
MODELO FAY-HERRIOT ESPACIO-TEMPORAL PARA LA ESTIMACIÓN DE LOS RESULTADOS DEL MÓDULO DE MATEMÁTICAS DE LA PRUEBA SABER-11 2018 CALENDARIO A DESAGREGADA POR DEPARTAMENTO

SPATIO-TEMPORAL FAY-HERRIOT MODEL FOR ESTIMATE THE RESULTS OF THE MATHS MODULE OF THE SABER-11 TEST 2018 CALENDAR A DISAGGREGATED BY DEPARTMENT

Mayron Esteban Gutiérrez Pérez.^a
mayrongutierrez@usantotomas.edu.co

Wilmer Pineda Rios.^b
wilmerpineda@usantotomas.edu.co

Resumen

El examen de estado SABER-11 es una prueba de evaluación de la educación media aplicada a todos los estudiantes en proceso de grado de educación media en Colombia, los resultados son utilizados por distintas entidades estatales, así como colegios y estudiantes, al ser una prueba de tipo censal, los resultados a menudo tardan un tiempo importante en ser procesados. El presente trabajo académico pretende estimar los resultados del módulo de Matemáticas de la prueba Saber-11 del año 2018 Calendario A a partir de un diseño probabilístico utilizando el modelo de Fay-Herriot (1979), añadiendo los factores espacial y temporal propuestos por Esteban et al (2016). Los resultados muestran que con una muestra pequeña es posible estimar de manera precisa los resultados desagregados por departamento, además, de estimar también los casos de no respuesta. –

Palabras clave: sae, saber-11, SAR, AR, spacio-temporal, matemáticas.

Abstract

SABER-11 is a test of evaluation of secondary education applied to all students in the process of middle school education in Colombia, the results are used by different state entities, as well as schools and students, being a test of census type, the results often take an important time to be processed. The present academic work aims to estimate the results of the Maths module of the Saber-11 test of 2018 Calendar A from a probabilistic design using the Fay-Herriot model (1979), adding the spatial and temporal factors proposed by Esteban et al (2016). The results show that with a small sample it is possible to estimate in a precise way the results disaggregated by department, in addition, to also estimate the cases of non-response. –

Keywords: sae, saber-11, SAR, AR, spatio-temporal, maths.

1. Introducción

El examen de estado para la evaluación de la educación media Saber-11 es una prueba estandarizada realizada por el ICFES (Instituto Colombiano para la evaluación de la educación), cuyo objetivo es

^aEstudiante de Estadística Universidad Santo Tomás Bogotá

^bDocente de Estadística Universidad Santo Tomás Bogotá

comprobar el grado de desarrollo en cinco competencias consideradas como fundamentales para el gobierno colombiano (lectura crítica, matemáticas, ciencias naturales, sociales y ciudadanas e inglés) de los estudiantes que cursan el último grado de la educación media en Colombia. Este examen es de vital importancia para el monitoreo de la calidad de los establecimientos educativos en el país, así como un insumo importante para el desarrollo de políticas estatales que contribuyan al mejoramiento de la educación en Colombia. Por otro lado, el examen permite a estudiantes e instituciones brindar una herramienta de auto-evaluación de destrezas y valores adquiridos durante su proceso de educación.

Hoy en día, existen técnicas estadísticas que permiten estimar indicadores y cifras de manera precisa tomando una muestra representativa de la población objetivo, como por ejemplo, las técnicas de muestreo. Estas permiten ahorrar recursos logísticos, económicos y de tiempo. En el campo de la evaluación de políticas públicas, estimar indicadores a través de un diseño muestral permite ahorrar recursos y obtener más rápidamente indicadores importantes para la toma de decisiones (Gutiérrez et al, 2019).

Sin embargo, las estimaciones desagregadas por municipios o departamentos presentan el problema de tener alta variabilidad, esto debido a que el gran número de estratos aportan individualmente varianza al indicador general (Rao, 2003). Fay y Herriot (1979), proponen un modelo que estima indicadores, totales y proporciones en áreas pequeñas. Posteriormente, varios autores, como Benavent y Morales (2016), Sverchkov (2018), Rao y Molina (2015), y Patresi (2016) modelan las correlaciones espaciales y temporales entre estas áreas para mejorar la calidad de las estimaciones.

El propósito de este trabajo es estimar los resultados del módulo de matemáticas las pruebas Saber-11 del periodo 2018 Calendario A desagregados por departamento a través de un modelo Fay Herriot espacio-temporal basándose en la información de los periodos 2015-2017. Esto para verificar que los estimadores espacio temporales de Fay-Herriot estiman los resultados desagregados por departamento con una buena precisión, incluso en caso de no respuesta.

2. Marco de referencia

2.1. Prueba Saber-11

El examen de estado para la educación media Saber-11 es una prueba estandarizada realizada anualmente a estudiantes que estén finalizando el último grado de la educación media en Colombia, su objetivo es comprobar el grado de desarrollo de las competencias de los estudiantes de educación media. Se compone de 5 pruebas: Matemáticas, lectura crítica, competencias ciudadanas, ciencias naturales e inglés. Estas 5 competencias son evaluadas en 254 preguntas de opción múltiple con única respuesta. A continuación se muestra una descripción general de cada una de las pruebas.

La prueba de matemáticas evalúa las competencias de los estudiantes para enfrentar situaciones que pueden resolverse con el uso de herramientas matemáticas (ICFES, 2018), para evaluar las áreas de álgebra, cálculo y estadística, se introduce al evaluado en contextos familiares, laborales, comunitarios y científicos. Las competencias internas evaluadas son interpretación, formulación, ejecución y argumentación.

La prueba de lectura crítica evalúa la habilidad de comprender, interpretar y evaluar textos, esta prueba consta de 41 preguntas, cada pregunta está acompañada de un texto, estos textos pueden ser cuentos, novelas, ensayos, artículos de prensa, y fragmentos de textos argumentativos.

La prueba de sociales y competencias ciudadanas evalúa la capacidad del estudiante para usar conceptos básicos de las ciencias sociales que permiten la comprensión de problemáticas y fenómenos sociales, políticos, económicos, culturales y geográficos, como también los principios básicos de la constitución y la estructura y funcionamiento del sistema político colombiano (ICFES, 2018). Esta prueba se enfoca principalmente en la interpretación y análisis de perspectivas y en el pensamiento reflexivo.

