Typology, classification and characterization of farms for agricultural production planning

C. J. Álvarez-López*, J. A. Riveiro-Valiño and M. F. Marey-Pérez

Departamento de Ingeniería Agroforestal. Universidad de Santiago de Compostela. 27002 Lugo. Spain

Abstract

Agricultural activity in Galicia, North West Spain, is carried out on farms that are characterized by a diversity of land use and production models, a variety of farm sizes, and considerable geographical dispersion. Any attempt of Agricultural Production Planning aimed at characterizing production models requires a method of analysing data and obtaining technical-economic results from farms in the different areas. Models based on average statistical data are limited because they represent farms that do not exist in reality. This study develops a methodology to characterize and group dairy farms into different types according to the following basic variables: land use, size classes and production systems. The information used in this study was microdata from the 1999 Census of Agriculture. The methodology developed was also applied to microdata from the 1989 Census of Agriculture, thus obtaining significant information about the evolution of agricultural activity. The tools used in the analysis were Microsoft Access and Excel, and an application that was developed using Microsoft Visual Basic. The methodology presented can be used to analyse the evolution of the sector or to model future trends.

Additional key words: agricultural census, dairy farm types, Galicia, rural development.

Resumen

Tipificación, clasificación y caracterización de explotaciones agrarias para la ordenación productiva agraria

Galicia (NO de España) presenta una agricultura con explotaciones caracterizadas por su diversidad de aprovechamientos y modelos productivos, gran variedad dimensional y dispersión territorial. Cualquier intento de ordenación productiva agraria para caracterizar los modelos productivos en las distintas áreas territoriales necesita disponer de una metodología de análisis de datos y obtención de resultados técnico-económicos a partir de las explotaciones existentes. Dado que los análisis basados en datos estadísticos medios son limitados al representar explotaciones inexistentes en la realidad, se ha desarrollado una metodología que agrupa y caracteriza explotaciones lecheras en tipologías de acuerdo con sus variables básicas: aprovechamientos que las integran, rangos dimensionales y sistemas de producción. Nuestra fuente de información son los microdatos del Censo Agrario de 1999 y la metodología desarrollada también se aplicó a los microdatos del censo de 1989, obteniéndose información significativa sobre la evolución de la actividad agraria. Las herramientas utilizadas fueron Microsoft Access y Excel, además de desarrollar una aplicación utilizando Microsoft Visual Basic. La metodología expuesta puede ser usada para analizar la evolución del sector o para modelizar tendencias futuras.

Palabras clave adicionales: censo agrario, desarrollo rural, Galicia, tipologías de explotaciones lecheras.

Introduction[1]

Agricultural production planning is a process for the spatial organization of agricultural and forestry production that allocates particular uses to preferential land areas in an attempt to attain sustainable development of a territory by optimizing the agricultural production systems according to environmental concerns, socio-economic and structural conditions (Riveiro *et al.,* 2005). Since the 1930s (Storie, 1933), mathematical

[1] Abbreviations used: C (size classes), CU (cow units), ESU (economic size units), EU (European Union), fadn (farm accounting data network), GIS (geographic information system), LU (labour force per unit), MDM (mixture of distributions model), SGM (standard gross margin), TF (type of farming), UAA (average power of tractors per unit area).

methods have been developed to allocate agricultural uses to land areas. Since then, different research lines have been followed (Rossiter, 1996). There are two main lines of research that differ in the unit of analysis. In the first line, the land is analyzed and the analytical results are extrapolated to farms. In this approach, the first methods used were multi-criteria evaluations based on mathematical programming (Voogd, 1983). Later, cellular automata (Parker *et al.*, 2003) and heuristic models were used because of lower computing costs and the versatility of the solutions from such models (Nalle *et al.*, 2002; Boyland *et al.*, 2004). Other authors combined the mathematical basis with social participation (De Wit and Van Keulen, 1988; Leitner *et al.*, 2002; Snyder, 2003). In the second line of research, the farm is taken as the unit of analysis (Loftsgard and Heady, 1959) and the results are used to determine crop areas allocated to each species (Duloy and Norton, 1983; Hwang *et al.*, 1994). In this approach, the results can be applied to a whole territory (Glenj and Tipper, 2001) or used for strategic land use planning (Carsjens and Van Der Knaap, 2002).

Given that the previous methods were developed to model the behaviour of single farms, it is necessary to determine the representativeness and characteristics of farms for a given territory or sector (Thenail and Baudry, 2004). The models used to classify farms simplify into three approaches (Kostov and McErlean, 2006). The first is to choose an average farm and assume that the other farms are linearly related to the characteristics of the chosen farm. In the second approach, more than one farm is required to adequately represent the farm population (and to minimize variability). Usually, multivariate statistical classification techniques are used to conduct such a selection, mainly hierarchical clustering (Everitt, 1993). The difficulty with this type of classification lies in deciding the number of clusters and the representativeness of each cluster. In the last few years, models have been developed that solve the problem of the number of groups, such as the mixture of distributions model (MDM) (Kostov and McErlean, 2006). A third approach consists in determining the evolution of farms from official statistical indicators, mainly based on economic criteria. Reidsma *et al.* (2006) used data from the European Farm Accounting Data Network (FADN) to model the behaviour of European agriculture. At the regional level, the farms present in the network seem to have both geographical and economic representativeness (Judez and Chaya, 1999). Chatelier *et al.* (2000) proposed a system for reclassification of dairy farms based on the classes proposed by the FADN. Such a reclassification system conforms better to the variety of possible production processes and to future sustainability.

With a view to grouping relatively homogeneous farms into disjoint classes, the European Union (EU) uses two essential farm characteristics for classifying agricultural holdings: type of farming (TF) and economic size. Both the TF and the economic size are determined on the basis of standard gross margin (SGM). Economic size is determined by the total SGM of farms; TFs are defined in terms of the relative importance of the different enterprises on the farm, measured as a proportion of each enterprise's SGM to the farms' total SGM (RICA, 1988). The economic size of farms is expressed in terms of economic size units (ESU). One ESU corresponds to €1,200 of farm SGM.

