



INFORME ESTADO DE LA NACIÓN EN DESARROLLO HUMANO SOSTENIBLE

Metodología Primera Encuesta Ambiental

**Patrones y percepciones ciudadanas sobre medio ambiente y condiciones para el
cambio**

Junio, 2018



Cuestionario

El cuestionario fue elaborada por el equipo técnico del Programa Estado de la Nación (PEN), y consultado con la Escuela de Estadística de la Universidad de Costa Rica y el equipo del Proyecto Creación de Capacidades de los Acuerdos Multilaterales Ambientales (AMA), Ministerio de Ambiente y Energía (Minae)/ Programa de Naciones Unidas para el Desarrollo (PNUD), adicionalmente se hizo revisión de la literatura (ver sección de Revisión de literatura: variables y modelo) y de la experiencia nacional e internacional para confirmar los temas y variables que debían incluirse para alcanzar los objetivos.

El cuestionario consta de las siguientes secciones:

- I. Características sociodemográficas y políticas
- II. Huella ambiental de las personas –acciones
- III. Conocimiento ambiental
- IV. Prioridad valorativa sobre el tema ambiental
- V. Responsabilidades sobre la gestión ambiental
- VI. Percepción sobre la situación ambiental y el riesgo inminente
- VII. Disposición y viabilidad para la adopción de prácticas ambientales sostenibles

Las preguntas fueron principalmente de registro cerrado. Las preguntas abiertas fueron codificadas para su procesamiento. La entrevista tuvo una duración promedio de 50 minutos por entrevistado. Se elaboró un manual de trabajo de campo con las instrucciones del procedimiento y del cuestionario.

Unidad de análisis

Se llevó a cabo una encuesta cara-a-cara domiciliar con una muestra de 1.200 residentes de la GAM y fuera de la GAM, de 18 a 65 años de edad. En este estudio interesaba explorar la disposición a optar por comportamentales responsables no solo de los hogares sino de las personas por lo que la unidad de análisis fueron los individuos y no los hogares. A pesar de lo anterior, aleatoriamente el 49% de las entrevistas se realizaron a la jefatura del hogar, cuando se suma a las parejas, el subgrupo representó el 71% de la muestra (n=850). Esto permitió que en el análisis sobre la disposición de realizar acciones en el hogar (sobre las cuales jefaturas y parejas tienen incidencia), se pudiera hacer una submuestra.

Marco muestral y selección de la muestra

Como marco muestral se utilizó la cartografía resultante del Censo de Población y Vivienda de junio de 2011 del Instituto Nacional de Estadística y Censos (INEC). La muestra fue seleccionada por el INEC.

La selección de las unidades primarias de muestreo (UPM) fue aleatoria y estratificada con los siguientes cuatro estratos:

1. GAM (Región Central, que incluye Gran Área Metropolitana y Resto de la Región Central), urbano.

2. GAM, rural
3. Fuera GAM (que incluye Región Chorotega, Región Pacífico Central, Región Brunca, Región Huetar Atlántica, Región Huetar Norte), urbano
4. Fuera GAM, rural.

Si bien la representación en la GAM (urbano+rural, primeros dos estratos) en el Censo 2011 fue del 62% de los hogares, se definió trabajar con 700 observaciones de ese grupo, es decir con el 58% de la muestra de 1.200 entrevistas. El resto de las entrevistas se realizaron fuera de la GAM.

La división territorial administrativa del país sigue el orden de mayor a menor tamaño de la región (6 regiones), zona (urbana-rural), provincia, cantón, distrito. Los distritos se subdividen en unidades primarias de muestreo (UPMs), para las cuales el INEC proporciona el número de viviendas y la cantidad de población.

Para la solicitud de la muestra al INEC, el PEN calculó el número de unidades primarias de muestreo (UPMs) proporcionales a la cantidad de hogares en los cuatro estratos según el Censo 2011, y se estimó que en cada UPM se pudieran completar 15 entrevistas, es decir se estableció una cuota por UPM, independientemente de la cantidad de personas de 18 a 65 años de edad que habitaran en los segmentos. En promedio las UPM contaban con aproximadamente 140 viviendas con 3,2 personas (de cualquier edad) por vivienda.

Para la selección de UPM por estrato se respetó la proporción de hogares por zona (urbano/rural) y dos grandes estratos (GAM/Fuera GAM)ⁱ. Finalmente se solicitaron 80 UPM con la distribución que se observa en el cuadro 1. Cabe mencionar que se hizo una valoración del costo ante distintas cantidades de UPM para completar la muestra de interés, entre mas UPM seleccionadas el costo tendía a incrementarse, esto influyó en la cantidad final de 80 UPM con selección de 15 viviendas en cada una. El INEC proporcionó los mapas de los 80 segmentos seleccionados y la información básica del marco muestral (división territorial, cantidad de viviendas y cantidad de población).