La prueba de ciencias naturales evalúa el uso comprensivo del conocimiento científico, la capacidad de comprender argumentos y modelos y la indagación en componentes químicos, biológicos, físicos, y

tecnológicos. Esta prueba no evalúa conocimientos específicos, sino competencias, enfocando al evaluado en situaciones no cotidianas en las que se requieran habilidades científicas.

Por último, la prueba de inglés evalúa la competencia comunicativa en la segunda lengua inglesa del evaluado. Se realiza por medio de grupos de preguntas organizada según el marco común europeo, a través de lectura, gramática y léxico.

2.2. Muestreo estratificado

El muestreo estratificado es una técnica de muestreo probabilístico propuesta por Sarndal (1992), que consiste en dividir a toda la población en diferentes estratos. Luego, selecciona aleatoriamente a los sujetos finales de los diferentes estratos en forma proporcional al tamaño de muestra a través de un muestreo aleatorio simple. Es importante tener en cuenta que los estratos no deben superponerse ya que esto negaría el concepto de muestreo estratificado como un tipo de muestreo probabilístico.

Para hacer las estimaciones utilizando el diseño de muestreo planteado, se utilizarán las siguientes expresiones tomadas de (Sarndal, 1992):

la estimación del promedio esta dada por

$$\bar{Y}_u = \sum_{h=1}^H W_h \bar{Y}_{u_h} \quad (1)$$

Donde $\bar{Y}_{u_h}$ es la media aritmética de cada estrato (en este caso, cada uno de los departamentos seleccionados; y $W_h = \frac{N_h}{N}$ es el tamaño relativo de cada estrato.

La varianza estimada está definida como:

$$\hat{V}_{st}(\hat{t}_\pi) = \sum_{h=1}^H \hat{V}_h(\hat{t}_{\pi h}) \quad (2)$$

Donde $\hat{V}_h(\hat{t}_{\pi h})$ es la varianza estimada de $\hat{t}_{\pi h}$

2.3. Estimación en dominios (Áreas pequeñas)

Según Corcuera(2012): las técnicas de estimación en dominios o áreas pequeñas persiguen la obtención de estimadores de las medias y de los totales de las variables poblacionales de ciertas áreas, haciendo uso para ello de unos datos muestrales recogidos atendiendo a un diseño muestral en el que dichas áreas no recibieron una consideración específica, sino que son entidades contenidas en los estratos del diseño muestral. También se realiza un fuerte uso de la información auxiliar disponible, tanto la referente a los dominios como a las unidades muestrales.

El modelo básico de área asume que la media en la subpoblación de la variable objeto de interés, $\bar{Y}_i'$, o una cierta función de ésta, $\theta_i = g(\bar{Y}_i')$, está en relación con las variables de x_i a través del modelo lineal con efectos aleatorios ϑ_i (Corcuera, 2012):

$$\theta_i = x_i \beta + \vartheta_i \quad \text{con} \quad 1, \dots, m \quad (3)$$

donde β es el p-vector de parámetros de la regresión y las ϑ_i son incorrelacionadas con media cero y varianza σ_ϑ (Corcuera, 2012)

2.4. Modelo de Fay-Herriot

El modelo de Fay-Herriot (1979) es generalmente usado para la estimación de áreas pequeñas. Fue propuesto inicialmente para estimar el ingreso per-cápita de lugares con poblaciones menores a 1000 habitantes, este modelo es útil cuando se conoce información auxiliar acerca de las áreas y no es posible por aspectos logísticos o presupuestales realizar una muestra representativa para cada área en la población objetivo. Puede ser expresado en términos de un modelo de efectos mixtos como:

$$y_i = x_i' \beta + u_i + e_i \quad (4)$$

con $i = 1, 2, \dots, m$ cada una de las áreas

Donde x_i' es un vector de covariables conocidas, β es un vector desconocido de regresores, u_i son los efectos aleatorios del área y e_i representa el error de muestreo. Se asume que u_i y e_i son independientes con $u_i \sim N(0, A)$ y $e_i \sim N(0, D_i)$. Las varianzas A son conocidas, sin embargo, las varianzas muestrales D_i se asumen conocidas (Fay y Herriot, 1979).

La inferencia sobre las q áreas esta basado en el supuesto de que $q < m$. En adición a esto, el modelo $\hat{y}_i = y_i + e_i$ con $i = 1, 2, \dots, q$ indica como la muestra estimada y_i está relacionada con el valor desconocido y el error de muestreo $e_i \sim N(0, D_i)$. Combinando el muestreo y los modelos poblacionales se obtiene (Hansen, 2011):

$$\hat{y}_i = x_i' \beta + u_i + e_i, \quad (5)$$

con $i = 1, 2, \dots, q$ cada una de las áreas desconocidas

2.5. Mejor predictor lineal insesgado (BLUP)

Sea X una matriz de $N \times p$ de variables auxiliares, y sea Z una matriz de $n \times q$ de variables cualitativas, se supone que el vector de observaciones $Y_{N \times 1}$ cumple el modelo

$$Y = X\beta + Zu + e \quad (\text{Morales, 2015}) \quad (6)$$

donde β y u son vectores de parámetros, y e es el vector de errores aleatorios. Suponemos que los errores son independientes y normales con media cero y matriz de varianzas y covarianzas $\text{var}(e) = E(ee') = \sigma_e^2 W^{-1}$ con $W = \text{diag}(w_1, \dots, w_N)$, $w_i > 0$, los pesos w_i , son constantes positivas conocidas, pudiendo ser funciones de variables auxiliares o inversas de probabilidades de inclusión (Morales, 2015)

Usando mínimos cuadrados generalizados, los estimadores para $\hat{\beta}$ y $\hat{u}$ son para una muestra de $s \subset U$ de tamaño $n \leq N$, de la siguiente manera

$$\hat{\beta} = (X_s' P_z X_s)^{-1} X_s' P_z y_s \quad (7)$$

$$\hat{u} = G_z Z_s' W_s y_s - G_z Z_s' W_s X_s \hat{\beta} \quad (8)$$

Con $G_z = (Z_s' W_s Z_s)^{-1}$ y $P_z = W_s - W_s Z_s G_z Z_s' W_s$ (Morales, 2015)

El predictor BLU o estimador BLUP del total es:

$$\hat{y} = 1_s' y_s + 1_r' \hat{y}_r = 1_s' y_s + 1_r' Z_r \hat{u} + 1_r' X_r \hat{\beta} \quad (9)$$