A particular farm is assigned to a particular TF when at least 2/3 of farm's total SGM is contributed by that TF. Each TF is made up of four levels of disaggregation: the first level is represented by one digit and corresponds to TF; the second level is represented by two digits and corresponds to principal TF, the third level is represented by three digits and corresponds to a particular TF, and the fourth level is represented by four digits and corresponds to subdivisions of a particular TF.

Table 1 shows the distribution of 15,000 commercial dairy farms in Galicia included in the 1999 Census of Agriculture and classified in terms of TFs and ESU. The table reflects the difficulty of establishing such a classification, given the different TFs, up to 25, and the variation in economic size, which yields an average of 12.66 ESU with a SD of 15.45.

The dairy sector is the main agricultural sector in Galicia, and the most complex in terms of farm structure (agriculture and livestock are integrated), geographical dispersion (the sector is present in dissimilar land areas), size classification, and complexity of production.

Barbeyto (1998) conducted an analysis of the Autonomous Community of Galicia. He used data from farms that belonged to different farm management associations, and classified dairy farms into two groups, a head group and a tail group, in terms of farm net yield per litre of milk. Such a classification can be applied to more than two groups and enables a comparison between different technical and economic indicators for farms with different levels of economic efficiency. This study contributes an innovative approach in that all the farms are analyzed taking into consideration all

Table 1. Classification of commercial dairy farms in Galicia, in terms of types of farming (TFs) and economic size units (ESUs).

TF	Description	ESUs								
		< 10	10-20	20-30	30-40	40-50	50-60	60-70	>70	Total
1430	General field cropping: field vegetables	0	0	0	1	1	0	0	0	2
2011	Market garden vegetables-outdoor	0	0	0	0	0	0	1	1	2
2021	Flowers and ornamentals-outdoor	0	0	0	0	0	0	0	1	1
2022	Flowers and ornamentals-under glass	0	0	0	1	0	0	1	0	2
4110	Grazing livestock: dairying	4,375	6,986	1,753	413	117	60	26	38	13,768
4120	Grazing livestock: dairying and cattle rearing	111	205	35	7	1	1	0	5	365
4310	Grazing livestock: cattle-dairying, rearing and fattening combined	83	424	38	7	2	0	0	1	555
4420	Grazing livestock: sheep and cattle combined	0	0	2	1	0	0	0	0	3
4440	Grazing livestock: various grazing livestock	2	16	1	0	1	2	0	0	22
5011	Monogastrics: pig rearing	0	0	0	1	0	2	1	3	7
5012	Monogastrics: pig fattening	0	0	1	3	3	3	2	8	20
5013	Monogastrics: pig rearing and fattening combined	0	0	1	0	3	1	3	5	13
5021	Monogastrics: laying hens	0	0	0	0	0	1	0	0	1
5022	Monogastrics: poultry-meat production	0	1	1	1	0	0	0	0	3
5032	Monogastrics: pigs, poultry and other monogastrics combined	0	0	0	0	1	1	1	1	4
6050	Mixed cropping: mainly field crops	0	0	0	0	0	0	0	1	1
6062	Mixed cropping: mainly permanent crops	0	0	0	1	0	0	0	0	1
7110	Mixed livestock: mainly dairying	17	53	4	3	1	0	0	0	78
7210	Mixed livestock: monogastrics and dairying	0	13	15	17	7	2	3	12	69
7220	Mixed livestock: monogastrics and non-dairy grazing	0	0	0	0	0	1	0	1	2
7230	Mixed livestock: monogastrics with various livestock	0	0	1	2	2	0	0	2	7
8110	Mixed crops-livestock: field crops and dairying	0	7	2	0	0	0	0	0	9
8120	Mixed crops-livestock: dairying and field crops	4	35	3	1	1	0	0	1	45
8220	Mixed crops-livestock: permanent crops and grazing livestock	0	2	1	0	0	0	0	0	3
8232	Mixed crops-livestock: various mixed holdings	2	10	2	1	1	0	1	0	17
	Total	4,594	7,752	1,860	460	141	74	39	80	15,000

Source: 1999 Census of Agriculture (INE, 2003).

the variables that affect their technical, economic and environmental viability (Zhen and Routray, 2003). Thus, farms are aggregated into groups based on the values of each variable and the representativeness of each group is related to the number of farms in which a variable occurs.

Another contribution in Galicia was made through the «Study of Agricultural Production Planning» (Xunta de Galicia, 2004). A model was developed for analysis of different agricultural and forestry land use in Galicia, both current and potential, and implemented as spreadsheets based on Excel 2000® (Riveiro *et al.*, 2005). Such

a model produces technical-economic results that can be integrated with other indicators (structural, social and environmental), enabling the development of a decision-making system that allocated priority activities to each territorial area in the region.

This study aims to develop a systematic process for determining different types, production sizes, production systems and locations of farms based on data from the Census of Agriculture (Table 2). The process must allow for the introduction of more specificity in any production planning process. In addition, this study attempts to characterize the evolution of the agricultural activity and to predict production trends by analysing information from the consecutive 1989 and 1999 Censuses of Agriculture. The novelty of the work lies in establishing different groups of holdings considering the total population according to existing data from the agricultural census, which allows accurate characterization.

Material and methods

The working universe of data, considered in this study, was microdata from the 1989 and 1999 Censuses

Table 2. Farm survey fields considered in the 1999 Census of Agriculture (INE, 2003)

Farm identification
— Holder.
— Legal entity and management of the farm.
— Total area and farming system.
— Type of tenure.
— Irrigation.
— Land uses.
— Arable crops and fallow land.
— Kitchen gardens.
— Woody crops.
— Greenhouse and mushrooms.
— Set-aside of arable land under the support system of the EU.
— Successive secondary crops.
— Combined cropping.
— Machinery.
— Livestock.
— Storage facilities for natural fertilisers of animal origin.
— Family labour force.
— Non-family labour force.
— Working days of farm work undertaken by persons not employed directly by the holder.
— Production marketing.