Cuadro 1

Distribución de los hogares y proporción de UPMs, por zona según región

Región/Zona	Proporción hogares. En porcentajes			Proporción de UPMs para lograr 15 entrevistas por UPM. En cantidad de UPMS y cantidad de entrevistas		
	GAM	Resto	Total	GAM	Resto	Total
Urbano	94,7	50,7	73,2	44	17	61
Rural	5,3	49,3	26,8	3	16	19
Total	58,3	41,7	100,0	47	33	80
Entrevistas				700	500	1.200

Fuente: Elaboración propia con base en datos de Censo 2011.

En una segunda etapa, la selección de la vivienda (USM, unidad secundaria de muestreo) la hizo el equipo de UNIMER en campo con los mapas de las UPMs proporcionados por el INEC. En cada UPM debían hacerse 15 entrevistas, una sola persona por hogar: un hombre o una mujer de 18 años a 65 años de edad que residiera en la vivienda

seleccionada en el momento de realizar el trabajo de campo y que fuera costarricenses o residentes con 10 años o más de vivir en Costa Rica. La persona fue seleccionada si se encontraba en el hogar y si cumplía con las características de las cuotas pendientes en el segmento.

En cada mapa se seleccionó un punto de partida para comenzar a hacer un barrido en dirección a las manecillas del reloj por cada una de las cuadras, es decir, se señaló el orden en que debía recorrerse. Como parte del protocolo una vez que se lograba una entrevista efectiva que cumpliera con alguna de las cuotas definidas para la UPM se saltaba vivienda siguiente inmediata y se reiniciaba el barrido. En caso de rechazos se continuaba el barrido de viviendas.

La razón por la cual no se hizo un salto sistemático de viviendas, sino solamente en aquellas en las que se lograban entrevistas efectivas, era para no agotar los segmentos ya que no se contaba con segmentos sustitutos. En caso de agotar el segmento, UNIMER pedía autorización al PEN para ampliarlo manteniendo las características sociodemográficas de la UPM original. Esto ocurrió en 5 de las 80 UPMs y la principal razón, aunque no la única, fue que las zonas tenían un uso de suelo más comercial que residencial por lo cual, luego de recorrerlos no se completaban las cuotas, específicamente esto ocurrió en las UPMs ubicadas en: Pavas, Guadalupe, San Pedro, Curridabat y Belén. En las UPMs en las que se hizo la extensión, esta no superó los 2 kilómetros de la UPM original.

Con la cuota de 15 observaciones por UPM se estableció también la cantidad de entrevistas a realizar por sexo (femenino, masculino), edad (18-24, 25-34, 35-44, 45-54,55-65), nivel socioeconómico (5 niveles según definición de UNIMER), la actividad de la persona (activo o inactivo), y nivel educativo (estudios primarios, secundarios, universitarios), proporcionales a su distribución en zonas de GAM y fuera de la GAM según el Censo 2011, de manera que cada UPM representara la diversidad de la población y las características no se concentraran por UPM (cuadro 2).

Cuadro 2
Distribución de la muestra
(absolutos y porcentajes)

Sexo	Absolutos	Relativos	Censal 2011
			(población de 18 a 65 años)
Hombre	588	49,0	49,0
Mujer	612	51,0	51,0
Edad			
18 a 24 años	254	21,2	21,8
25 a 34 años	316	26,3	26,3
35 a 44 años	253	21,1	21,1
45 o más	377	31,4	31,4
Nivel socioeconómico			
Medio-alto/alto	167	13,9	
Medio	518	43,2	
Medio-bajo/bajo	515	42,9	

Zona			
GAM	700	58,3	
Fuera GAM	500	41,7	
Nivel educativo			
Primaria	464	38,7	37,7
Secundaria	440	36,7	35,4
Superior	296	24,7	24,1
Total	1.200	100,0	100,0

Fuente: Elaboración propia con base en datos de la muestra y del Censo 2011.

Por lo anteriormente explicado, la base de datos no contiene ponderadores y no debe llevarlos porque la muestra no fue probabilística, se establecieron cuotas para representar la estructura poblacional fijada al seleccionar la muestra por lo que resulta autoponderada por estos criterios.

Para el análisis de los grupos se consideró que el tamaño de las muestras no fuera menor a cien observaciones y por eso, por ejemplo, como se observa en el cuadro 2, se hacen cuatro intervalos de edades, o los niveles socioeconómicos se dividieron en tres y no se hicieron mayores desagregaciones.