2.6. Modelo de Fay-Herriot espacio-temporal

Muchas extensiones del modelo Fay-Herriot han sido propuestas en la literatura. Singh et al. (2005), Petrucci y Salvati (2006). Patresi y Salvati (2008), consideraron extensiones del modelo Fay-Herriot asumiendo que los efectos de área u_i siguen un proceso espacial autorregresivo de orden 1, ó $SAR(1)$. Patresi y Salvati(2008) asumen que el vector de efectos de área $(u_1, \dots, u_m)'$ siguen un proceso $SAR(1)$ con varianza σ_u^2 , autocorrelación espacial ρ y matriz de proximidad $W = w_{d,e}$ con $d = 1, \dots, m$, es decir (Esteban et al, 2016):

$$u_d = \rho \sum_{e \neq d} w_{ed} u_e + \varepsilon_d \quad (10)$$

con $|\rho| < 1$, $\varepsilon_d \sim N(0, \sigma_u^2)$ y $d = 1, \dots, m$

La matriz W es obtenida de la matriz de proximidades original W^0 , cuyos elementos de la diagonal son iguales a 0 y los elementos restantes son iguales a 1 cuando las dos áreas correspondientes a la fila y los índices de la columna son considerados vecinos y 0 en otro caso (Esteban et al, 2016).

Por otro lado, los datos históricos ofrecen información que puede ser usada para mejorar significativamente los estimadores en el sentido que permiten observar su comportamiento en instantes anteriores. Choudry y Rao (1989) extendieron el modelo básico de Fay Herriot incluyendo varios instantes de tiempo y considerando una estructura autocorrelacionada para los errores de muestreo. El modelo considerado se muestra a continuación:

$$y_{dt} = X'_{dt} \beta + u_d + e_{dt} \quad (11)$$

con $d = 1, \dots, m$, y $t = 1, \dots, T$, donde y_{dt} y x_{dt} son, la variable respuesta y el vector de variables auxiliares para el área d y el instante de tiempo t , con $\mu_{dt} = X'_{dt} \beta + u_d$ siendo la característica objetivo para la misma área e intervalo de tiempo. Para cada dominio d , se asume que los errores $[e_{dt}]_{t=1}^T$ siguen un proceso autorregresivo de orden 1, $AR(1)$ (Choudry y Rao, 1989), esto es:

$$e_{dt} = \rho e_{dt-1} + \varepsilon_{dt} \quad (12)$$

con $|\rho| < 1$ y $\varepsilon_{dt} \sim N(0, \sigma_e^2)$

Este modelo no permite variaciones de tiempo en las características del área μ_{dt} que no estén explicadas por las variables auxiliares. Por este motivo, Esteban et al (2016) consideraron el modelo:

$$y_{dt} = X'_{dt} \beta + u_{dt} + e_{dt}, \quad (13)$$

con $d = 1, \dots, m$, $t = 1, \dots, T$, donde, para cada dominio d , los efectos aleatorios $[u_{dt}]_{t=1}^T$ se asumen que siguen un proceso estocástico $AR(1)$ y los efectos aleatorios $[e_{dt}]_{t=1}^T$ son i.i.d $N(0, \sigma_e^2)$

2.7. EBLUP y MSE del modelo Fay-Herriot espacio temporal

Se proponen los estimadores EBLUP dado que en el campo real, no es posible conocer los valores y_s de los datos. Una fórmula adaptada por Esteban et al. (2016) de los modelos de Prasad y Rao (1990) para el MSE de $\tilde{y}_{dt}^{eblop}$ con $y_{dt} = a'y$ es

$$MSE(\tilde{y}_{dt}^{eblop}) = g_1(\theta) + g_2(\theta) + g_3(\theta) \quad (14)$$

Donde $\theta = (\sigma_1^2, \sigma_2^2, \theta)$

$$g_1(\theta) = a' Z T Z' a \quad (15)$$

$$g_2(\theta) = [a' X - a' Z T Z' V_e^{-1} X] Q [X' a - X' V_e^{-1} Z T Z' a] \quad (16)$$

$$g_3(\theta) \simeq tr[(\bar{V} b') V (\bar{V} b')' E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)']] \quad (17)$$

y $Q = (X' V^{-1} X)^{-1}$, $T = V_u - V_u Z' Z V_u$ y $b' = a' Z V_u Z' V^{-1}$, el $MSE(\bar{y}_{dt}^{eblup})$ puede ser estimado por el estimador:

$$mse_{dt}^0 = mse(\bar{y}_{dt}^{eblup}) = g_1(\hat{\theta}) + g_2(\hat{\theta}) + 2g_3(\hat{\theta}) \quad (18)$$

2.8. Condiciones de regularidad

Singh et al (2005), establecen que el modelo Fay-Herriot espacio-temporal tiene algunas condiciones de regularidad que se deben cumplir, las cuales se citan a continuación:

- los elementos de $X, t = 1, 2, \dots, T$ están uniformemente distribuidos tal que $X_t^T \Sigma_t^{-1}(\Psi) X_t = [O(m)]_{p \times p}$, donde $\Sigma_t(\Psi) = \sigma_v^2 A^{-1}(\Psi) + R_t$, Σ y A son funciones de Ψ , R es una matriz diagonal de orden m que puede ser expresada como $R = diag(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)$ donde los σ_i^2 s son las varianzas muestrales conocidas de cada área.
- m y T son finitos
- $\hat{\Psi}$ es el estimador de Ψ que satisface $\hat{\Psi} - \Psi = O_p(m^{-1/2})$, $\hat{\Psi}(-y) = \hat{\Psi}(y)$ para todo $h \in R^p$ y para todo y .

3. Resultados

3.1. Descripción del problema

El examen Saber-11 es una prueba que realiza el estado colombiano a todos los estudiantes que cursan el último grado de la educación media. Esta prueba es un referente estatal para la evaluación de la calidad de la educación y la creación de políticas públicas que mejoren la educación media. Adicionalmente, los resultados del examen Saber-11 sirven a los colegios y estudiantes para evaluar su desempeño y en algunos casos como incentivo para ingresar a la educación superior. El carácter periódico del Saber-11, permite observar los avances temporales y establecer el impacto de los programas y medidas tomadas por instituciones, alcaldías y el mismo estado colombiano en cada una de las competencias evaluadas.

Las principales dificultades del desarrollo y posterior análisis de esta prueba para la toma de decisiones son el presupuesto y logística requerida para la aplicación del examen, y el tiempo que requiere el procesamiento de los datos recolectados, lo que implica que los resultados sean costosos y tardan un tiempo considerable desde la fase de recolección hasta la toma de decisiones.

