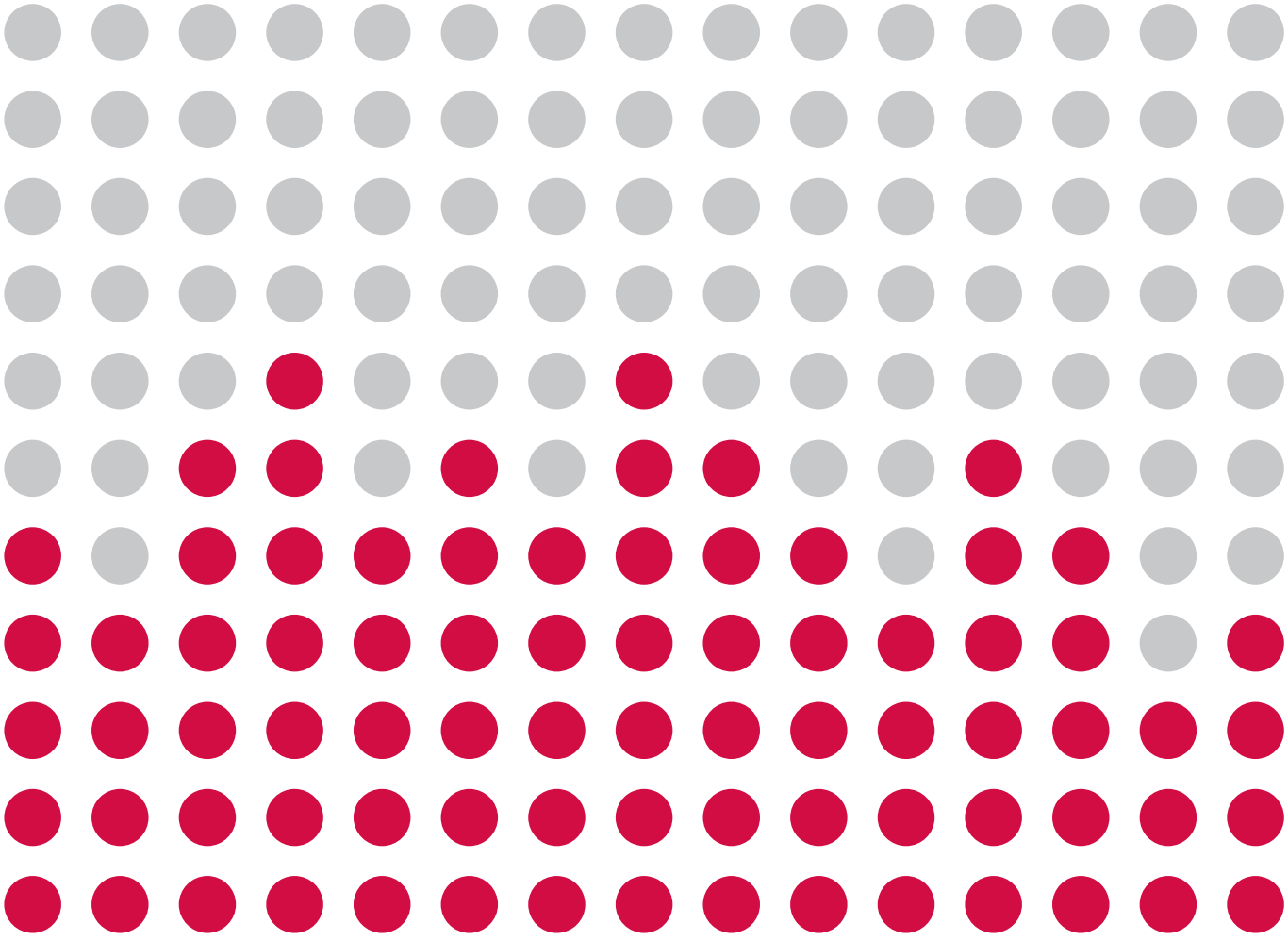
ESTADÍSTICAS LABORALES

ENCUESTA DE POBLACIÓN ACTIVA

METODOLOGÍA PARA LA ESTIMACIÓN EN PEQUEÑAS ÁREAS DE CANARIAS

istac | INSTITUTO CANARIO DE ESTADÍSTICA



ESTADÍSTICAS LABORALES

ENCUESTA DE POBLACIÓN ACTIVA

METODOLOGÍA PARA LA ESTIMACIÓN EN PEQUEÑAS ÁREAS DE CANARIAS

istac | INSTITUTO CANARIO DE ESTADÍSTICA

ESTADÍSTICAS LABORALES

Encuesta de Población Activa
Metodología para la estimación en pequeñas áreas de Canarias.

Diseño, elaboración y edición:
Instituto Canario de Estadística

Luis Doreste Silva 101 – Planta 7.
35004 Las Palmas de Gran Canaria.
Tlf.: 928 290 062 – Fax: 928 243 354.

Rambla de Santa Cruz, 149, Edif. Mónaco. Planta baja.
38001 Santa Cruz de Tenerife.
Tlf.: 922 922 801 – Fax: 922 475 004.

Servidor Web: http://www.gobiernodecanarias.org/istac
E-mail: istac@gobiernodecanarias.org

Depósito Legal: G.C. 1126-2008
Licencia:
Este documento se distribuye bajo Licencia de Reconocimiento 3.0 de Creative Commons.



Texto legal:
http://creativecommons.org/licenses/by/3.0/es/legalcode.es

Presentación

La demanda de estadísticas oficiales con un gran detalle en la desagregación, tanto en el campo de la estadística económica como en el de la estadística social y laboral, no deja de crecer. En consecuencia, la necesidad de disponer sistemáticamente de datos publicados para dominios pequeños, se ha venido consolidando en los últimos años entre los objetivos de los sistemas de estadística oficiales, al mismo tiempo que se han venido desarrollando diversas líneas de investigación sobre la utilización de estimadores en áreas pequeñas como forma de superar la limitación de la escasez de muestra en las operaciones estadísticas por muestreo y cuya información tiene interés en pequeños dominios.

En esa línea, el Instituto Canario de Estadística (ISTAC) conjuntamente con el Departamento de Estadística, Investigación Operativa y Computación de la Universidad de La Laguna, ha venido ejecutando desde el curso 2004/2005 el proyecto CANAREA para la determinación de estrategias de estimación a nivel insular y comarcal de las principales variables de la Encuesta de Población Activa (EPA) del Instituto Nacional de Estadística (INE). Asimismo, las actividades de dicho proyecto se enmarcan dentro del conjunto de acciones planteadas en el seno de Grupo de Trabajo de Áreas Pequeñas del INE.

En el presente documento se presentan sintéticamente las estrategias de estimación investigadas así como la metodología establecida para la fijación de las estimaciones insulares y comarcales de las principales variables de la EPA.

Esther del Toro Cáceres
DIRECTORA DEL INSTITUTO CANARIO DE ESTADÍSTICA

Sumario

1. Introducción /9
1.1. Requerimientos de información sobre Pequeñas Áreas de Canarias **/9**
1.2. Pequeñas Áreas de Canarias: islas, comarcas y municipios **/9**
1.2.1. Provincia de Las Palmas (tres islas y 12 comarcas) **/9**
1.2.2. Provincia de S/C de Tenerife (cuatro islas y 15 comarcas) **/10**

2. La Encuesta de Población Activa (EPA) /12
2.1. Descripción de la Encuesta de Población Activa **/12**
2.2. Estimadores utilizados en la Encuesta de Población Activa **/13**

3. Estimadores de pequeñas áreas evaluados /13
3.1. Estimadores basados en el diseño **/13**
3.2. Estimadores basados en el modelo **/15**

4. El estimador sintético-región /16
4.1. Regiones estáticas **/16**
4.2. Regiones dinámicas **/17**
4.2.1. Precisión constante **/17**
4.2.2. Precisión variable **/17**

5. Calibrado a posteriori de las Estimación en pequeñas áreas y estimación de la población inactiva /18

6. Estimación en pequeñas áreas de la población actual: un input necesario /18

7. Métodos de estimación del Error Cuadrático Medio /20
7.1. Método Jack-knife **/20**
7.2. Método Bootstrap **/20**

8. Valoración de estimadores sobre simulaciones de muestras EPA a partir del Censo 2001 /21
8.1. Indicadores de rendimiento **/21**
8.2. Suplementación de la muestra EPA **/22**
8.3. Resultados sobre los estimadores evaluados **/22**
8.4. Evaluación del comportamiento de los métodos de estimación de los MSE **/22**
8.5. Conclusiones **/23**

9. Aplicación a muestras reales de EPA /24
9.1. Resultados sobre las muestra reales de EPA de 2005 a 2007 **/25**

Bibliografía /33

Índice de tablas /33

1. Introducción

1.1. Requerimientos de información sobre Pequeñas Áreas de Canarias

Tal como señala el Consejo Económico y Social de Canarias[1], la fragmentación del territorio y la insularidad, hacen que resulte difícil hablar de un Mercado Laboral único en Canarias, **de manera que el marco geográfico natural en el que interactúan la oferta y la demanda de trabajo es la isla**. Además, el análisis territorializado del mercado laboral muestra la existencia, dentro de cada territorio insular, de Mercados Locales de Trabajo con características propias.

Por otra parte, el pasado 1 de febrero de 2007 se aprobó el Reglamento (CE) Nº 105/2007 de la Comisión, publicado en el Diario Oficial de la Unión Europea del 10 de febrero de 2007, que modifica los anexos del Reglamento (CE) Nº 1059/2003 del Parlamento Europeo y del Consejo, publicado en el Diario Oficial de la Unión Europea del 21 de junio de 2003, por el que se establece una nomenclatura común de unidades territoriales (NUTS).

En lo que respecta a España, el citado Reglamento Nº 105/2007 modifica el nivel NUT 3 de la Nomenclatura. En particular y en el caso de Canarias, son las islas las que pasan a considerarse el nivel NUT 3 de la clasificación en lugar de la provincia. Este cambio se introdujo atendiendo expresamente la petición que el Instituto Nacional de Estadística (INE) formuló a la Comisión. El INE, a su vez, recogía las demandas que en ese sentido y reiteradamente venía realizando el Instituto Canario de Estadística (ISTAC).