Table 3. Distribution of farms in Galicia according to size classes. The size of a dairy farm is measured in terms of the numbers of cows on the farm

Classes	Class values	Number of farms	
		1989	1999
Not considered	< 10 dairy cows	75,926	22,981
C_1	10 to 24 dairy cows	13,160	10,318
C_2	25 to 39 dairy cows	1,284	3,239
C_3	40 to 54 dairy cows	245	984
C_4	55 to 69 dairy cows	53	241
C_5	≥ 70 dairy cows	92	218
Total considered	> 10 dairy cows	14,834	15,000
Total dairy farms	> 0 dairy cows	90,760	37,981

Source: 1989 and 1999 Census of Agriculture (INE, 1989, 1999).

of Agriculture (INE, 1989, 1999) that corresponded to Galicia. In 1999 there were 270,053 farms that distributed over 696,691 ha of utilised agricultural area.

From among all the farms, dairy farms were chosen. From the 1999 census, we selected farms that sold milk (the data is in the census). From the 1989 census, we selected farms with 10 or more dairy cows (there was no information on farms that sold milk). This selection involved 15,000 dairy farms from the 1999 census and 14,834 from the 1989 census (Table 3).

An operational procedure was developed for grouping farms with common characteristics (type and size) and for determining the spatial distribution of farms (location) in a systematic manner. After the different groups were established, each group could be characterized separately, further information could be searched, and representative samples could be selected to carry out targeted surveys (focused on production processes, yield and consumption).

Tools used in the analysis were Microsoft Access and Excel, and an application developed using Microsoft Visual Basic. The application generated the possible combinations of land uses automatically and gave the composition and number of farms under each combination. This process was conducted in successive stages, and gave intermediate or final results.

Figure 1 gives the procedure for generation of TFs. It consisted of establishing groups of farms composed of a particular combination of productive land uses that caused differences in the technical-economic results of the farm as compared to other combinations. A database was built from the microdata included in the most recent census of agriculture (1999). This database contained information about all the farms and showed a
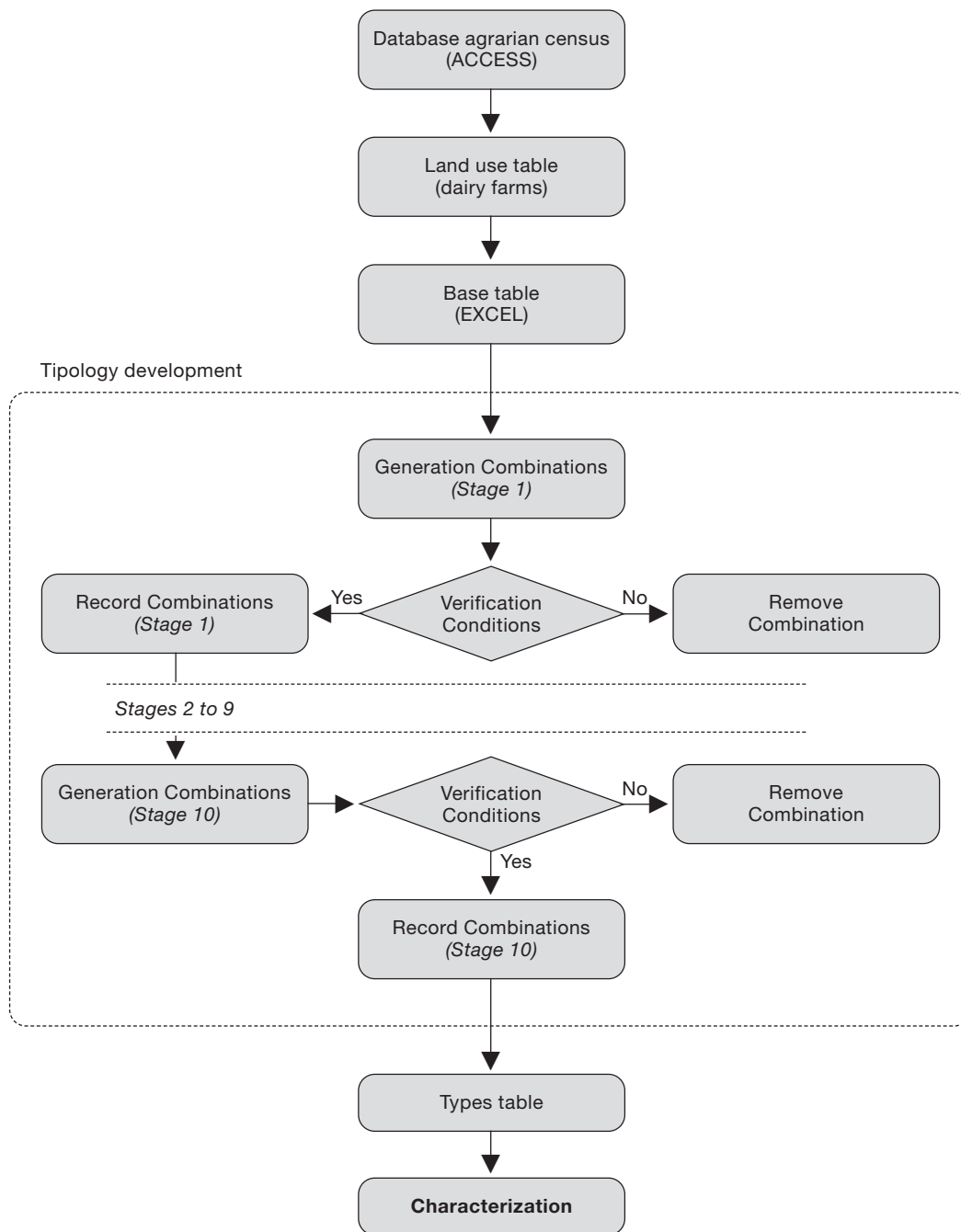
Figure 1. Procedure for the generation of farm types.

standardised structure (Table 2), in keeping with its further use. The database was queried to extract the necessary information about all the commercial farms that contained a specific reference, to define the problem group. Farms in which the reference land use of dairy was present were selected. The reference land use must be present on all farms, although it does not need to be the main land use. The selection of the farms was carried out by making a Microsoft Access query in the database. The data obtained was exported to Microsoft Excel, and a «land use table» generated.

