APLIKASI SMALL AREA ESTIMATION (SAE) METODE PSEUDO-EBLUP PADA OFFICIAL STATISTICS DI INDONESIA

Studi Kasus: Estimasi Pengeluaran Rumah Tangga di Provinsi Jawa Timur Tahun 2016

Application of Small Area Estimation (SAE) Method Pseudo-EBLUP in Official Statistics of Indonesia

Deo Dedianto¹, Ika Yuni Wulansari²

Politeknik Statistika STIS^{1,2}, Jakarta 13330 Email: ¹deodedianto08@gmail.com

Abstrak

Small Area Estimation (SAE) merupakan metode estimasi yang berbasis pemodelan. Dalam praktiknya, SAE yang sering digunakan adalah model level area, karena ketersediaan variabel pendukung hanya tersedia pada level area. Salah satu metode estimasi area kecil adalah Empirical Best Linear Unbiased Prediction (EBLUP). Ketika estimasi EBLUP tidak menyertakan penimbang sampling yang tepat dan spesifik, maka diasumsikan pengambilan sampel dilakukan secara Simple Random Sampling (SRS) yang mana mengasumsikan bahwa setiap unit memiliki peluang yang sama untuk terpilih sebagai sampel. Akan tetapi, desain penarikan sampel Probability Proportional to Size Sampling (PPS) cenderung lebih baik dan efisien digunakan dibandingkan SRS karena setiap unit memiliki peluang yang berbeda untuk terpilih sebagai sampel, tergantung ukuran (size) yang digunakan. Salah satu indikator/ statistik ofisial yang sering menjadi perhatian adalah pengeluaran rumah tangga sebagai indikator proksi dalam pengukuran kemiskinan rumah tangga. Penelitian ini menerapkan metode Pseudo-EBLUP untuk estimasi pengeluaran rumah tangga di Provinsi Jawa Timur tahun 2016. Berdasarkan hasil penelitian, Pseudo EBLUP dapat mengakomodir perbedaan variasi antar area sehingga dihasilkan estimator yang lebih robust dalam estimasi pengeluaran rumah tangga dibandingkan metode EBLUP. Di samping itu, metode Pseudo EBLUP-PPS dengan memasukkan sampling weight pada estimasi langsung memberikan hasil yang lebih baik daripada Pseudo EBLUP-SRS karena menghasilkan estimator yang lebih efisien.

Kata kunci: EBLUP, Pseudo-EBLUP, Sampling Weight, Small Area Estimation (SAE)

Abstract

Small Area Estimation (SAE) is a modeling-based estimation method. In practice, SAE that is often used is a level area model, because the availability of auxiliary variables is only available at area level. One of small area estimation method is Empirical Best Linear Unbiased Prediction (EBLUP). When the EBLUP estimation does not include an appropriate and specific sampling weight, it is assumed that the sample is selected by Simple Random Sampling (SRS) which assumes that each unit has the same probability to be selected as a sample. However, the design of Proportional Proportion to Size Sampling (PPS) sampling tends to be better and more efficient to use than SRS because each unit has different probability to be selected as a sample, depending on the size used. One of important indicator / official statistics is household expenditure as a proxy indicator of household poverty measurement. This study using Pseudo-EBLUP method to estimate household expenditure in East Java Province 2016. Based on the results of the study, Pseudo EBLUP accommodate differences in variation between areas so it produced more robust estimator to estimate household expenditure compared to the EBLUP method. In addition, Pseudo EBLUP-PPS by including sampling weight in direct estimation provides better results than Pseudo EBLUP-SRS because it produces a more efficient estimator.

Keywords: EBLUP, Pseudo-EBLUP, Sampling Weight, Small Area Estimation (SAE)

PENDAHULUAN

Small Area Estimation (SAE) merupakan metode estimasi yang berbasis pemodelan. SAE adalah metode estimasi pada area atau level dengan meminjam kekuatan variabel penyerta untuk meningkatkan efektifitas sample size (Rao, 2003). Variabel yang akan dijadikan sebagai variabel penyerta diharapkan merupakan nilai populasi dari karakteristik yang akan diestimasi. Jenis SAE ada 2 (dua), yaitu model unit level dan area level. Untuk area level menggunakan model dasar dari Fay-Harriot yaitu EBLUP. Selama ini, penggunaan estimasi EBLUP mengabaikan sampling weight yang tepat. Ketika estimasi EBLUP tidak menyertakan sampling weight yang tepat, diasumsikan pengambilan dilakukan secara SRS. Pengambilan sampel ini mengasumsikan bahwa setiap unit sampling yang terpilih memiliki peluang yang sama.