Entre las obligaciones estadísticas que comporta la mencionada modificación están las relativas a la transmisión de series históricas de datos para las islas, entre las que se encuentran las series históricas desde el año 2000 sobre población y su relación con la actividad (ocupados, parados e inactivos) por sexo y grupos de edad (<25 y >=25 años) a suministrar a partir de la operación Labour Force Survey (Encuesta de Población Activa); debiendo ofrecer al EUROSTAT un dato provisional referido al año 2007 en mayo de 2008 y la serie completa con datos definitivos en enero de 2009.

Por todo ello, para la Comunidad Autónoma de Canarias es de interés poder suministrar información trimestral sobre el comportamiento de la actividad, la ocupación y el paro tanto a nivel de islas como de comarcas, información debidamente diferenciada según sexo.

1.2. Pequeñas Áreas de Canarias: islas, comarcas y municipios

En la actualidad el territorio insular se divide en 27 comarcas estadísticas, 12 en la provincia de Las Palmas y 15 en la de S/C de Tenerife. Es sabido que cuanto más pequeñas sean estas áreas, y probablemente con menos información directa por parte de la muestra seleccionada, más complejo y menos precisas serán las estimaciones que se puedan suministrar, salvo que sea posible utilizar información externa con alto nivel de fiabilidad. La conformación en pequeñas áreas en la Comunidad Autónoma de Canarias es la que a continuación se enumera:

1.2.1. Provincia de Las Palmas (tres islas y 12 comarcas)

En la provincia oriental canaria se han establecido 12 comarcas estadísticas. El ratio obtenido es de 2,8 municipios por comarca, lo que supone la conformación de agrupaciones de pequeña o mediana dimensión.

Lanzarote: La isla de Lanzarote está formada por 3 comarcas de diferente dimensión territorial pero de gran coherencia desde la perspectiva geográfica y funcional. Estas comarcas son: Este (LZ-E), Norte (LZ-N) y Suroeste (LZ-SO).

1 Dictamen sobre "Movilidad geográfica del empleo en Canarias". Consejo Económico y Social de Canarias.

Lanzarote	
Comarcas: 3 Municipios: 7 Ratio: 2,3	
LZ-Este Centro-Oriental	Arrecife San Bartolomé Tías
LZ-Norte Norte	Haría Teguise
LZ-Suroeste Sur-Occidental	Tinajo Yaiza

Fuerteventura: La isla de Fuerteventura está formada por 3 comarcas que se reparten el amplio espacio insular atendiendo sobre todo a criterios geográficos, aunque también se ha considerado la actual dinámica funcional del territorio y sus aprovechamientos tradicionales. Estas comarcas son: Centro (FV-C), Norte (FV-N) y Sur (FV-S).

Fuerteventura	
Comarcas: 3 Municipios: 6 Ratio: 2	
FV-Centro Centro	Antigua Betancuria
FV-Norte Norte	La Oliva Puerto del Rosario
NFV-Sur Sur	Pájara Tuineje

Gran Canaria: La isla de Gran Canaria, en razón de su complejidad geográfica, los procesos recientes de ocupación del territorio, así como su centralidad funcional y administrativa, está constituida por 6 espacios comarcales bien definidos, integrados por un número variable de municipios (de 2 a 5). Estas comarcas son: Área Metropolitana (GC-Me), Centro Norte (GC-CN), Noroeste (GC-NO), Oeste (GC-O), Sur (GC-S) y Sureste (GC-SE).

Gran Canaria	
Comarcas: 6 Municipios: 21 Ratio: 3,5	
GC-Área Metropolitana Área Metropolitana	Arucas Las Palmas de Gran Canaria Santa Brígida Telde
GC-Centro Norte Medianías de Barlovento	Firgas Teror Valleseco Vega de San Mateo Valsequillo
GC-Noroeste Norte-Noroeste	Agaete Gáldar Moya Santa María de Guía
GC-Oeste Centro-Occidental	Artenara San Nicolás de Tolentino Tejeda
GC-Sur Sur	Mogán San Bartolomé de Tirajana
GC-Sureste Sur-Oriental	Agüimes Ingenio Santa Lucía de Tirajana

1.2.2. Provincia de S/C de Tenerife (cuatro islas y 15 comarcas)

En la provincia occidental canaria se han establecido 15 comarcas. El ratio obtenido es de 3,31 municipios por comarca, lo que supone la conformación de agrupaciones de pequeña o mediana dimensión.

Tenerife: La isla de Tenerife está constituida por 8 espacios comarcales, integrados por un número variable de municipios (de 3 a 5). Responde a una división bastante generalizada y aceptada de la isla, en la que destacan tres grandes áreas: metropolitana, barlovento y sotavento. A partir de este primer planteamiento, surgen agrupaciones menores bastante operativas y coherentes, tanto desde la perspectiva territorial como desde la óptica administrativa. Estas comarcas son: Abona (TF-Ab), Acentejo (TF-Ac), Área Metropolitana (TF-Me), Daute (TF-D), Icod (TF-I), Suroeste (TF-SO), Valle de Güimar (TF-VG) y Valle de La Orotava (TF-VO).

Tenerife	
Comarcas: 8 Municipios: 31 Ratio: 3,9	
TF- Abona Abona	Arico Fasnia Granadilla de Abona San Miguel Vilaflor
TF-Acentejo Acentejo	El Sauzal La Matanza de Acentejo La Victoria de Acentejo Santa Úrsula Tacoronte
TF-Área Metropolitana Área Metropolitana	El Rosario La Laguna S/C de Tenerife Tegueste
TF-Daute Daute	Buenavista del Norte El Tanque Garachico Los Silos
TF-Icod Icod	La Guancha Icod de los Vinos San Juan de la Rambla
TF- Suroeste Isora	Adeje Arona Guía de Isora Santiago del Teide
TF-Valle de Güímar Valle de Güímar	Arafo Candelaria Güímar
TF-Valle de La Orotava Valle de La Orotava	La Orotava Los Realejos Puerto de la Cruz

La Gomera: La isla de La Gomera está formada por 2 comarcas que responden a la tradicional división barlovento-sotavento de este ámbito geográfico. Quizás este caso sea el más conflictivo por la propia configuración territorial del municipio de Vallehermoso, que recorre la isla de Norte a Sur. Estas comarcas son: Norte (LG-N) y Sur (LG-S).

La Gomera	
Comarcas: 2 Municipios: 6 Ratio: 3,0	
LG-Norte Norte	Agulo Hermigua Vallehermoso
LG-Sur Sur	Valle Gran Rey San Sebastián de La Gomera Alajeró

La Palma: La isla de La Palma está formada por 4 comarcas, que responde, sobre todo, a una división funcional del territorio insular. La propuesta parte de la distribución convencional de la isla: norte, comarca capitalina y Valle de Aridane; a partir de aquí, se han fragmentado el primero de los ámbitos geográficos atendiendo a criterios básicamente operativos. Estas comarcas son: Capitalina (LP-Ca), Noreste (LP-NE), Noroeste (LP-NO), Valle de Aridane (LP-VA)

La Palma	
Comarcas: 4 Municipios: 14 Ratio: 3,5	
LP-Capitalina Capitalina	Breña Alta Breña Baja S/C de La Palma Villa de Mazo
LP-Noreste Noreste	Barlovento Puntallana San Andrés y Sauces
LP-Noroeste Noroeste	Garafía Puntagorda Tijarafe
LP-Valle de Aridane Valle de Aridane	El Paso Fuencaliente Los Llanos de Aridane Tazacorte

El Hierro: La isla de El Hierro forma una unidad dada su dimensión territorial y el número de municipios que la componen. A partir de 1 de Enero de 2008 se incorpora un nuevo municipio a la isla, El Pinar de El Hierro, segregado de Frontera.

El Hierro	
Comarcas: 1 Municipios: 3 Ratio: 3,0	
EH-El Hierro El Hierro	Frontera El Pinar de El Hierro Valverde

2. La Encuesta de Población Activa (EPA)

2.1. Descripción de la Encuesta de Población Activa

Las características técnicas de la Encuesta de Población Activa se describen en el documento metodológico "Diseño de la Encuesta y Evaluación de la calidad de los datos: Informe Técnico", de septiembre de 2005 y en el "Resumen metodológico"[2], documentos publicados en su momento por el área de Diseño de Muestras y Evaluación de Resultados del Instituto Nacional de Estadística (INE). Por lo tanto a continuación tan solo se realiza una descripción somera referente a su diseño muestral, estimadores y ámbito de publicación de datos, para la mejor compresión del resto del presente documento.

La EPA es una encuesta continua, con periodicidad trimestral, cuyo objetivo principal es conocer la actividad económica del país en lo relativo al componente humano. Su diseño está orientado a proporcionar información de las categorías poblacionales con respecto al mercado de trabajo, así como obtener clasificaciones de estas categorías según distintas variables. La encuesta está diseñada para dar resultados detallados a nivel nacional, considerando como población económicamente activa la constituida por las personas de 16 y más años. Para las comunidades autónomas y las provincias se ofrece información sobre las principales características al nivel de desagregación que permiten los coeficientes de variación de los estimadores.

El diseño muestral que utiliza la encuesta es bietápico con estratificación de las unidades de primera etapa, las secciones censales. Las unidades de segunda etapa están constituidas por las viviendas familiares y los alojamientos fijos, recogiéndose información de todas las personas que tengan su residencia habitual en las mismas. Las secciones censales son agrupadas en estratos dentro de cada provincia, 7 en la de Las Palmas y 8 en la de S/C de Tenerife, de acuerdo con la importancia demográfica del municipio al que pertenecen. El número de secciones actualmente seleccionadas en Las Palmas y S/C de Tenerife es de 104 y 91, respectivamente.

2 El documento puede encontrarse en http://www.ine.es/daco/daco43/resumetepa.pdf

Dentro de cada sección seleccionada, son elegidas 18 viviendas por muestreo aleatorio simple, renovándolas parcialmente cada trimestre de encuesta en la sexta parte de las secciones.

