



**METODOLOGÍA DE  
DISEÑO MUESTRAL**

Código: DM-ENA-DIM-01  
Versión : 012  
Página 1  
Fecha: 01-07-08

# **Corporación Colombia Internacional**

## **Sistema de Información del Sector Agropecuario y Pesquero Colombiano**

### **Oferta agropecuaria**



## **COMPONENTE DE MUESTREO**

**METODOLOGÍA DEL DISEÑO DE LA ENCUESTA NACIONAL AGROPECUARIA**

**Bogotá D.C, Julio 2010**



|  |           |
|--|-----------|
| <b>1. GENERALIDADES .....</b>                              | <b>4</b>  |
| <b>2. OBJETIVOS.....</b>                                   | <b>6</b>  |
| <b>2.1 OBJETIVO GENERAL.....</b>                           | <b>6</b>  |
| <b>2.2 OBJETIVOS ESPECIFICOS.....</b>                      | <b>6</b>  |
| <b>3 DISEÑO ESTADISTICO .....</b>                          | <b>7</b>  |
| <b>3.1 ENCUESTAS CON MULTIPLES MARCOS DE MUESTREO.....</b> | <b>7</b>  |
| <b>3.2 UNIVERSO.....</b>                                   | <b>7</b>  |
| <b>3.3 POBLACION MUESTREADA.....</b>                       | <b>8</b>  |
| <b>3.4 COBERTURA .....</b>                                 | <b>8</b>  |
| <b>3.5 UNIDADES ESTADISTICAS .....</b>                     | <b>8</b>  |
| <b>3.5.1 UNIDAD DE MUESTREO.....</b>                       | <b>8</b>  |
| <b>3.5.2 UNIDAD DE OBSERVACIÓN.....</b>                    | <b>9</b>  |
| <b>3.5.3. UNIDAD DE ANÁLISIS.....</b>                      | <b>9</b>  |
| <b>3.6 MARCO ESTADISTICO .....</b>                         | <b>10</b> |
| <b>3.6.1 MACROESTRATIFICACIÓN.....</b>                     | <b>11</b> |
| <b>3.6.2. MICROESTRATIFICACIÓN .....</b>                   | <b>11</b> |
| <b>3.6.3 DOMINIOS DE ESTUDIO.....</b>                      | <b>12</b> |
| <b>3.6.4 MARCO DE LISTA .....</b>                          | <b>13</b> |
| <b>3.7 PARÁMETROS A SER ESTIMADOS .....</b>                | <b>13</b> |
| <b>3.8 DISEÑO MUESTRAL .....</b>                           | <b>14</b> |



|  |           |
|--|-----------|
| <b>3.8.1 MÉTODO DE MUESTREO.....</b>                               | <b>14</b> |
| <b>3.8.2 TAMAÑO DE MUESTRA. ....</b>                               | <b>15</b> |
| <b>3.8.3 SELECCIÓN DE LA MUESTRA.....</b>                          | <b>18</b> |
| <b>3.8.4 MÉTODO DE ESTIMACIÓN .....</b>                            | <b>18</b> |
| PROCEDIMIENTOS DE ESTIMACIÓN .....                                 | 18        |
| METODOLOGÍA DE ESTIMACIÓN .....                                    | 20        |
| ESTIMACIÓN DE LAS DIFERENCIAS Y SU SIGNIFICACIÓN ESTADÍSTICA ..... | 22        |
| ESTIMACIÓN DE LOS ERRORES MUESTRALES .....                         | 23        |
| <b>3.9 LA ESTRATEGIA DE CRITICA E IMPUTACION.....</b>              | <b>25</b> |
| <b>3.9.1 DEFINICIÓN DE LA CRITICA ESTADÍSTICA: .....</b>           | <b>25</b> |
| <b>3.9.2 PROCESO DE CRITICA ESTADÍSTICA: .....</b>                 | <b>25</b> |



## **1. GENERALIDADES**

Ante las dificultades técnicas y financieras para la realización del III Censo Nacional Agropecuario en Colombia y la ausencia de información objetiva para el sector, se firma en 1983 un convenio entre el Ministerio de Agricultura y la FAO, con el fin de implementar en el país un sistema de estadísticas agropecuarias por muestreo para conseguir información oportuna y confiable que permitiera un eficiente planeamiento del sector.

Para probar la aplicación de esta metodología, se llevaron a cabo cinco encuestas piloto entre noviembre de 1983 y abril 1986, donde se combinó la agricultura comercial en tierras planas, con la agricultura tradicional de las zonas montañosas, con tierras dedicadas a la ganadería. Se ensayaron todos los aspectos logísticos y operativos de la encuesta, el tamaño y contenido de los cuestionarios y los costos asociados.

Entre abril y mayo de 1988 se realizó la Primera Encuesta Nacional Agropecuaria (PENAGRO), por muestreo, en 23 departamentos, la cual permitió obtener una medición de variables estructurales y coyunturales del sector, lo que no sucedía desde 1970, cuando se realizó el II Censo Nacional Agropecuario, considerado de baja calidad por recoger la información por convocatoria y no cumplir el plan intercensal de encuestas.

Este ejercicio de la PENAGRO no tuvo la continuidad requerida por diversas razones, sin embargo, entre el año de 1995 y el año 2005 el sistema fue permanente a través del DANE. En el año 2006 la Corporación Colombia Internacional (CCI) asumió la responsabilidad en el manejo del sistema de información estadística agropecuaria, alcanzando importantes logros en materia de calidad y oportunidad de los datos generados para los años 2006 y 2007.

El sistema de información estadística del sector agropecuario hace parte del sistema nacional y debe facilitar la labor de los diferentes usuarios de los datos que permitan los ejercicios de planificación, gerencia y toma de decisiones, debe proveer una vasta gama de datos con calidad, confiabilidad y oportunidad. Debe entregar herramientas para la realización de nuevos estudios e investigaciones estadísticas, lo cual redundara en una adecuada formulación de política y toma de decisiones muy sólidas.

Algunas de las actividades que se pueden desarrollar con base en el sistema de información, pueden ser:

- Analizar el desempeño del sector agropecuario en la economía mundial
- Formular metas del desarrollo agropecuario que sean consistentes con los objetivos del mismo.
- Estimar la demanda y la oferta de productos agropecuarios
- Realizar pronóstico de la producción agropecuaria



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 5  
Fecha: 01-07-08

La encuesta realizada por el DANE en 1995 cubrió los 32 departamentos del país. Con base en esta medición se procedió a la estimación semestral y anual de:

- Superficie total y aprovechamiento de la tierra, nacional y departamental.
- Principales cultivos permanentes y transitorios, área, producción y rendimiento, nacional y departamental.
- Inventario de ganado bovino, según orientación del hato, nacional y departamental.
- Inventario de otras especies pecuarias, nacional y departamental.
- Producción de leche el día anterior a la entrevista, nacional y departamental.
- Superficie en pastos, uso y manejo, nacional y departamental.
- Distribución de las unidades de observación (PSM), según tamaño y uso del suelo, nacional y departamental.

Estos resultados fueron aceptables a nivel nacional, no así a nivel departamental ya que la disponibilidad de recursos limitó el tamaño de la muestra de la encuesta.

En el año 1996 se implementó la optimización del diseño muestral a través de la focalización global del universo por producto o grupo de ellos, actualización del marco de muestreo y la reducción del tamaño de los segmentos.

De acuerdo con lo anterior la encuesta correspondiente al primer semestre de 1996 se orientó fundamentalmente a la estimación del área, la producción y el rendimiento de los cultivos transitorios de clima cálido y frío.