Sampling design PPS cenderung lebih banyak dipilih dibandingkan SRS karena lebih efisien. Zhou dan Jianjian (2012) dalam penelitiannya telah membuktikan bahwa efisiensi PPS sampling lebih tinggi dibanding SRS sampling, bahkan pada kasus sampel kecil. Pada PPS sampling, setiap unit memiliki peluang yang berbeda untuk terpilih sebagai sampel. Untuk menyertakan penimbang sampel dalam SAE maka salah satunya menggunakan pendekatan Pseudo-EBLUP. Hasil estimasi EBLUP dengan Pseudo EBLUP akan berbeda pada koefisien β. Sedangkan, estimasi langsung Pseudo-EBLUP akan lebih baik daripada EBLUP karena estimasi **EBLUP** sudah memenuhi benchmarking property, dengan asumsi $\sum_{i=1}^{m} N_{i} \hat{\mu}_{iw}^{H}$, dimana N adalah jumlah populasi dan $\tilde{\mu}_{iw}^H$ hasil estimasi Pseudo EBLUP, sama dengan regresi estimator pada survey untuk estimasi total yaitu \hat{Y}_w + $(X - \hat{X}_w)^T \hat{\beta}_w$, dimana \hat{Y}_w adalah estimasi total variabel dependen dengan penimbang sampling, \hat{X}_w estimasi total variabel penyerta dengan penimbang sampling dan $\hat{\beta}_w$ estimasi koefisien regresi dengan penimbang sampling (Rao, 2015). Hingga saat ini belum terdapat package yang digunakan untuk melakukan perhitungan Pseudo-EBLUP pada area level. Dalam penelitian ini dilakukan simulasi dengan cara membuat fungsi Pseudo-EBLUP area level pada software R. Estimasi dilakukan pada variabel pengeluaran rumah tangga di Provinsi Jawa Timur tahun 2016. Variabel pengeluaran rumah tangga dipilih karena merupakan salah satu indikator yang sangat penting dalam pengukuran kemiskinan.

METODE

EBLUP Estimator

Estimator BLUP mengasumsikan bahwa sampel diambil secara SRS sehingga tidak menyertakan design weight. Estimator BLUP menggunakan δ_d sebagai true value yang didekati dengan estimasi langsung δ_d^{DIR} . Estimasi langsung memiliki perbedaan sample size pada setiap area sehingga estimator langsung memiliki error yang berbeda di setiap area. Persamaannya sebagai berikut :

$$\hat{\delta}_d^{DIR} = \delta_d + e_d , e_d \stackrel{ind}{\sim} N(0, \sigma_e^2) ,$$

$$d = 1, \dots, D$$
 (1)

Dimana σ_e^2 adalah *sampling variance* dari estimasi langsung dari setiap area d. Selanjutnya δ_d diasumsikan berhubungan linier dengan variabel penyerta, dengan persamaan berikut:

$$\delta_d = \mathbf{x}_d' \boldsymbol{\beta} + v_d , v_d \stackrel{iid}{\sim} N(0, \sigma_v^2) ,$$

$$d = 1, \dots, D$$
(2)

Dimana v_d independen dengan e_d untuk setiap area d. Dari persamaan 1 dan 2 di atas dapat di substitusi sebagai berikut:

$$\hat{\delta}_{d}^{DIR} = \mathbf{x}_{d}^{\prime} \boldsymbol{\beta} + v_{d} + e_{d}$$

$$v_{d} \stackrel{iid}{\sim} N(0, \sigma_{v}^{2}), e_{d} \stackrel{ind}{\sim} N(0, \sigma_{e}^{2}),$$

$$d = 1, \dots, D$$
(3)

Estimasi komponen varians, σ_e^2 dan σ_v^2 didapatkan melalui *fitting-of-constant* metode REML. Dari persamaan 3 untuk mendapatkan estimator BLUP digunakan rumus sebagai berikut :

$$\tilde{\delta}_d = \mathbf{x}_d' \tilde{\beta}(A) - \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left\{ \tilde{\delta}_d^{DIR} - \mathbf{x}_d' \tilde{\beta}(A) \right\} \tag{4}$$

Kemudian untuk mendapatkan β sebagai berikut :

$$\widetilde{\boldsymbol{\beta}}(A) = \left[\sum_{d=1}^{D} (\sigma_e^2 + \sigma_v^2)^{-1} \boldsymbol{x}_d \boldsymbol{x}_d'\right]^{-1} \times \left[\sum_{d=1}^{D} (\sigma_e^2 + \sigma_v^2)^{-1} \boldsymbol{x}_d \delta_d^{DIR}\right]$$
(5)