Tampoco fue posible, ni fue la intención, hacer análisis por cantón, ni se puede concluir con ese nivel de desagregación. Pero, por ejemplo, sí se verificó conocimiento de si la municipalidad (sea cual fuere) de la comunidad de la persona entrevistada recogía la basura separada y se cruzó por la variable de reciclaje en el hogar, con el fin de explorar relaciones con tamaños de muestra aceptables. No se verificó si en efecto la municipalidad de la comunidad recogía la basura separada, el objetivo era preguntar si la persona tenía conocimiento sobre el servicio en su comunidad y la relación que este tenía con el comportamiento y la intención, por lo tanto, no se hicieron imputaciones. Toda la información proviene de las respuestas de las personas consultadas.

Trabajo de campo

Aun cuando UNIMER cuenta con personal experimentado, antes de iniciarse el trabajo de campo, se realizó una capacitación para el equipo supervisor y para las personas que iban a realizar las encuestas. Durante el mismo se discutió el funcionamiento del instrumento y se revisaron las funciones de los supervisores, encuestadores y de los directores del proyecto. Posteriormente, se procedió a realizar una prueba piloto. Los objetivos de la misma fueron evaluar el cuestionario y la habilidad de los encuestadores para administrarlo. Como resultado de la prueba, se hicieron las correcciones necesarias al instrumento.

El trabajo de campo se llevó a cabo con encuestadores organizados en equipos de 4 personas cada uno. Cada grupo contó con su respectivo supervisor de campo, quien fungió como jefe de grupo y responsable de garantizar la calidad de la información recogida. La dirección y coordinación de todo este personal estuvo a cargo de los directores del proyecto, del director de trabajo de campo y monitoreado por el equipo técnico del PEN. Durante la realización del trabajo de campo, una importante proporción

de los cuestionarios fueron sometidos a verificación personal por parte de los supervisores. Las entrevistas eran grabadas en los dispositivos, por lo que una vez en el sistema, parte de las entrevistas fueron sometidas a una supervisión de audio, principalmente para verificar los datos de identificación y en general la realización de la entrevista.

El horario del trabajo de campo fue de 10:30 a.m. a 6:30 p.m. para lograr la cuota definida de personas activas/inactivas y que las activas no tuvieran únicamente trabajos que se desempeñaran desde el hogar.

El avance del trabajo de campo por parte del equipo técnico del PEN según las cuotas por UPM era fácilmente controlable porque la encuesta se hizo con dispositivos móviles que llevaban la cuenta de las entrevistas realizadas y las pendientes por realizar.

Por razones de seguridad del personal del equipo de campo, en algunos segmentos en los que no era recomendable trabajar con dispositivos móviles, 67 entrevistas de las 1.200 se hicieron en papel y posteriormente se digitaron en la base de datos, esto ocurrió en UPMs de Alajuelita (San Felipe), Los Guidos y Pavas.

Antes de la entrevista, se entregaba una carta informativa y de consentimiento a la persona entrevistada con nombre y teléfono de contacto en el PEN.

Procesamiento de datos

El procesamiento de la información se realizó con los paquetes estadísticos SPSS y con STATA. Se elaboró un manual de verificación de inconsistencias para la limpieza de la base de datos y un plan de cruces bivariados y de creación de variables. Se generaron todos los cuadros descriptivos con cruzados con las variables de interés para cada pregunta del cuestionario.

Para el análisis se elaboraron tablas de frecuencias de doble entrada, con estimaciones de frecuencia y promedios, y se hicieron pruebas de *Chi Cuadrado* para conocer la relación entre variables que se definieron como de interés del estudio. En las pruebas de Chi-Cuadrado si el valor de p es pequeño ($p < 0,05$) se debe rechazar la hipótesis nula de independencia, asumiendo que ambas variables están asociadas estadísticamente (al 95% de confianza o nivel de significancia del 0.005). El valor del estadístico se podrá aproximar por una distribución Chi-cuadrado cuando el tamaño muestral n sea grande ($n > 30$), y todas las frecuencias esperadas sean iguales o mayores a 5 (en ocasiones se agruparon varias categorías a fin de que se cumpla este requisito).

Adicionalmente se realizaron *pruebas t* entre columnas de dos colas 95% de confianza. Los resultados se basan en pruebas bilaterales que asumen varianzas iguales, aplicado solapamiento en grupos independientes (*overlapping*).

Para la definición de los segmentos por disposición de realizar acciones responsables con el ambiente, se llevó a cabo un análisis factorial y de conglomerados para agrupar a

los entrevistados con base en sus actitudes y disposición de acción. Se trabajó con 10 frases actitudinales y 14 frases de disposición de acción.

El análisis factorial es un método multivariado que tiene como objetivo descubrir las dimensiones subyacentes en un conjunto de variables. Este análisis posibilita el agrupamiento de atributos –constituidos por las frases propuestas en cada pregunta– que tienen aspectos comunes entre sí en un número más pequeño de factores o dimensiones. Ello facilita la identificación de grupos y con ello, la clasificación de casos.