La encuesta del segundo semestre de este mismo año se focalizó para la medición del área, la producción y el rendimiento de la actividad agropecuaria predominante en los tres pisos térmicos para los 32 departamentos del país, es decir, tanto los cultivos transitorios, como los permanentes, el área en pastos y el inventario de especies pecuarias. Con la unión de estas dos mediciones se estimaron una mayor cantidad de productos agrícolas que los estimados en el año anterior, con una minimización de los errores de muestreo y mayor nivel de desagregación.

Dadas las restricciones presupuestales a que fue sometido el proyecto para el año 1997, la investigación fue reorientada para obtener estimaciones nacionales de una menor cantidad de productos que los estimados en los años anteriores. Se realizó una sola medición y una submuestra con lo cual se logró finalmente mantener las estimaciones del uso del suelo, los principales cultivos transitorios y permanentes en universos más reducidos, así como también la actividad pecuaria en los 23 departamentos agrícolas más representativos.

En el año 1998, no se obtuvieron resultados por problemas de recursos económicos y su manejo.

Para la encuesta de 1999 se implementó el muestreo en dos fases y en la del año 2000 el muestreo dual. En el año 2002 se aplicó el muestreo dual, mediante la realización de censos para papa en Cundinamarca y Boyacá, para el cultivo de soya en el Meta y Valle del



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 6  
Fecha: 01-07-08

Cauca y para la cebolla larga en la zona de Aquitania (Boyacá) y Buesaquillo (Nariño), los otros productos se estimaron con el muestreo de áreas.

La CCI a partir del año 2007 esta realizando dos operativos, uno en el primer semestre para estimar el áreas sembrada de los principales cultivos transitorios y otro en el segundo semestre donde se estiman las áreas sembradas y cosechadas de los principales cultivos transitorios y permanentes, además de realizar el inventario vacuno y de otras especies, también se estima el área ocupada en pastos y el numero de árboles frutales dispersos

## 2. OBJETIVOS

### 2.1 OBJETIVO GENERAL

Suministrar, a través de La Encuesta Nacional Agropecuaria (ENA), los datos necesarios para estimar las áreas y la producción de los cultivos de interés y el inventario de las diferentes especies a nivel departamental y nacional para su posterior uso en la elaboración y diseño de estudios y políticas que requieran de información actualizada y consolidada del sector agropecuario.

### 2.2 OBJETIVOS ESPECIFICOS

- Estimar el área, la producción y el rendimiento de los principales cultivos transitorios y permanentes.
- Estimar la superficie ocupada por pastos y forrajes según su variedad
- Estimar el inventario de ganado bovino según sexo, especie, orientación del hato y grupo etario
- Estimar el inventario de otras especies según sexo.
- Estimar la cantidad de vacas ordeñadas y la producción de leche y su destino según orientación de la raza y el tamaño del hato
- Estimar el uso de la tierra y la distribución según el tamaño de los pedazos de segmentos de muestreo (PSM).
- Estimar el número de árboles frutales dispersos

### 3 DISEÑO ESTADÍSTICO

#### 3.1 ENCUESTAS CON MÚLTIPLES MARCOS DE MUESTREO

Un estimador de una encuesta con múltiples marcos de selección es la suma de un estimador de la muestra de áreas más el total de datos recolectados de la lista de explotaciones especiales. Si la lista es objeto de muestreo, el estimador de múltiples marcos será la suma de los estimadores correspondientes a ambas muestras, el estimador de la muestra de áreas y el estimador de la muestra de lista.

El estimador de la muestra de áreas no considera los PSM correspondientes a las Unidades de Producción Agropecuaria de la lista de explotaciones especiales, para evitar duplicaciones. Por lo tanto, todas las Unidades de Producción Agropecuaria con PSM incluidos en los segmentos muestrales deben compararse con la lista de explotaciones especiales, para eliminar duplicidades con la muestra de áreas.

Para una variable de la encuesta, si se enumera por completo la lista de explotaciones especiales la varianza del estimador basado en múltiples marcos es igual a la varianza del estimador del marco de áreas únicamente. Si se selecciona una muestra de la lista, la varianza de esta muestra se suma a la varianza de la muestra de áreas, para obtener la varianza de la encuesta con múltiples marcos.

#### 3.2 UNIVERSO

Está constituido aproximadamente por 50.600.000 hectáreas, resultado de la exclusión de las áreas no agropecuarias, urbanas, semiurbanas, parques naturales, bosques, áreas erosionadas y cuerpos de aguas, del área total del país 114'174.800 hectáreas.

Cuadro1:

| Universo de estudio en cantidad de UPM y superficie en hectáreas, según cobertura vegetal |         |                 |                   |
|---|---------|-----------------|-------------------|
| Cobertura   | Térmico | Cantidad de UPM | área (ha)         |
| Pastos  | p1      | 10.098          | 12.867.943        |
|   | p2      | 3.392           | 4.199.909         |
|   | p3      | 2.690           | 3.208.726         |
| <b>Total Pastos</b>   |         | <b>16.180</b>   | <b>20.276.578</b> |
| Transitorios  | T1      | 5.414           | 4.911.702         |
|   | T2      | 57              | 31.309            |
|   | T3      | 3.106           | 1.638.970         |
| <b>Total Transitorios</b>   |         | <b>8.577</b>    | <b>6.581.981</b>  |
| Vegetación de Sabana  | VS      | 3.303           | 16.108.138        |
| Vegetación Xerofítica   | VX      | 195             | 1.095.052         |
| Permanentes   | Pe      | 9.424           | 6.602.744         |
| <b>Total general</b>  |         | <b>37.679</b>   | <b>50.664.493</b> |



### 3.3 POBLACION MUESTREADA

La ENA mide las áreas, la producción de los cultivos de interés, lo mismo que el inventario pecuario en 22 departamentos del país: Antioquia, Atlántico, Bolívar, Boyacá, Caldas, Casanare, Cauca, Cesar, Córdoba, Cundinamarca, Huila, Guajira, Magdalena, Meta, Nariño, Norte de Santander, Quindío, Risaralda, Santander, Sucre, Tolima y Valle del Cauca. Se excluyen los departamentos de: Arauca, Amazonas, Guaviare, Caquetá, Guainía, Putumayo, Choco, San Andrés, Vichada y Vaupés. Igualmente se excluyen las áreas plantadas en palma de aceite, caña de azúcar y banano de exportación, los parques naturales, los bosques y áreas seminaturales, la vegetación de paramo, los cuerpos de agua, los espacios abiertos sin o con poca vegetación y las áreas artificiales. El área efectivamente muestreada es de 37.900.546 hectáreas, que mediante el proceso de recuperación se llevan al universo agropecuario.

### 3.4 COBERTURA

En muestreo de conglomerados se observan 22 departamentos agropecuarios que reúnen más del 80% de la producción nacional.

### 3.5 UNIDADES ESTADISTICAS

Las unidades estadísticas se clasifican en: unidad de muestreo, unidad de información, unidad de análisis y unidad de respuesta.

#### 3.5.1 Unidad de muestreo

Parte que se selecciona de la población objetivo para la realización de la encuesta. El proyecto OFERTA AGROPECUARIA utiliza:

Para el muestreo de conglomerados tres unidades de muestreo.

**La unidad primaria de muestreo (UPM)** que es una partición de la superficie, delimitada por accidentes físicos naturales o culturales fácilmente identificables sobre el terreno y demarcados sobre una fotografía aérea, un mapa, una imagen de satélite o cualquier otro material que los sustituya adecuadamente.

**El segmento de muestreo (SM):** que es la unidad resultante de dividir las UPM seleccionadas en la primera etapa de selección. Es una porción pequeña de tierra perteneciente a una UPM.

**El pedazo de segmento de muestreo (PSM):** que corresponde a una Unidad de Producción (UP) o parte de la misma que se encuentra dentro de un segmento de muestreo (SM) seleccionado en la segunda etapa de selección. Corresponde a la unidad productora o parte de la misma, que se encuentra dentro de un SM seleccionado; puede o no coincidir



con el predio y está a cargo de un productor agropecuario o de un administrador independiente de la condición de tenencia.

