



# Diseño muestral

 ESPAC 2020

Mayo, 2021





## Documento de diseño muestral de la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC, 2020).





**Dirección**

Dirección de Infraestructura Estadística y Muestreo

**Unidad**

Gestión de Diseño Muestral

**Elaborado por:**

Pilar Carrillo

**Revisado por:**

Javier Núñez  
Francisco Céspedes

**Aprobado por:**

María Isabel García





## Contenido

Introducción.....	5
Objetivo .....	6
1. Determinación del Marco de Muestreo .....	6
1.1. Marco de Áreas .....	6
1.2. Marco de Lista.....	11
2. Diseño Muestral.....	13
2.1. Universo de la investigación .....	13
2.2. Población Objetivo .....	13
2.3. Unidades de muestreo y análisis.....	13
2.4. Cobertura geográfica .....	13
2.5. Tipo de muestreo .....	14
2.6. Tamaño de muestra.....	14
2.7. Selección de la muestra.....	15
2.8. Cobertura muestral .....	16
3. Cálculo de factores de expansión .....	16
4. Estimación de características .....	17
5. El error de muestreo .....	19
Bibliografía .....	20

## Lista de tablas

Tabla 1. Estratos del Marco de Muestreo para Investigaciones Agropecuarias.....	8
Tabla 2. Tamaño de las Unidades Primarias de Muestreo .....	10
Tabla 3. Distribución del número de segmentos del marco de áreas por provincia y por estrato ESPAC – 2019 .....	10
Tabla 4. Número de segmentos y superficie del Marco de Áreas de ESPAC 2019 ..	11
Tabla 5. Fuentes de información consideradas para el Marco de Lista ESPAC-2020 .....	12
Tabla 6. Código de estrato por producto .....	13
Tabla 7. Tasas de muestreo por estrato .....	14
Tabla 8. Tamaños de muestra por provincia y por estrato ESPAC – 2020.....	14
Tabla 9. Cobertura muestral.....	16

## Lista de figuras

Figura 1. Generación de las Unidades Mínimas de Estratificación (UME) .....	7
Figura 2. Límites provinciales del Marco de Muestreo para Investigaciones Agropecuarias. Un ejemplo para las provincias de Manabí y Esmeraldas .....	8
Figura 3. Áreas para estratificación .....	9
Figura 4. Distribución del número de segmentos por estrato .....	11
Figura 5. Zonas de selección del estrato 1b de la provincia de Santo Domingo de los Tsáchilas.....	15
Figura 6. Selección de la muestra del estrato 3 .....	16





# Introducción

El Instituto Nacional de Estadística y Censos (INEC) es la institución pública que coordina, norma y evalúa la producción de la información estadística oficial proveniente del Sistema Estadístico Nacional, con el propósito de entregar a la sociedad y al Estado información de calidad, pertinente, veraz y oportuna, que contribuya al desarrollo nacional.

La necesidad de generar herramientas de decisión para la planificación y formulación de política agropecuaria en el país es permanente, por lo que es prioritario contar con estadísticas agropecuarias con regularidad, tales como la producción de cultivos y de ganado. Esta información se recoge generalmente mediante registros administrativos y/o encuestas por muestreo.

Con este antecedente y con la finalidad de estimar los principales indicadores agropecuarios y ambientales, el INEC realiza anualmente la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC), la cual es el pilar fundamental del actual Sistema de Estadísticas Agropecuarias Nacionales de Ecuador.

El proceso que se realiza para la generación del diseño muestral consta de los siguientes puntos:

En la primera etapa:

- Definición del marco de muestreo
- Las unidades de muestreo y análisis
- Cobertura geográfica
- Cobertura temática

En la segunda etapa:

- Metodología para el diseño muestral
- Determinación de los dominios de estudio
- Determinación del tamaño de la muestra
- Selección de la muestra

Por último, se realiza:

- Validación de la cobertura obtenida en campo
- El cálculo de los factores de expansión
- Estimación de características
- Error de muestreo
- Declaración del plan de muestreo





# Objetivo

Elaborar el diseño muestral de la “Encuesta de Superficie y Producción Agropecuaria Continua 2020”, utilizando técnicas de muestreo probabilístico que permitan emitir resultados confiables de las principales variables que se investigará en dicha encuesta.

## 1. Determinación del Marco de Muestreo

Kish [3] manifiesta que el marco de muestreo es la piedra angular en torno a la cual los procesos de selección deben ser diseñados. Por tal motivo, es indispensable que el marco se encuentre completo, validado y depurado antes de iniciar el cálculo y la selección de la muestra.

El Marco de Muestreo de ESPAC, concierne a un **Marco de Muestreo Múltiple**, el cual se define como una combinación de un Marco de Áreas y un Marco de Lista (marco dual). Las estimaciones basadas en muestras de áreas tienden a ser sensibles a valores extremos, por lo que para conseguir estabilidad se incluye una lista de unidades de observación más extrema, denominado Marco de Lista.

Un Muestreo de Marcos Múltiples es más eficiente que un Marco de Áreas, puesto que proporciona estimaciones más exactas y el costo adicional de elaborar el Marco de Lista es muy bajo. En encuestas agropecuarias se usa frecuentemente un marco dual, el cual es completo, en el sentido de que cualquier unidad de observación de la población objetivo está en al menos uno de sus dos componentes.

### 1.1. Marco de Áreas

Un marco de áreas se construye a partir de la división del área a ser investigada en pequeños pedazos, generalmente conocidos como segmentos.

