission DG's, the National Statistical Offices and the National Ministry of Agriculture.

Publication of the main results involves selecting the appropriate data and presenting and disseminating them on various media: paper, CD-ROM and diskette. The data selected may be a selection of main data or a selection in response to a specific request.

A booklet containing the main results is being prepared.

10. Data transfer

Using this facility, a set of data can be extracted from the tabular database and transferred via the telecommunications network to a suitable location close to the user, or to a PC. The user is then

free to process the information to suit his own purposes and use specialized software to analyse the data.

11. CD-ROM

CD-ROM (Compact Disc Read Only Memory) is a form of publication on an electronic medium. It can provide access to all the tabular results disseminated by EUROFARM, without the need for special data-transmission links or an infrastructure beyond the basics of PC and CD-ROM reader. The software supplied with the CD-ROM enables information to be consulted in the form it appears in the BDT, data to be extracted in the format used by other PC software, and basic analyses to be carried out. The results of the surveys on fruit trees (FRUCTUS) and winegrowing (VITIS) will also be available on CD-ROM.

12. Monitoring the system

The EUROFARM software has a facility for analysing the use made of the system. Statistics on use of the information system will show:

- subjects accessed
- functions used
- access time
- attempts to violate protected files
- access to protected files.