Y para el muestreo de elementos.

**La unidad productora**, es la superficie continua de tierra compuesta manejada por un productor agropecuario.

**El lote** es una fracción de terreno de un PSM o unidad productora que está separado del resto del PSM por límites fácilmente reconocibles y en el cual se produce una variedad específica de cultivo, asociación de cultivos, barbecho o descanso.

### 3.5.2 Unidad de observación

Es el elemento básico hacia el cual se dirige el levantamiento de los datos. La unidad mas utilizada es La Unidad de Producción (UP), sin embargo ésta no es la única unidad de observación posible de una variable. Para el muestreo de áreas se utilizan las siguientes unidades de observación .

**El pedazo de segmento de muestreo PSM:** para medir el uso de la tierra de la explotación y a través de éste, estimar las variables pecuarias.

**El lote:** para medir las variables agrícolas. El lote es una parte de terreno del PSM o explotación que está separado del resto del PSM por límites fácilmente reconocibles y en el cual se produce una variedad específica de cultivo, asociación de cultivos, o que se encuentra en barbecho o descanso.

**La unidad productora:** Es un terreno, parte de un terreno o grupo de terrenos en los cuales se desarrolla actividad agropecuaria bajo la dirección de un mismo productor, independiente de la condición de tenencia, forma jurídica, tamaño o ubicación. Para el caso de la CCI, solo se tendrán en cuenta como parte de la UP, los terrenos contiguos al SM y ubicados dentro del departamento donde se encuentra el Segmento de Muestreo (SM).

### 3.5.3. Unidad de análisis

La unidad de análisis corresponde a la entidad mayor o representativa de lo que va a ser objeto específico de estudio en una medición y se refiere al qué o quién es objeto de interés en una investigación. Las unidades de análisis pueden corresponder a las siguientes categorías o entidades:

- Productores Agropecuarios
- Grupos de productores
- Poblaciones completas
- Unidades geográficas determinadas
- Unidades de Producción



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 10  
Fecha: 01-07-08

En el caso del sector agropecuario la unidad de análisis puede ser la Unidad de Producción (UP), el producto agropecuario: arroz, algodón, papa, café, cacao, bovinos, etc.

### 3.6 MARCO ESTADISTICO

El marco de muestreo es dual, es decir se armoniza la información del marco de lista con la cobertura del marco de áreas logrando mayor eficiencia en el diseño. Para su construcción, se divide el área total a ser investigada en pedazos pequeños, llamados unidades primarias de muestreo (UPM), mediante límites naturales y culturales fácilmente identificables sobre el terreno; luego del marco de lista se georeferencian las unidad productoras para evitar el traslape con las áreas.

Los estimadores de múltiples marcos combinan los estimadores de la muestra de áreas con los estimadores de la muestra de lista para cada variable de la encuesta. Si se selecciona una muestra del marco de lista, los resultados obtenidos de las explotaciones seleccionadas se expanden y se suman a la estimación de la muestra de áreas. Si el marco de lista es objeto de una enumeración completa, los resultados obtenidos de la lista se suman a la estimación de la muestra de áreas sin contribución adicional a la varianza total.

Todas las UP del marco de lista se retiran del marco de áreas. Es decir, ningún PSM de la muestra seleccionada de segmentos que corresponda a las explotaciones del marco de lista debe ser considerado para obtener las estimaciones de la muestra de áreas.

La lista de explotaciones se actualiza anualmente antes de la recolección de los datos de campo para la encuesta. La recolección anual o semestral de datos para la lista de las explotaciones especiales se realiza por entrevista directa con el productor o el encargado.

Para tal efecto se cumplen cuatro etapas básicas: La estratificación; (Macroestratificación y Microestratificación); la formación de las unidades primarias de muestreo (UPM); la división de estas en segmentos de muestreo (SM) si son seleccionadas en la muestra y la georeferenciación de las unidades productoras.

La estratificación presenta una fundamentación conceptual en los aspectos geográficos de la actividad agropecuaria del país y está estructurada de acuerdo con reglas de tipo estadístico y de costos de recolección de datos en campo. Clasifica las unidades de muestreo a investigar en el universo de estudio, en diferentes grupos, que se caracterizan por ser homogéneos internamente y heterogéneos entre sí, utilizando información de alguna (s) variable (s) correlacionada(s) directa o indirectamente con las características de la unidad de observación o variable objeto de estudio. Las variables directamente correlacionadas con el objeto de estudio son el área y la producción de los cultivos agrícolas, el uso de la tierra, la geofoma del terreno, los pisos térmicos y las indirectamente correlacionadas, como el tamaño y cantidad de predios o unidades de producción agropecuaria.

La estratificación tal como se ha definido para Colombia, presenta dos niveles de homogeneización del territorio nacional, macro y micro. En el primer nivel priman los factores permanentes y/o estables en el tiempo, como las geofomas del terreno, el clima y las grandes clases de sistemas de producción agrícola, pecuaria y silvícola.



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 11  
Fecha: 01-07-08

En este primer nivel se separan diferentes sistemas de producción propios de geoformas y climas específicos, como por ejemplo, la zona cafetera ubicada en clima templado de las vertientes de las cordilleras Oriental, Central y Occidental; igual situación se logra con los sistemas de producción de cultivos transitorios comerciales en climas cálidos de los valles interandinos, las planicies y el piedemonte.

En el segundo nivel priman factores dinámicos en el desarrollo agropecuario del país tales como la cobertura y uso de la tierra, clases y tipos de cultivos y tamaño de los predios. En este nivel se separan las clases y tipos de cultivos en función de la intensidad del uso de la tierra.

### 3.6.1 Macroestratificación

La macroestratificación del marco de áreas de Colombia es un nuevo concepto de estratificación definido por la Encuesta Agropecuaria, sustentado en factores de tipo agroecológico, así como el departamento es un concepto de estratificación fundamentado en la división político - administrativa del país. Ambos niveles espaciales configuran niveles de estimación, los cuales presentan información de especial importancia para el país.

La división político administrativa del país, define 32 departamentos y 1.117 municipios. El diseño de la muestra contempla estimaciones a niveles departamental y nacional. El área nacional corresponde a 114.174.800 hectáreas. Los departamentos presentan áreas que van desde 4.814 ha., departamento de San Andrés y Providencia, hasta 11.324.575 ha. en departamento del Amazonas.

La macroestratificación del marco de áreas, involucra parámetros generales de tipo agroecológico, estableciendo espacios geográficos relativamente homogéneos, los cuales permiten separar las principales clases de coberturas vegetales tanto naturales (áreas boscosas) como culturales (agroecosistema), dentro de cada uno de los 32 departamentos.

### 3.6.2. Microestratificación

Una vez establecidos los macroestratos a nivel departamental y municipal, se pasa a la microestratificación, donde se realizan delimitaciones más detalladas de los usos de la tierra y particularmente para los principales cultivos que rigen el diseño de la encuesta y que la información auxiliar y la escala de los instrumentos permiten (levantamientos de uso de la tierra y cobertura vegetal).

Mediante la microestratificación se clasifican las UPM a investigar en el universo, en diferentes grupos por intensidad de uso del suelo.

El muestreo agrícola de áreas emplea como una de las variables principales de microestratificación “el uso actual de la tierra”, por cuanto este rasgo geográfico y socio-cultural está altamente correlacionado con las variables agrícolas y pecuarias de interés de esta investigación.



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 12  
Fecha: 01-07-08

Dentro de los conceptos de “intensidad del uso de la tierra”, entendido como la superficie ocupada por un tipo de cultivo, clases de cultivos o grupo de cultivos o cobertura vegetal antrópica o natural, se tienen en cuenta otras variables de estratificación, como por ejemplo, el tamaño de los predios, dado que en Colombia hay una distribución geográfica del tamaño de los predios muy correlacionada con los tipos de agricultura y usos de la tierra.