The «land use table» contains all dairy farms with thresholds that sell milk and farms with 10, or more, dairy cows. Fields that were used refer to the different land uses or activities on each farm, the farm size, and the reference or code of each farm (to retrieve the rest

of the data from the census after completing classification into TFs). The land use table reduces the computing time from days to hours. By using the land use table, the computer does not have to process a large volume of unnecessary data during TF generation.

The threshold is used as the primary classification parameter for removing non-commercial farms from the classification. Given the frequency of multi-crop farms in Galicia, particularly in small farms, this process is necessary to obtain a useful long-term methodology.

The process, described above, is essential because multi-crop and self-supply farms are not viable, which has caused a dramatic decrease in the number of farms in the last few years while total production levels have been maintained in Galicia. For example, Table 3 shows figures for the evolution of the dairy sector in Galicia between 1989 and 1999. The use of this threshold is essential to reflecting the current situation of the dairy production sector based on the average data of all farms, including self-supply farms. However, this threshold can be modified or removed.

The data for the variables, considered as reference variables, during the process of TF generation (> 50 fields) were extracted from the «land use table», which was the base table. Then, farm distribution was analysed for each land use according to production size, and a minimum threshold defined based on this distribution. The minimum threshold was considered by the system for generation of the different TFs. In addition, the minimum number of farms required to form a TF was established, with a view to generating representative TFs, i.e., types that are composed of a minimum number of farms.

The critical thresholds and parameters that act as dividers for the classification process are variable, i.e. the user may establish the critical thresholds and parameters deemed appropriate according to the situation under analysis. In this study, the critical thresholds and parameters that act as dividers for the classification process were established based on the information available, with a view to choosing only those farms where each land use or activity was of a minimum commercial size, such that self-supply farms, without commercial significance, were not included in the classification.

The study established minimum thresholds for other land uses or activities, included in the census, that could be combined with dairying: permanent grassland, multi-annual forage crops, corn silage, other pastures, annual forage crops, cattle rearing, beef cattle, potatoes, hardwood, wheat, softwood and mixed forest species.

The thresholds were established to guarantee the commercial size of the farm. The methodology of Riveiro *et al.* (2005) was used to determine the economic size of a farm based on economic yield. The values used to establish the thresholds and parameters required are defined, for the different regions of Galicia in Xunta de Galicia (2004), considering a farm is a dairy farm only if the land use or enterprise combined with dairying does not exceed 50% of farm production.

The minimum value adopted for corn silage, alfalfa or other pasture species was 0.5 ha because any farm with 10 or more cows must have an area of more than 0.5 ha under these crops for feeding cattle. Additional livestock activities such as beef cattle breeding, which were present on many farms, were only considered when there were 5 or more beef cows on the farm.

The minimum number of farms per type was set at 10, and the minimum percentage of farms out of the initial group of farms was set at 10%. By establishing these constraints, which can be modified, unrepresentative TFs could be eliminated at the start of the process by setting a minimum percentage, or at the end of the process by using an absolute value.

Having established these conditions, the process of combination generation was started. The process was implemented automatically in successive stages. The number of stages varied depending on the complexity of the TFs obtained, with a maximum of 10 stages, which was considered an appropriate value for this research.

The results of each stage are stored in spreadsheets, where fields that contain the characteristics of primary and secondary land uses or activities on each farm are qualitative. Conversely, fields that define farm size are quantitative, based on surface area or the number of cows.

Using 10 stages potentially allows grouping of farms, i.e., once all the farms with 10 or more dairy cows were selected, these farms were combined with the other land uses or activities present on the farms that were included in the census, a total of 12 TFs. The next stage would consist in combining each combination of two land uses or activities (dairy cattle + any of the land uses mentioned above) with the other 11 possible combinations. This would yield a total of 132 combinations (12*11). The third stage would consist in relating each of the 132 combinations obtained with the other 10 possible combinations, which would yield a total of 1320 (12*11*10) combinations at this stage. Potentially completing 10 stages would yield 239,500,800 (12*11*10*9*8*7*6*5*4*3) TFs. Assuming that all

Table 4. Number of possible combinations

Order (h)	Elements (n)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Combinations	12	66	220	495	792	924	792	495	220	66
Cumulatives	12	78	298	793	1,585	2,509	3,301	3,796	4,016	4,082

those TFs could exist, the process would become a farm enumeration rather than a classification.

Statistically, the above process is reduced by considering combinations (no element is repeated in the set considered and sets with a different order of elements are not counted). The number of possible combinations is estimated by the following equation:

$$\left(\frac{n}{h}\right) = C_{n,h} = \left(\frac{n!}{h!\,(n-h)!}\right)$$

thus obtaining the following results in Table 4 for 12 land uses, where stage 1 shows the 12 combinations of two elements included in Table 5 and the highest number of combinations for stage 10 corresponds to 4,082 different TFs.

Theoretically, considering the highest potential land uses considered in the census (50) and the highest level of order (10), a maximum combination potentiality is up to $8.22 * 10^9$, and a maximum cumulatives potentiality is $1.08 * 10^{10}$. As the agricultural census only includes 250,073 farms, such a level of disaggregation is excessive. Moreover, Excel does not allow for more than 65,536 rows (combination record). Therefore, in theory, going beyond stage 5 or 6 would be useless.

The work developed for classifying dairy farms becomes redundant at the sixth stage, in which the number of TFs generated (924) covers the total number of farms (15,000). Therefore, why is it interesting to study further stages? It is interesting because of a characteristic of the groupings made by the application: such groupings are not disjoint, i.e., farms can meet the conditions of various TFs at each stage. For this reason, it is essential to achieve higher disaggregation levels to group farms into TFs according to the most significant factors. In this case, after having achieved a high level of TFs, it was established that the most significant factors were the presence or absence of corn silage, and the presence of beef cattle as complementary activities. As the census did not include any beef cattle farms that used corn silage, three TFs were established.