El conjunto de segmentos conforma el Marco de Áreas, en cuya construcción se utiliza material cartográfico (mapas, imágenes de satélite y fotografías aéreas) para delimitar las unidades de muestreo y medir su superficie.

Para la construcción del Marco de Muestreo de Áreas para investigaciones agropecuarias, se empieza definiendo una malla cuadrada homogénea sobre el territorio continental ecuatoriano cuyos elementos tienen un área de 576 ha., de tal forma que esta forme un recubrimiento finito del mismo. A los elementos de la malla cuadrada homogénea los denominaremos Unidad Mínima de Estratificación (UME).<sup>1</sup>

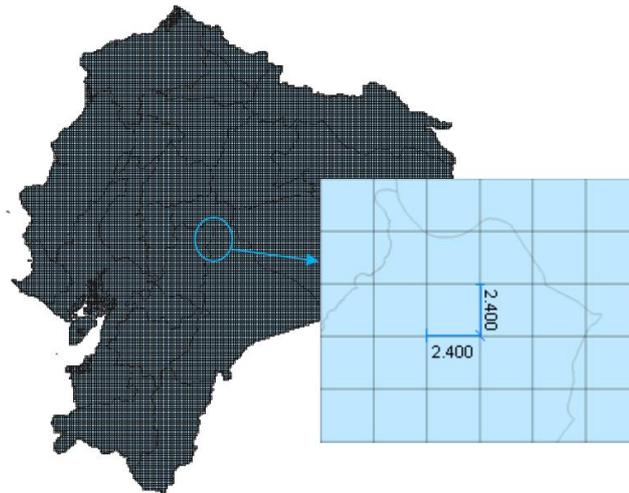
---

<sup>1</sup> Reciben este nombre porque las UME serán estratificadas en función a la intensidad de uso de suelo, procedimiento que será explicado más adelante.





**Figura 1.** Generación de las Unidades Mínimas de Estratificación (UME)



Fuente: INEC [4]

Para la construcción del marco de áreas se debe excluir zonas del territorio ecuatoriano en las cuales la producción agropecuaria es ínfima y que no afecta significativamente a los objetivos de la operación estadística. Entre estas zonas se encuentran:

- Segmentos completamente contenidos en superficies mayores a los 3.000 metros sobre el nivel del mar (msnm).
- Segmentos completamente contenidos en superficie sin uso agropecuario tales como reservas naturales, selva amazónica, entre otros.
- Segmentos completamente contenidos en superficies urbanas.
- Segmentos con intersección parcial o total con sectores amanzanados pertenecientes al Marco de Muestreo generado para el Sistema Integrado de Encuestas de Hogares (SIEH).<sup>2</sup>

Se puede garantizar que el marco de muestreo es completo puesto que el conjunto de todas las UME forman un recubrimiento finito del territorio continental ecuatoriano. Además, puesto que cada UME ocupa una superficie única del mismo, se entiende que no existe intersección entre cualquier par de UME. Esto implica que dentro del marco de muestreo podemos identificar de manera única cada UME, evitando así la duplicidad de elementos del mismo. Cabe destacar que esta propiedad se mantiene en el caso de dividir cualquier UME.

Las UME ubicadas sobre las zonas fronterizas y costeras del territorio ecuatoriano abarcan superficie de otros países y del Océano Pacífico, respectivamente. Para evitar que población ajena o sin información agropecuaria (en el caso de las UME costeras) formen parte del marco de muestreo, todo subconjunto territorial de cada UME que no forme parte de la superficie continental ecuatoriana no será investigado. Tomando en cuenta esta consideración, se asegura que el marco de muestreo no cuente con población ajena al objetivo de la investigación.

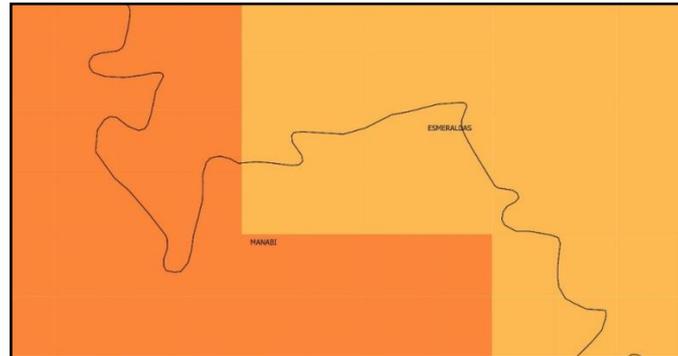
<sup>2</sup> El Sistema Integrado de Encuestas de Hogares (SIEH) es una estrategia de producción de información acerca de las características demográficas y socioeconómicas de los hogares y las personas, basado en el diseño y ejecución coordinada de un conjunto de encuestas que comparten un cierto marco conceptual y metodológico.





Por otro lado, puesto que el requerimiento de información contempla la estimación de información a nivel provincial, los límites provinciales se han ajustado al perfil de las UME. En otras palabras, dentro del marco de muestreo para investigaciones agropecuarias, cada provincia tiene límites poligonales definidos en el porcentaje de cada UME que pertenece a cada provincia (Ver Figura 2).