Esta relación se refleja en el tamaño de las UPM. Es así como en las regiones de los Llanos Orientales donde predominan los grandes predios de más de 10.000 ha., las UPM son de 5.000 ha. y en zonas de minifundio del altiplano cundiboyacense, por ejemplo, para los estratos de cultivos (áreas intensivamente cultivadas y sistemas de producción rotacional) las UPM son de tamaños entre 200 y 500 ha.

En resumen, en la construcción del marco de muestreo de áreas de Colombia, dentro de cada departamento se han definido, según la intensidad de uso de la tierra, 7 micro-estratos; los tres primeros micro-estratos hacen referencia a los cultivos transitorios y permanentes; los demás, a pastos (introducidos), sabanas (naturales), malezas y rastros; bosques; eriales y cuerpos de agua; urbano y semiurbano.

### 3.6.3 Dominios de estudio.

Se considera como un nivel adicional de estratificación definido en función de tipos de cultivos concentrados geográficamente. En cuanto a los cultivos permanentes se encuentran los dominios de palma de aceite en 7 departamentos que cubren un área aproximada de 400.000 ha. ubicados entre 0 y los 1.000 msnm; el banano y plátano de exportación en los departamentos de Antioquia y Magdalena 42.000 ha. y caña de azúcar en el Valle geográfico del río Cauca con 220.000 ha.

Los principales cultivos transitorios comerciales que presentan altos grados de desarrollo tecnológico y se concentran en determinadas áreas geográficas, planas, cercanas a las cabeceras urbanas, con infraestructura de riego y buenas vías de acceso y donde se dan diferentes clases de sistemas de producción: rotación de cultivos, de barbecho y descanso, solos, asociados y mixto, concentrados en el piso térmico cálido, como arroz, sorgo, algodón, maíz, soya y ajonjolí (Denominado ASA), constituyen otro dominio de estudio. Así mismo, en el clima frío, existen los dominios de cultivos transitorios que presentan una distribución muy definida por sus exigencias agronómicas y biofísicas, como es el caso de las hortalizas, la papa, la cebolla; y cereales, como el trigo y la cebada con un alto grado de concentración espacial y altitudinal (la papa en el 90% se concentra en cuatro departamentos y en clima frío) y en clima cálido - templado cacao, caña panelera, uva y piña.

Existe otro grupo de cultivos transitorios con una distribución espacial y altitudinal general, es decir, que se encuentran en todas partes o en un determinado piso térmico, pero de forma generalizada. Es el caso del maíz, el frijol, la yuca y algunas hortalizas y frutales, que se desarrollan bajo múltiples sistemas de producción y se presentan mezclados con las otras clases de coberturas vegetales. Generalmente son cultivos que ocupan pequeñas áreas dentro de las unidades productoras y se destinan al autoconsumo en mayor proporción.

Los pastos y bosques (naturales y plantados) constituyen la cobertura vegetal con mayor área, 44 millones y 57 millones de ha., respectivamente. Los pastos al interior de la frontera agrícola se presentan en todos los pisos térmicos y constituyen un estrato muy evidente desde el punto de vista de gran clase de cobertura vegetal. Los bosques compactos presentan un alto grado de concentración en el área exterior a la frontera agrícola. En este subuniverso y en los correspondientes a los parques naturales nacionales, zonas de páramo, áreas urbanas y semiurbanas, eriales y cuerpos de agua, hasta el momento esta investigación no adelanta ninguna estimación estadística.

### **3.6.4 Marco de lista**

Los marcos de lista están formados a menudo por explotaciones o direcciones de productores. La unidad de información es la UP (Unidad de Producción) y la unidad de muestreo de última etapa es el lote o la explotación.

Los marcos de muestreo de lista proceden generalmente de censos agropecuarios, de censos de población y vivienda, de asociaciones de agricultores, catastros, registro de contribuyentes o de otras fuentes. Suele haber información sobre el tamaño de las explotaciones, cultivos, ganado y otras características de la explotación. Parte de esta información permite estratificar los marcos.

Para que el marco de lista sea útil debe ser actualizada de manera constante y sistemática, lo cual es costoso y requiere tiempo y personal considerables.

Los diseños de muestreo de lista incluyen a menudo por lo menos un estrato de explotaciones especiales cuya enumeración es completa, o del que se toma una muestra utilizando una fracción de muestreo relativamente grande. Estas explotaciones especiales se definen como aquellas que contribuyen en proporción de manera significativa al valor estimado de importantes variables de la encuesta.

Por marcos de lista la ENA investiga los productos siguientes:

Cebolla larga, Algodón, Soya, Maíz tecnificado, Sacrificio de aves, Piscicultura para diez departamentos (Antioquia, Valle del Cauca, Tolima, Huila, Meta, Córdoba, Cundinamarca, Boyacá, Santander y Casanare) .

## **3.7 PARÁMETROS A SER ESTIMADOS**

- **Agrícolas:** Área, superficie en hectáreas dedicadas a los principales cultivos transitorios, permanentes y pastos y se expresa en hectáreas; producción producto obtenido de un cultivo expresado en toneladas y el rendimiento es la producción expresada como toneladas por hectárea.
- **Pecuarios:** Inventario vacuno y otras especies pecuarias, expresado en cantidad de cabezas, la producción de leche en litros de leche el día anterior a la entrevista,

capacidad de carga expresada en unidades de gran ganado (UGG) por hectárea y el área dedicada a pastos en hectáreas.

- **Socioeconómicos:** para el muestreo de conglomerados, la cantidad de PSM o pedazos de segmento y para los censos por producto, el número de unidad productoras, expresada en cantidad de unidades productivas, siendo estas, los terrenos dedicados al cultivo bajo la responsabilidad de un mismo productor y adicionalmente cantidad de productores por cultivo.

### 3.8 DISEÑO MUESTRAL

El diseño muestral comprende todos los aspectos relacionados con la obtención de una muestra, representativa de la población respecto a la características objeto de estudio, y con la decisión sobre la forma de inferir los resultados poblacionales.

El hecho de inferir los resultados de la población a partir de la investigación de una parte de la misma introduce el error de muestreo. Con objeto de que el usuario de la información tenga conocimiento de la calidad de la misma, el diseño muestral proporciona una medida de la precisión de las estimaciones.

El diseño muestral abarca el método de selección de la muestra, método de estimación y formulación y cálculo de errores.

#### 3.8.1 Método de muestreo

El diseño de muestreo de la ENA es el Muestreo de Marcos Múltiples (MMM) que combina el Muestreo del Marco de Áreas (MMA) en dos etapas de selección con el Muestreo de Marcos de Lista (MML).

En el Muestreo de Marco de Áreas se selecciona en la primera etapa las Unidades Primarias de Muestreo (UPM) , mediante EST\_PPT.

En la segunda etapa se escoge mediante MAS un Segmento de Muestreo (SM) dentro de cada Unidad Primaria de Muestreo.

- **Probabilística.** Todas las unidades de muestreo de La población objetivo, tienen una probabilidad de selección conocida y superior a cero.
- **Estratificada.** La población objetivo está clasificada en función de variables altamente correlacionadas con los parámetros a ser estimados, que comprenden para muestreo de conglomerados:

Macroestratos, según cobertura vegetal y piso térmico, se clasifica en:

- T1 Subuniverso con uso del suelo mayoritariamente para cultivos transitorios de clima cálido.