In addition, the possibility of working with 10 stages would be interesting for classifying land uses or activities that are present in a larger number of farms or in wider geographical areas.

Analysis of farms grouped in each TF was conducted by considering different size classes. By default, the application considers five size classes (C_1 to C_5, Table 5), which allows the user to search for curves or trends

Table 5. Percentage of farms for each first-order combination with two land uses

Land use	C_1		C_2		C_3		C_4		C_5	
	1989	**1999**	**1989**	**1999**	**1989**	**1999**	**1989**	**1999**	**1989**	**1999**
Permanent grassland	97.2	90.8	95.5	87.3	96.3	81.6	92.4	78.5	85.9	75.7
Multi-annual forage crops	21.4	34.8	27.3	40.5	20.8	43.4	15.1	37.8	23.9	33.2
Corn silage	0.0	36.5	0.0	45.4	0.0	47.9	0.0	55.3	0.0	62.4
Other pastures	28.2	18.0	21.9	13.2	23.7	12.6	100.0	11.8	19.6	12.4
Annual forage crops	23.6	5.3	26.4	5.8	22.5	6.5	22.6	9.4	32.6	8.4
Cattle rearing	33.6	62.1	54.1	85.7	55.1	88.5	58.5	86.2	66.3	87.6
Beef cattle	3.5	10.1	4.1	5.9	3.3	6.8	7.6	10.2	12.0	8.0
Potatoes	16.4	5.8	13.9	4.5	12.2	3.8	7.6	6.1	5.4	4.4
Hardwood	5.2	2.7	7.3	2.3	2.9	1.2	5.7	3.3	1.1	4.00
Wheat	9.4	27.4	11.5	26.7	7.8	27.3	13.2	24.8	8.7	21.2
Softwood	30.2	29.2	28.7	29.5	22.9	30.0	30.2	27.2	25.0	21.2
Mixed forest species	8.5	11.2	9.0	14.0	10.2	16.6	9.4	17.1	9.8	17.3

Data from 1989 and 1999 censuses.

that represent a different behaviour depending on the farm size, thus simplifying the process. The different size classes were established on the basis that most Galician farms are family farms. It was assumed that the reference farm was a farm with about 40 dairy cows, in which the labour force corresponded to the labour force available in an average family (Maseda *et al.*, 2004).

Spreadsheets were also used to characterize production sectors. Through the use of macros, spreadsheets allow the recovery of data for all the farms included in a representative group and to estimate values of the variables. There are many possibilities for data recovery. Thus, the user can recover data for farms included in one or more TFs, in one or several classes, in one municipality, in several municipalities or in the whole region.

Finally, an analysis can be performed to determine the variability of the dairy cattle sector among the different farm groups, classified according to TF, size, class and location. Such an analysis is conducted by retrieving from the land use table (dairy cattle table) the data for all the variables considered in the agriculture census for farms assigned to each group. The number of farms of each type, and class, in each municipality was automatically determined, so is possible to study their spatial distribution (Fig. 2).

Moreover, evolution of the sector can be determined based on the information available from two consecutive censuses. The 1989 Census of Agriculture was analysed using the TFs and size classes defined for the year 1999. Analysis of the two censuses aimed at characterizing the evolution of the sector in terms of TFs, size classes and location.

Results

Based on the reference farm, the five size classes shown in Table 3 were built (C_1, C_2, C_3, C_4 and C_5). Farms with less than 10 diary cows were not considered as it was considered that such farms were not economically viable.

Table 3 shows the number of farms in both Censuses. The number of farms has decreased, particularly in the case of the smallest farms, while the number of large farms has increased.

Considering the constraints established, a total of 12 first-order combinations (with two land uses) were obtained. The combinations obtained met the pre-established minimum conditions and resulted from combining dairy cattle with permanent grasslands, other pastures, wheat, potato, corn silage, annual forage crops, multi-annual forage crops, beef cattle, cattle rearing, hardwood, softwood and mixed forest uses.

Table 5 shows the structure of land uses and size classes generated by the system for first-order combinations, and the number of farms in which first-order combinations were present. Table 5 also shows the results based on the 1989 Census of Agriculture, which did not consider corn silage corn (this crop was not considered as an independent entry in the 1989 Census
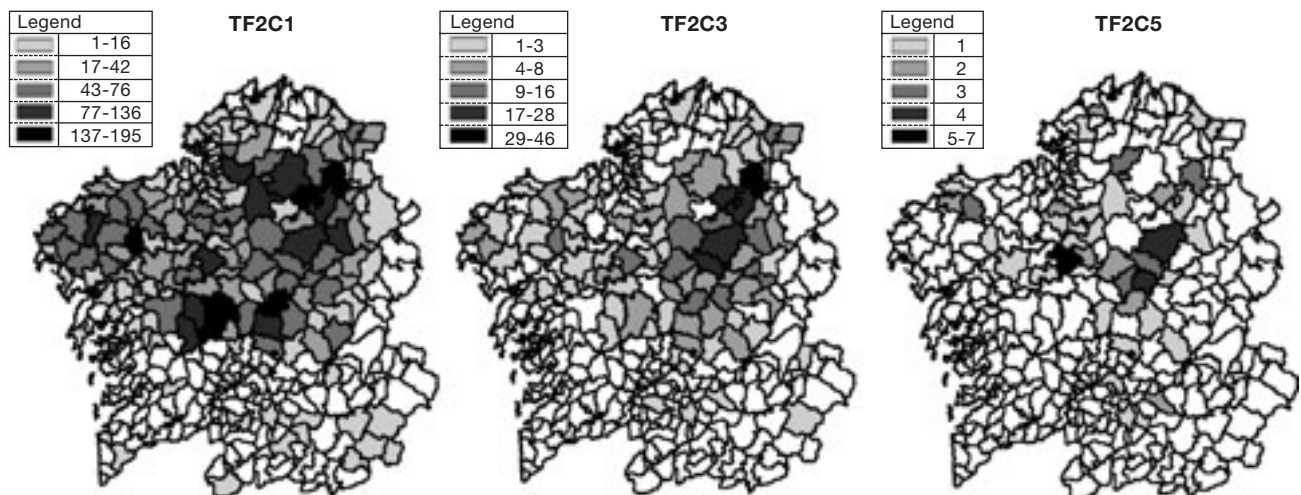