Un modelo que estime los resultados desagregados por departamento de las pruebas Saber-11 con una muestra representativa, podría ayudar a obtener resultados de la prueba más rápidamente para tomar decisiones más rápidas con respecto a las políticas educativas departamentales, así mismo, puede ser un primer paso para la reestructuración de la prueba, dado que permite visualizar alternativas más rápidas y económicas para la aplicación de la prueba.

3.2. Resultados de la prueba de matemáticas

Las bases de datos a tener en consideración son las correspondientes a los resultados de las pruebas Saber 11 para colegios de calendarios A de Colombia en los años 2015 a 2018, estos años comprenden los periodos en los cuales la prueba ya estaba estandarizada. Las variables a consideración para realizar el modelo los puntajes correspondientes a las pruebas descritas a continuación:

Prueba de Matemáticas: evalúa las competencias de los estudiantes para enfrentar situaciones que pueden resolverse con el uso de herramientas matemáticas. (ICFES, 2018)

Prueba de Lectura crítica: evalúa el conjunto de habilidades para comprender, interpretar y evaluar textos que pueden encontrarse en la vida cotidiana y en ámbitos académicos no especializados. (ICFES, 2018)

Prueba de Sociales y Ciudadanas: evalúa conocimientos y habilidades del estudiante que le permiten comprender el mundo social desde la perspectiva propia de las ciencias sociales y situar esta comprensión como referente del ejercicio de su papel como ciudadano. (ICFES, 2018)

Prueba de Ciencias Naturales: evalúa las competencias en las áreas de química, biología, física, explicación de fenómenos, indagación, uso comprensivo del conocimiento científico y ciencia, tecnología y sociedad

Prueba de Inglés: evalúa la competencia comunicativa en lengua inglesa del estudiante de Saber 11, con el fin de dar cuenta de lo que es capaz de hacer, según lo expresado en el marco común Europeo, a través de tareas de lectura, gramática y léxico. (ICFES, 2018)

las bases de 2015 y 2016 contienen 452.248 y 548.110 registros de resultados de estudiantes, la base de datos del año 2017 contiene 552.088 registros, la base del año 2018 contiene 563.678 registros, correspondientes a un total de 4.679 colegios en todos los periodos. Para construir el modelo Fay-Herriot, se tomará, como información auxiliar, los resultados de los departamentos de los años 2015-2017. Nótese que el número de estudiantes por departamentos no son los mismos en ambos años, sin embargo, esto no es una restricción para generar el modelo puesto que aquí el objetivo es el resultado del departamento.

A continuación, se muestran un mapa con los resultados agregados por departamento en los años 2017 y 2018 y una tabla que contiene los mismo resultados. En el mapa (Figura 1) se puede observar, claramente la correlación espacial que hay entre algunos departamentos. En primera instancia, se realizarán las estimaciones según un modelo de Fay-Herriot, seguido a esto, se realizarán usando los componentes espacial y espacio-temporal.

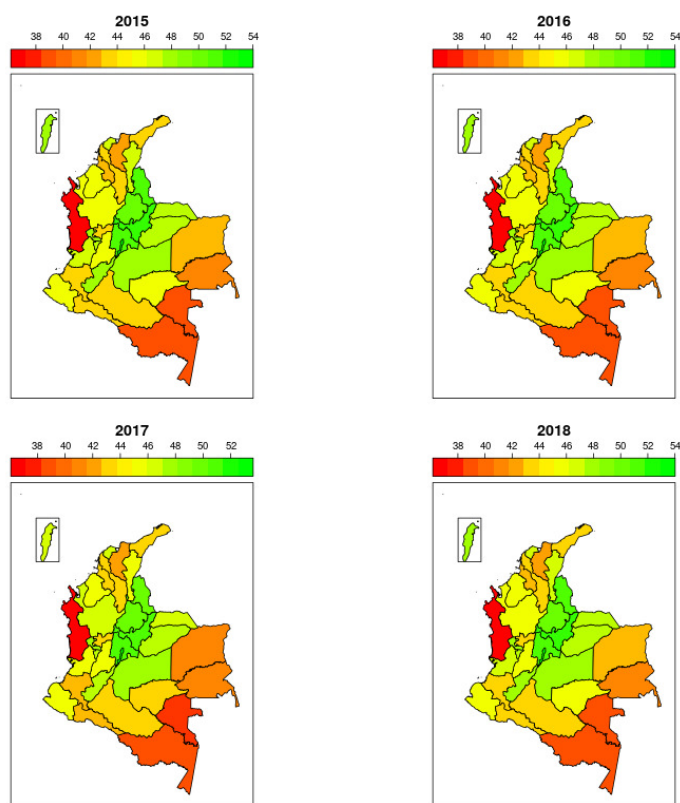


Figura 1: Resultados históricos módulo de matemáticas

Se observa que existe una correlación en los promedios obtenidos en cada departamento, esto por la similitud entre los colores entre los resultados de cada departamento en cada una de los años estudiados. Los departamentos de Chocó y la región sur y oriente del país (Amazonas, Vaupés, Guainía y Vichada) obtienen los resultados promedio más bajos

3.3. Correlación entre variables

El modelo Fay Herriot y el Fay Herriot espacio temporal obtienen estimaciones más precisas en el caso donde las covariables estén altamente correlacionadas, como primer paso, es necesario analizar la correlación espacial y temporal entre los resultados de cada una de las pruebas, con el fin de establecer la pertinencia en la elección de covariables.

El correlograma que muestra la correlación entre la prueba de matemáticas y las covariables se presenta a continuación:

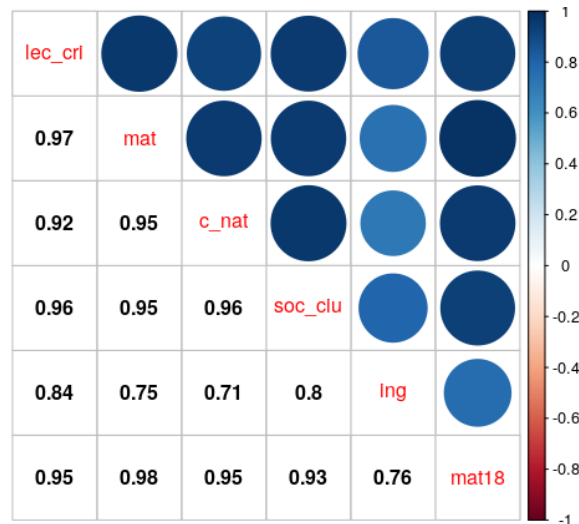


Figura 2: Correlograma entre los módulos

Como lo muestra la Figura 2, la correlación entre los resultados del módulo de matemáticas y los otros módulos es alta en todos los años, superando el 0.9 en casi todos los casos. La correlación con la prueba de inglés es la más baja, sin embargo, es mayor a 0.7. Se realizó adicionalmente un análisis de la correlación entre cada periodo de la prueba de matemáticas, con el fin de identificar si existe una correlación temporal entre los resultados promedio obtenidos por cada departamento en la prueba de matemáticas, los resultados se muestran a continuación:

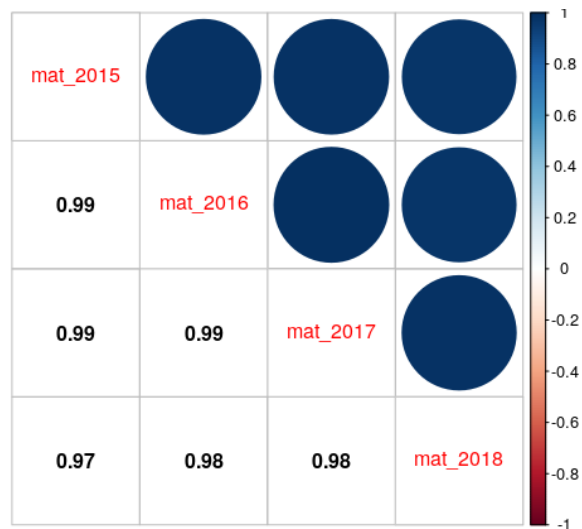


Figura 3: Correlograma de los resultados módulo matemáticas

Se observa que los resultados promedio obtenidos por cada departamento en el módulo de matemáticas son altamente correlacionados de acuerdo al año, lo que muestra una alta correlación temporal

El test de auto correlación espacial de Moran (1950) mide la auto correlación de la posible variable de respuesta en función de las ubicaciones de los departamentos, con hipótesis nula: existe aleatoriedad espacial (No existe correlación espacial). Se evaluó el test de Moran para los datos de respuesta con un nivel de insignificancia de 5 %, los resultados obtenidos fueron:

Moran I statistic standard deviate = 3.6735, p-value = 0.0001196

Esto indica que a un nivel de significancia del 5%, se rechazó la hipótesis nula de aleatoriedad en las vecindades para la Prueba de Matemáticas 2018 Calendario A. Por lo que se concluye que existe correlación espacial

3.4. Diseño de muestreo

Dado que las características de la estimación por Fay Herriot no requieren un diseño de muestreo, sino que se basa en las estimaciones obtenidas por las variables auxiliares para la estimación de los parámetros, se propone estudiar los resultados con un diseño estratificado, en el cual, se escoja un porcentaje representativo de cada departamento.

El diseño propuesto entonces será tomar una muestra representativa de acuerdo al número de colegios presentes en cada departamento, de acuerdo a la fórmula de afijación proporcional, que consiste en asignar a cada estrato un número de unidades muestrales proporcional al tamaño. Se toma un tamaño de muestra de 1.000 colegios (21.37% del total de colegios). La siguiente tabla muestra la muestra a considerar en cada departamento

Departamento	Total	Muestra	Departamento	Total	Muestra
Antioquia	499	107	Norte de Santander	105	22
Atlántico	203	43	Quindío	43	9
Santafé de Bogotá	499	107	Risaralda	75	16
Bolívar	238	51	Santander	259	55
Boyacá	150	32	Sucre	109	23
Caldas	85	18	Tolima	168	36
Caquetá	48	10	Valle del Cauca	363	78
Cauca	232	50	Arauca	25	5
Cesar	87	19	Casanare	79	17
Córdoba	204	44	Putumayo	59	13
Cundinamarca	420	90	San Andrés y Providencia	5	1
Chocó	36	8	Amazonas	10	2
Huila	162	35	Guainía	4	1
La Guajira	61	13	Guaviare	23	5
Magdalena	152	32	Vaupés	9	2
Meta	112	24	Vichada	8	2
Nariño	147	31			

Tabla 1: Distribución de muestra por departamento

Para los departamentos de San Andrés y Guainía la muestra seleccionada es de 1 colegio, en este caso, se tomarán estos dos departamentos como áreas con No respuesta para ser estimados posteriormente por medio del modelo Fay Herriot con ayuda de las variables auxiliares.

3.5. Estimación de los resultados por HORVITZ-THOMPSON

La figura 4 muestra los resultados de las estimaciones del diseño por Horvitz thompson de los resultados del módulo de matemáticas para el año 2018.

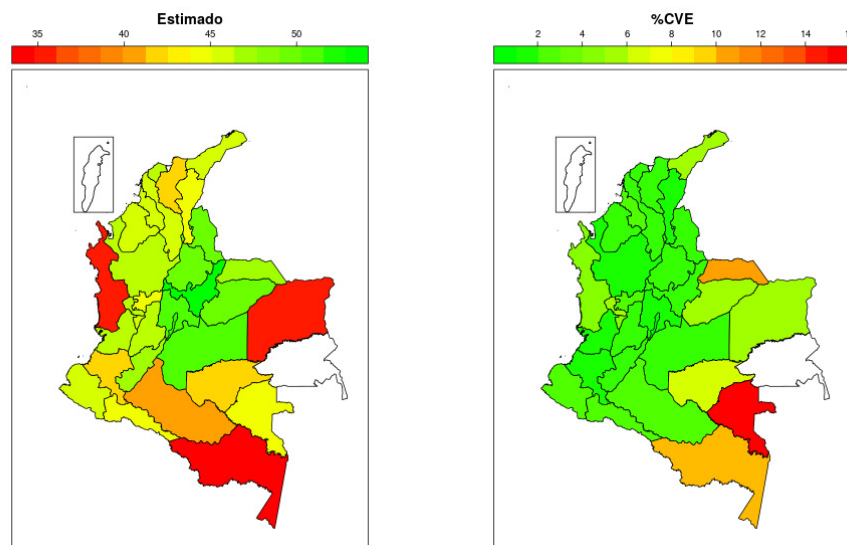


Figura 4: Estimación por Horvitz-Thompson del promedio del modulo de Matemáticas con %cve

Se observa los resultados faltantes para los departamentos de San Andrés y Guainía, dado que el estimador de Horvitz Thompson no considera la estimación en la no respuesta, también se observan coeficientes de variación mayores a 10 % en algunos departamentos, lo que demuestra que el diseño tiene estimaciones con una variación considerablemente alta en los departamentos donde se tomaron muestras pequeñas. La siguiente tabla muestra un comparativo entre los resultados estimados y los reales