- T2 Subuniverso con uso del suelo mayoritariamente para cultivos transitorios de clima medio.
- T3 Subuniverso con uso del suelo mayoritariamente para cultivos transitorios de clima frío.
- P Subuniverso con uso del suelo mayoritariamente para cultivos permanentes
- p1 Subuniverso con uso del suelo mayoritariamente para cultivo de pastos de clima cálido.
- p2 Subuniverso con uso del suelo mayoritariamente para cultivo de pastos de clima templado.
- p3 Subuniverso con uso del suelo mayoritariamente para cultivo de pastos de clima frío.

*Dominios:*

- 101 Área con más de un 70% en Palma Africana
- 102 Área con más de un 70% en Banano de Exportación
- 103 Área con más de un 70% en Caña de Azúcar
- 104 Área con más de un 70% en Arroz, Sorgo o Algodón
- 106 Área con más de un 70% en Papa
- 107 Área con más de un 70% en Cebolla
- 109 Área con más de un 70% en Caña Panelera

y microestratos dentro de los macroestratos (según la intensidad del uso del suelo).

- 10 Superficie intensamente cultivada, entre un 70% y un 100%
- 20 Superficie medianamente cultivada, entre un 30% y un 69%
- 30 Superficie poco cultivada, entre un 10% y un 29%
- 40 Superficie cultivada en pastos del 90 al 100%
- 50 Son tierras cubiertas del 90 al 100% por bosques
- 60 Superficie sin uso agropecuario son tierras cubiertas del 90 al 100% por eriales, zonas erosionadas, afloramientos rocosos
- 70 Superficie agrourbana

- **Áreas.** Desde el punto de vista estadístico, las áreas son conglomerados desiguales de unidades de superficie, con diversidad de usos agropecuarios. El diseño contempla tres tipos de áreas o conglomerados: UPM, SM y PSM.
- **Bietápica.** En la primera etapa se seleccionan probabilísticamente UPM dentro de cada estrato y en la segunda se selecciona un SM dentro de cada UPM seleccionada en la primera etapa.

### **3.8.2 Tamaño de muestra.**

Se calcula para conglomerados, mediante la fórmula de un diseño muestral estratificado de conglomerados desiguales, para los cultivos más importantes dentro de cada macroestrato por departamento.

El tamaño de la muestra se determina para los productos más importantes dentro de cada estrato por cultivo y departamento, así:

$$n_{ih} = \frac{\left( \sum_h \frac{N_h \sigma_h}{b_h} \right)^2}{\frac{\epsilon^2}{4} + \sum_h \frac{N_h \sigma_h^2}{b_h}} \quad (1)$$

donde:

$n_{ih}$  = Tamaño de muestra para el cultivo i en el macroestrato h por departamento, en hectáreas.

$N_h$  = Tamaño del universo de estudio (hectáreas) en el microestrato h-ésimo de un departamento.

$\sum N_h = N$  universo total de un departamento.

$h = T_1, T_2, T_3, P, p_1, p_2, p_3$

$b_h$  = área promedio por SM en el estrato h

$\sigma_h^2$  = Varianza natural del área de un determinado producto, en el estrato h-ésimo.

$\frac{\epsilon^2}{4} = \frac{d^2}{t^2} = V$  = varianza deseada o esperada en la formula de Cochran

En el cuadro 2 se muestra los tamaños de muestra en hectáreas y en número de segmentos para cada una de las encuestas realizadas, en el periodo del año 1995 al año 2010.

*Cuadro2. Distribución de tamaños de muestra, según año de realización de la Encuesta Nacional Agropecuaria*

<sup>1</sup> COCHRAN, G. (1977): "Técnicas de Muestreo", John Wiley and Sons. New York (Edición en Español): Cia. Editorial Continental S.A. Mexico. 1982).



Tamaño de la muestra para la ENA de 1995 a 2010

| Año  | Primer trimestre (hortalizas) |                | Primer semestre |                | Segundo semestre |                |
|------|-------------------------------|----------------|-----------------|----------------|------------------|----------------|
|      | ha                            | cantidad de SM | ha              | cantidad de SM | ha               | cantidad de SM |
| 1995 |                               |                | 1.120.000       | 5.569          | 1.044.568        | 4.187          |
| 1996 |                               |                | 348.634         | 1.719          | 1.044.568        | 5.187          |
| 1997 |                               |                |                 |                | 1.355.439        | 6.187          |
| 1998 |                               |                |                 |                | 960.000          | 6.680          |
| 1999 |                               |                |                 |                | 1.700.000        | 10.000         |
| 2000 |                               |                |                 |                | 1.080.000        | 6.853          |
| 2001 |                               |                |                 |                |                  | 6.216          |
| 2002 |                               |                |                 |                | 718.986          | 5.276          |
| 2003 |                               |                |                 |                | 744.576          | 5.419          |
| 2004 |                               |                |                 |                | 750.230          | 5.534          |
| 2005 |                               |                |                 |                | 502.548          | 4.352          |
| 2006 |                               |                |                 |                | 827.516          | 5.539          |
| 2007 |                               |                | 244.090         | 2.283          | 877.911          | 5.796          |
| 2008 |                               |                | 256.253         | 2.353          | 945.062          | 5.820          |
| 2009 |                               |                | 296.296         | 2.582          | 947.392          | 5.820          |
| 2010 | 1735                          | 95             | 293.528         | 2.622          | 934.092          | 5.870          |

Fuente: CCI-ENA

Para el muestreo de elementos el tamaño se calcula por producto así:

$$n = \frac{\left( \sum \frac{N_h \sigma_h}{W_h} \right)^2}{V + \frac{1}{N} \sum N_h \sigma_h^2} \quad (2)$$

$n$  =Tamaño de muestra de un departamento en cantidad de unidades productoras

$N_h$  =Tamaño del universo de estudio (número de unidad productoras) en el estrato h-ésimo de un departamento.

$\sum N_h = N$ , es el universo total de unidad productoras de un departamento.

$h = 1,2,\dots,H$ , son los diferentes estratos en el departamento.

$\sigma_h^2$  = Varianza natural del área (x) en el estrato h-ésimo.

<sup>2</sup> COCHRAN, G. (1977): "Técnicas de Muestreo", John Wiley and Sons. New York (Edición en Español): Cia. Editorial Continental S.A. Mexico. 1982).

$W_h = \frac{N_h}{N}$  Ponderación que controla el variable tamaño de los estratos,

con una asignación óptima supuesta para un tamaño de muestra fijo n

$$n_h = \frac{nW_h \sigma_h}{\sum_h w_h \sigma_h}$$

### **3.8.3 Selección de la muestra**

En el Marco de Muestreo de Áreas se seleccionan en la primera etapa las Unidades Primarias de Muestreo (UPM) mediante EST\_PPT de áreas.

En la segunda etapa y con igual probabilidad de selección, se escoge mediante MAS un Segmento de Muestreo (SM) dentro de cada Unidad Primaria de Muestreo.

### **3.8.4 Método de Estimación**

Se denomina así, la metodología estadística para garantizar estimaciones insesgadas de los parámetros de estudio. Hay procedimientos básicos para todas las encuestas; procedimientos para la comparación de los resultados de dos encuestas, y procedimientos para la consolidación de resultados anuales a partir de dos encuestas semestrales (1995 a 1997 también 2007, 2008 y 2009) y actualmente procedimientos para la estimación proveniente del muestreo dual.

Con el fin de garantizar estimaciones de los parámetros de interés a nivel de los universos nacional y departamental, se ponderan los resultados muestrales de cada segmento por el recíproco de su probabilidad de selección, denominado Factor de Expansión.

#### **Procedimientos de estimación**

##### **➤ Procedimientos básicos**

Para la estimación de los indicadores de interés a nivel de los universos nacional y departamental, se ponderaron los resultados muestrales, de cada segmento por el recíproco de su probabilidad de selección, denominado Factor de Expansión.