2.2. Estimadores utilizados en la Encuesta de Población Activa

Hasta el año 2001 se utilizaron **estimadores de razón** tomando como variable auxiliar las Proyecciones Demográficas de población elaboradas por el INE, siendo la expresión del estimador de una determinada característica Y en un trimestre de encuesta la siguiente

$$\hat{Y} = \sum_{h} \frac{P_h}{p_h} \sum_{i=1}^{n_h} y_{hi}$$

extendiéndose el sumatorio h a los estratos de una provincia, una Comunidad Autónoma o al total nacional, y donde:

$P_h =$ Proyección de la población, que reside en viviendas familiares, en el estrato h, referida a la mitad del trimestre.

$p_h =$ Número de personas que habitan en las viviendas de la muestra, en el estrato h, en el momento de la entrevista.

$n_h =$ Número de viviendas en las secciones de la muestra en el estrato h.

$y_{hi} =$ Valor de la característica investigada de un individuo de la vivienda i-ésima, del estrato h.

A partir del primer trimestre de 2002, se aplican **técnicas de reponderación** a los estimadores con objeto de ajustar las estimaciones de la encuesta a la información procedente de fuentes externas. Por lo tanto, en la actualidad los estimadores que se emplean son los estimadores directos de razón con pesos obtenidos como cociente entre la proyección de la población, elaboradas por el INE, y el número de personas en la muestra en los distintos estratos. Estos pesos son modificados con técnicas de reponderación, utilizando la macro CalMar programada en SAS por el Institut Nacional de la Statistique et des Études Économiques de Francia, para calibrar a las siguientes variables auxiliares: a) población de 16 y más años por grupo de edad y sexo, 22 grupos, a nivel de comunidad autónoma; b) población de 16 y más años por comunidad autónoma y nacionalidad española o extranjera; c) población de 16 y más años por provincia; d) población menor de 16 años por grupo de edad y sexo, 6 grupos, a nivel de comunidad autónoma y; e) población menor de 16 años por provincia. Estos pesos serán denotados para una persona observada j como w_j.

3. Estimadores de pequeñas áreas evaluados

Los distintos estimadores evaluados en las diferentes fases del proyecto CANAREA ejecutadas desde el año 2005 se engloban dentro de dos grupos estándares de estimadores: **estimadores basados en el diseño y basados en modelos**[3]. A continuación comentamos las características de los estimadores aplicados.

3.1. Estimadores basados en el diseño

Sea w_j el peso de la persona j en la muestra EPA. Dado un dominio d, área pequeña*sexo, y un grupo de edad g (<25; 25-54; >54 años), se proponen los siguientes estimadores de la variable objetivo y (variable dicotómica, p.e., 1 sí parado, 0 no parado):

3 El proyecto europeo "Enhancing Small Area Estimation Techniques to meet European Needs" -Eurarea- en el que participó el Instituto Nacional de Estadística, desarrolla en detalle los estimadores basados en modelos, ver la web del proyecto en: http://www.ine.es/docutrab/eurarea/eurarea_05.pdf.

a. **Directo:**

$$\hat{y}_d^{direct} = \frac{\sum_{j \in EPA_d} w_j y_j}{\sum_{j \in EPA_d} w_j} N_d$$ y de su media $\hat{\bar{y}}_d^{direct} = \frac{\hat{y}_d^{direct}}{N_d}$ con N_d el tamaño poblacional del dominio d.

b. **Posestratificado:** la estimación de un dominio d se obtiene como suma ponderada de los estimadores directos de la media en los grupos de edad:

$$\hat{y}_d^{post} = \sum_g \hat{\bar{y}}_{dg}^{direct} N_{dg}$$

c. **Sintético Básico:** la estimación de un dominio d se obtiene como suma ponderada de los estimadores directos de la media en los grupos de edad pero de la provincia:

$$\hat{y}_d^{sint} = \sum_g \hat{\bar{y}}_g^{direct} N_{dg}$$

d. **Sintético-Región:** la estimación de una dominio d se obtiene como suma ponderada de los estimadores directos de la media en los grupos de edad en una región r, construida principalmente como unión de dominios homogéneos:

$$\hat{y}_d^{sint_r} = \sum_g \hat{\bar{y}}_{r,g}^{direct} N_{dg} = \sum_g \frac{\sum_{j \in EPA_{r,g}} w_j y_j}{\sum_{j \in EPA_{r,g}} w_j} N_{dg}$$

En particular, cuando la región r coincida con una de las provincias, dicho estimador será el estimador sintético básico y si cada región coincide con su propia comarca, el estimador será el posestratificado.

En el caso de que las regiones se hayan construido con información de dominios de nivel inferior -por ejemplo, para la estimación de islas por sexo se utilice regiones basadas en las comarcas por sexo- se empleará la expresión anterior sustituyendo g por $g \times c$ indicando no sólo grupo de edad, sino grupo de edad por dominio utilizado en la construcción de la región. Así, por ejemplo, si se desea estimar la población ocupada de Lanzarote, debemos calcular la suma ponderada de las correspondientes tres comarcas con sus respectivos grupos de edades de cada estimación media directa, estas últimas basadas en las regiones r que se estuvieran utilizando.

Adicionalmente, como Eurostat también solicita información a nivel de isla por sexo y grupo de edad (<25 y ≥25 años), se consideró la opción de mantener la estructura de regiones y estimador sintético-región por comarcas, ya calculado para tres grupos de edades, y construir el correspondiente a cada comarca sumando únicamente dentro de aquellos grupos de edades en los que se desea obtener estimación. Así la estimación para los grupos de edades viene dada por:

$$\hat{y}_{d(<25)}^{sint_r_d} = \hat{\bar{y}}_{(<25),r_d}^{direct} N_{d(<25)} = \frac{\sum_{j \in EPA_{(<25),rd}} w_j y_j}{\sum_{j \in EPA_{(<25),rd}} w_j} N_{d(<25)}$$, y

$$\hat{y}_{d(\geq 25)}^{sint_r_d} = \sum_{g'} \hat{\bar{y}}_{g',r_d}^{direct} N_{dg'} = \sum_{g'} \frac{\sum_{j \in EPA_{g',rd}} w_j y_j}{\sum_{j \in EPA_{g',rd}} w_j} N_{dg'}$$

Donde g' indica los grupos de edades de 25-54 y >54 años.

e. **Compuesto dependiente del diseño:** la estimación se obtiene como combinación convexa del estimador posestratificado y del sintético, tanto básico como región, esto es,

$$\hat{y}_d^{dep} = \lambda_d \, \hat{y}_d^{post} + (1 - \lambda_d)\hat{y}_d^{sint} \text{ , con } \lambda_d = 1 \text{ si } \hat{N}_d \geq \alpha N_d \text{ y } \lambda_d = \hat{N}_d / (\alpha N_d)$$

en otro caso. El parámetro α vendrá fijado a 2/3, 1, 1,5 y 2, dando lugar a los estimadores que denominaremos, respectivamente, compuesto1, compuesto2, compuesto3 y compuesto4. $\hat{N}_d$ es el total estimado según Horvitz-Thompson de la población en el dominio d, esto es,

$$\hat{N}_d = \sum_{j \in EPA_d} w_j$$

3.2. Estimadores basados en el modelo

El trabajo desarrollado en el proyecto CANAREA para este tipo de estimadores se centró en los estimadores basados en modelos a nivel de área[4]. Estos modelos suelen ser utilizados cuando el número de pequeñas áreas es suficientemente alto, como puede ser el caso de la Comunidad Autónoma de Canarias. La utilización de información auxiliar a nivel de pequeña área diferente de la suministrada por la propia encuesta, y disponible en los Institutos de Estadística, puede introducirse fácilmente en ellos. Los modelos a nivel de individuos no han sido desarrollados en el proyecto pues, si bien estos modelos podrían llegar a tener un rendimiento superior a los de nivel de área, la utilización de información auxiliar a nivel de individuo, diferente a la suministrada por la encuesta, requiere de una tarea de identificación de todos los individuos muestreados inabordables por los Institutos autonómicos desde el punto de vista práctico.

Los modelos explícitos de áreas que se han utilizado son los **modelos lineales generales mixtos** que introducen efectos aleatorios específicos para las pequeñas áreas con la idea de contemplar variaciones entre ellas que no queden explicadas por las variables auxiliares introducidas en el modelo. En el proyecto Eurarea se identifica como modelo B y en la literatura de pequeñas áreas como modelo de Fay-Herriots. Denotando por $\theta_d = g(\hat{Y}_d)$, con $\hat{Y}_d$ generalmente la estimación según el estimador directo o su media y $g(\cdot)$ una función especificada, habitualmente la identidad o el logaritmo, y asumiendo que se disponen de p variables auxiliares a nivel de dominio, $x_d = (x_{1d}, \ldots, x_{pd})'$, se considera el modelo siguiente:

$$\hat{\theta}_d = x'_d \beta + \upsilon_d + e_d \, ,$$

Donde $\beta = (\beta_1, \ldots, \beta_p)'$ es el vector de coeficientes de regresión, υ_d son los efectos aleatorios específicos de cada dominio, asumidos independientes e idénticamente distribuidos con media igual a cero y varianza σ_υ^2 -es habitual además considerar la normalidad- y e_d son los errores muestrales considerados independientes con media dado θ_d igual a cero y varianza ψ_d. Además es usual asumir que ambos errores son independientes y que ψ_d es conocida. En particular, se ha fijado que $\psi_d = \sigma_e^2 / w_d$, probando varias opciones a la hora de elegir w_d (w_d=N_d, ó w_d=n_d, ó w_d=$\hat{N}_d$).

Nos centramos en los estimadores EBLUP (Empirical Best Linear Unbiased Predictor) y Sintético-regresión. El estimador sintético-regresión se define en función de los efectos fijos del modelo de regresión como $\hat{Y}_d^{sint-regr} = x_d' \, \hat{\beta}$ y el EBLUP como un estimador compuesto entre el sintético-regresión y un estimador basado en el diseño, el estimador directo, esto es:

4 Ver Rao 2003, pag. 76, Jiang and Lahiri 2006 y la documentación del proyecto Eurarea: http://www.ine.es/docutrab/eurarea/eurarea_05.pdf

$$\hat{Y}_d^{EBLUP} = \gamma_d \hat{y}_d^{directo} + (1 - \gamma_d)\hat{Y}_d^{sint-regr}$$

Con γ_d la ratio de varianza para un área específica, $\gamma_d = \hat{\sigma}_v^2 \big/ (\hat{\sigma}_v^2 + \hat{\sigma}_e^2 / w_d)$. En la estimación de β y de la varianza se utilizó el método de Fisher-Scoring, partiendo de valores iniciales generalmente calculados por mínimos cuadrados pesados asumiendo un modelo de efectos fijos.