**Figura 2.** Límites provinciales del Marco de Muestreo para Investigaciones Agropecuarias. Un ejemplo para las provincias de Manabí y Esmeraldas



Fuente: INEC [4]

### 1.1.1. Estratificación del Marco de Áreas

La construcción del marco de áreas inicia con la estratificación del territorio nacional, el cual se divide en dos estratos primarios: (i) las zonas urbanas densamente pobladas y (ii) el resto del territorio, incluyendo áreas agropecuarias, forestales, agua y otras áreas rurales con baja densidad de población. Del estrato (ii) se separan únicamente los cuerpos de agua. La superficie restante se estratifica utilizando los siguientes criterios:

- **Distribución Político Administrativa del Ecuador actualizada a junio 2018 hasta el nivel provincial:** Puesto que el requerimiento de información contempla estimaciones a nivel provincial, cada provincia debe ser tratada como un dominio de estudio independiente. Y puesto que el marco de muestreo es nacional, cada provincia es tratada como un dominio de estudio independiente dentro del marco de muestreo.
- **Intensidad de uso de suelo:** Para la selección de la muestra, se tomó la variable “superficie de suelo con uso agropecuario” como variable de diseño de la muestra, y se estratifica el marco de muestreo en función al porcentaje de superficie con uso agropecuario que tiene cada UME, asignándolos en 3 estratos de la siguiente manera:

**Tabla 1.** Estratos del Marco de Muestreo para Investigaciones Agropecuarias

Estrato	% de la UME con uso agropecuario	Superficie de la UME con uso agropecuario (ha.)
Estrato 1	60% - 100%	345,6 – 576
Estrato 2	20% - 60%	115,2 – 345,6
Estrato 3	0% - 20%	0 – 115,2

Fuente: INEC [4]

Los límites entre estratos son los estándares recomendados en [FAO (1996,1998)] citado en INEC [4]: áreas con una proporción mayor del 60% (Estrato 1); áreas con





una proporción de superficie cultivada entre 20% y 60% (Estrato 2) y áreas con una proporción de superficie cultivada inferior al 20% (Estrato 3).

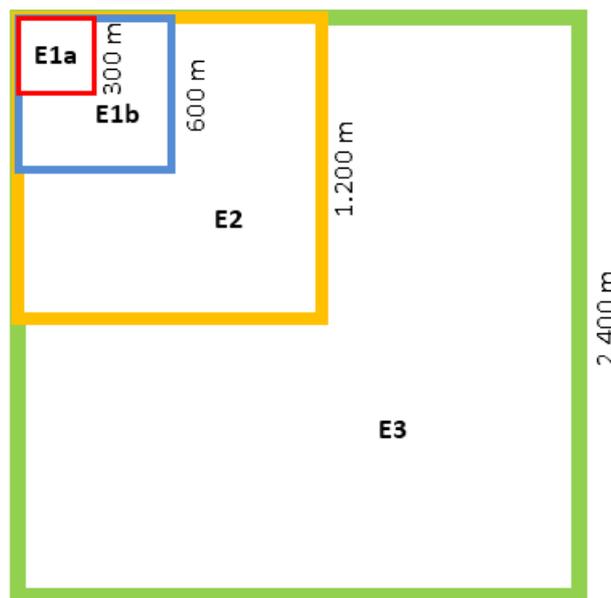
### 1.1.2. Segmentación y definición de tamaño por estrato

Gracias a la experiencia en la recolección de información en campo, cubrir una superficie de 576 hectáreas resulta sumamente dificultoso por varios factores, entre los que se destaca el alto número de entrevistas a realizar; razón por la cual, las UME en su estructura original no fueron consideradas como Unidades Primarias de Muestreo (UPM).

Además, puesto que las características de producción de la región natural Sierra (producción intensiva) son distintas a las características de las regiones naturales Costa y Amazonía (producción extensiva), se decidió diferenciar las UME del Estrato 1 entre estas regiones naturales, dando como resultados los Estratos 1a y 1b, pertenecientes a las regiones naturales Sierra y Costa-Amazónica, respectivamente. Esta diferenciación fue construida considerando las regiones biogeográficas del Ecuador, establecidos por la Dirección de Cartografía Estadística del INEC.

Considerando todo lo expuesto anteriormente, para construir UPM más acordes a las características de producción, se procedió a dividir sucesivamente cada UME en función al estrato al que pertenece. Mientras que las UME del Estrato 3 no fueron divididas, las UME del Estrato 2, Estrato 1b y 1a fueron divididas en 4, 16 y 64 partes iguales, respectivamente (Ver Figura 3).

**Figura 3. Áreas para estratificación**



Fuente: INEC [4]

Con todas estas consideraciones, los tamaños de las UPM, quedan definidos de la siguiente manera:





**Tabla 2. Tamaño de las Unidades Primarias de Muestreo**

Estrato	Tamaño UPM (ha.)
1a	9
1b	36
2	144
3	576

Fuente: INEC [4]

De ahora en adelante, las UPM se denominarán segmentos. Es importante mencionar que para la ESPAC 2020, no se realizaron actualizaciones del Marco de Muestreo de Áreas, por lo tanto, tiene la misma conformación del Marco de Muestreo del año 2019, es decir, consta de 357.599 segmentos, repartidos a nivel provincial y por estrato, como se muestra en la tabla 3, recordando que para este ejercicio se consideró únicamente la superficie continental del Ecuador.

Cabe indicar que la determinación de los límites provinciales es competencia de la Secretaría Técnica del Comité Nacional de Límites Internos, ante esto y de acuerdo a las resoluciones hechas hasta Diciembre 2018 en la División Político Administrativa del territorio, las zonas no delimitadas que se presentaban en años anteriores, se anexaron a las provincias constituidas, donde el territorio conocido como Manga del Cura pasa a ser parte de la provincia de Manabí; Las Golondrinas, pertenece a Imbabura y El Piedrero forma parte de las provincias de Guayas y Cañar.