#### **Factor de expansión**

**Primera fase del diseño muestral**

$F_{hi\alpha}$  = **Factor de Expansión del segmento  $\alpha$  de la  $i$ -ésima UPM del estrato  $h$  de un departamento**

$$F_{hi\alpha} = \frac{1}{f_{hi\alpha}}$$

$f_{hi\alpha}$  = **Probabilidad final de selección del segmento  $\alpha$  de la UPM  $i$  del estrato  $h$  de un departamento**

$$f_{hi\alpha} = \left( \frac{x_{hi}}{X_h} A_h \right) \left( \frac{x_{hi\alpha}}{x_{hi}} a_{hi} \right)$$

Es decir la probabilidad de selección de la UPM  $i$  en el estrato  $h$  y la probabilidad de selección del segmento  $\alpha$  en la UPM  $i$

**En donde:**

$X_h$  = **área planimetrada del estrato  $h$**

$x_{hi}$  = **área planimetrada de la UPM  $i$  en el estrato  $h$**

$x_{hi\alpha}$  = **área planimetrada del  $\alpha$ -iésimo segmento seleccionado en la UPM  $i$  del estrato  $h$**

$A_h$  = **cantidad de UPM seleccionadas en el estrato  $h$**

$a_{hi}$  = **cantidad de SM seleccionados en la UPM  $i$  del estrato  $h$**

**Segunda fase de muestreo con una submuestra traslapada totalmente**

$F_{hi\alpha\alpha^1\beta}$  = **Factor de Expansión en el PSM  $\beta$  del SM  $\alpha^1$  del SM  $\alpha$  de la UPM  $i$  del estrato  $h$  de un departamento (o del país) y grupo de cultivo determinado.**

$$F_{hi\alpha\alpha^1\beta} = \frac{1}{f_{hi\alpha\alpha^1\beta}}$$

$f_{hi\alpha\alpha^1\beta}$  = **Probabilidad final de selección del PSM  $\beta$ , en el SM  $\alpha^1$  del SM  $\alpha$ , de la UPM  $i$  del estrato  $h$  de un departamento y cultivo determinado.**

$$f_{hi\alpha\beta} = f_{hi\alpha} * f_{\alpha^1\beta}$$

$f_{hi\alpha}$  = Probabilidad final de selección en la primera fase del segmento  $\alpha$  de la UPM  $i$  del estrato  $h$  de un departamento, explicado inicialmente y ,

$$f_{\alpha^1\beta} = \frac{\sum_{\alpha^1} x_{hi\alpha\alpha^1}}{\sum_{\alpha} x_{hi\alpha}} * \frac{\sum_{\beta} x_{hi\alpha\alpha^1\beta}}{\sum_{\alpha^1} x_{hi\alpha\alpha^1\beta}}$$

En donde:

$x_{hi\alpha\alpha^1}$  = área del SM  $\alpha^1$  seleccionado de los SM  $\alpha$ , de la UPM  $i$ , del estrato  $h$

$x_{hi\alpha\alpha^1\beta}$  = área del PSM  $\beta$  seleccionado en el SM  $\alpha^1$  seleccionado de los  $\alpha$  SM seleccionados en la primera fase, de la UPM  $i$  del estrato  $h$ , para un cultivo determinado.

### Metodología de estimación

La estimación de área es:

$$(Fx)_D = \sum_h \sum_i \sum_{\alpha} x_{hi\alpha} F_{hi\alpha}$$

La estimación de rendimiento es:

$$\left(\hat{F}\right)_D = \frac{\sum_h \sum_i \sum_{\partial} \sum_{\alpha^1} \sum_{\beta} r_{hi\alpha\alpha^1\beta} x_{hi\alpha\alpha^1\beta} F_{hi\alpha\alpha^1\beta}}{\sum_h \sum_i \sum_{\partial} \sum_{\partial^1} \sum_{\beta} x_{hi\alpha\alpha^1\beta} F_{hi\alpha\alpha^1\beta}}$$

La estimación de producción es:

$$(Fy)_D = \sum_h \sum_i \sum_{\alpha} y_{hi\alpha} F_{hi\alpha}$$

### Procedimientos para la consolidación de resultados anuales

Para generar los resultados de cada año, se realizan dos encuestas: una en el primer semestre para la obtención de resultados de los principales cultivos semestrales de papa, arroz, trigo, cebada, zanahoria y hortalizas, entre otras. Una segunda encuesta se realiza en el segundo semestre para la obtención de resultados de cultivos semestrales, anuales, permanentes, área en pastos, uso del suelo, inventario ganadero, producción de leche y número de PSM.

### Estimación anual del uso de la tierra, área en pastos



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 21  
Fecha: 01-07-08

Las estimaciones se generan a partir de la encuesta de los segmentos referidos al segundo semestre, inferidos al universo agropecuario. Se obtienen datos a nivel nacional y departamental.

### **Estimación anual de área y producción para cultivos transitorios**

La estimación anual del área y la producción de los cultivos transitorios, se realiza con al suma de las estimaciones generadas en al muestra de primer semestre mas las generadas con al muestra de segundo semestre inferidas a nivel nacional y departamental.

Respecto a las áreas sembradas, la estimación anual resulta de sumar las estimaciones de cada una de las mediciones semestrales realizadas y ajustados con un indicador de cobertura obtenido a través de estimaciones anteriores de la misma ENA.

En cuanto a las áreas cosechadas en el año, estas son calculadas con base en las mediciones de la encuesta de segundo semestre del 2009 que reporta las áreas cosechadas en el primer semestre y sumadas con aquellas áreas reportadas en la encuesta del primer semestre del 2010, que fueron sembrados y cosechados en este mismo semestre. Para el caso del segundo semestre, se suman las estimaciones de la encuesta del primer semestre que serían cosechadas en el segundo semestre con las reportadas en la encuesta del segundo semestre que fueron sembrados y cosechados en este mismo semestre.

La producción es obtenida a través de la multiplicación de las áreas cosechadas en cada semestre con el rendimiento estimado en cada una de las mediciones de los semestres.

## **Estimación anual de área y producción para cultivos permanentes**

La estimación del área plantada y la producción se genera para el nivel departamental y nacional, a partir de la muestra del segundo semestre.

## **Estimación del inventario ganadero y de la producción de leche**

La estimación del inventario ganadero y la producción de leche el día anterior a la entrevista, se genera a partir de la muestra de segundo semestre. A nivel de muestra, la información se tomó referida a toda la unidad productora y fue ponderada para el PSM.

## **Estimación de las diferencias y su significancia estadística**

Para la estimación de la diferencia entre un año y otro y su significancia estadística se aplica una prueba, por medio de la cual se estiman las diferencias relativas, los errores estándar de la diferencia absoluta y el resultado del test de significancia estadística al 95% de confiabilidad, para las variables uso de la tierra, a nivel nacional y departamental, al igual que la superficie cultivada y la producción para algunos productos a nivel nacional.

Si el resultado de la prueba es SI se puede afirmar que existe una diferencia estadísticamente significativa, lo cual confirma la existencia de un fenómeno de cambio agropecuario (con un 10% de riesgo a equivocarse).

Si por el contrario la diferencia NO es estadísticamente significativa, ello puede tener diferentes interpretaciones:

- a. Si no se esperan cambios en los indicadores comparados, o los cambios esperados son mínimos, el resultado de la prueba de significancia simplemente lo ratifica, independientemente de la magnitud del error estándar.
- b. Si se esperaban cambios importantes; pero el error de muestreo (para el 95%) es aún más grande que la diferencia observada, dicho cambio simplemente no puede ser validado con los resultados observados en los indicadores (aunque sea obvio), y para lograrlo en el futuro se requiere un mayor tamaño de muestra y un diseño muestral más eficiente.
- c. Si se tiene incertidumbre sobre los cambios, y era indispensable arribar a una conclusión agropecuaria, pero el error de muestreo es demasiado grande, debe también en el futuro, aumentarse el tamaño de la muestra y mejorar el diseño.