4. El estimador sintético-región

Es conocido que la variabilidad de los estimadores sintéticos es en general de las más pequeñas, puesto que utilizan muestra indirecta grande; pero en estos estimadores también suele ocurrir que una elevada proporción del error cuadrático medio es debida al sesgo. Teniendo en cuenta tal consideración se abordó, tal como indica Rao (2003, pág. 46), la evaluación del rendimiento del estimador sintético-región, construyendo para ello regiones como uniones de comarcas que presenten un comportamiento homogéneo en cuanto a las variables objetivo a estudiar, con la finalidad de reducir el sesgo de la estimación y muy probablemente su error cuadrático.

Después de unas primeras pruebas, y apreciando un comportamiento diferenciador dependiendo del sexo así como del tipo de variable objetivo que se pretendía estimar, se construyeron distintas regiones homogéneas para cada una de las siguientes combinaciones de variables objetivo: hombre ocupado, mujer ocupada, hombre parado y mujer parada. La definición de dichas regiones homogéneas se abordó por una parte utilizando información del Censo de Población y Viviendas 2001 y por otra usando información estadística basada en registros administrativos, exactamente datos de la Estadística del Movimiento Laboral Registrado (INEM) y de la Estadística de Empleo Registrado (ISTAC), con las siguientes características:

- En el caso de información censal se emplearon los totales comarcales por grupo de edad (16 - 24; 25 - 54; > 54 años) de las propias variables objetivo relativizadas con respecto al total comarcal por sexo observadas en el Censo de Población y Viviendas de 2001. La ventaja de este tipo de información es que existen variables homónimas de las que se quieren hacer las regiones y presenta, como desventajas, que las revisiones sólo podrán realizarse cada diez años y la diferencia de definición de conceptos entre Censo y EPA.

- En el caso de variables auxiliares provenientes de registros administrativos se han utilizado inicialmente como posibles candidatas los tamaños poblacionales comarcales por grupo de edad y sexo, empleo registrado total -agrario, industria, construcción y servicio-, empleo registrado por cuenta ajena -agrario, industria, construcción y servicio-, empleo registrado por cuenta propia -agrario, industria, construcción y servicio-, paro registrado total -hombre, mujer, mayor 25 años, agrario, industria, construcción y servicio- y contratos registrados -hombres, mujeres, indefinidos, temporal, agrario, industria, construcción y servicio-. Todas ellas han sido relativizadas al tamaño de la población de la comarca, y aquellas en las que se conocía la información por sexo, al tamaño de la comarca por sexo. La ventaja de este tipo de variables estriba en que las regiones pueden actualizarse en el momento que se crea oportuno, si bien, presenta como desventajas su elección y la fijación de un periodo de revisión.

Partiendo de la información auxiliar anteriormente citada se definieron dos tipos de regiones homogéneas, regiones estáticas y regiones dinámicas, con los siguientes objetivos y características:

4.1. Regiones estáticas

Utilizando técnicas de análisis cluster -clusters de K medias o jerárquicos- sobre las variables auxiliares se construyen K regiones de tal manera que todas las comarcas de una misma región comparten su información. Las pruebas realizadas indican que el número de regiones a emplear no es conveniente que supere cinco, siendo incluso válido y apropiado únicamente dos.

4.2. Regiones dinámicas

En la construcción de las regiones estáticas se considera que todas las comarcas que componen una determinada región compartirán su información, sin tener en cuenta que para algunas de las comarcas no es necesaria tanta información indirecta. En este apartado se introduce la posibilidad de construir regiones borrosas (fuzzy), formadas por unión de aquellas comarcas que más próximas están a la que se desea estimar. De esta forma la aportación de información indirecta para la estimación de una comarca podrá variar en cada realización de la encuesta, y por lo tanto esto hace que dichas regiones dinámicas dependan de la comarca a estimar y de la realización particular de la EPA. Así si, en una realización dada, una determinada comarca posee mucha muestra directa, la utilización de la información suministrada por las comarcas próximas será menor que la de esa misma comarca si tuviese una muestra directa menor. También se producirán situaciones donde una comarca A puede prestar su información a una comarca B, no necesitando esta última de la información de la comarca A.

Para la construcción de las regiones dinámicas no es preciso definir cual es el método de generación de los cluster, sino simplemente indicar cual es la distancia entre comarcas, y más concretamente el orden de cercanía. Y para ello, se han evaluado distintas distancias: euclídea, de Manhattan y de Chebychev, no apreciándose diferencias importantes entre ellas.

Ahora bien, una vez fijados los órdenes de cercanía entre comarcas para cada una de las variables objetivo que se pretenden estimar, surge el problema de fijar cuántas de ellas prestarán su información a la estimación, y por lo tanto, resulta importante fijar un método de parada en la utilización de la información de las comarcas vecinas. Para ello, se ha considerado una versión simplificada del cálculo del tamaño muestral necesario que debería tener una comarca para obtener una estimación con una precisión fija. En particular, se ha empleado el tamaño muestral bajo muestreo aleatorio simple, a partir de la ecuación siguiente:

$$n = \frac{z_{\alpha/2}^2 P(1-P)N}{\varepsilon^2(N-1) + z_{\alpha/2}^2 P(1-P)}$$

Donde $z_{\alpha/2}$ se ha fijado a 1.96, P a 0.20 y ε la precisión variando entre el 1 y el 4%. Estos valores se han obtenido después de un estudio para los datos provinciales. Hay que recalcar que la interpretación de la precisión en la ecuación anterior varía con respecto a su interpretación estándar, donde toda la muestra es tomada dentro de la población que se desea estimar y, por tanto, con el mismo peso. Nótese que valores pequeños de ε, muestras grandes, puede interpretarse como que a una comarca dada se le posibilita la utilización de muestra proveniente de comarcas más alejadas y, al contrario, valores grandes de ε, muestras pequeñas, como que utilizará un número pequeño de comarcas quizás porque tiene pocos vecinos relativamente próximos. De esta forma, una vez definidas las matrices de distancias, una comarca utilizará la información de las comarcas más próximas hasta que la muestra alcance el nivel muestral obtenido con la precisión dada. La forma en la que se fija la precisión ha dado lugar a dos métodos diferentes.

4.2.1. Precisión constante

El valor de la precisión, ε, se fija a una cantidad dada independientemente de la comarca, variable objetivo a estudio y sexo. En general, valores entre el 1 y el 4% se han comprobado que son adecuados y, en particular, en torno al 2% los que mejores resultados presentan.

4.2.2. Precisión variable

En este caso, la precisión puede variar entre comarcas, variable objetivo y sexo. Para el método de fijación se han probado distintas posibilidades. Una primera es utilizar aquellas precisiones que produzcan menores errores cuadráticos medios en un procedimiento de simulación sobre el último Censo de Población y Viviendas realizado. Una segunda posibilidad consiste en elegir aquellas que minimizan los errores cuadráticos medios de cada estimación de variable objetivo por comarca y sexo, después de una aplicación de distintos valores de precisión constante en la realización de la EPA que se está analizando. Esta última posibilidad será la empleada, después de probar con siete valores distintos de precisiones entre el 1 y el 4% con saltos de 0.5%.

5. Calibrado a posteriori de las Estimación en pequeñas áreas y estimación de la población inactiva

Una vez obtenidas las estimaciones $\hat{Y}_d$ para cada uno de los dominios pequeños, éstas deben ser corregidas con la finalidad de que sean consistentes con los datos publicados por el INE para cada provincia. Para ello se utiliza el factor de calibrado

$$F_{yp} = \hat{y}_p^{epa} \Big/ \sum_{d \in P} \hat{Y}_d$$

siendo $\hat{y}_p^{epa}$ la estimación según la EPA para la provincia P. Asimismo, para la estimación de la población inactiva se procedió tal como sigue:

1. Se calibraron las estimaciones de población ocupada y parada para cada provincia y sexo.

2. Se obtuvo la población inactiva para cada dominio d como la diferencia entre el total poblacional[5] de dicho dominio, N_d, menos la suma de población ocupada y parada estimada.

También fue probada la posibilidad de realizar la estimación de la población inactiva, no por diferencia, tal como se ha comentado en el párrafo anterior, sino por los mismos procedimientos de estimación que la ocupada y la parada. Para ello, finalmente se aplicaba un procedimiento tipo CalMar, donde para cada provincia y sexo, las estimaciones de las tres variables objetivo en las comarcas se calibraban para que se cumpliera que:

- la suma para cada variable objetivo en las distintas comarcas produzca la estimación provincial de dicha variable objetivo, y

- la suma de las estimaciones para cada comarca en las tres variables objetivo produzca el valor poblacional de dicha comarca.

Las pruebas realizadas no mostraron diferencias importantes en el rendimiento global en la utilización de ambos métodos, así que por simplicidad se ha optado por el primero de ellos.

Adicionalmente para obtener la consistencia de las estimaciones a nivel de isla por grupo de edad (<25 y ≥25) con las de nivel de isla totales y las de la provincia por grupo de edad, tanto por sexo como por variable objetivo, se calibraron utilizando un procedimiento tipo CalMar, similar al descrito en la sección 6.

6. Estimación en pequeñas áreas de la población actual: un input necesario

Tal como señalamos en la introducción, el pasado 1 de febrero de 2007 se aprobó el Reglamento (CE) Nº 105/2007 de la Comisión, publicado en el Diario Oficial de la Unión Europea del 10 de febrero de 2007, que modifica los anexos del Reglamento (CE) Nº 1059/2003 del Parlamento Europeo y del Consejo, publicado en el Diario Oficial de la Unión Europea del 21 de junio de 2003, por el que se establece una nomenclatura común de unidades territoriales (NUTS). En lo que respecta a España, el citado Reglamento Nº 105/2007 modifica el nivel NUT 3 de la Nomencla-

5 Los totales poblaciones insulares son ofrecidos por la Estimación de la Población Actual (ePOBa) del INE a principio de cada año, para la estimación del resto de totales poblacionales –trimestrales y comarcales- el ISTAC ha procedido según lo descrito en el apartado 6 del presente documento.

tura. En particular y en el caso de Canarias, son las islas las que pasan a considerarse el nivel NUT 3 de la clasificación en lugar de la provincia.