**Tabla 3. Distribución del número de segmentos del marco de áreas por provincia y por estrato ESPAC – 2020**

PROVINCIA	E1a	E1b	E2	E3	TOTAL
Azuay	29.963	190	1.008	189	31.350
Bolívar	17.038	252	601	51	17.942
Cañar	9.176	691	503	94	10.464
Carchi	10.154	0	405	90	10.649
Cotopaxi	23.425	412	660	123	24.620
Chimborazo	17.749	60	816	155	18.780
El Oro	12.375	3.354	657	83	16.469
Esmeraldas	1.905	16.830	1.713	296	20.744
Guayas	275	15.880	1.620	302	18.077
Imbabura	14.455	98	661	105	15.319
Loja	19.674	0	1.271	270	21.215
Los Ríos	2.144	14.884	447	6	17.481
Manabí	0	26.210	2.886	366	29.462
Morona Santiago	9.790	2.898	2.235	521	15.444
Napo	6.835	127	1.485	291	8.738
Pastaza	4.586	367	2.454	726	8.133
Pichincha	26.430	373	914	196	27.913
Tungurahua	9.522	0	263	52	9.837
Zamora Chinchipe	3.556	669	754	167	5.146
Sucumbíos	1.994	4.874	2.404	468	9.740
Orellana	324	2.311	1.959	422	5.016
Santo Domingo de los Tsáchilas	12.514	1.815	286	0	14.615
Santa Elena	0	181	193	71	445
<b>TOTAL</b>	<b>23.3884</b>	<b>92.476</b>	<b>26.195</b>	<b>5.044</b>	<b>357.599</b>

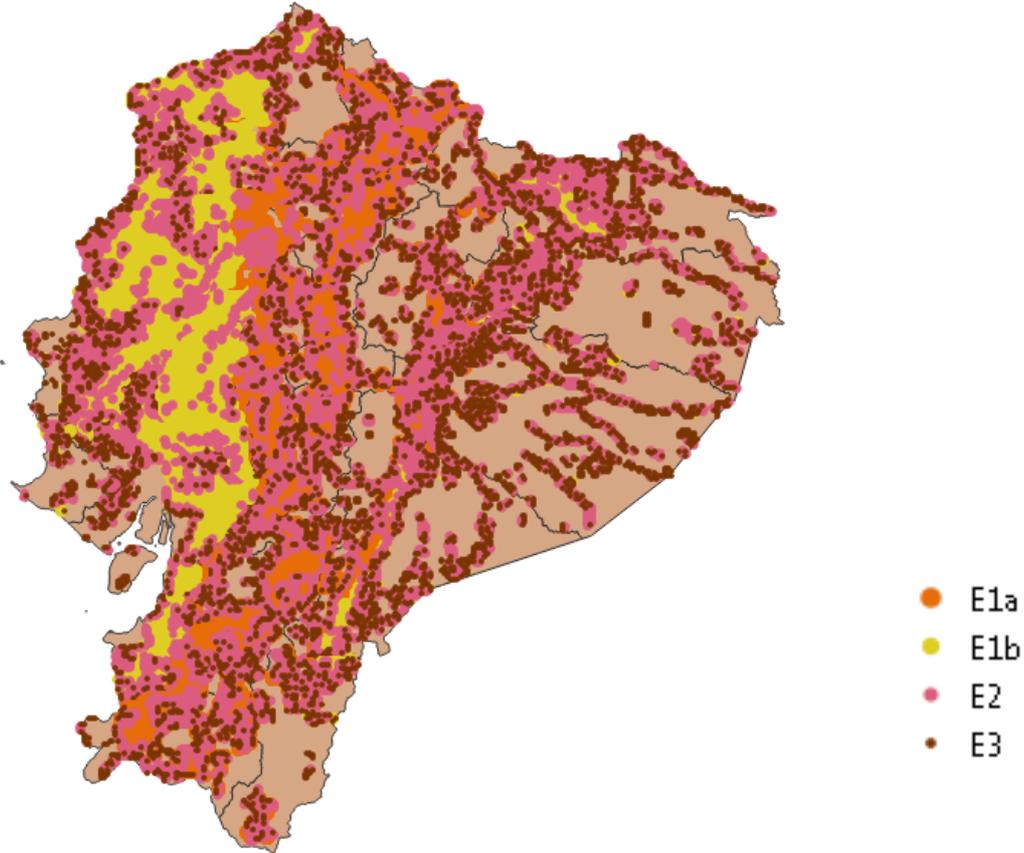
Fuente: Marco de Áreas ESPAC 2020

En la figura que se muestra a continuación, se observa la distribución del número de segmentos muestrales por estrato en el territorio continental ecuatoriano:





Figura 4. Distribución del número de segmentos por estrato



Fuente: Marco de Áreas ESPAC 2020

Por otra parte, en la siguiente tabla se muestra un resumen del número de segmentos con su superficie respectiva del Marco de Áreas de ESPAC 2019:

Tabla 4. Número de segmentos y superficie del Marco de Áreas de ESPAC 2020

Estrato	Total	Superficie (ha.)	Total (ha.)	Superficie (%)
E1a	233.884	9	2.104.956	17,38%
E1b	92.476	36	3.329.136	27,49%
E2	26.195	144	3.772.080	31,14%
E3	5.044	576	2.905.344	23,99%
<b>Total</b>	<b>357.599</b>		<b>12.111.516</b>	<b>100,00%</b>

Fuente: Marco de Áreas ESPAC 2020

## 1.2. Marco de Lista

Para la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC 2020), con el objetivo de mejorar la calidad de las estimaciones de las variables de superficie y producción de algunos cultivos, se elaboró un directorio de investigación, con los principales productores de ciertos productos.