### **Estimación de los errores muestrales**

Se entiende como el grado de aproximación con que se estiman las características del universo de estudio a través de una muestra probabilística tomada aleatoriamente, utilizando como base las fórmulas desarrolladas a partir de las clásicas series de Taylor del muestreo de conglomerados desiguales. Con base en este error se estiman los límites de confianza entre los cuales están los valores del universo investigado. La estimación de los valores del universo y de los límites de confianza de los mismos, se denomina inferencia estadística.

$$ESrel = \frac{\sqrt{Var(\hat{y}_r)}}{\hat{y}_r}$$

**donde:**

*ESrel* : error estándar relativo

*Var*( $\hat{y}_r$ ) : varianza estimada de la producción estimada

$\hat{y}_r$  = producción estimada como área estimada por el rendimiento estimado

### Error estándar del área estimada a nivel departamental

$$\begin{aligned}
 es(Fx)_D &= \sqrt{var(Fx)_D} \\
 &= \sqrt{\sum_h^H \frac{I}{(a_h - 1)} \left[ a_h \sum_{\alpha} (1 - f_{hi\alpha}) (F_{hi\alpha} x_{hi\alpha})^2 - \left( \sum_{\alpha} \sqrt{(1 - f_{hi\alpha})} F_{hi\alpha} x_{hi\alpha} \right)^2 \right]}
 \end{aligned}$$

Para el cálculo del error estándar de la estimación para el universo de estudio sobre área, se partió de un diseño Estratificado de áreas con muestreo proporcional al tamaño.

En donde:

$F_{hi\alpha}$  = factor de expansión del segmento  $\alpha$ , de la UPM  $i$ , del estrato  $h$

$x$  = área en la muestra

$Fx$  = área estimada a nivel departamental

$F_{hi\alpha} x_{hi\alpha}$  = área de un segmento de la muestra ponderado por el factor de expansión de dicho segmento

$a_h$  = Cantidad de segmentos seleccionados en el estrato  $h$

### Error estándar del rendimiento estimado a nivel departamental

Para el cálculo del error estándar de la estimación para el universo de estudio sobre rendimiento, se partió de un diseño efectivo en dos etapas estratificado de conglomerados desiguales, en una submuestra traslapada totalmente.

El super-índice 1 indica que los cálculos se efectúan en la submuestra.

En donde:

$$VAR(\hat{r})_D = \frac{I}{(\sum F^1)^2} \left[ \sum_h^H \frac{I}{(a_h^1 - 1)} \left( \frac{a_h^1 \sum_1^I \sum_{\alpha} \sum_{\alpha'} \sum_{\beta} (1 - f_{hi\alpha\alpha'\beta}) (F_{hi\alpha\alpha'\beta} r_{hi\alpha\alpha'\beta})^2 - \left( \sum_1^I \sum_{\alpha} \sum_{\alpha'} \sum_{\beta} \sqrt{(1 - f_{hi\alpha\alpha'\beta})} F_{hi\alpha\alpha'\beta} r_{hi\alpha\alpha'\beta} \right)^2 \right) \right]$$

$$es(\hat{r})_D = \sqrt{var(\hat{r})_D}$$

con:



$a_h^1$  : cantidad de PSM en el estrato  $h$

$$\sum F^1 = \sum_h \sum_i \sum_{\partial} \sum_{\alpha^1} \sum_{\beta} F_{hi\alpha^1\beta}$$

### Error estándar de la producción estimada a nivel departamental

$$\begin{aligned} \text{es}(Fy)_D &= \sqrt{\text{var}(Fy)_D} \\ &= \sqrt{\sum_h \frac{1}{(a_h - 1)} \left[ a_h \sum_{\alpha} (1 - f_{hi\alpha}) (F_{hi\alpha} y_{hi\alpha})^2 - \left( \sum_{\alpha} \sqrt{(1 - f_{hi\alpha})} F_{hi\alpha} y_{hi\alpha} \right)^2 \right]} \end{aligned}$$

En donde:

$F_{hi\alpha}$  = factor de expansión del segmento  $\alpha$ , de la UPM  $i$ , del estrato  $h$

$y$  = Variable producción

$Fy$  = producción estimada a nivel departamental

$F_{hi\alpha} y_{hi\alpha}$  = producción de un segmento de la muestra ponderado por el factor de expansión de dicho segmento

$a_h$  = Cantidad de segmentos seleccionados en el estrato  $h$

### 3.9 LA ESTRATEGIA DE CRITICA E IMPUTACION

**3.9.1 Definición de la crítica estadística:** La crítica es el proceso de identificar los valores erróneos o faltantes, la imputación es el proceso de corregirlos. La relación entre estos dos procesos es muy estrecha tanto que generalmente se asocia a un único proceso.

La crítica estadística es el proceso por el cual se garantiza que la información de la encuesta es aceptable en termino de completitud, consistencia y corrección. La evaluación de si los datos de la encuesta son o no aceptables depende de la utilidad que se haga de ellos. Desde un punto de vista general la calidad de la información esta asociada al grado de varianza de los resultados, la magnitud de los errores debidos al operativo y la pertinencia de los métodos de observación y medición.

#### 3.9.2 Proceso de crítica estadística

**Durante la recolección:** existe una crítica que realiza el encuestador, quien revisa y corrige a medida que realiza la entrevista. En estos casos el encuestador se basa en reglas preestablecidas pero también en la experiencia y el sentido común y juicio a cerca de los que es aceptable.

La siguiente forma es la realizada por el supervisor. Entre cuyas funciones esta la de revisar una parte de las entrevistas para encontrar errores.

En el proceso de imputación se ejecutan procesos sistemáticos, mediante los cuales se asignan valores plausibles a las variables para las cuales no se obtuvo información en campo, de acuerdo con parámetros técnicos definidos según el comportamiento de éstas en el sector y en la región.

Los datos expandidos y agrupados a nivel de estrato son acumulados a los niveles superiores, como región, departamento y país. Cada uno de estos niveles suministra la información necesaria para la realización de la fase subsiguiente de análisis de consistencia. En la última fase, se hace una revisión de los resultados de cada una de las variables investigadas, verificando su consistencia frente a otras variables de la encuesta o frente a parámetros técnicos dados por la misma investigación como por otras fuentes.

La publicación de los resultados está condicionada por el error estándar relativo de las estimaciones de la variable área, este valor depende esencialmente del tamaño de la muestra.

### **3.9.3 Imputación y Cobertura**

En el proceso de imputación se ejecutan procesos sistemáticos, mediante los cuales se asignan valores plausibles a las variables para las cuales no se obtuvo información en campo, de acuerdo con parámetros técnicos definidos según el comportamiento de éstas en el sector y en la región.

El proceso de cobertura consiste en cubrir adecuadamente el país cumpliendo con la meta de estimar todas las áreas y Unidades de Producción que conforman el universo agropecuario de la nación. Para lograr este propósito se hace necesario imputar datos de la ENA, a partir de encuestas o censos realizados en años anteriores. Dicho proceso implica determinar las áreas a nivel de estrato, región o departamento que no son objeto de selección de muestra o realización de censos por parte de la ENA actual. Una vez hecho este proceso se procede a cubrir estas áreas con los datos de las encuestas o censos más recientes y de esta manera generar información a nivel nacional de acuerdo con los datos reportados por el INSTITUTO GEOGRAFICO AGUSTIN CODAZZI.

### **3.10 ERRORES NO MUESTRALES**

Los errores no muestrales son aquellos que no son fruto del diseño de la muestra sino de factores ajenos a la misma tales como falta de cobertura, mal diligenciamiento del cuestionario, la no-respuesta total o la respuesta parcial.



