

# ***tipificación de explotaciones agropecuarias incluidas en el censo general agropecuario 2000***

Ing. Agr. Alfredo Hernández\*

Juan José Goyeneche\*\*

Laura Nalbarte\*\*

---

Este trabajo forma parte del plan de análisis que DIEA viene realizando de los resultados del Censo General Agropecuario 2000 (CGA). En esta ocasión se propone una clasificación en grandes grupos homogéneos de las 52.111 explotaciones comerciales del CGA 2000<sup>1</sup> a partir de la utilización de técnicas multivariadas de tipificación o análisis de *cluster*. Para ello se usan variables de composición del uso del suelo y existencias animales.

Este es un nuevo aporte a la larga tradición de tipificación de empresas agropecuarias que se han llevado a cabo en el Ministerio de Ganadería, Agricultura y Pesca con el objetivo primordial de contribuir al mejor conocimiento de la realidad agropecuaria.

Seguidamente se presentan los puntos clave de la metodología aplicada y una primera aproximación de los resultados obtenidos.

## **1. METODOLOGÍA, ETAPAS CONSIDERADAS**

Las etapas seguidas para obtener los distintos grupos de explotaciones son:

- De las 52.111 explotaciones se separa un grupo de 3.083 que presentan valores extremos en algunas variables. Las restantes 49.028 son clasificadas según el uso del suelo. Este paso determina 16 grupos que llamamos “primarios”.
- Dentro de cada grupo primario se usarán técnicas multivariadas de tipificación para obtener grupos “secundarios”, considerando el conjunto de variables que definen las principales actividades de producción.
- Eventualmente, de existir grupos secundarios similares dentro de dos grupos primarios distintos, se considera la posibilidad de unirlos en un único grupo secundario.
- El conjunto final de grupos secundarios obtenido constituyen la base para análisis posteriores: posible identificación de sistemas de producción,

---

\* DIEA-MGAP e \*\* Instituto de Estadística, Facultad de Ciencias Económicas.

<sup>1</sup> Previamente se excluyeron 5.020 explotaciones no comerciales del total de las censadas.

estudios en profundidad para casos especiales, elaboración de marcos muestrales, etc.

## 2. VARIABLES TIPIFICADORAS

Para cada una de las explotaciones consideradas se definieron las siguientes variables tipificadoras (en paréntesis se presentan los nombres de las variables):

- % de hectáreas de bosque natural, tierras de rastrojo, campo natural y tierras improductivas (*campo*).
- % de hectáreas utilizadas en cultivos forrajeros, praderas artificiales, campo natural sembrado en cobertura y fertilizado (*mejoramiento*).
- % de hectáreas utilizadas en bosques artificiales (*nap2*).
- % de hectáreas totales de frutales (*nap345*) o por grupo de especies: citrus (*nap3*), frutales de hoja caduca (*nap4*) y viña (*nap5*).
- % de hectáreas totales de huerta (*huerta*) o por separado: papa(*nap6\_1*), huerta a campo (*nap6\_2*) y huerta protegida (*nap6\_3*).
- % de hectáreas con arroz y tierra arada (*arroz*).
- % de hectáreas con cereales (excluido arroz) y tierra arada<sup>2</sup> (*cereal*).
- total de vacunos de carne por hectárea (*tva\_ca*).
- total de vacas de leche por hectárea (*tva\_le*).
- total de ovinos por hectárea (*tovinos*).
- total de aves (*aves*).
- total de porcinos (*porcinos*).

## 3. ANÁLISIS PRELIMINAR DE LOS DATOS Y DETECCIÓN DE VALORES EXTREMOS (*outliers*)

En el análisis de las variables que se consideran en la tipificación se encontraron algunas características comunes muy resaltadas: gran cantidad de casos con valores de cero, fuerte asimetría positiva con presencia de valores extremos. La lista de variables analizadas se presentan en el Cuadro 1.

¿Cómo definir los valores atípicos (*outliers*)? Se consideró un criterio estadístico para las distribuciones de las variables medidas en valores absolutos: mediante transformaciones de Box-Cox se normalizan los datos y se considera *outlier* las observaciones que se alejan de la media más de 3.09 veces el desvío estándar<sup>3</sup>. Tal como se esperaba, debido a la asimetría de las

---

<sup>2</sup> Si la superficie de arroz es positiva, la tierra arada se suma al arroz, en caso contrario, la tierra arada se agrega a los cereales.

<sup>3</sup> El valor de 3.09 es el cuantil 99.9% en la distribución Normal estándar.

distribuciones sólo se detectaron valores atípicos en el extremo superior de la distribución.