Los productos considerados para el marco de lista, son aquellos que actualmente son de importancia macroeconómica y determinados como sensibles o estratégicos para instituciones como el Banco Central del Ecuador, Ministerio de Agricultura, Secretaría Nacional de Planificación y Desarrollo; Vicepresidencia de la República, resultando necesario la elaboración de un marco de lista para los siguientes productos:





- Arroz
- Banano
- Brócoli
- Café
- Cacao
- Caña de azúcar para azúcar
- Maíz duro seco
- Mango
- Maracuyá
- Palma Africana
- Ganado Vacuno
- Ganado Porcino
- Flores
- Aves

Debido a la complejidad de generación de un marco utilizando una fuente homogénea, para todos estos productos citados, se elaboró el directorio de investigación utilizando varias fuentes de información que cumplió con las siguientes características:

- Directorio actualizado
- Cobertura del directorio superior al 85% de superficie y/o producción a nivel nacional, según las referencias entregadas por MAG
- Directorio validado (superficie, producción, existencia de animales)

Una vez analizada la consistencia de información disponible proveniente de varias instituciones, se elaboró un directorio para cada producto y se estableció un parámetro de aporte dentro de actividad agrícola o pecuaria. La determinación del parámetro de aporte para cada producto dentro de su actividad es de vital importancia y debe ser construido con cautela, debido a que el objetivo es mejorar las estimaciones alcanzando niveles aceptables que reflejen la realidad de las actividades agropecuarias, es decir, no sobreestimar las variables de interés.

En la tabla 5 se detallan las fuentes de información consideradas para el directorio del Marco de Lista:

**Tabla 5. Fuentes de información consideradas para el Marco de Lista ESPAC-2020**

Producto	Fuente del directorio	Elaboración
Arroz	ESPAC 2013	INEC
Banano	Catastro bananero 2012-2013	MAG
Brócoli	ESPAC 2013	INEC
Café	ESPAC 2013	INEC
Cacao	ESPAC 2013	INEC
Caña de azúcar para azúcar	ESPAC 2013	INEC
Maíz duro seco	ESPAC 2013	INEC
Mango	Directorio de productores, empaques y exportadores de mango 2014	Fundación Mango Ecuador
Maracuyá	ESPAC 2013	INEC
Palma Africana	Censo palmicultor 2005	MAG
Ganado Vacuno (cabezas)	Registro de productores (vacunación fiebre aftosa) 2013	MAG – AGROCALIDAD
Ganado Porcino (cabezas)	Censo porcicultor 2010	MAG
Flores	Catastro florícola 2010	MAG
Planteles Avícolas	Censo avícola 2006	MAG – CONAVE

Fuente: INEC [4]

Los directorios se consideraron de forma individual y una vez esquematizada la información en función de su tamaño, se realizó una selección por producto. Para mantener un control y seguimiento de los directorios, a cada producto se le asignó un código de estrato, como se muestra en la siguiente tabla:





*Tabla 6. Código de estrato por producto*

Estrato	Producto
51	Arroz
52	Banano
53	Brócoli
54	Cacao
55	Café
56	Caña de azúcar para azúcar
57	Maíz duro seco
58	Mango
59	Maracuyá
60	Palma Africana
61	Ganado vacuno
62	Ganado porcino
63	Flores
64	Aves

Fuente: INEC [4]

## 2. Diseño Muestral

### 2.1. Universo de la investigación

El universo de investigación para la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC - 2020) está conformado por el conjunto de todos los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano.

### 2.2. Población Objetivo

La población objetivo comprende todos los terrenos con superficie agropecuaria en el territorio continental ecuatoriano.

### 2.3. Unidades de muestreo y análisis

En el caso de la ESPAC, la unidad elemental de muestreo es el segmento, también denominado como Unidad Primaria de Muestreo (UPM). Por otro lado, las unidades de análisis son todos los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano.

### 2.4. Cobertura geográfica

La cobertura geográfica de la ESPAC 2020 está definida por los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano, por lo que se excluye la región Insular.

El nivel de estimación es uniforme para cada provincia del Ecuador, manteniendo de esta manera representatividad provincial, por cuanto cada una de estas es considerada como un dominio de diseño.





## 2.5. Tipo de muestreo

El tipo de muestreo corresponde a un Muestreo Probabilístico Estratificado por conglomerados y por réplicas, con selección aleatoria.

## 2.6. Tamaño de muestra

En cada dominio (provincia) se calcula un número determinado de segmentos en relación a una tasa de muestreo específica por estrato. Con el propósito de mejorar la eficiencia en las estimaciones, estas tasas de muestreo se establecieron en función al número de segmentos del marco, es decir, si el número de UPM del Marco de Áreas es mayor, la tasa de muestreo será menor, y viceversa.