Figure 2. Distribution of number of Type 2 farms, Classes 1, 3 and 5 (1999 Census). Farm distribution was per municipality (there are 315 municipalities in Galicia).

and was included in annual forage crops). Between 1989 and 1999, significant changes occurred, among which, was a decrease in permanent grassland, other pastures or annual forage crops compared to multi-annual forage crops or corn silage, and an increase in cattle rearing. There was an increase forestry land use on dairy farms with hardwood and mixed forest species, compared to softwood, which barely maintained the same values.

These results are not the final stage of the classification algorithms. Rather, they are illustrative of the results that can be obtained from agricultural censuses once farms are grouped according to the TFs obtained from the method.

Analysis of data from the 1999 census shows that pasture species, corn, and beef cattle breeding (as an activity that is integrated in some dairy farms) must be considered as they influence the productive process of holdings, and their high frequency. The following three TFs result from first-order combinations:

— TF_1. Dairy farms in which milk production is associated with land and based on pasture and forage crops without presence of corn crops, regardless of the presence of other complementary activities or land uses such as cattle rearing, potato crops or forest land (among the most common ones). In this type of farm, beef cattle breeding stock cannot exceed 4 animals.

— TF_2. Identical to TF_1, but with corn crops present.

— TF_3. Dairy farms in which milk production is associated with the presence of beef cattle breeding stock of ≥ 5 animals, irrespective of the presence of forage or other complementary activities or land uses such as cattle rearing or fattening, potato crops or forest land, among the most common ones.

These TFs were applied to the data from the 1989 and 1999 censuses to obtain the number of farms included in each size class according to each TF (Table 6). Analysis of Table 6 shows a clear difference between the 1989 and the 1999 censuses with respect to farm number in each TF. As per the evolution of land use, the 1999 census incorporates corn silage (which was not considered in the 1989 census), and shows a less significant increase in the presence of beef cattle, which occurs mainly in the intermediate size classes (C_2 and C_3).

More specifically, the comparison of data between both censuses reveals that the TFs established do not coincide. TF_1 and TF_3 would better fit production systems observed in 1989. The census data confirms that milk production in Galicia around 1989 was mainly associated

Table 6. Distribution of farm type in Galicia according to size class for the 1989 and 1999 census of agriculture (corn silage was not considered in the 1989 census)

Size classes	TF_1		TF_2		TF_3	
	1989	1999	1989	1999	1989	1999
C_1	9,983	3,904	—	5,245	457	1,169
C_2	937	1,458	—	1,582	52	199
C_3	190	461	—	480	8	43
C_4	41	131	—	103	4	7
C_5	62	129	—	71	11	18
Total	11,213	6,083	—	7,481	532	1,436

with production of forage crops based on grassland species that provided fresh forage during the spring-autumn period and dry forage (hay) as a winter reserve. A smaller number of farms combined this diet with annual forage crops such as corn, which was consumed green at the end of the summer, with turnips during winter or, in some cases in early spring and with forage from autumn sown grain crops (e.g. wheat, rye, oats). This was the prevailing animal diet pattern around 1970. At that time farm systems were less mechanized, and grazing was more relevant, as confirmed by the higher labour force levels, lower mechanization level and greater levels of permanent grasslands and other pastures.

Considering evolution of the different TFs, only some aspects of TF_1 and TF_3 from the 1999 Census, would admit comparison with the potentially homologous TFs from the 1989 Census. Retrospectively, TF_1, for 1989, grouped all farms (or most of them) based on pasture and forage crop production, with an increasing presence of silage from these crops and similar management systems, which substantially coincides with the homologous type from the 1999 Census. Although not comprehensively studied, the TF with the highest similarity in terms of structure and evolution of management systems is TF_3 (which includes beef cattle). The forage base is similar, the management systems are extensive or semi-extensive, and most of them maintain grazing.

The spatial distribution of the different TFs was analysed at a municipal scale based on the three TFs obtained and on the corresponding size classes. By exporting these values to a GIS application or to an Excel application, maps can be generated for the whole of Galicia with the corresponding number of farms. Figure 2 is an example and shows the spatial distribution of part of the results included in Table 6.
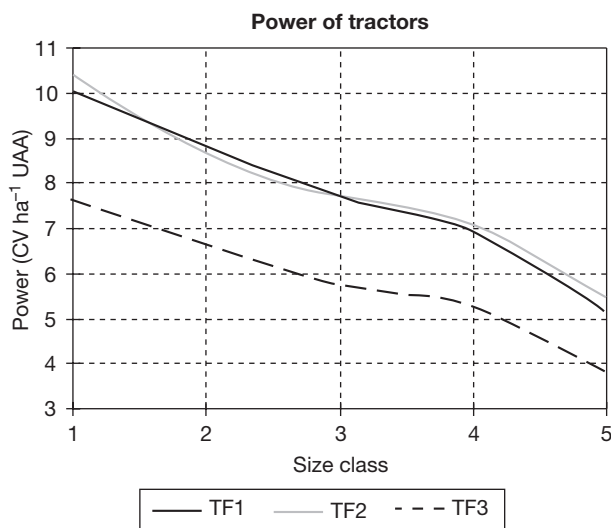
Power of tractors

Figure 3. Average tractor power referred to the unit area (UAA) of the different size classes of farmas (data from 1999 census).

Considering average tractor power (Fig. 3) per unit area (UAA) as an indicator of farm mechanization level, it can be concluded that the farms belonging to TF_1 and TF_2 had a similar mechanization level, while mixed farms (TF_3) had lower values, with differences that were more evident for the smallest farms. The mechanization level per UAA and, consequently, per CU decreases with increased farm size.