Pseudo-EBLUP Estimator

Estimator Pseudo-EBLUP menyertakan design weight yang memenuhi sifat design-consistency, yaitu error variansnya sama pada setiap area i. Estimator Pseudo EBLUP memenuhi sifat benchmarking tanpa memerlukan adjustment. Pertama, model design weight pada level area didapatkan sebagai berikut $\widetilde{\boldsymbol{w}}_{ij} = \boldsymbol{w}_{ij} / \sum_{k=1}^{n_i} \boldsymbol{w}_{ij}$. Metode Pseudo-EBLUP berawal dari model unit level sebagai berikut :

$$\begin{split} \tilde{\boldsymbol{y}}_{ij} &= \sum_{j=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij} \boldsymbol{y}_{ij} = \sum_{j=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij} (\boldsymbol{x}_{ij}^T \boldsymbol{\beta} + \boldsymbol{v}_i + \boldsymbol{e}_{ij}) \\ &= \bar{\boldsymbol{x}}_{iw}^T \boldsymbol{\beta} + \bar{\boldsymbol{v}}_{iw} + \bar{\boldsymbol{e}}_{iw} \qquad (6) \\ \text{Dimana} \qquad \tilde{\boldsymbol{x}}_{iw} &= \sum_{k=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij} \, \boldsymbol{x}_{ij}, \qquad \boldsymbol{v}_{iw} = \\ \sum_{k=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij} \, \boldsymbol{v}_i \, \text{dan} \, \boldsymbol{e}_{iw} &= \sum_{k=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij} \, \boldsymbol{e}_{ij}. \end{split}$$
Kemudian dari persamaan 6 untuk mendapatkan estimasi Pseudo-EBLUP, didapatkan estimasi sebagai berikut:

Dimana
$$\boldsymbol{\varphi}_{iw} = \bar{\boldsymbol{x}}_{iw}^T \boldsymbol{\beta} + \gamma_{iw}(\bar{\boldsymbol{y}}_{iw} - \bar{\boldsymbol{x}}_{iw}^T \boldsymbol{\beta})$$
 (7)
$$\begin{aligned}
\widetilde{\boldsymbol{\mu}}_{iw}^H &= \bar{\boldsymbol{x}}_{iw}^T \boldsymbol{\beta} + \gamma_{iw}(\bar{\boldsymbol{y}}_{iw} - \bar{\boldsymbol{x}}_{iw}^T \boldsymbol{\beta}) & \text{(7)} \\
\widetilde{\boldsymbol{\mu}}_{iw} &= \boldsymbol{\sigma}_v^2 \boldsymbol{\varphi}_{iw} / (\boldsymbol{\sigma}_v^2 \boldsymbol{\varphi}_{iw} + \boldsymbol{\sigma}_e^2 \boldsymbol{\varphi}_{iw}) \\
\widetilde{\boldsymbol{\mu}}_{iw} &= \sum_{j=1}^{n_i} \widetilde{\boldsymbol{w}}_{ij}^2.
\end{aligned}$$

Persamaan 7 bergantung pada pada β , σ_e^2 dan σ_v^2 . Untuk mendapatkan penduga β dengan penimbang digunakan persamaan berikut:

$$\widetilde{\boldsymbol{\beta}}_{w}(\sigma_{v}^{2}, \sigma_{e}^{2}) = \left[\sum_{i=1}^{m} \sum_{j=1}^{n_{i}} w_{iw} \boldsymbol{x}_{ij} (\boldsymbol{x}_{ij} - \gamma_{iw} \bar{\boldsymbol{x}}_{iw})^{T} \right]^{-1} \times \left[\sum_{i=1}^{m} \sum_{j=1}^{n_{i}} w_{iw} (\boldsymbol{x}_{ij} - \gamma_{iw} \bar{\boldsymbol{x}}_{iw}) y_{iw} \right]$$

Estimasi MSE Pseudo-EBLUP menggunakan persamaan berikut :

$$mse_B(\hat{\mu}_i^H) \approx g_{1i} + g_{2i} + 2g_{3i}$$
 (9)

Persamaan 9 (Rao, 2015) dapat dihitung menggunakan pendekatan MSE bootstrap dengan persamaan di bawah ini:

$$mse_{B1}(\hat{\mu}_i^H) = B^{-1} \sum_{b=1}^B [\hat{\mu}_{i*}^H(b) - \hat{\mu}_{i*}(b)]^2$$
(10)

Dimana $\hat{\mu}_{i*}^H(b)$ merupakan estimasi bootstrap Pseudo-EBLUP dan $\hat{\mu}_{i*}(b)$ adalah estimasi bootstrap langsung.