**Tabla 7. Tasas de muestreo por estrato**

Estrato	Tasa de muestreo
E1a	1,50%
E1b	1,50%
E2	1,20%
E3	4%

Fuente: INEC [4]

Cabe señalar que, en casos donde la muestra sea menor a 10, se debe ajustar por los mínimos muestrales<sup>3</sup>, de manera que se puede calcular los errores de muestreo asociados a las estimaciones. En la tabla 8 se muestran los tamaños de muestra para ESPAC 2020 por provincia y por estrato:

**Tabla 8. Tamaños de muestra por provincia y por estrato ESPAC – 2020**

PROVINCIA	E1a	E1b	E2	E3	TOTAL
Azuay	450	10	15	10	485
Bolívar	255	10	10	10	285
Cañar	140	10	10	10	170
Carchi	155	0	10	10	175
Cotopaxi	355	15	15	10	395
Chimborazo	270	10	15	10	305
El Oro	185	50	10	10	255
Esmeraldas	30	255	30	15	330
Guayas	10	240	20	15	285
Imbabura	220	10	10	10	250
Loja	295	0	20	15	330
Los Ríos	35	215	10	6	266
Manabí	0	395	45	20	460
Morona Santiago	150	45	30	20	245
Napo	105	10	20	15	150
Pastaza	70	10	30	30	140
Pichincha	400	10	20	10	440
Tungurahua	145	0	10	10	165
Zamora Chinchipe	55	10	10	10	85
Sucumbíos	30	75	30	20	155
Orellana	10	35	25	20	90
Santo Domingo de los Tsáchilas	190	40	10	0	240
Santa Elena	0	10	10	10	30
<b>TOTAL</b>	<b>3.555</b>	<b>1.465</b>	<b>415</b>	<b>296</b>	<b>5.731</b>

Fuente: Muestra ESPAC 2020

<sup>3</sup> El número mínimo de segmentos a considerar en la muestra es de 10 por estrato y por provincia, excepto donde el número de segmentos del Marco de Áreas sea menor a 10, en cuyo caso en la muestra se incluirán todos aquellos segmentos.



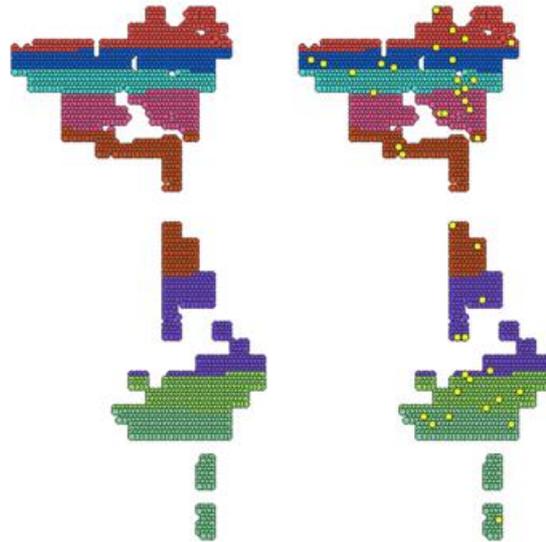


## 2.7. Selección de la muestra

La selección de los segmentos que forman parte de la muestra, de acuerdo al tamaño establecido, se realiza independientemente en cada uno de los estratos de los dominios. Para tal efecto, se construyen Zonas de selección, las cuales son conjuntos de segmentos muestrales construidas en cada estrato de las provincias, estos estratos implícitos son creados con el objetivo de seleccionar de manera uniforme y eficiente la muestra.

Dado que el muestreo de ESPAC es replicado, se escogen 5 segmentos dentro de cada Zona de selección<sup>4</sup> (estrato implícito), los cuales corresponden a las réplicas del año 2020.<sup>5</sup> En la figura que se presenta a continuación, se muestra las Zonas de selección del estrato 1b de la provincia de Santo Domingo de los Tsáchilas, con el escogimiento de muestra respectiva.

*Figura 5. Zonas de selección del estrato 1b de la provincia de Santo Domingo de los Tsáchilas*



Fuente: Marco de Áreas y Muestra ESPAC 2020

Según las asistencias técnicas proporcionadas por la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), y con la finalidad de optimizar el tamaño de muestra y mejorar la eficiencia de las estimaciones, la selección de la muestra de los segmentos del estrato 3 se realiza en dos etapas. En la primera etapa, se seleccionan los segmentos del estrato 3 en función al tamaño de muestra calculada; mientras que, en la segunda etapa, se dividieron estas UPM en cuatro áreas de igual superficie (144 Ha.) y se escoge dos subsegmentos de 144 Ha. cada uno; dentro de cada segmento muestral seleccionado en la primera etapa, como se presenta en la siguiente figura:

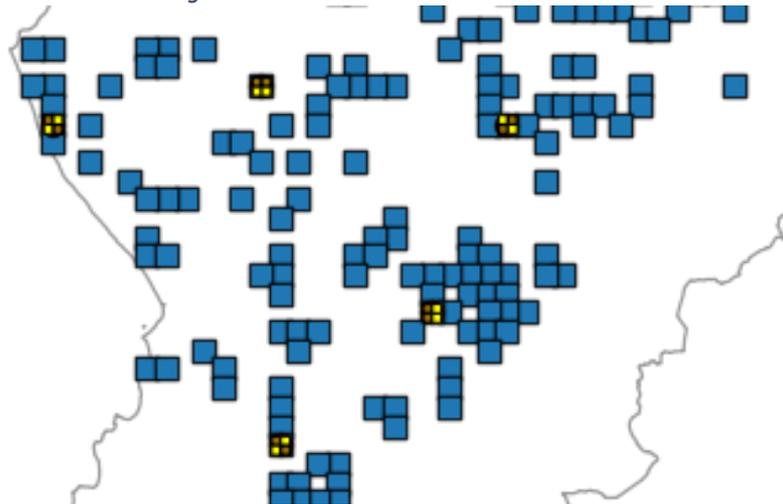
4 Para el cálculo de los errores de muestreo, debe existir mínimo dos Zonas de selección por provincia y por estrato, por tal motivo los mínimos muestrales son 10 segmentos, 5 por cada Zona de selección.