Departamento	Puntaje Estimado	%CVE	Puntaje Real	Diferencia Absoluta	Departamento	Puntaje Estimado	%CVE	Puntaje Real	Diferencia Absoluta
Antioquia	47,03	1,51	45,95	1,09	Atlántico	46,54	2,31	46,55	0,00
Bogotá	52,57	1,02	52,92	0,35	Bolívar	46,19	2,80	44,80	1,39
Boyacá	52,88	2,93	51,88	0,99	Caldas	43,95	3,91	44,93	0,98
Caquetá	41,20	3,50	44,92	3,72	Cauca	43,81	1,83	44,01	0,20
Cesar	44,25	1,82	46,52	2,27	Córdoba	45,84	2,34	45,49	0,35
Cundinamarca	51,22	1,64	51,05	0,16	Chocó	35,54	4,47	37,25	1,71
Huila	48,77	2,18	48,71	0,06	La Guajira	46,02	5,75	44,08	1,94
Magdalena	42,77	2,19	42,64	0,13	Meta	50,66	2,77	49,04	1,62
Nariño	45,57	3,03	45,84	0,26	Norte de Santander	49,30	2,69	51,40	2,10
Quindío	45,96	4,63	47,78	1,82	Risaralda	47,12	5,36	48,48	1,36
Santander	50,12	2,40	50,21	0,09	Sucre	46,35	3,52	43,62	2,73
Tolima	46,74	2,23	46,02	0,72	Valle del Cauca	46,82	1,71	46,50	0,32
Arauca	48,75	11,16	48,26	0,49	Casanare	49,09	5,43	48,09	0,99
Putumayo	44,74	3,06	44,14	0,61	San Andrés y Providencia			48,86	
Amazonas	34,76	10,40	39,55	4,79	Guainía			41,53	
Guaviare	43,60	7,97	45,99	2,39	Vaupés	43,88	15,02	39,79	4,09
Vichada	35,26	5,41	43,50	8,24					

Tabla 2: Comparación Estimación por HT vs Real

Para los departamentos de Caquetá, Amazonas, Vaupés y Vichada se encontraron diferencias entre el promedio estimado y el puntaje real de más de tres puntos. Adicionalmente, la tabla 2 muestra que los departamentos de Arauca, Amazonas y Vaupés tienen un coeficiente de variación estimado mayor a 10 %.

3.6. Estimación de los resultados por modelo de Fay-Herriot

Para estimar los resultados de matemáticas del año 2018 en los distintos departamentos, usando los resultados de la prueba Saber-11 para el año 2017 utilizaremos el siguiente modelo:

$$\text{Matematicas} = \beta_1 \text{Sociales} + \beta_2 \text{Ciencias} + \beta_3 \text{Lectura} + \beta_4 \text{Ingles} + e_i \quad (19)$$

Para el caso de los diseños basados en modelos, una aproximación al CVE es la tasa de error estimada (EER por sus siglas en inglés), propuesta por Zhu et al (2003) como una medida para indicar el cambio esperado de un modelo. Y su formula es $EER = \frac{\sqrt{mse}}{\bar{y}}$

La siguiente imagen muestra las estimaciones de la prueba por medio del modelo de Fay-Herriot

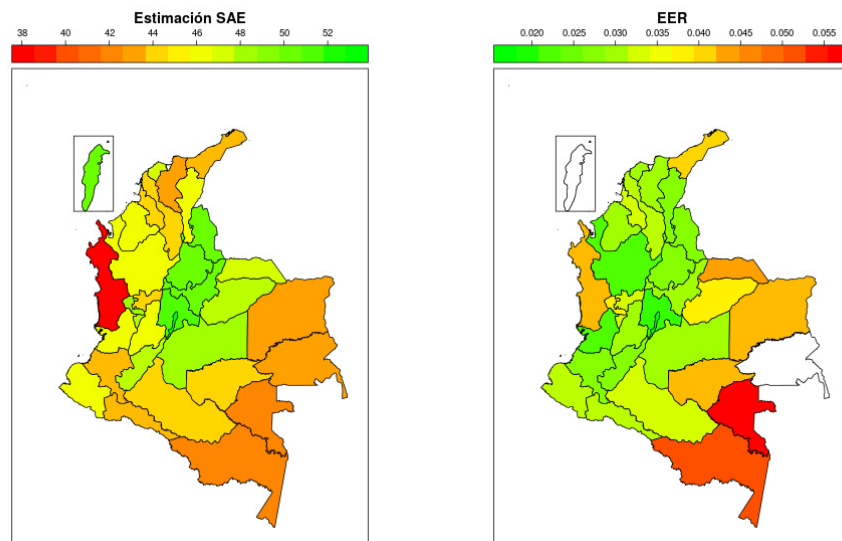


Figura 5: Estimación por Fay-Herriot con EER

Se observa una estimación para los departamentos con no respuesta, adicionalmente, los EER, que son equivalentes al CVE, no superan en ningún caso el 6 % de variación. A continuación se detallan las diferencias entre el puntaje estimado por Fay-Herriot y el puntaje real desagregado por departamento.

Departamento	Puntaje Esti- mado	%EER	Puntaje Real	Diferencia Absolu- ta	Departamento	Puntaje Esti- mado	%EER	Puntaje Real	Diferencia Absolu- ta
Antioquia	46,39	1,28	45,95	0,44	Norte de Santander	49,93	2,04	51,40	1,47
Atlántico	46,74	1,81	46,55	0,20	Quindío	48,39	2,98	47,78	0,60
Bogotá	52,85	0,91	52,92	0,07	Risaralda	49,24	3,28	48,48	0,76
Bolívar	45,29	2,10	44,80	0,50	Santander	49,92	1,87	50,21	0,28
Boyacá	50,74	2,17	51,88	1,14	Sucre	45,28	2,48	43,62	1,66
Caldas	45,45	2,67	44,93	0,52	Tolima	46,10	1,76	46,02	0,08
Caquetá	45,47	2,47	44,92	0,56	Valle del Cauca	46,69	1,42	46,50	0,18
Cauca	44,17	1,50	44,01	0,16	Arauca	47,00	4,79	48,26	1,26
Cesar	46,22	1,50	46,52	0,30	Casanare	47,82	3,30	48,09	0,28
Córdoba	45,90	1,83	45,49	0,41	Putumayo	44,14	2,24	44,14	0,01
Cundinamarca	50,93	1,37	51,05	0,12	San Andrés y Providencia	50,05		48,86	1,19
Chocó	38,54	2,93	37,25	1,30	Amazonas	42,07	4,65	39,55	2,52
Huila	48,41	1,73	48,71	0,30	Guainía	43,08		41,53	1,55
La Guajira	44,60	3,42	44,08	0,52	Guaviare	45,18	4,09	45,99	0,81
Magdalena	43,61	1,74	42,64	0,97	Vaupés	41,83	5,39	39,79	2,04
Meta	49,21	2,08	49,04	0,17	Vichada	42,91	3,30	43,50	0,59
Nariño	46,04	2,23	45,84	0,20					