Pada penelitian ini persamaan 8 dimodifikasi. Modifikasi dilakukan untuk mengestimasi level area. Variabel penyerta yang digunakan adalah data PODES 2014 dan variabel pengeluaran rumah tangga dari SUSENAS 2016. Variabel pengeluaran rumah tangga dan variabel penyerta telah diagregasi menggunakan sampling weight tiap area sehingga tidak ditimbang lagi saat menghitung $\tilde{\beta}_w$, tetapi menggunakan penimbang γ_{iw} . Selanjutnya, estimasi $\tilde{\beta}_w$ diperoleh dengan persamaan sebagai berikut: Guadarrama Sanz (2017)

$$\widetilde{\boldsymbol{\beta}}_{w}(\sigma_{v}^{2}, \sigma_{e}^{2}) = \left[\sum_{i=1}^{I} \boldsymbol{x}_{i}^{T} \boldsymbol{\gamma}_{iw} \boldsymbol{x}_{i}\right]^{-1} \times \left[\sum_{i=1}^{I} \boldsymbol{x}_{i}^{T} \boldsymbol{\gamma}_{iw} \bar{\boldsymbol{y}}_{iw}\right]$$
(11)

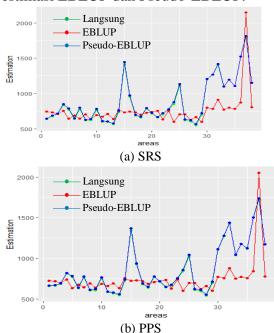
Dimana x adalah variabel penyerta yang diambil dari PODES 2014, dan \bar{y} adalah estimasi langsung dari rata-rata pengeluaran rumah tangga.

HASIL DAN PEMBAHASAN

Variabel diestimasi adalah yang pengeluaran rumah tangga yang diambil dari data hasil survey Badan Pusat Statistik (BPS) yaitu SUSENAS 2016. SUSENAS BPS menggunakan two stages design sampling. Estimasi langsung dilakukan dengan menggunakan penimbang rumah tangga berdasarkan weight SUSENAS. Variabel dependen yang digunakan adalah pengeluaran rumah tangga. Sedangkan, variabel penyerta menggunakan data dari Potensi Desa (PODES) yang memiliki korelasi tinggi dengan variabel pengeluaran rumah tangga, serta menggunakan prosedur model selection dalam model linier.

Estimasi Pengeluaran Rumah Tangga

Dari hasil olah data, diperoleh hasil estimasi pengeluaran rumah tangga dengan membandingkan hasil estimasi langsung SRS dan PPS, EBLUP dan Pseudo-EBLUP. Estimasi langsung SRS dan PPS dapat dijadikan benchmarking, karena estimasi langsung dilakukan pada level kabupaten yang mana sampel telah memenuhi syarat kecukupan sampel untuk estimasi. Kemudian hasil estimasi langsung dibandingkan untuk mengetahui ketepatan estimasi EBLUP dan Pseudo-EBLUP.



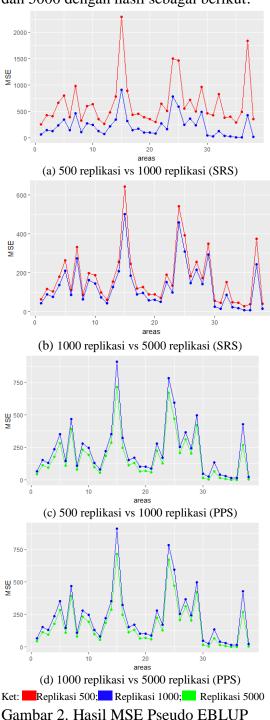
Gambar 1. Hasil Estimasi Langsung, EBLUP dan Pseudo EBLUP

Dapat dilihat pada Gambar 1, estimasi langsung PPS/SRS dengan Pseudo EBLUP-SRS/PPS memberikan pola hasil yang hampir sama. Di sisi lain, estimasi EBLUP-SRS/PPS memberikan pola hasil yang dengan langsung berbeda estimasi PPS/SRS. Perbedaan terutama terjadi pada kota. dimana terdapat wilavah/ area variabilitas pengeluaran yang tinggi. Berdasarkan hal ini, dapat disinyalir bahwa Pseudo EBLUP lebih estimasi dibandingkan dengan **EBLUP** dalam mengakomodir perbedaan variasi antar area sehingga diperoleh estimasi yang lebih baik.