CUADRO 1  
Variables analizadas, límites obtenidos y número de explotaciones  
y porcentaje acumulado

Variable	Límite	Número de casos	Porcentaje del total de la variable
Bosque artif. (nap2)	98	817	83.8%
Citrus (nap3)	201	23	58.9%
Frutales h. Caduca (nap4)	25	62	29.0%
Viña (nap5)	17	95	43.4%
Papa (nap6_1)	13	115	79.1%
Huerta campo (nap6_2)	13	295	28.3%
Huerta protegida (nap6_3)	3.1	30	22.7%
Arroz	942	42	38.4%
Cereal	828	88	23.4%
Porcinos	77	442	55.4%
Aves	5442	389	86.4%
Ovinos	3781	459	22.2%
Vacunos de carne	2710	391	18.6%
Vacunos de leche	520	232	26.9%
Total de <i>outliers</i> <sup>4</sup>		3083	

Esto es por ejemplo, el límite de 201 para *citrus* implica que explotaciones con áreas superiores a la indicada son considerados extremos.

Cabe señalar que en varios casos las 3.083 explotaciones que consideramos *outliers* concentran un porcentaje importante del total nacional de las variables estudiadas. La clasificación de este grupo primario en secundarios se hará en forma separada al resto, atendiendo a su complejidad y variedad y a la relevancia que para los principales rubros tiene dentro del total nacional.

#### 4. CONFORMACIÓN DE GRUPOS PRIMARIOS

Una vez separados los *outliers* se procede a formar grupos primarios según los componentes de uso del suelo con las 49.028 explotaciones restantes.

Un primer grupo son las explotaciones que sólo tienen *campo* y *bosques artificiales* en el uso del suelo. Un segundo grupo son las explotaciones con

<sup>4</sup> Una determinada explotación puede resultar como atípica en una o más de las variables consideradas.

*campo, bosques y mejoramientos*, etc. El nombre y una descripción de los grupos obtenidos se presenta a continuación.

- x02Nb. campo y bosques art.
- x04Nbm. campo, bosques art. y mejoramientos
- x05\_c. sólo cereales
- x06Nbmc. campo, bosques art., mejoramientos y cereales
- x07\_a. sólo arroz
- x09Nbmc. campo, bosques art., mejoramientos, cereales y arroz
- x10\_h. sólo huerta
- x11\_f. sólo frutales
- x12\_hf. sólo huerta y frutales
- x13N\_hf. campo, huerta y frutales
- 

Una vez considerados los 10 grupos anteriores quedan 4580 explotaciones que tienen combinaciones de huerta y/o frutales con los rubros considerados extensivos. La clasificación de este “resto” de explotaciones se hizo de la siguiente forma:

- x15\_1a. tienen arroz
- x15\_2h. declaran huerta para la venta
- x15\_3\_f. declaran frutales para la venta
- x15\_4hf. declaran huerta y frutales para la venta
- x15\_resto. otras explotaciones no mencionadas anteriormente

En los Cuadros A1 y A2 del Anexo se presentan los totales de algunas variables para los 16 grupos primarios, los 15 recién mencionados y el grupo de *outliers*. Como puede apreciarse los grupos primarios son muy disímiles entre sí en cuanto a número de explotaciones y superficie. Los grupos x00, x02, x04 y x06 representan el 83% de las explotaciones y el 96% de la superficie y concentran la casi totalidad de los ovinos y los vacunos de carne. Por su parte, los vacunos de leche están concentrados en los grupos x00, x04 y x06 (y prácticamente ausentes del x02 donde no hay presencia de mejoramientos). Similar concentración aparece en algunas otras variables: el total del arroz está en los grupos x00, x07 y x09 (por definición de los mismos), y el 94% de los cereales está en los grupos x00 y x06.

La apertura de estos grupos primarios en secundarios se hace atendiendo a las variables derivadas de las características del grupo primario que se considere. Las técnicas multivariadas producen grupos secundarios de variada homogeneidad, aparecen grupos claramente “avícolas”, o “lecheros”, mientras que otros grupos presentan actividades combinadas. Esta heterogeneidad es desde un punto de vista agroeconómico, ya que estadísticamente tienen el mismo grado de homogeneidad. No obstante, el punto de vista conceptual, que podría catalogarse de subjetivo, es imprescindible para darle valor real a la tipificación. En el caso del CGA 2000 se prefirió que el criterio agroeconómico fuese el paso final (y eventualmente validador último) en el proceso tipificatorio.

#### **4. CONCLUSIONES**

El proceso de tipificación consta básicamente de dos etapas:

- conformación de grupos primarios ya sea por tener valores atípicos o por su uso del suelo y
- separación de estos grupos primarios a través de técnicas de *cluster* multivariados en grupos secundarios.

El procedimiento permite que las explotaciones agropecuarias comerciales del país queden clasificadas en grupos “homogéneos” que servirán de base a estudios en profundidad de los distintos sectores, como punto de partida en la construcción de marcos muestrales para posibles encuestas, y para una caracterización de las explotaciones agropecuarias. Los estudios en profundidad sobre estos grupos secundarios deberán atender sus especificidades desde una perspectiva agroeconómica, tecnológica, social, productiva, dependiendo del tipo de clasificación que se desee.