Tabla 3: Comparación Estimación por Fay-Herriot vs Real

En la Tabla 3 se puede observar que todas las estimaciones de los promedios del módulo de matemáticas estan por debajo de las tres unidades, incluyendo los departamentos de no respuesta. Las estimaciones por medio del modelo de Fay Herriot resultan ser mas precisas y con menos variación estimada que las estimaciones por Horvitz-Thompson.

3.7. Estimación de los resultados por Fay-Herriot espacio-temporal

Finalmente, se estimaron los resultados para el módulo de matemáticas del año 2018 basados en los resultados de las otras pruebas en los años 2015 a 2017 teniendo en cuenta la matriz de distancias entre los centroides de las capitales de cada departamento. Esto a través de un modelo Fay Herriot espacio-temporal con factor temporal $AR(1)$. Los resultados de las estimaciones se muestran a continuación:

$$Matematicas = \beta_1 Sociales + \beta_2 Ciencias + \beta_3 Lectura + \beta_4 Ingles + AR(1) + e_i + u_i \quad (20)$$

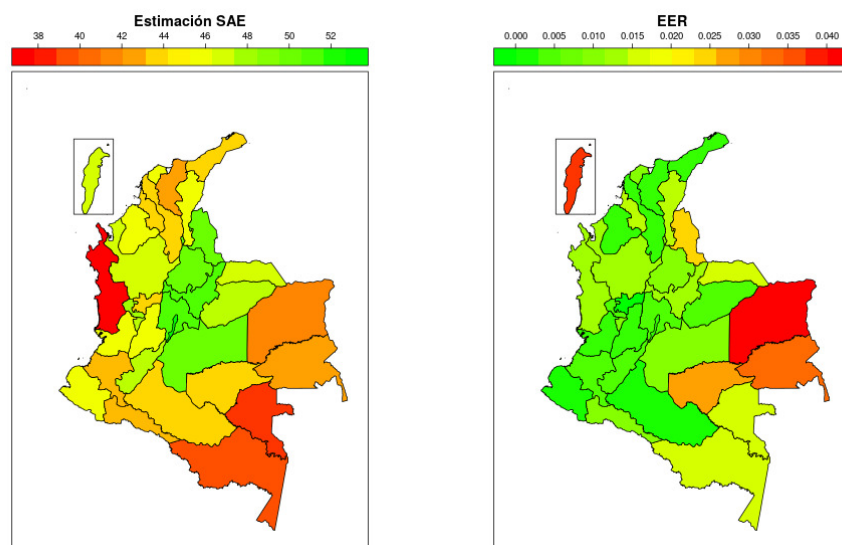


Figura 6: Estimación por Fay-Herriot ST con EER

Se observa que el máximo EER obtenido es del 4%, lo que muestra que la variación estimada por el modelo es bastante baja, las estimaciones del modelo también fueron más exactas comparadas con las estimaciones obtenidas por Horvitz-Thompson y Fay-Herriot. La siguiente tabla muestra los resultados desagregados por departamento.

Departamento	Puntaje Estimado	%EER	Puntaje Real	Diferencia Absoluta	Departamento	Puntaje Estimado	%EER	Puntaje Real	Diferencia Absoluta
Antioquia	46,58	1,36	45,95	0,63	Norte de Santander	50,23	2,32	51,40	1,16
Atlántico	46,12	0,92	46,55	0,42	Quindío	48,24	0,94	47,78	0,45
Bogotá	52,73	0,36	52,92	0,19	Risaralda	48,87	0,80	48,48	0,39
Bolívar	45,02	0,49	44,80	0,22	Santander	49,71	1,00	50,21	0,49
Boyacá	51,28	1,17	51,88	0,60	Sucre	44,35	1,64	43,62	0,73
Caldas	44,93	0,00	44,93	0,00	Tolima	45,88	0,30	46,02	0,14
Caquetá	44,91	0,01	44,92	0,00	Valle del Cauca	46,30	0,42	46,50	0,20
Cauca	43,68	0,75	44,01	0,33	Arauca	47,33	1,97	48,26	0,93
Cesar	45,85	1,47	46,52	0,67	Casanare	47,75	0,73	48,09	0,35
Córdoba	45,27	0,48	45,49	0,22	Putumayo	43,70	0,99	44,14	0,43
Cundinamarca	50,73	0,63	51,05	0,32	San Andrés y Providencia	47,00	3,95	48,86	1,86
Chocó	37,79	1,43	37,25	0,54	Amazonas	40,28	1,80	39,55	0,72
Huila	48,39	0,65	48,71	0,32	Guainía	42,99	3,40	41,53	1,46
La Guajira	44,24	0,36	44,08	0,16	Guaviare	44,64	3,02	45,99	1,35
Magdalena	42,89	0,58	42,64	0,25	Vaupés	39,06	1,87	39,79	0,73
Meta	49,54	1,01	49,04	0,50	Vichada	41,82	4,00	43,50	1,68
Nariño	45,78	0,12	45,84	0,06					

Tabla 4: Comparación Estimación por Fay-Herriot ST vs Real

En la tabla 4 se puede observar que las estimaciones para cada departamento son mucho más precisas cuando se agrega el componente espacio-temporal al modelo, esto debido a la alta correlación mostrada en las Figuras 2 y 3. También se puede observar la que variación estimada es menor al 5 % para cualquier departamento seleccionado. Estos factores convierten al modelo Fay Herriot espacio-temporal como el más preciso de los analizados en este trabajo.

4. Conclusiones

Este trabajo muestra que gracias a las fuertes correlaciones entre los resultados de las diferentes competencias evaluadas por la prueba Saber-11 en los años 2015 a 2017, con una muestra probabilística, es posible estimar de forma precisa los resultados de la prueba saber-11 basados en los resultados de las pruebas anteriores de las instituciones seleccionadas.