No existe una teoría a priori acerca de cómo agrupar los atributos. Para construir los factores sintéticos o dimensiones se utilizan las cargas factoriales que resultan del análisis estadístico, las cuales permiten intuir la estructura factorial de los datos.

Hay varios métodos para identificar o extraer los factores. En este caso se utilizó el Análisis de Componentes Principales y la rotación usada fue la de Normalización Varimax con Kaiser.

Como parte del procedimiento del SPSS, se chequea la adecuación muestral usando la medida de Kaiser-Meyer-Olkin (KMO) que predice si los datos van a factorizar bien, es decir, si los factores van a quedar claramente delimitados. El valor KMO varía entre 0 y 1 y tiene que ser superior a 0.6 para proceder con el análisis. En el caso del análisis factorial aplicado en este estudio se obtuvieron valores KMO superiores a 0.75 en las baterías de preguntas analizadas.

Siguiendo el criterio de Kaiser, se seleccionan los factores más importantes, que corresponden a aquellos componentes con valores característicos –conocidos como “eigenvalues”– de 1.0 o superiores. Con base en este criterio, se identificaron los factores.

Se utilizaron los puntajes factoriales obtenidos en la fase anterior y se procedió a correr el análisis de conglomerados que se basa en un algoritmo de clasificación que busca grupos relativamente homogéneos de objetos.

Una vez obtenidos los resultados de la clasificación del análisis de conglomerados se interpretaron los resultados y se identificaron los segmentos o conglomerados resultantes. Para cada uno se escogió un nombre basado en las principales características que se observaron.

Para el análisis de variables de interés según el modelo planteado (diagrama 1), se elaboraron los siguientes índices (aditivos):

- a. *Índice de acción hogar*: se elabora con base en las frases de la pregunta P9 del cuestionario se hace un conteo de respuestas "sí" para definir intervalos en tres categorías. Poco: 0 a 5, Bastante: 6 a 8, Mucho: 9 a 11 (máximo en 11 frases). Para definir los intervalos, se calculó la frecuencia de la sumatoria y se analizaron los puntos de inflexión, es decir, aquellos en los que la frecuencia mostraba un cambio en la proporción de respuestas siempre y cuando se

conformaran intervalos con suficiente cantidad de observaciones para poder hacer análisis.

- b. *Índice de acción personal*: se elabora con base en las frases de la pregunta P10 del cuestionario se hace un conteo de respuestas "Siempre" para definir intervalos en tres categorías. Poco: 0 a 3, Bastante: 4 a 6, Mucho: 7 a 10 (máximo en 10 frases). Al igual que para el índice de acción hogar los intervalos surgen de la revisión de las frecuencias simples de las sumatorias de los puntajes.
- c. *Índice de valoración*: se elabora con base en las frases de la pregunta P30 del cuestionario se hace un conteo de respuestas 1 y 2 "Menos importante que proteger el ambiente" e "igual de importante" para definir intervalos en tres categorías. Poco importante: 0 a 6, Importante: 7 a 10, Muy importante: 11 (máximo de 11 frases). Para definir los intervalos, se observó aquellos donde se presentará una inflexión en la proporción de respuestas y que los grupos se conformaran por suficiente cantidad de observaciones para poder hacer análisis. Para definir este índice primero se evaluó la alternativa de solamente sumando a quienes decían que era menos importante, pero esta forma no discriminaba los grupos de manera adecuada ya que se creaba un grupo que concentraba el 57% de las observaciones.
- d. *Índice de pesimismo*: se construye con las preguntas P45, P46, P47 del cuestionario se hace un conteo de respuestas "peor" y "muy de acuerdo/de acuerdo" para posteriormente definir intervalos en tres categorías. Nada: 0, Algo: 1, Pesimista: 2 a 7 (7 máximo de frases). Para definir este índice primero se evaluó la alternativa de cruzar y revisar traslapes entre P45 y P46, sin embargo, uno de los grupos concentraba el 64% de las observaciones por lo tanto no era posible discriminar diferencias.
- e. *Índice ideológico*: se basa en las preguntas P6 del cuestionario. Se denomina Social si 5 o más frases con código 1, Balance si 5 o más frases con código 3, Mercado si 3 o más frases con código 2, Sin posición si 4 o más frases con código 9 (un solo caso, se revisó y se decidió sumarlo al grupo "Social"). Para la definición de cantidad de frases por grupo también se probó inicialmente hacer intervalos iguales, sin embargo, se decidió la su forma actual porque las áreas de política muestran una importancia particular para los costarricenses. Por ejemplo, Salud, Educación, Seguridad (tres frases) difícilmente serían relegadas, a diferencia por otras (las menos) como Deporte.

ⁱ En el informe, las regiones GAM y fuera GAM se denominan zonas.