Estimasi MSE

Selanjutnya hasil estimasi langsung PPS dan SRS, EBLUP-SRS/PPS dan Pseudo EBLUP-SRS/PPS dibandingkan MSE-nya. MSE Pseudo-EBLUP didapatkan dengan metode *bootstrap*. Estimator yang memiliki MSE terkecil adalah estimator yang paling efisien. Selanjutnya, untuk mendapatkan hasil

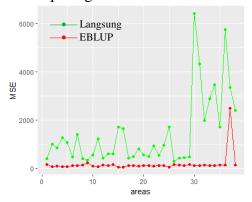
estimasi MSE dari metode Pseudo EBLUP-SRS dan Pseudo EBLUP-PPS digunakan metode *bootstrap*. Pada penelitian ini dilakukan 3 kali *bootstrap* yaitu 500, 1000, dan 5000 dengan hasil sebagai berikut:



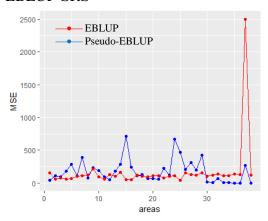
Gambar 2. Hasil MSE Pseudo EBLUP dengan metode sampling SRS dan PPS, bootstrap 500, 1000 dan 5000

Pada Gambar 2 dapat dilihat bahwa MSE terbesar yaitu hasil *bootstrap* dengan 500 replikasi, dan yang terkecil adalah MSE *bootstrap* dengan 5000 replikasi. Hal ini menunjukan bahwa perhitungan MSE

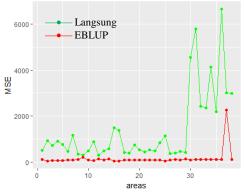
menggunakan bootstrap konsisten dan semakin besar replikasi semakin kecil MSE nya. Sehingga bootstrap yang dipakai pada penelitian ini adalah bootstrap dengan 5000 replikasi. Dengan demikian, hasil perbandingan MSE yang digunakan dapat dilihat pada gambar di bawah ini:



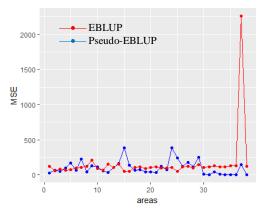
Gambar 3a. Hasil MSE Langsung SRS vs EBLUP-SRS



Gambar 3b. Hasil MSE EBLUP-SRS vs Pseudo EBLUP-SRS



Gambar 3c. Hasil MSE Langsung PPS vs EBLUP-PPS



Gambar 3d. Hasil MSE EBLUP-PPS vs Pseudo EBLUP-PPS

Dapat dilihat dari Gambar 3, MSE estimasi langsung paling besar, kemudian MSE EBLUP dan yang paling kecil adalah MSE Pseudo EBLUP pada PPS dan SRS. Hal ini menunjukan bahwa estimasi SAE EBLUP lebih efisien dibandingkan estimasi langsung, sedangkan estimasi Pseudo EBLUP lebih efisien dibandingkan dengan EBLUP. Kemudian MSE Pseudo EBLUP-PPS lebih kecil daripada MSE Pseudo EBLUP-SRS. Hal ini menunjukkan bahwa hasil Pseudo-EBLUP dengan menggunakan estimasi langsung yang mengakomodir sampling weight sesuai dengan metode survei (yaitu PPS) lebih efisien.

Unit analisis dalam penelitian (kabupaten/kota) yang ada di Provinsi Jawa Timur sebanyak 29 kabupaten (1-29) dan 9 kota (29-38). Dari hasil estimasi EBLUP-**SRS** dan **EBLUP-PPS** menunjukkan perbedaan estimasi antara daerah kabupaten dan kota. Pada daerah perkotaan estimasi EBLUP memiliki variasi yang lebih besar daripada daerah kabupaten, hal ini dapat dilihat dari munculnya satu titik ekstrim dari plot MSE di area kota. Perbedaan terutama terjadi pada wilayah/ area kota, dimana terdapat variabilitas pengeluaran yang tinggi. Sedangkan untuk estimasi Pseudo EBLUP-SRS dan Pseudo EBLUP-PPS tidak ada perbedaan antara hasil estimasi pada daerah pedesaan perkotaan. Berdasarkan hal ini, dapat disimpulkan bahwa estimasi EBLUP lebih baik dibandingkan dengan EBLUP dalam mengakomodir perbedaan variasi antar area sehingga diperoleh estimasi yang lebih robust.