5 En el año 2014, se seleccionaron 5 segmentos correspondientes a las réplicas 1, 2, 3, 4 y 5; en el año 2015, réplicas 2, 3, 4, 5 y 6; en el año 2016, réplicas 3, 4, 5, 6 y 7; en el año 2017, réplicas 4, 5, 6, 7 y 8; en el año 2018, réplicas 5, 6, 7, 8 y 9; y en el año 2019, réplicas 6, 7, 8, 9 y 10. Para el año 2020, las réplicas son 7, 8, 9, 10 y 11.





Figura 6. Selección de la muestra del estrato 3



Fuente: Marco de Áreas y Muestra ESPAC 2020

Es importante señalar que únicamente en la Provincia de Los Ríos, se investigaron todos los segmentos del estrato 3, es decir, 6 segmentos como consta en la tabla 8.

## 2.8. Cobertura muestral

Una vez levantada la información de ESPAC 2020, se identifica que de los 5.731 segmentos que se enviaron en la muestra para ser investigados, 5.703 de ellos fueron efectivos, es decir, se tiene su información completa; mientras que existieron segmentos con rechazo total, rechazo parcial, subdivididos<sup>6</sup> y fuera del territorio, los cuales se muestran en la tabla que se presenta a continuación:

Tabla 9. Cobertura muestral

DESCRIPCIÓN	Segmentos
Rechazo total	38
Rechazo parcial	76
Subdivididos	14
Fuera del territorio	19
<b>Total</b>	<b>147</b>

Fuente: Muestra ESPAC 2020

## 3. Cálculo de factores de expansión

El objetivo de una encuesta por muestreo es realizar inferencias acerca de la población a partir de la información contenida en la muestra. En este sentido, se utilizan factores de expansión para llevar la información de la muestra a niveles poblacionales. Se define como el inverso de la probabilidad de incluir un segmento en una determinada Zona de selección.

La fórmula de cálculo de los factores de expansión teóricos de ESPAC para los segmentos seleccionados del Marco de Áreas es:

$$fexp_{teórico} = \frac{N_h}{n_h}$$

<sup>6</sup> La subdivisión de un segmento se realiza cuando existe un número de productores considerablemente grande, por lo que se divide al segmento en 4 superficies de igual tamaño y se levanta la información de una de ellas.





Donde:

$N_h$  es el número de segmentos en el Marco de Áreas por zona de selección  $h$ ,  
 $n_h$  es el número de segmentos en la muestra por zona de selección  $h$ .

Dado que la selección de la muestra en algunos segmentos del estrato 3 se realizó en dos etapas para algunos de ellos, es necesario realizar un ajuste al factor de expansión para estos segmentos que no se investigaron por completo (se escogió dos subsegmentos de 144 Ha. cada uno).

Para estos segmentos en cuestión, el ajuste toma un valor de 2, es decir, el factor teórico se multiplica por 2, con la finalidad de que las variables o características investigadas se expandan a la información del Marco de Muestreo de Áreas.

Debido a que en ocasiones existen segmentos con rechazo total, rechazo parcial o que fueron subdivididos, es necesario calcular un factor de expansión ajustado por cobertura, como se detalla a continuación:

$$fexp_{ajustado} = fexp_{teórico} * \alpha_1 * \alpha_2 * \alpha_3$$

Donde:

- $\alpha_1$  es el ajuste por rechazo total, resultado de dividir el número de segmentos en la muestra por zona de selección  $h$  sobre el número de segmentos efectivos (completos) en la zona de selección  $h$ ,
- $\alpha_2$  es el ajuste por rechazo parcial, resultado de dividir la superficie total sobre la superficie investigada del segmento,
- $\alpha_3$  es el ajuste por subdivisión, el cual es igual a 4 si el segmento fue subdividido una vez, e igual a 16 si el segmento fue subdividido dos veces.

El factor de expansión final es el resultado de multiplicar el factor de expansión ajustado por cobertura por un delta y por un ponderador, de forma que se expanda la superficie de campo de los terrenos a la superficie del Marco de Áreas, como se indica a continuación:

$$fexp_{final} = fexp_{ajustado} * \delta * \theta$$

donde:

- $\delta$  es la corrección de la superficie de las provincias, resultado de dividir la superficie no cuadrada (determinada por los límites reales de las provincias) sobre la superficie expandida de la encuesta.
- $\theta$  es la división de la superficie de cartografía sobre la superficie declarada de campo.

Por otra parte, es indispensable mencionar que el factor de expansión para los terrenos del Marco de Lista es igual a 1, ya que se levanta la información de todos los productores de este Marco, siendo equivalente a un censo.

## 4. Estimación de características

Según el Manual sobre Marcos Maestros de Muestreo para las Estadísticas Agropecuarias [2], la estimación del Total Poblacional para determinada





característica o variable Y, se obtiene a partir del estimador de marco dual de Hartley, acorde a la siguiente fórmula:

$$\hat{Y}_H = \hat{Y}_a + \rho \hat{Y}_{ab} + (1 - \rho) \hat{Y}_L$$

Donde:

- $\hat{Y}_H$ : es el estimador de Hartley para el total poblacional para una determinada característica Y,
- $\hat{Y}_a$ : es el estimador del total poblacional de la característica Y correspondiente al Marco de Áreas,
- $\hat{Y}_{ab}$ : es el estimador del total poblacional de la característica Y correspondiente a la intersección entre el Marco de áreas y el Marco de Lista,
- $\hat{Y}_L$  es el estimador del total poblacional de la característica Y correspondiente al Marco de Lista,
- $\rho$  es el coeficiente de la combinación lineal entre el Marco de Áreas, Marco de Lista y su intersección.