Las estimaciones a través del modelo de Horvitz-Thompson y del modelo de Fay-Herriot fueron precisas para los departamentos grandes, sin embargo, estas estimaciones no obtienen las mejores aproximaciones en departamentos donde la muestra de colegios es pequeña. Agregar el componente espacio-temporal al modelo Fay-Herriot mejora notablemente las estimaciones en todos los casos.

Adicionalmente, se observa como los modelos Fay-Herriot y Fay-Herriot espacio temporal mejoran las estimaciones de un diseño de muestreo estratificado añadiendo información auxiliar.

Existen otros diseños basados en modelos diferentes al Fay-Herriot para estimar resultados desagregados por áreas, es indicado plantearse en trabajos futuros la evaluación de este trabajo comparada con otros modelos, como el planteado por Battese, Harter y Fuller (1988) y verificar la pertinencia de otros posibles modelos, así mismo, se podría evaluar los factores espacial y temporal con modelos diferentes al SAR y AR(1) propuestos por Esteban et al (2016).

Este es un trabajo con intenciones académicas, sin embargo, existen pruebas estatales desarrolladas por el ICFES que son de carácter muestral, como el SABER 3,5, y 9, además de otras como las pruebas PISA, para las cuales la aplicación de estas metodologías resulta también pertinente y son una posible opción para realizar estimaciones de áreas precisas.

5. Agradecimientos

En primer y más importante lugar, agradezco a mis padres Wilson Gutiérrez y Yury Pérez, porque su constante e incondicional apoyo permitieron realizarme como profesional y como persona. Sus valores y consejos me han guiado en cada momento de mi vida. Agradezco profundamente su ayuda sentimental y económica en los momentos más difíciles.

A mi tutor Wilmer Pineda, su constante orientación, su paciencia y sus conocimientos me ayudaron a llevar este trabajo hasta su finalización. Un agradecimiento especial al profesor Cristian Tellez, sus enseñanzas en la clase de muestreo 2 y el trabajo que presentamos en el VI IWAS fueron de gran ayuda para darle orientación a este trabajo.

A todos mis profesores, y en especial a los profesores Yesid Rodríguez, Alex Zambrano, Deisy Camargo, José Zea y Dagoberto Bermudez. Son un ejemplo como profesores, profesionales y personas éticas, muchas gracias por su paciencia y sobre todo por las invaluable lecciones dentro y fuera del aula de clases.

Finalmente, a mi amiga Eva Salcedo, que estuvo en todo momento siempre presente en infinitas horas de estudio, traspasadas, trabajos, tareas, eventos, publicaciones, comunicaciones, momentos de reflexión y celebraciones, Su apoyo y consejo desde el primer día de clases y en todo momento fue la mejor de las tutorías, no existe ningún premio que demuestre la excelente persona que es para mí. A mis hermanos Wilson Gutiérrez y Paula Gutiérrez y a mis amigos Viviana Castro y David Aroca, que siempre me brindaron su apoyo en los momentos difíciles y también en los más fáciles.

Referencias

- Auguie, B. (2017), *gridExtra: Miscellaneous Functions for "Grid" Graphics*. R package version 2.3.
*<https://CRAN.R-project.org/package=gridExtra>
- Batesse, G.E., H. R. & Fuller, W. (1988), 'An error-components model for prediction of county crop areas using survey and satellite data', *Journal of the American Statistical Association* **83**(1), 28–36.
- Bharat Bhushan Singh, S. & Kundu (2005), 'Spatio-temporal models in small area estimation', *Survey Methodology* **31**(2), 183–195.
- Corcuera, J. (2002), *Técnicas de estimación en áreas pequeñas*, Eustat.
- Cressie, N. (1993), *Statistics for Spatial Data*, John Wiley and Sons, Inc.
- Esteban, M.D., M. D. P. A. S. L. (2012), 'Small area estimation of poverty proportions under area-level time models.', *Journal of the American Statistical Association* **56**(1), 2840–2855.
- Fay, R. & Herriot, R. (1979), 'Estimation of income from small places: An application of james-stein procedures to census data.', *Journal of the American Statistical Association* **74**(1), 269–277.
- Gutierrez, Esteban y Tellez, C. (2019), *UNA APLICACION DEL MODELO FAY HERRIOT PARA LA ESTIMACION DE LA MEDIA DE LOS RESULTADOS POR DEPARTAMENTO DEL MODULO DE MATEMATICAS PARA LA PRUEBA SABER 11 DEL SEGUNDO PERIODO DEL 2018*. Presentado en el VI IWAS. Bogota Colombia.
*Trabajo pendiente de publicacion
- ICFES (2018), *Tomado de www.ism.ac.jp*, ICFES.
- Marhuenda, Y., M. I. M. D. (2013), 'Small area estimation with spatio temporal fay herriot models', *Computational Statistics and Data Analysis* **58**(1), 308–325.
- Molina, I. & Marhuenda, Y. (2015), 'sae: An R package for small area estimation', *The R Journal* **7**(1), 81–98.
*<https://journal.r-project.org/archive/2015/RJ-2015-007/RJ-2015-007.pdf>
- Molina, I., S. N. P. M. (2009), 'Bootstrap for estimating the mse of the spatial eblup', *Computational Statistics* **24**(1), 441–458.
- Morales, D. (2015), *Estimación en áreas pequeñas: Métodos basados en Modelos*, Universidad Miguel Hernandez de Elche.
- Moran, P. (1950), 'Notes on continuous stochastic phenomena', *Biometrika* **37**(1), 17–23.
- Pratesi, M., S. N. (2008), 'Small area estimation: the eblup estimator based on spatially correlated random area effects.', *Statistical Methods and Applications* **17**(1), 113–141.
- Rao, J. (2003), *Small Area Estimation.*, New York: John Wiley and Sons.
- Roger S. Bivand, Edzer Pebesma, V. G. R. (2013), *Applied spatial data analysis with R, Second edition*, Springer.
- Rojas, H. A. G. (2018), *TeachingSampling: Selection of Samples and Parameter Estimation in Finite Population*. R package version 3.4.2.
*<https://CRAN.R-project.org/package=TeachingSampling>
- Sarndal, C.-E., Swensson, B. & Wretman, J. (1992), *Model Assisted Survey Sampling*, Springer.
- Wei, T. & Simko, V. (2017), *R package corrplot: Visualization of a Correlation Matrix*. (Version 0.84).
- Wickham, H. (2017), *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1.
*<https://CRAN.R-project.org/package=tidyverse>