Pseudo-EBLUP pada estimasi langsung SRS dan PPS paling tepat untuk digunakan dalam mengestimasi pengeluaran. Tetapi, jika dilihat lebih dalam estimasi menggunakan lagi Pseudo EBLUP-PPS MSE nya lebih kecil daripada Pseudo EBLUP-SRS. Sehingga dapat disimpulkan bahwa untuk mendapatkan hasil estimasi yang tepat dari Pseudo EBLUP perlu memperhitungkan sampling weight.

KESIMPULAN

Dari hasil pembahasan di atas terdapat beberapa poin kesimpulan diantaranya adalah estimasi SAE lebih efisien daripada estimasi langsung. Hal ini dapat dilihat dari MSE estimasi langsung lebih besar daripada MSE estimasi EBLUP dan Pseudo EBLUP. Selanjutnya estimasi Pseudo EBLUP lebih efisien daripada EBLUP, karena MSE yang dihasilkan Pseudo EBLUP lebih kecil daripada EBLUP. Kemudian. estimasi Pseudo EBLUP dapat mengakomodir perbedaan variasi antar area sehingga dihasilkan estimator yang lebih robust. Dan yang terakhir adalah Pseudo EBLUP-PPS lebih kecil MSE nya daripada Pseudo EBLUP-SRS. Dengan demikian, dapat disimpulkan bahwa metode Pseudo EBLUP dengan memasukkan sampling weight estimasi langsung memberikan hasil yang lebih baik karena menghasilkan estimator yang lebih efisien dan robust.

DAFTAR PUSTAKA

- Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting Linear Mixed-Effects Models Using Ime4. *Journal of Statistical Software*, Vol. 67 (1), 1-48. doi: http://dx.doi.org/10.18637/jss.v067.i01
- Guadarrama Sanz, María & Molina, Isabel & Rao, J.N.K. 2016. A comparison of small area estimation methods for poverty mapping. *Statistics in Transition new series*, Vol. 17 (1), 41-66. 10.21307/stattrans-2016-005.
- Guadarrama Sanz, María & Molina, Isabel & Rao, J.N.K. 2017. Small Aea Estimation of General Parameters under Complex Sampling Designs. *Computational*

- Statistics & Data Analysis, Vol. 121, 20-40
- Heng Penga, Ying Lu. 2012. Model Selection in Linear Mixed Effect Models. *Journal of Multivariate Analysis*, Vol.109, 109-129.
- Lumley, Thomas S. 2015. *Complex Surveys A Guide to Analysis Using R*. New Jersey: John Wiley & Sons.Inc.
- Molina, Isabel & Rao, J.N.K. 2015. *Small Area Estimation*. New Jersey: John Wiley & Sons.Inc.
- Susana Rubin-Bleuer, Leon Jang, Serge Godbout. 2016. The Pseudo-EBLUP Estimator for a Weighted Average with an Application to the Canadian Survey of Employment, Payrolls and Hours. *Journal of Survey Statistics and Methodology*, Vol. 4 (4), 417–435. https://doi.org/10.1093/jssam/smw013
- You, Y., & Rao, J. 2002. A Pseudo-Empirical Best Linear Unbiased Prediction Approach to Small Area Estimation Using Survey Weights. *The Canadian Journal of Statistics / La Revue Canadienne De Statistique*, Vol. 30 (3), 431-439. Retrieved from http://www.jstor.org/stable/3316146
- W. González-Manteiga, M. J. Lombardía, I. Molina, D. Morales & L. Santamaría. 2008. Bootstrap mean squared error of a small-area EBLUP. *Journal of Statistical Computation and Simulation*, Vol. 78(5), 443-462, DOI: 10.1080/00949650601141811
- Zhou, Qingyuan & Luo, Jianjian. 2012. Empirical Test of Efficiency Comparison between PPS Estimation and Simple Random Sampling. *Advanced Science Letters*, Vol. 5, 437-440. 10.1166/asl.2012.3166
- Zimmermann, Thomas, & Burgard, Jan Pablo, & Munnich, Ralf. 2014. The Impact of Sampling Designs on Small Area Estimates for Business Data. *Journal of Official Statistics*, Vol. 30, (4), 749–771. http://dx.doi.org/10.2478/JOS-2014-0046