Según el Manual sobre Marcos Maestros de Muestreo para las Estadísticas Agropecuarias (2015), "se puede utilizar cualquier valor para  $\rho$  tal que  $0 \leq \rho \leq 1$ . Uno de ellos es una opción óptima ( $\hat{\rho}_o$ ), en el sentido de que reduce al mínimo la varianza  $Var(\hat{Y}_H)$ . En la práctica, el valor  $\hat{\rho}_o$  puede ser muy cercano a 0, en estos casos es conveniente y ventajoso adoptar el estimador de examen preliminar (elegir  $\rho = 0$ )."

Por tanto, la expresión del estimador de examen preliminar resultante es:

$$\hat{Y}_S = \hat{Y}_a + \hat{Y}_L$$

$$\hat{Y}_S = \sum_{i \in S_a}^{N_a} fexp_{final_a} * y_{i_a} + \sum_{i \in S_L}^{N_L} fexp_{final_L} * y_{i_L}$$

Donde:

- $fexp_{final_a}$  es el factor de expansión final de los terrenos i del Marco de Áreas,
- $fexp_{final_L}$  es el factor de expansión final de los terrenos i del Marco de Lista,
- $y_{i_a}$  es la característica o variable que se está investigando del terreno i del Marco de Áreas,
- $y_{i_L}$  es la característica o variable que se está investigando del terreno i del Marco de Lista.

La fórmula anterior puede simplificarse como:

$$\hat{Y}_S = \sum_{i \in S}^N fexp_{final} * y_i$$

Donde:

- $fexp_{final}$  es el factor de expansión final tanto para el Marco de Áreas como para el Marco de Lista del terreno i,
- $y_i$  es la característica o variable que se está investigando tanto para el Marco de Áreas como para el Marco de Lista del terreno i.





## 5. El error de muestreo

Según el Manual sobre Marcos Maestros de Muestreo para las Estadísticas Agropecuarias [2], el error de muestreo asociado a la estimación de una característica es calculado a través del estimador de marco dual de examen preliminar de Hartley:

$$\hat{Y}_S = \hat{Y}_a + \hat{Y}_L$$

Dado que los dominios son mutuamente excluyentes y que los estimadores utilizan información de diferentes marcos, la varianza del estimador de examen preliminar es:

$$Var(\hat{Y}_S) = Var(\hat{Y}_a) + Var(\hat{Y}_L)$$

donde:

$Var(\hat{Y}_a)$  es el estimador de la varianza del estimador del total poblacional de la variable  $Y$  del Marco de Áreas,

$Var(\hat{Y}_L)$  es el estimador de la varianza del estimador del total poblacional de la variable  $Y$  del Marco de Lista.

La varianza del estimador del total poblacional del Marco de Lista  $Var(\hat{Y}_L)$  es igual a 0, debido a que se levanta la información de todos los productores, es decir, el tamaño de muestra es igual al de la población, anulando la varianza.

En cuanto a lo mencionado, la varianza del estimador de examen preliminar sería:

$$Var(\hat{Y}_S) = Var(\hat{Y}_a)$$

El estimador insesgado de la varianza del estimador del total  $Var(\hat{Y}_a)$  se calcula a partir de un muestreo probabilístico aleatorio simple de conglomerados en una etapa, de acuerdo a la siguiente fórmula:

$$Var(\hat{Y}_S) = \sum_{h=1}^H M_h^2 (1 - f_h) \frac{\hat{S}_h^2}{m_h}$$

donde:

$M_h$  es el número de segmentos en el Marco de Áreas en la zona de selección  $h$ ,

$m_h$  es el número de segmentos en la muestra en la zona de selección  $h$ ,

$f_h$  es la fracción global de muestreo.

$$f_h = \frac{m_h}{M_h}$$

$\hat{S}_h^2$  es la varianza estimada de la característica  $Y$  en los segmentos de la zona de selección  $h$

$$\hat{S}_h^2 = \frac{1}{m_h - 1} \sum_{i=1}^m (y_i - \bar{y}_h)^2$$





Ante lo mencionado, el coeficiente de variación para una estimación de una determinada característica se define como la división entre la raíz de la varianza del estimador del total poblacional (error estándar) sobre la suma de los totales poblacionales estimados del Marco de Áreas y del Marco de Lista:

$$CV(\hat{Y}_S) = \frac{\sqrt{Var(\hat{Y}_S)}}{\hat{Y}_S}$$

$$CV(\hat{Y}_S) = \frac{\sqrt{Var(\hat{Y}_S)}}{\hat{Y}_a + \hat{Y}_L}$$

## Bibliografía

- [1] Ambrosio, L. (2017). Curso de Muestreo con Marcos Múltiples: Estimación con marcos sencillos. Muestreo Aleatorio Simple sin información auxiliar. Escuela Técnica Superior de Ingenieros Agrónomos. Madrid: España.
- [2] Global Strategy. (2015). Manual sobre Marcos Maestros de Muestreo para las Estadísticas Agropecuarias: elaboración del Marco, Diseño de la Muestra y Estimación.
- [3] Kish, L. (1972). Muestreo de Encuestas. México: Trillas.
- [4] Instituto Nacional de Estadística y Censos [INEC] (2015). Metodología de la Encuesta de Superficie y Producción Agropecuaria Continua ESPAC 2014. Quito.





**INEC**

Buenas cifras,  
**mejores vidas**



@ecuadorencifras



@ecuadorencifras



@InecEcuador



t.me/euadorencifras



INEC/Ecuador



INECEcuador



INEC Ecuador