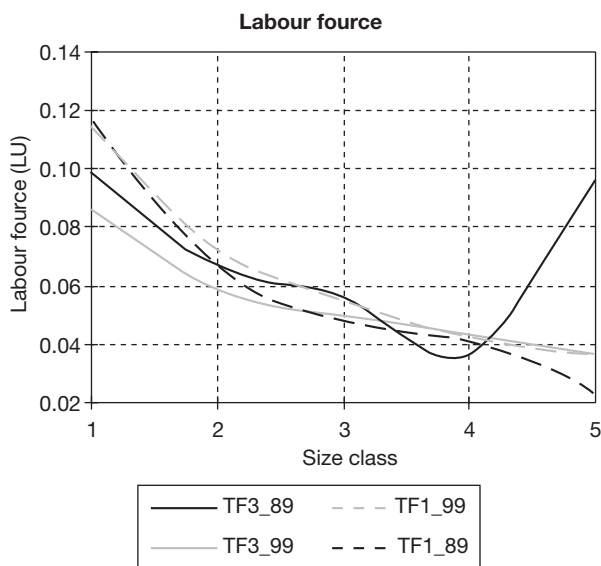The labour force occupied on farms (Fig. 4) per cow unit present on the farm (LU) decreases with decreased

Labour fource

Figure 4. Farm labour force by cow units (data from 1989 and 1999 censuses).

farm size. The values for TF_1 farms are higher than those for TF_3. Differences between the two TFs are most evident for C_1 farms, while the differences disappear on C_5 farms. Comparison between the 1989 and 1999 censuses suggests similar trends for this variable. The trends in 1989 are more erratic, perhaps due to the lower number of farms included in each TF, there were only 23 farms in TF_3, classes C_3-C_5, in 1989. However, values for 1989 are higher than for 1999. This could be explained by lower technology use on the farms.

Discussion

In 1999, the forage base on farms in Galicia was focused on production of annual or multi-annual pasture crops that are transformed into silage or, to a lesser extent, into hay for consumption over the whole year. On an increasing number of farms, pastures were combined with corn silage for year round consumption. These crops are used as ingredients of rations that can be adjusted by using concentrates according to farmers' needs, thus reducing, or avoiding, seasonality of milk production. This is associated with management systems, which show a high level of mechanization, less use of labour per unit (per animal), and reduction or cessation of grazing. Currently, the production system shows important differences, even with respect to the data in the 1999 census.

The methodology for establishment of a typology, classification and characterization of farms presented in this study is implemented in different phases and is based on automated analysis of data contained in the census of agriculture. This methodology enables the user to group, at the municipal level, all farms that show the same production pattern in a previously established size class. The spatial distribution of each reference pattern is determined based on these groups of similar elements. Determination of spatial distribution of TFs is essential to choose a reduced number of farms that are representative of each type and to give updated information about the production process (operations, raw materials, labour) and structure (size, machinery, facilities, land) of the farms. This information is useful for characterizing the production patterns (type-class combinations) that are representative of each activity. Such production patterns complete the information of the agricultural production-planning model developed.

The methodology presented here, can be used to analyse the evolution of a sector or to model future

trends. The comparison between the 1989 and 1999 censuses reveals changes in both farm size and production structure. The dairy sector in Galicia has restructured. The main TFs, from the 1999 Census, did not necessarily evolve from the TFs present in 1989. For example, the TF based on the production of corn silage emerged as a new TF during the 1990s. Before that different winter forages, such as turnips or spring cereals, were grown. With the increase in the average farm size, these systems evolved towards grass silage, which could be mechanized. Grass silage was used as a reserve feed source for winter. This pattern has further evolved towards the production of corn silage, which is now common on most farms, as suggested by the evolution of the different land uses and by the variations observed in the farms included in the different size classes.

This model can be adapted to other regions and countries by making slight modifications in the structure of the software application used. In addition, the model can be further applied to specific geographical areas such as parishes, municipalities or agricultural regions to adjust the different TFs and size classes according to specifications of the physical environment.

The methodology proposed in this study obtains information from a very significant statistical source, the agricultural censuses. This valuable source is under utilised, perhaps because of a lack of tools to handle the large volume of data they contain. The use of census data is usually restricted to data abstracts, in which data are grouped and systematized, thus providing some information. The traditional trend is to work with census abstracts that provide characterization of a production system or region by grouping data (Reidsma *et al.,* 2006). This approach changes dramatically if micro-data are used, i.e. when there are individual data (the results of the census) available for each farm included in the census.

Individual data allows the division of farms into differential groups, which shows a clear potential for the use of the information in agricultural censuses. This approach is completely innovative.

This approach was applied to dairy farms in Galicia, and allowed the establishment of relational TFs with the dairy production system or the presence of complementary activities and/or land uses. Grouping farms into homogeneous groups enables one to: i) study the location of the different groups to determine if there is any relationship with environmental factors that can explain the differences; ii) obtain census data that are grouped for each TF to present differences among groups, i.e. to characterize each TF; and iii) carry out a field survey that actually represents the dairy sector by stratifying the population by types and classes.

These analyses can be applied to agricultural production planning at different levels: public administration, individual farmers or farmers' associations.

Further, the frequency of agricultural censuses (every 10 years) and the frequency of approval by the EU, allows for analyses and prospects across time and space.

The methodology, explained above, is compared with traditional classifications based on TFs and ESUs (Andersen *et al.,* 2007). Table 1 shows the large number of TFs (25 for farms that sell cow milk) and the large variation in ESUs (from 6 to more than 300). Traditional classifications do not allow the user to make a specific analysis of the sector or to fully use information from agricultural censuses. Moreover, these classifications cannot be applied to agricultural production planning.