Como consecuencia del citado Reglamento el INE debe ofrecer estimaciones anuales de la población actual a nivel insular, datos que se proporcionan a través de la operación estadística "Estimación de la Población Actual" (ePOBa). La información que ofrece dicha operación es la estimación anual del total poblacional de las islas por sexo y grupo de edad (<25, 25-54, >54 años). Sin embargo, las Estimación en pequeñas áreas de la Encuesta de Población Activa requieren adicionalmente estimaciones trimestrales de totales poblacionales a nivel comarcal.

Para obtener las estimaciones de datos poblacionales en los diferentes trimestres se parte de los datos insulares anuales ofrecidos por la ePOBa siendo modificados siguiendo un método similar al CalMar de tal manera que se minimicen las perturbaciones al cuadrado bajo las restricciones siguientes:

- los datos ajustados concuerden con las estimaciones poblacionales por sexo a nivel de provincia suministradas por la correspondiente realización trimestral de la EPA, y

- los datos ajustados concuerden con los totales poblacionales por grupo de edad y sexo a nivel de la Comunidad Autónoma de Canarias.

Esto es, para cada sexo, sean:

a. o_{ig} y e_{ig} los total poblacional observados y ajustados, respectivamente, del grupo de edad g (g=1,2,3) en la isla i (i=1,...,7).

b. C_g el total poblacional de la C.A. en el grupo de edad g y P_p (p=1,2) el total estimado poblacional en la provincia p. Se verifica que $P_1+P_2=C_1+C_2+C_3$.

c. $o_{.g}$ el total poblacional observado en la C.A. en el grupo de edad g, $o_{.g} = \sum_{i=1}^{7} o_{ig}$.

d. t_g^p el total poblacional observado en la provincia p en el grupo de edad g y $t_.^p$ el total poblacional observado en la provincia p. Para la provincia de Las Palmas se tendrá que

$t_g^1 = \sum_{i=1}^{3} o_{ig}$ y $t_.^1 = \sum_{g=1}^{3} t_g^1$.

Esto hace que se deba resolver el siguiente problema de optimización:

$$Min \sum_{i,g} \frac{(e_{ig} - o_{ig})^2}{2o_{ig}}$$

s.a. $\sum_{i=1}^{7} e_{ig} = C_g$ para g=1,2,3

$$\sum_{i=1}^{3} \sum_{g=1}^{3} e_{ig} = P_1$$

Después de utilizar el método de los multiplicadores de Lagrange se obtiene que:

$$e_{ig} = \begin{cases} o_{ig}\left(\frac{C_g}{o_{.g}} + \alpha\left(1 - \frac{t_g^1}{o_{.g}}\right)\right) & \text{si } i \leq 3 \\ o_{ig}\left(\frac{C_g}{o_{.g}} - \alpha\frac{t_g^1}{o_{.g}}\right) & \text{si } i > 3 \end{cases}$$

Con

$$\alpha = \frac{P_1 - \sum_{g=1}^{3} C_g t_g^1 / o_{.g}}{t_{.}^1 - \sum_{g=1}^{3} (t_g^1)^2 / o_{.g}}$$

Una vez obtenidas las estimaciones insulares trimestrales, se usan los datos proporcionados por el Padrón Municipal de Habitantes a 1 de enero del respectivo año, con información de comarcas por grupo de edad y sexo, para estimar los datos comarcales mediante reparto proporcional.

7. Métodos de estimación del Error Cuadrático Medio

Se han estudiado dos métodos de remuestreo para la estimación del Error Cuadrático Medio (MSE), el método Jack-knife y el método Bootstrap. A continuación se detallan sus cálculos y se ejemplifica su uso bajo el estimador sintético-región. Ambos han sido evaluados en un proceso de simulación observando que el estimador jack-knife es el que más se aproxima a los valores reales.

7.1. Método Jack-knife

Este método consiste en tomar tantas submuestras como clusters -secciones censales en el caso de la EPA- existen en la muestra (ver Tukey, 1958). Cada submuestra se forma así con la muestra original al dejar fuera una sección censal cada vez. Denotemos por $\hat{Y}_{d(hi)}^{sint_r_d}$ el estimador sintético-región obtenido con la submuestra que elimina la sección censal i del estrato h, y donde los pesos han sido modificados (denotamos por $w_{j(hi)}$) de la siguiente forma:

$$w_{j(hi)} = \begin{cases} w_j & \text{si el individuo } j \text{ no está en el estrato } h \\ 0 & \text{si el individuo } j \text{ está en la sección } i \text{ del estrato } h \\ \dfrac{n_h}{n_h - 1} w_j & \text{si el individuo } j \text{ está en el estrato } h \text{ pero no en la seccón } i \end{cases}$$

Con n_h el número de secciones en el estrato h $(h=1,...,H)$. El estimador jack-knife del error cuadrático medio del estimador sintético-región es:

$$\hat{MSE}_{JK}(\hat{Y}_d^{sint_r_d}) = \sum_{h=1}^{H} \left(\frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{Y}_{d(hi)}^{sint_rd} - \hat{Y}_{d(h.)}^{sint_rd})^2 + ((n_h - 1)(\hat{Y}_{d(h.)}^{sint_r_d} - \hat{Y}_d^{sint_r_d}))^2 \right)$$

Con $\hat{Y}_{d(h.)}^{sint_r_d} = \sum_{i=1}^{n_h} \hat{Y}_{d(hi)}^{sint_r_d} / n_h$

7.2. Método Bootstrap

Existen diversas versiones para introducir este método de remuestreo. En particular, se ha utilizado la propuesta dada por Rao y Wu (1988) para muestreo simple estratificado. Otra posibilidad es dada por Efron y Tibshirani (1988) y Shao y Tu (1995, Cap. 6), que requiere la generación de muestras con un diseño bietápico. Los pasos a seguir en la implementación llevada a cabo son:

a. Dado un estrato h con n_h secciones, se extrae una submuestra con n_h-1 secciones mediante muestreo aleatorio simple con reemplazamiento, repitiendo este procedimiento de forma independiente para cada estrato.

b. Para cada una de las submuestras r $(r=1,...,R)$ se construye unos nuevos pesos, $w_{hij}^r = w_{hij} n_h / (n_h - 1) m_i^r$ con m_i^r el número de veces que la sección i es seleccionada en la submuestra. A partir de estos pesos se obtiene el estimador sintético-región, esto es, $\hat{Y}_{dr}^{sint_r_d}$.

Este procedimiento se repetirá R veces, y finalmente la estimación bootstrap del error cuadrático medio es obtenida como:

$$M\hat{S}E_B(\hat{Y}_d^{sint_r_d}) = \frac{1}{R-1}\sum_{r=1}^{R}(\hat{Y}_{dr}^{sint_r_d} - \hat{Y}_d^{sint_r_d})^2$$

Los valores de R probados han variado entre $200,\ 500,\ 1000$ y $5000,$ no encontrando variaciones a partir de los valores de $R=500$.

8. Valoración de estimadores sobre simulaciones de muestras EPA a partir del Censo 2001

Con la intención de realizar una valoración de los distintos estimadores propuestos, así como de los métodos de calibrado y estimación de errores cuadráticos medios, se construyó un simulador de muestra EPA partiendo como población de los datos del Censo de Población y Viviendas de 2001. Ya que los datos que se desean estimar serán conocidos bajo estas condiciones se pueden valorar todos los métodos propuestos. En particular todas las conclusiones se han basado en los resultados de 1000 muestras simuladas EPA, valorando los indicadores de rendimiento que se detallan a continuación.

8.1. Indicadores de rendimiento

Para cada uno de los estimadores y variables objetivo estudiados se han obtenido los siguientes indicadores ($Sim=1000$, D el número de áreas pequeñas, Y_d los valores reales a estimar e $\hat{y}_d(k)$ la estimación en la simulación k):

a. Sesgo relativo *(SR)*: $SR_d(\hat{y}) = \frac{1}{Sim}\sum_{k=1}^{Sim}\frac{\hat{y}_d(k) - Y_d}{Y_d}100$

b. Media del sesgo relativo absoluto *(SRAM)*: $SRAM(\hat{y}) = \frac{1}{D}\sum_d |SR_d(\hat{y})|$

c. Error cuadrático medio relativo *(REMC)*:

$$REMC_d(\hat{y}) = \left(\frac{1}{Sim}\sum_{k=1}^{Sim}\left(\frac{\hat{y}_d(k) - Y_d}{Y_d}\right)^2\right)^{1/2}100$$

d. Media del error cuadrático medio relativo *(REMCM)*:

$$REMCM(\hat{y}) = \frac{1}{D}\sum_d REMC_d(\hat{y})$$

Nótese que $REMC$ es el coeficiente de variación estimado, si bien, cuando se calcula para una muestra EPA dada, en el denominador aparecerá el valor estimado y no el real.

8.2. Suplementación de la muestra EPA

En las primeras fases de la investigación se valoró la posibilidad de que se incluyera muestra suplementaria para cada una de las comarcas y que se combinara con la información proveniente de la muestra EPA. La muestra suplementaria se obtuvo por muestreo uniforme en comarca o islas con selección de las viviendas por muestreo aleatorio simple. Estos estudios de simulación mostraron que el estimador sintético-básico (el estimador sintético basado en la provincia) producía los menores $REMC$ en la mayoría de las comarcas, incluso después de duplicar la muestra, si bien también, con los mayores SR.

Los estimadores compuestos suavizaban estos sesgos a costa de aumentar la variabilidad. Estos resultados dieron pie a la construcción de los estimadores sintéticos-región ya comentados, los cuales intentan resolver los problemas de sesgos con la intención de no tener que utilizar muestra suplementaria, y a su vez reducir los costos económicos para la obtención de las estimaciones y también facilitar la generación de series pasadas de estimaciones.