## METODOLOGÍA DE DISEÑO MUESTRAL

Código: DM-ENA-DIM-01  
Versión : 012  
Página 27  
Fecha: 01-07-08

Los errores de no-respuesta total o parcial y de respuesta se manejan así: Para minimizar el error de no-respuesta total, se selecciona un porcentaje diferencial de sobre muestra por departamento. Se ejecuta procesos sistemáticos de imputación, mediante el cual se asignan valores plausibles a las variables de las cuales no se obtuvo información en campo, de acuerdo a parámetros técnicos definidos según el comportamiento de estas en el sector y en la región.

Los datos expandidos y agrupados a nivel de estrato son acumulados a los niveles superiores como región, departamento y país. Cada uno de estos niveles suministra la información necesaria, para la realización de la fase subsiguiente de análisis de consistencia. En esta última fase, se hace una revisión de los resultados de cada una de las variables investigadas, verificando su consistencia frente a otras variables de la encuesta o frente a parámetros técnicos dados tanto como por la misma investigación como por otras fuentes.

La publicación de los resultados está condicionada por el error estándar relativo de las estimaciones de la variable área. Este valor depende esencialmente del tamaño de la muestra, que como ya se mencionó, ha sido siempre inferior al requerido de acuerdo con el diseño muestral.

### BIBLIOGRAFÍA

**COCHRAN, G. (1977): “Técnicas de Muestreo”, John Wiley and Sons. New York (Edición en Español): Cia. Editorial Continental S.A. Mexico. 1982).**

**HOUSEMAN, E.E. (1975): “ El Muestreo de Areas en la Agricultura”, SRS, USDA No. 20 (Edición en Español: SRS. USDA. No 210-S, 1975).**

**KISH, I. (1965): “Muestreo de Encuestas” John Wiley and Sons, New York (Edición en Español: Editorial Trillas, Mexico. 1972).**

**OTAÑEZ, G. (1988). Metodología del Muestreo Agrícola de Áreas en Colombia. PNUD/FAO, Bogota- Colombia.**

**OTAÑEZ, G. (Septiembre 1986): Manual para la Construcción del Marco de Muestreo de Áreas y Selección de la Muestra”, FAO, Bogotá – Colombia.**



## ANEXO 1

### **AJUSTES A LA METODOLOGIA DE LA ENA PARA EL AÑO 2010**

Para el desarrollo de la metodología en el año 2010 se conserva el diseño inicial. Se rota un 10% de la muestra y se hacen ajustes al estrato de hortalizas en los departamentos de Cundinamarca, Boyacá y Nariño y al estrato de cebolla larga en los departamentos de Boyacá y Nariño.

En total se procede a rotar el 9.5% de los Segmentos de Muestreo que hacen parte de la muestra de seguimiento y que corresponden a la replica cuatro (4). De los cuales 380 segmentos se rotaron aleatoriamente y 183 se rotaron forzosamente, los primeros representan el 6.4% y el complemento el 3.1%.

Los segmentos de Muestreo de inclusión forzosa, incluye Segmentos de Muestreo que en la ultima encuesta no fueron encuestados por problemas de orden publico o porque el productor se negó a suministrar información, también por problemas de ubicación en el campo en el sentido de estar ubicados a grandes distancias de vías carreteables.

Los segmentos rotados aleatoriamente corresponden a segmentos de la replica 4 en los departamentos de Antioquia, Bolívar, Caldas, La guajira, Magdalena, Norte de Santander, Risaralda y Santander, a los que por el sistema les correspondió la rotación este año.

Para mejorar los resultados de hortalizas y de cebolla larga se construyo un estrato especial, para cada producto. Para hortalizas se construyo estrato especial en los departamentos de Boyacá, Cundinamarca y Nariño. Para el cultivo de cebolla larga se construyo estrato especial en los departamentos de Boyacá y Nariño. Es-

ta mejora en el Marco de Áreas repercute en el aumento del tamaño de muestra para estos estratos y por tanto una mayor precisión (esrel del 5%) en los resultados de estos cultivos.

| Rotación  | Concepto           | Cantidad | %     |
|-----------|--------------------|----------|-------|
| Nuevos    | Cebolla larga      | 45       | 0,8   |
|           | Hortalizas de hoja | 50       | 0,9   |
|           | nueva              | 77       | 1,3   |
| Rotados   | rotada             | 376      | 6,4   |
|           | Cambio forzoso     | 104      | 1,8   |
| sin rotar |                    | 5218     | 88,9  |
|           | Total general      | 5870     | 100,0 |

Un total de 5218 SM que representan el 88.9% se mantienen en la muestra sin sufrir efectos de la rotación, los cuales pertenecen a la muestra de seguimiento y están asignados a las replicas 1, 2, 3 y 5 respectivamente.

La mayor rotación aleatoria de la muestra se da en el departamento de Antioquia (34,2%), seguido de el departamento de Santander(23.9%). Por el contrario, el departamento con menor numero de segmentos rotados aleatoriamente es La Guajira con el 4.5%.

**Cuadro 2. Rotación de la muestra, según departamento. Año 2010**

| departa       | Cebolla larga | Hortalizas de hoja | Nueva | Sub total | Rotada | Cambio forzoso | Sub total | sin rotar | Total |
|---------------|---------------|--------------------|-------|-----------|--------|----------------|-----------|-----------|-------|
| Antioquia     |               |                    | 7     | 7         | 127    | 13             | 140       | 507       | 654   |
| Atlantico     |               |                    | 8     | 8         |        |                | 0         | 37        | 45    |
| Bolivar       |               |                    | 10    | 10        | 28     |                | 28        | 112       | 150   |
| Boyaca        | 30            | 8                  |       | 38        |        | 15             | 15        | 536       | 589   |
| Caldas        |               |                    | 2     | 2         | 28     | 1              | 29        | 112       | 143   |
| Casanare      |               |                    | 5     | 5         |        | 6              | 6         | 119       | 130   |
| Cauca         |               |                    | 8     | 8         |        |                | 0         | 191       | 199   |
| Cesar         |               |                    |       | 0         |        | 4              | 4         | 233       | 237   |
| Cordoba       |               |                    | 5     | 5         |        | 2              | 2         | 231       | 238   |
| Cundinamarca  |               | 31                 |       | 31        |        | 12             | 12        | 647       | 690   |
| Guajira       |               |                    | 9     | 9         | 17     |                | 17        | 71        | 97    |
| Huila         |               |                    |       | 0         |        | 5              | 5         | 271       | 276   |
| Magdalena     |               |                    | 10    | 10        | 25     |                | 25        | 89        | 124   |
| Meta          |               |                    | 2     | 2         |        | 18             | 18        | 260       | 280   |
| Nariño        | 15            | 11                 |       | 26        |        | 1              | 1         | 421       | 448   |
| Norte de Stdr |               |                    |       | 0         | 38     | 17             | 55        | 156       | 211   |
| Quindio       |               |                    | 2     | 2         |        |                | 0         | 65        | 67    |
| Risaralda     |               |                    |       | 0         | 22     |                | 22        | 80        | 102   |
| Santander     |               |                    | 3     | 3         | 91     | 3              | 94        | 358       | 455   |
| Sucre         |               |                    |       | 0         |        | 2              | 2         | 217       | 219   |
| Tolima        |               |                    |       | 0         |        | 2              | 2         | 298       | 300   |
| Valle         |               |                    | 6     | 6         |        | 3              | 3         | 207       | 216   |
| Total         | 45            | 50                 | 77    | 172       | 376    | 104            | 480       | 5218      | 5870  |

El mayor número de segmentos a rotar forzosamente se presenta en los departamentos del Meta, Antioquia, Norte de Santander y Boyacá. En Risaralda no se presenta ninguna rotación forzosa y en este orden los departamentos con menor número de rotación forzosa son: Nariño, Sucre y Tolima con escasamente 1 segmento para el primero y 2 segmentos para los dos siguientes.