From the data farms with less than 10 cows have dramatically decreased during the last few years. Such farms are self-supply/subsistence farms and are not economically viable. This situation persists today. According to values for 2003, from 27,106 farms in Galicia, 14,183 have 10 or more cows and 12,923 have less than 10 cows (7% of the total number of cows and < 8% of total yield). Milk production in Galicia has increased every year despite the dramatic decrease in farms with less than 10 cows. Based on this, 10 cows was consider an appropriate minimum value.

The results of the study can be extrapolated because agricultural censuses published in Spain and in all the member states of the EU are harmonized. Therefore, the methodology proposed in this study can also be developed for those countries.

Acknowledgement

References

ANDERSEN E., ELBERSEN B., GODESCHALKC F., VERHOOG D., 2007. Farm management indicators and farm typologies as a basis for assessments in a changing policy environment. J Environ Manage 82(3), 353-362.

BARBEYTO F., 1998. Explotaciones de vacuno de leche en Galicia. Manejo técnico y resultados económicos. Xunta de Galicia, Santiago de Compostela. [In Spanish].

BOYLAND M., NELSON J., BUNNELL F.L., 2004. Creating land allocation zones for forest management: a simulated annealing approach. Can J Forest Res 34, 1669-1682.

CARSJENS G.J., VAN DER KNAAP W., 2002. Strategic land-use allocation: dealing with spatial relationships and fragmentation of agriculture. Landscape Urban Plan 58, 171-179.

CHATELIER V., COLSON F., FUENTES M., VARD T., 2000. Grazing livestock holdings in European Union. Productions Animales 13(3), 201-203.

DE WIT C.T., VAN KEULEN H., 1988. Application of interactive multiple goal programming techniques for analysis and planning of regional agricultural development. Agr Syst 26, 211-230.

DULOY J.H., NORTON R.D., 1983. CHAC: a programming model for Mexican agriculture. In: The book of CHAC: programming studies for Mexican agriculture (Norton R.D., Solis M.L., eds). The Johns Hopkins University Press, Baltimore, MD. pp. 23-57.

EVERITT B., 1993. Cluster analysis, 3rd ed. Halsted, NY.

GLENG J.J., TIPPER R., 2001. A mathematical programming model for improvement planning in a semi-subsistence farm. Agr Syst 70, 295-317.

HWANG S.W., ALWANG J., NORTON G.W., 1994. Soil conservation practices and farm income in the Dominican Republic. Agr Syst 46, 59-77.

INE, 1989. Censo Agrario de 1989. Base de microdatos. Instituto Nacional de Estadística, Madrid, Spain. Available at http://www.ine.es/prodyser/micro_censoag.htm. [In Spanish].

INE, 1999. Censo Agrario de 1999. Base de microdatos. Instituto Nacional de Estadística, Madrid, Spain. Available at http://www.ine.es/prodyser/micro_censoag.htm. [In Spanish].

INE, 2003. Censo Agrario de 1999. Instituto Nacional de Estadística, Madrid. Spain. [In Spanish].

JUDEZ L., CHAYA C., 1999. Effects of geographical stratification in a farm accountancy data network on the accuracy of the estimates. J Agric Econ 50(3), 388-399.

KOSTOV P., McERLEAN S., 2006. Using the mixtures-of-distributions technique for the classification of farms into representative farms. Agr Syst 88, 528-537.

LEITNER H., McMASTER R.B., ELWOOD S., McMASTER S., SHEPPARD E., 2002. Models for making GIS available to community organizations: dimensions of difference and appropriateness. In: Community participation and geographic information systems (Craig W.J., Harris T.M., Weiner D., eds). Taylor and Francis, London. pp. 37-52.

LOFTSGARD L.D., HEADY E.O., 1959. Application of dynamic programming models for optimum farm and home plans. J Farm Econ 41, 51-62.

MASEDA F., DÍAZ F., ÁLVAREZ C.J., 2004. Family dairy farms in Galicia (N.W. Spain): classification by some family and farm factors relevant to quality of life. Biosyst Eng 87(4), 509-521.

NALLE D.J., ARTHUR J.L., SESSIONS J., 2002. Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. Forest Sci 48(1), 59-68.

PARKER D.C., MANSON S.M., JANSEN M.A., HOFFMANN M.J., DEADMAN P., 2003. Multi agent system for the simulation of land-use and land-cover change: a review. Annals of the Association of American Geographers 93(2), 314-337.

REIDSMA P., TEKELENBUG T., VAN DEN BERG M., ALKEMADE R., 2006. Impacts of land-use change on biodiversity: an assessment of agricultural biodiversity in the European Union. Agr Ecosyst Environ 114(1), 86-102.

RICA, 1988. Recopilación de reglamentos, disposiciones de aplicación, informaciones. Fascículo II, La ficha de explotación. Red de Información Contable Agrícola, CECA-CEE-CEEA, Brussels, Luxembourg. [In Spanish].

RIVEIRO J.A., ÁLVAREZ C.J., MIRANDA D., PEREIRA J.M., 2005. Profitability and production requirements for land use allocation of farming and forestry land. Biosyst Eng 90(4), 477-484.

ROSSITER D.G., 1996. A theoretical framework for land evaluation. Geoderma 72, 165-202.

SNYDER K., 2003. Tools for community design and decision-making. In: Planning support systems in practice (Geertman S., Stillwell J., eds). Springer-Verlag, Berlin. pp. 99-120.

STORIE R.E., 1933. An index for rating the agricultural value of soils. University of California, Berkeley, CA.

THENAIL C., BAUDRY J., 2004. Variation of farm spatial land use pattern according to the structure of the hedgerow network (bocage) landscape: a case study in northeast Brittany. Agr Ecosyst Environ 101(1), 53-72.

VOOGD H., 1983. Multicriteria evaluation for urban and regional planning. Pion Limited, London.

XUNTA DE GALICIA, 2004. Estudio de ordenación productiva agraria de 21 comarcas. Consellería de Política Agroalimentaria y Desarrollo Rural, Santiago de Compostela, Spain. 2025 pp. [In Spanish].

ZHEN L., ROUTRAY J.K., 2003. Operational indicators for measuring agricultural sustainability in developing countries. Environ Manage 32(1), 34-46.