8.3. Resultados sobre los estimadores evaluados

Para los diferentes estimadores descritos en la sección 3 se recogen en la tabla A los resultados de los indicadores de rendimiento $REMCM$ y $SRAM$. Las regiones utilizadas para los estimadores sintético-región estáticos como dinámicos que se muestran corresponden a las construidas con variables del Censo de Población y Viviendas de 2001. También se ha incluido la posibilidad de combinación del estático y el dinámico, utilizando las regiones del estático en aquellos casos donde los valores del $REMC$ por comarcas mejoraban al dinámico en más de dos puntos.

Como valores de referencias indicar que las estimaciones provinciales en estas simulaciones ofrecen unos valores de $REMCM$ para población ocupada de 2,32 y 3,76 y para la parada de 7,53 y 7,32 para hombres y mujeres, respectivamente. La utilización de variables auxiliares provenientes de registros administrativos para la construcción de las regiones que emplean los estimadores sintético-región también han sido evaluadas, obteniendo resultados similares a los proporcionados con las variables censales, si bien puede ser un aspecto fundamental a la hora del rendimiento de estos estimadores.

8.4. Evaluación del comportamiento de los métodos de estimación de los MSE

Los dos métodos de estimación de los errores cuadráticos medios comentados en la sección 7, Jack-knife y Bootstrap, fueron evaluados comparándolos con los valores de $REMCM$ obtenidos en el proceso de simulación. En particular, tanto la raíz cuadrada de $\hat{MSE}_{JK}$ como la de $\hat{MSE}_{B}$ fueron divididos por Y_d (el valor real en el área d) promediando los resultados obtenidos para cada una de las 1000 muestras EPA simuladas.

Para el estimador sintético-región se obtiene que el método Jack-knife es el que mejor ajusta los valores de los coeficientes de variación y que el método Bootstrap implementado los subestima.

TABLA A. Media de Errores Cuadráticos Medios, REMCM, y de Sesgos Relativos Absolutos

	POBLACIÓN OCUPADA				POBLACIÓN PARADA			
	Hombres		Mujeres		Hombres		Mujeres	
	REMCM	SRAM	REMCM	SRAM	REMCM	SRAM	REMCM	SRAM
Clásicos								
Directo	11,77	1,47	19,90	2,83	43,40	3,51	47,02	2,23
Posestratificado	10,50	1,40	18,29	2,39	44,42	3,85	46,84	2,14
Sintético-básico	6,76	6,07	13,77	13,03	21,96	19,15	20,08	17,71
Compuestos								
Compuesto 1	10,46	1,42	18,20	2,33	44,07	3,97	46,51	2,17
Compuesto 2	9,90	1,52	17,34	2,50	40,70	5,38	42,87	3,07
Compuesto 3	8,35	2,24	14,95	4,38	33,61	8,15	34,37	6,27
Compuesto 4	7,25	2,98	13,36	6,36	28,44	10,54	28,24	9,00
Sintético-región								
Estático 2-2-2-2	4,82	3,79	8,96	7,39	16,54	12,66	14,08	8,31
Estático 3-3-3-3	4,98	3,19	8,37	6,18	17,97	9,95	14,76	8,31
Estático 4-4-4-4	4,95	2,80	8,01	5,24	15,35	6,98	15,88	7,24
Estático 5-5-5-5	5,67	2,87	9,70	5,51	15,51	5,53	18,19	7,23
Sintético-región dinámico con precisión constante								
1,00%	5,33	4,47	10,28	9,23	17,02	13,14	16,28	12,78
1,50%	4,62	3,30	8,39	6,75	15,13	9,15	14,12	7,91
1,75%	4,56	3,19	8,21	5,98	15,04	8,09	13,89	7,24
2,00%	4,61	3,25	7,91	5,44	15,28	6,94	14,46	7,17
2,50%	4,80	2,93	8,25	5,37	16,70	6,58	16,29	6,63
3,00%	4,86	2,53	8,48	5,24	17,74	5,88	17,53	6,46
Sintético-región dinámico con precisión variable	4,22	2,67	6,73	4,26	13,61	6,70	13,53	7,45
Combinación sintético-región dinámico con precisión variable y estático	4,22	2,67	6,42	3,83	13,19	6,73	13,33	7,39

SRAM obtenidos en el estudio de simulación de generación de 1000 muestras EPA, para población ocupada y parada por sexo con los estimadores directo, posestratificado, sintético-básico, sintético-región estático (con dos, tres, cuatro y cinco cluster utilizando información censal), sintético-región dinámico con precisión constante del 1, 1.5, 1,75, 2, 2.5 y 3%, sintético-región dinámico con precisión variable (elige la precisión constante por comarca y variable objetivo que minimiza el REMC) y sintético-dinámico con precisión variable combinado con el estático 2-2-2-2.

8.5. Conclusiones

Después de todos los estudios de simulación realizados se concluye que un estimador adecuado para las variables objetivo en las pequeñas áreas planificadas es el estimador sintético-región dinámico con precisión variable. Este tipo de estimador tiene las siguientes ventajas:

a. Posee en general bajas variabilidades, lo que ayudará a la estabilización de la serie trimestral.

b. Posibilita que áreas que posean muestra directa suficiente utilicen poca o ninguna muestra indirecta y, al contrario, áreas con escasa muestra directa se aprovechen de información indirecta, empezando por las que aportan las áreas más próximas. Aunque los estimadores compuestos se basan en esta idea, el hecho de separar el proceso de asignación de información del de estimación, da mayores grados de libertad.

c. Permite conocer la estructura de asignación de información que se está compartiendo para cada una de las comarcas, pudiendo en cualquier instante cambiar o añadir nuevas variables auxiliares que ayuden a mejorar las estimaciones.

d. Posibilita la modificación de la regla de parada a la hora de selección de comarcas -elección de precisión- lo que permite trabajar con estimadores con errores cuadráticos

medios relativamente bajos pero con altos sesgos, o errores cuadráticos medios más elevados pero insesgados (en el límite el estimador posestratificado).

Las desventajas que tiene son:

a. Necesita definir una matriz de órdenes de cercanía para cada una de las comarcas y variable objetivo, además de indicar el periodo de vigencia de tales matrices. Pruebas llevadas a cabo sobre realizaciones reales de EPA muestran que periodos inferiores a cinco años en general no serán necesarios y, en particular, se prevé hacer revisiones coincidentes con las publicaciones de los Censos de Población y Vivienda.

b. Los estimadores sintéticos no son insesgados y pueden producir sesgos relativamente grandes en comparación con sus errores cuadráticos medios. En los procesos de simulación realizados este problema era prácticamente inexistente puesto que los errores cuadráticos medios que producían estos estimadores eran los más pequeños de todos los estimadores evaluados.

9. Aplicación a muestras reales de EPA

A continuación se realiza una breve descripción, ficha técnica, detallando los métodos que serán utilizados en la obtención de las estimaciones de los totales de población ocupada, parada e inactiva, así como de las tasas de actividad, paro y empleo.

Estimador:

Sintético-región con precisión dinámica. Se utiliza la edad –tres grupos- como variable de estratificación.

Regiones homogéneas:

Se emplea la distancia euclídea sobre las variables estandarizadas en función del tamaño comarcal por sexo y se construyen matrices de cercanía para hombres ocupados, mujeres ocupadas, hombres parados y mujeres paradas. Las variables consideradas inicialmente son las proporcionadas por el Censo de Población y Viviendas desagrupadas por grupo de edad (<25, 25-54, >54 años).

Método de parada:

Construidas las matrices de cercanía, el método utilizado de parada empleará el tamaño muestral obtenido bajo muestreo aleatorio simple para una proporción ($P=0.20$ y $z_{\alpha/2}=1.96$) con precisión variable, $\mathcal{E}$, en función de comarca, variable objetivo y sexo obtenidas a partir de la minimización de los coeficientes de variación estimados cuando se utilizan los valores de precisión constante del 1, 1.5, 2, 2.5, 3, 3.5 y 4%.

Calibrado de datos poblacionales:

Los datos a nivel de isla proporcionados por la ePOBa son ajustados a las estimaciones provinciales por sexo y a los datos de la Comunidad Autónoma de Canarias por grupo de edad por sexo. Posteriormente los datos comarcales por grupo de edad y sexo del Padrón Municipal de Habitantes a uno de enero del correspondiente año serán ajustados, por reparto proporcional, a los datos de las islas por grupo de edad por sexo.

Calibrado de las estimaciones:

Las estimaciones comarcales de población ocupada y parada por sexo proporcionadas por el estimador sintético-región serán calibradas, por reparto proporcional, a las estimaciones provinciales proporcionadas por la EPA, y posteriormente calculada la población inactiva por diferencia con los totales comarcales por sexo. Para las estimaciones a nivel de isla por grupo de edad (<25 y ≥25 años) y sexo se sigue un proceso similar pero calibrando a los datos

provinciales para esos dos grupos de edades, y la población inactiva por diferencia a los datos comarcales por grupo de edad y sexo.

Estimación de los coeficientes de variación:

Se calcula el cociente entre la raíz cuadrada de la estimación del error cuadrático medio y la estimación. La estimación del error cuadrático medio es obtenida utilizando el método Jack-knife.

Publicación de datos:

Las estimaciones serán realizadas trimestralmente, si bien cuando se precisen datos anuales se promediarán los cuatro trimestres, tanto de las estimaciones como de los coeficientes de variación.

9.1. Resultados sobre las muestra reales de EPA de 2005 a 2007

Como ejemplo de la metodología anteriormente comentada se muestran las estimaciones y los coeficientes de variación obtenidos cuando se aplican a las muestras EPA reales comprendidas entre el primer trimestre de 2005 y el cuarto trimestre de 2007.

Tabla 1. Población ocupada de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**836,14**	**2,82**	**880,37**	**2,68**	**915,41**	**2,76**
Las Palmas	**432,49**	**3,90**	**463,81**	**3,76**	**480,48**	**4,02**
Lanzarote	54,04	3,96	60,39	3,85	64,13	4,10
Fuerteventura	37,02	3,96	42,23	3,84	45,09	4,08
Gran Canaria	341,44	3,89	361,19	3,75	371,27	4,02
S/C Tenerife	**403,64**	**4,07**	**416,57**	**3,82**	**434,93**	**3,73**
Tenerife	355,54	4,08	368,60	3,84	384,48	3,72
La Gomera	9,07	4,09	9,17	3,83	9,72	3,84
La Palma	35,02	4,12	34,78	3,83	36,57	3,91
El Hierro	4,01	4,17	4,02	3,82	4,17	3,96

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 2. Hombres ocupados de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**502,01**	**3,07**	**526,24**	**2,83**	**538,44**	**2,96**
Las Palmas	**259,96**	**4,17**	**276,93**	**3,95**	**285,88**	**4,23**
Lanzarote	32,70	4,20	36,25	4,11	38,27	4,27
Fuerteventura	22,86	4,21	25,89	4,11	27,58	4,30
Gran Canaria	204,40	4,17	214,79	3,93	220,03	4,24
S/C Tenerife	**242,06**	**4,53**	**249,30**	**4,06**	**252,56**	**4,09**
Tenerife	212,31	4,53	220,01	4,08	222,47	4,08
La Gomera	5,62	4,56	5,61	4,10	5,85	4,21
La Palma	21,58	4,61	21,13	4,07	21,68	4,23
El Hierro	2,55	4,66	2,55	4,08	2,56	4,33

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 3. Mujeres ocupadas de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**334,12**	**3,44**	**354,14**	**3,48**	**376,98**	**3,44**
Las Palmas	**172,54**	**4,70**	**186,87**	**4,80**	**194,61**	**5,09**
Lanzarote	21,34	4,91	24,14	4,98	25,85	5,45
Fuerteventura	14,16	4,95	16,34	5,02	17,51	5,51
Gran Canaria	137,04	4,67	146,40	4,82	151,24	5,05
S/C Tenerife	**161,59**	**5,04**	**167,26**	**5,03**	**182,37**	**4,59**
Tenerife	143,24	5,06	148,59	5,03	162,01	4,57
La Gomera	3,45	5,16	3,56	5,01	3,87	4,69
La Palma	13,44	5,10	13,65	5,26	14,88	4,90
El Hierro	1,46	5,33	1,46	5,52	1,61	6,16

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 4. Población ocupada de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**87,60**	**6,73**	**91,68**	**7,10**	**90,46**	**7,02**
Las Palmas	**47,43**	**9,14**	**49,52**	**10,22**	**48,51**	**10,35**
Lanzarote	5,97	9,52	6,67	10,75	6,74	11,18
Fuerteventura	4,11	9,64	4,75	10,86	4,93	11,15
Gran Canaria	37,35	9,07	38,11	10,13	36,84	10,21
S/C Tenerife	**40,17**	**9,97**	**42,15**	**9,68**	**41,95**	**9,23**
Tenerife	35,73	9,95	37,83	9,68	37,30	9,22
La Gomera	0,78	10,53	0,78	10,29	0,83	9,87
La Palma	3,31	10,49	3,19	10,39	3,45	9,65
El Hierro	0,36	10,55	0,35	10,51	0,37	9,91

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 5. Población ocupada de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**748,53**	**2,95**	**788,70**	**2,75**	**824,96**	**2,76**
Las Palmas	**385,06**	**3,99**	**414,28**	**3,82**	**431,97**	**4,09**
Lanzarote	48,07	4,05	53,72	3,93	57,39	4,15
Fuerteventura	32,90	4,05	37,48	3,91	40,16	4,17
Gran Canaria	304,09	3,98	323,08	3,82	334,43	4,10
S/C Tenerife	**363,47**	**4,36**	**374,41**	**3,97**	**392,98**	**3,64**
Tenerife	319,82	4,37	330,77	3,99	347,18	3,63
La Gomera	8,29	4,32	8,39	3,92	8,88	3,74
La Palma	31,71	4,38	31,59	3,93	33,12	3,80
El Hierro	3,65	4,40	3,66	3,91	3,80	3,83

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 6. Población parada de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**111,00**	**5,99**	**116,40**	**6,21**	**106,70**	**6,25**
Las Palmas	**63,64**	**7,69**	**63,89**	**8,70**	**61,14**	**9,00**
Lanzarote	6,46	12,15	6,45	14,34	6,43	13,68
Fuerteventura	4,68	9,49	5,07	11,87	5,00	12,60
Gran Canaria	52,51	8,32	52,38	9,21	49,70	9,35
S/C Tenerife	**47,35**	**9,53**	**52,51**	**8,82**	**45,56**	**8,30**
Tenerife	40,86	9,91	45,66	9,29	39,73	8,72
La Gomera	1,22	11,86	1,30	11,28	1,08	11,57
La Palma	4,74	9,80	5,00	8,84	4,30	9,22
El Hierro	0,54	14,38	0,55	12,49	0,46	13,33

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 7. Tasa de paro entre la población de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**11,73**	**5,44**	**11,68**	**5,66**	**10,44**	**5,61**
Las Palmas	**12,83**	**6,81**	**12,10**	**7,69**	**11,28**	**8,05**
Lanzarote	10,68	11,07	9,65	13,26	9,11	12,25
Fuerteventura	11,22	8,71	10,71	10,80	9,98	11,17
Gran Canaria	13,33	7,28	12,66	8,04	11,80	8,43
S/C Tenerife	**10,51**	**8,86**	**11,20**	**8,35**	**9,48**	**7,51**
Tenerife	10,32	9,26	11,03	8,84	9,37	7,92
La Gomera	11,81	10,44	12,43	9,92	9,97	10,30
La Palma	11,93	8,76	12,56	7,95	10,52	8,26
El Hierro	11,84	12,70	12,07	11,15	9,93	11,95

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 8. Hombres parados de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**54,67**	**8,11**	**54,82**	**8,33**	**50,14**	**9,42**
Las Palmas	**32,54**	**11,02**	**30,41**	**11,93**	**29,46**	**13,88**
Lanzarote	3,50	16,37	3,28	19,43	3,18	22,63
Fuerteventura	2,36	16,16	2,38	20,61	2,39	24,11
Gran Canaria	26,69	11,82	24,75	12,38	23,89	13,57
S/C Tenerife	**22,13**	**11,85**	**24,41**	**11,36**	**20,68**	**11,49**
Tenerife	18,84	12,63	20,96	12,19	17,85	12,35
La Gomera	0,66	14,69	0,70	13,80	0,57	16,26
La Palma	2,33	13,34	2,44	13,21	2,02	14,42
El Hierro	0,31	14,62	0,32	13,83	0,25	16,26

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 9. Mujeres paradas de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**56,33**	**7,91**	**61,58**	**7,61**	**56,55**	**7,53**
Las Palmas	**31,11**	**10,13**	**33,48**	**10,54**	**31,68**	**10,50**
Lanzarote	2,96	19,21	3,17	19,31	3,26	18,02
Fuerteventura	2,32	10,42	2,68	10,43	2,61	10,43
Gran Canaria	25,82	10,82	27,63	11,35	25,81	11,47
S/C Tenerife	**25,22**	**12,52**	**28,10**	**11,00**	**24,88**	**10,65**
Tenerife	22,02	12,87	24,70	11,43	21,88	11,04
La Gomera	0,56	17,08	0,61	17,19	0,51	16,31
La Palma	2,42	12,41	2,56	10,32	2,28	10,59
El Hierro	0,23	26,19	0,23	21,64	0,21	21,47

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 10. Tasa de paro entre hombres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**9,83**	**7,59**	**9,43**	**7,74**	**8,52**	**8,54**
Las Palmas	**11,13**	**9,98**	**9,89**	**10,80**	**9,34**	**12,34**
Lanzarote	9,66	15,18	8,28	18,25	7,63	20,60
Fuerteventura	9,35	15,12	8,42	19,24	7,95	22,09
Gran Canaria	11,56	10,59	10,33	11,09	9,79	12,15
S/C Tenerife	**8,38**	**11,56**	**8,92**	**10,97**	**7,57**	**10,86**
Tenerife	8,16	12,35	8,69	11,83	7,43	11,69
La Gomera	10,47	13,22	11,03	12,31	8,82	14,76
La Palma	9,73	12,32	10,34	12,03	8,52	13,29
El Hierro	10,72	13,17	11,19	12,50	8,92	14,73

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC)

Tabla 11. Tasa de paro entre mujeres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**14,44**	**7,36**	**14,81**	**7,12**	**13,05**	**7,12**
Las Palmas	**15,26**	**9,54**	**15,19**	**9,72**	**14,00**	**10,34**
Lanzarote	12,17	17,77	11,62	17,74	11,15	17,32
Fuerteventura	14,07	10,15	14,11	9,91	12,98	10,64
Gran Canaria	15,84	9,95	15,87	10,25	14,58	11,01
S/C Tenerife	**13,53**	**11,51**	**14,38**	**10,53**	**12,01**	**9,47**
Tenerife	13,35	11,87	14,25	10,92	11,90	9,82
La Gomera	13,91	14,98	14,56	15,04	11,65	14,41
La Palma	15,26	11,10	15,80	9,78	13,28	9,58
El Hierro	13,71	23,16	13,55	19,64	11,45	19,25

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC)

Tabla 12. Población parada de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**28,14**	**11,45**	**28,43**	**11,05**	**26,08**	**10,98**
Las Palmas	**15,62**	**14,36**	**16,77**	**15,05**	**15,35**	**15,06**
Lanzarote	1,57	22,23	1,53	21,47	1,76	18,61
Fuerteventura	1,18	18,89	1,26	17,75	1,32	16,54
Gran Canaria	12,88	14,83	13,98	15,63	12,26	16,40
S/C Tenerife	**12,52**	**18,35**	**11,66**	**15,70**	**10,73**	**15,81**
Tenerife	10,91	18,90	10,02	16,53	9,39	16,49
La Gomera	0,28	19,47	0,32	16,96	0,26	18,11
La Palma	1,20	17,62	1,18	14,99	0,98	16,41
El Hierro	0,14	23,77	0,14	18,70	0,10	22,79

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 13. Población parada de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**82,85**	**6,61**	**87,97**	**6,99**	**80,61**	**7,10**
Las Palmas	**48,02**	**8,79**	**47,12**	**10,11**	**45,78**	**10,03**
Lanzarote	4,89	13,28	4,92	16,57	4,67	15,61
Fuerteventura	3,50	10,44	3,81	14,11	3,68	14,48
Gran Canaria	39,63	9,49	38,40	10,54	37,43	10,32
S/C Tenerife	**34,83**	**10,12**	**40,85**	**9,55**	**34,83**	**9,84**
Tenerife	29,95	10,58	35,63	10,03	30,34	10,33
La Gomera	0,93	12,81	0,99	12,63	0,81	12,85
La Palma	3,55	10,35	3,82	9,77	3,32	10,28
El Hierro	0,40	15,06	0,41	13,68	0,36	14,90

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 14. Tasa de paro entre la población de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**24,36**	**9,13**	**23,66**	**9,33**	**22,36**	**9,47**
Las Palmas	**24,73**	**11,35**	**25,23**	**12,74**	**24,00**	**12,23**
Lanzarote	20,65	18,05	18,60	19,20	20,70	15,92
Fuerteventura	22,18	15,05	20,92	15,97	21,17	13,93
Gran Canaria	25,58	11,64	26,75	12,77	24,90	13,07
S/C Tenerife	**23,93**	**14,94**	**21,71**	**13,33**	**20,30**	**15,01**
Tenerife	23,54	15,48	21,00	14,14	20,03	15,61
La Gomera	26,83	14,52	28,85	13,00	23,99	15,17
La Palma	26,74	13,57	26,98	12,02	22,12	14,81
El Hierro	27,59	18,70	28,80	14,69	21,13	19,86

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias. Instituto Canario de Estadística (ISTAC).

Tabla 15. Tasa de paro entre la población de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**9,97**	**6,25**	**10,03**	**6,47**	**8,90**	**6,50**
Las Palmas	**11,09**	**7,97**	**10,21**	**9,15**	**9,57**	**9,22**
Lanzarote	9,24	12,40	8,39	15,51	7,52	14,28
Fuerteventura	9,61	9,92	9,22	12,98	8,39	13,22
Gran Canaria	11,53	8,49	10,62	9,46	10,05	9,55
S/C Tenerife	**8,75**	**9,98**	**9,84**	**9,13**	**8,14**	**8,96**
Tenerife	8,57	10,44	9,73	9,63	8,04	9,43
La Gomera	10,09	11,74	10,50	11,36	8,37	11,67
La Palma	10,06	9,73	10,78	8,91	9,11	9,33
El Hierro	9,91	13,92	10,03	12,27	8,68	13,42

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias. Instituto Canario de Estadística (ISTAC).

Tabla 16. Tasa de actividad entre la población de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**59,38**	**1,52**	**61,02**	**1,43**	**61,04**	**1,42**
Las Palmas	**60,78**	**1,99**	**63,16**	**1,94**	**63,27**	**1,86**
Lanzarote	63,29	3,89	66,60	3,82	66,71	4,20
Fuerteventura	66,28	3,80	70,05	3,79	69,57	4,16
Gran Canaria	59,89	3,77	61,94	3,69	62,06	3,88
S/C Tenerife	**57,92**	**2,31**	**58,79**	**2,07**	**58,71**	**2,14**
Tenerife	58,31	3,90	59,34	3,61	59,14	3,64
La Gomera	56,41	4,10	56,25	3,84	56,37	3,86
La Palma	55,14	4,01	54,76	3,69	55,72	3,82
El Hierro	53,70	4,27	53,17	3,83	53,68	4,03

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias. Instituto Canario de Estadística (ISTAC).

Tabla 17. Tasa de actividad entre hombres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**70,17**	**1,40**	**71,46**	**1,37**	**70,58**	**1,40**
Las Palmas	**71,32**	**1,95**	**72,79**	**1,98**	**72,53**	**1,90**
Lanzarote	73,09	4,16	75,45	4,10	74,55	4,70
Fuerteventura	74,61	4,12	77,41	4,20	76,65	4,75
Gran Canaria	70,71	4,13	71,87	3,94	71,72	4,29
S/C Tenerife	**68,94**	**2,02**	**70,02**	**1,87**	**68,46**	**2,04**
Tenerife	69,19	4,29	70,55	3,85	68,80	3,97
La Gomera	68,19	4,57	67,38	4,15	66,94	4,29
La Palma	67,14	4,49	66,13	4,04	66,04	4,21
El Hierro	66,39	4,64	65,87	4,07	65,01	4,41

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 18. Tasa de actividad entre mujeres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**48,71**	**2,38**	**50,68**	**2,16**	**51,57**	**2,14**
Las Palmas	**50,13**	**2,95**	**53,31**	**2,82**	**53,71**	**2,82**
Lanzarote	52,74	4,75	56,93	4,92	58,04	5,07
Fuerteventura	56,61	4,40	61,38	4,57	61,18	4,88
Gran Canaria	49,20	4,28	52,05	4,50	52,34	4,52
S/C Tenerife	**47,23**	**3,76**	**48,00**	**3,28**	**49,42**	**3,24**
Tenerife	47,79	4,76	48,60	4,54	49,97	4,47
La Gomera	44,40	5,25	44,98	5,14	45,78	4,81
La Palma	43,43	4,85	43,81	4,72	45,83	4,65
El Hierro	40,58	5,92	40,07	5,47	42,28	5,39

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 19. Tasa de actividad entre la población de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**49,64**	**6,30**	**52,28**	**6,40**	**51,12**	**6,30**
Las Palmas	**51,75**	**8,53**	**55,15**	**9,13**	**52,44**	**9,47**
Lanzarote	50,75	9,62	54,48	9,88	54,91	10,25
Fuerteventura	52,00	9,47	56,47	9,70	55,96	10,24
Gran Canaria	52,08	8,50	55,48	9,17	53,10	9,49
S/C Tenerife	**47,29**	**9,35**	**49,18**	**8,83**	**49,60**	**7,90**
Tenerife	47,15	9,36	49,00	8,87	48,36	7,91
La Gomera	48,08	10,36	50,31	9,53	49,97	9,04
La Palma	47,17	9,89	47,04	9,31	48,67	8,40
El Hierro	46,95	10,58	47,37	9,66	45,92	9,13

(p) Datos provisionales.
Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias.
Instituto Canario de Estadística (ISTAC).

Tabla 20. Tasa de actividad entre la población de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p)

	2005		2006		2007	
	Estimación	cv	Estimación	cv	Estimación	cv
Canarias	**61,05**	**2,81**	**62,46**	**2,66**	**62,60**	**2,70**
Las Palmas	**62,37**	**3,85**	**64,50**	**3,76**	**65,06**	**3,98**
Lanzarote	65,60	3,95	68,74	3,91	68,74	4,25
Fuerteventura	69,04	3,86	72,59	3,91	72,07	4,22
Gran Canaria	61,23	3,88	63,00	3,78	63,47	3,96
S/C Tenerife	**59,69**	**4,10**	**60,33**	**3,75**	**60,07**	**3,60**
Tenerife	60,21	4,10	61,03	3,77	60,81	3,60
La Gomera	57,55	4,23	57,04	3,93	57,20	3,77
La Palma	56,36	4,17	55,90	3,81	56,72	3,74
El Hierro	54,67	4,31	53,98	3,97	54,71	3,97

(p) Datos provisionales.

Fuente: Encuesta de Población Activa (EPA). Estimación en pequeñas áreas de Canarias. Instituto Canario de Estadística (ISTAC).

Bibliografía

Efron, B. and Tibshirani, R.B. (1988). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Enhancing Small Area Estimation Techniques to meet European Needs (EURAREA Proyect) (2005). http://www.ine.es/docutrab/eurarea/eurarea_05.pdf.

Jiang, J. and Lahiri, P. (2006). *Mixed model prediction and small area estimation*. Test 15 (1), pp. 1-96.

Institut Nacional de la Statistique et des Études Économiques (1993). *"La macro Calmar, Redressement d'un échantillon par calage sur marges"* O. Sautory. Document de travail de la Direction des Statistiques Démographiques et Sociales, INSEE, F9310.

Rao, J.N.K. (2003). *Small Area Estimation*. John Wiley & Sons, New Jersey.

Rao, J.N.K. and Wu, C.F.J. (1988). *Resampling Inference with Complex Survey Data*. Journal of the American Statistical Association, 83, pp. 231-241.

Shao, J. and Tu, D. (1995). The Jackknife and Bootstrap. Springer Verlag.

Tukey, J. (1958). *Bias and Confidence in not Quite Large Simples*. *Abstract, Annals Mathematics and Statistics*, 29, p. 614.

Índice de Tablas

TABLA A. Media de Errores Cuadráticos Medios, REMCM, y de Sesgos Relativos Absolutos **/23**

Tabla 1. Población ocupada de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/25**

Tabla 2. Hombres ocupados de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/26**

Tabla 3. Mujeres ocupadas de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/26**

Tabla 4. Población ocupada de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/26**

Tabla 5. Población ocupada de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/27**

Tabla 6. Población parada de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/27**

Tabla 7. Tasa de paro entre la población de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/27**

Tabla 8. Hombres parados de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/28**

Tabla 9. Mujeres paradas de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/28**

Tabla 10. Tasa de paro entre hombres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/28**

Tabla 11. Tasa de paro entre mujeres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/29**

Tabla 12. Población parada de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/29**

Tabla 13. Población parada de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/29**

Tabla 14. Tasa de paro entre la población de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/30**

Tabla 15. Tasa de paro entre la población de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/30**

Tabla 16. Tasa de actividad entre la población de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/30**

Tabla 17. Tasa de actividad entre hombres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/31**

Tabla 18. Tasa de actividad entre mujeres de 16 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/31**

Tabla 19. Tasa de actividad entre la población de 16 a 24 años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/31**

Tabla 20. Tasa de actividad entre la población de 25 y más años en Canarias y sus provincias e islas. Estimación en miles y coeficiente de variación en porcentajes. (p) **/32**

istac | INSTITUTO CANARIO
DE ESTADÍSTICA