

Penerapan Metode *Empirical Best Linear Unbiased Prediction* (EBLUP) pada Model Fay-Herriot *Small Area Estimation* (SAE)

Luthfatul Amaliana¹, Evellin Dewi Lusiana²

¹Jurusan Matematika, Universitas Brawijaya Malang

²Jurusan Manajemen Sumber Daya Perairan, Universitas Brawijaya Malang()

Info Artikel

Riwayat Artikel:

Diterima: 15 Mei 2017

Direvisi: 1 Juni 2017

Diterbitkan: 31 Juli 2017

Kata Kunci:

EBLUP,
Model Fay-Herriot,
SAE,
Small area

ABSTRAK

Small area digambarkan sebagai suatu area geografis kecil, seperti desa/kelurahan, kecamatan, atau kabupaten. Ukuran sampel yang kecil dari *small area* menyebabkan estimasi parameter secara langsung tidak mampu menghasilkan ketelitian yang cukup baik. Metode *Empirical Best Linear Unbiased Prediction* (EBLUP) pada *Small Area Estimation* (SAE) menjadi salah satu solusi untuk permasalahan estimasi pada model *small area*, yaitu model Fay-Herriot. Tujuan dari penelitian ini adalah mencari estimator model Fay-Herriot yang bersifat linier, *unbiased*, dan terbaik serta diaplikasikan pada kasus kemiskinan di Kabupaten Jember. Salah satu ukuran kebaikan estimator yaitu dari *Mean Square Error* (MSE) hasil estimasi. Dalam penelitian ini, metode *Lagrange* digunakan untuk memperoleh MSE EBLUP yang minimum. Hasil estimasi pengeluaran rumah tangga per kapita di desa di Kabupaten Jember menunjukkan bahwa hasil estimasi EBLUP lebih baik dibandingkan hasil estimasi langsung.

Copyright © 2017 SI MaNIs.
All rights reserved.

Corresponding Author:

Luthfatul Amaliana,
Jurusan Matematika,
Universitas Brawijaya Malang,
Jl. Veteran, Malang, Jawa Timur, Indonesia
Email: luthfatul@ub.ac.id

1. PENDAHULUAN

Small area didefinisikan sebagai suatu subpopulasi kecil atau gambaran suatu area geografis kecil yang dapat berupa desa/kelurahan, kecamatan, kabupaten, kelompok suku, maupun kelompok umur. Suatu teknik statistika yang memanfaatkan informasi tambahan (*auxiliary variables*), seperti data sensus dan atau catatan administratif *small area* tersebut, maupun catatan administratif *small area* lain yang memiliki karakteristik hampir sama dinamakan *small area estimation* [4].

Estimasi parameter pada *small area* dapat dilakukan secara langsung (*direct estimation*) dan tidak langsung (*indirect estimation*). Namun, ukuran sampel *small area* menjadi permasalahan dalam estimasi langsung yang sering menggunakan model desain penarikan sampel, *small area* yang diamati memiliki ukuran sampel kecil. Hal ini dikarenakan, estimator langsung tidak dapat menghasilkan ketelitian yang cukup atau dengan kata lain memiliki *standar error* estimasi yang besar. Berbeda dengan estimasi langsung, pada estimasi tidak langsung terdapat model penghubung antara *small area* yang satu dengan *small area* lainnya melalui *auxiliary variables*. Model penghubung inilah yang disebut sebagai model eksplisit atau model *small area*. Model *small area* termasuk dalam *General Linear Mixed Model* (GLMM), karena mengandung pengaruh acak *small area* yang menjelaskan variasi antar *small area* yang belum dijelaskan oleh *auxiliary variables* [1].

Model *small area* terbagi menjadi model area level dan model unit level, berdasarkan ketersediaan datanya. Kasus khusus paling sederhana dari model area level adalah model Fay-Herriot. Parameter pada model Fay-Herriot dapat diestimasi dengan metode Bayes dan non-Bayes. Namun, estimasi parameter dengan metode Bayes membutuhkan informasi *prior* yang jarang tersedia dalam data. Sedangkan metode estimasi non-Bayes dapat dilakukan dengan metode BLUP (*Best Linear Unbiased Prediction*) dan EBLUP (*Empirical Best Linear Unbiased*

Prediction). Perbedaan diantara kedua metode non-Bayes tersebut terletak pada informasi pengaruh acak *small area*, yang pada metode BLUP sudah diketahui, sedangkan pada metode EBLUP tidak diketahui dan harus diestimasi secara empiris. Dalam penelitian ini, parameter model Fay-Herriot akan diestimasi dengan metode EBLUP, dimana estimator EBLUP akan bersifat linier, tak-bias (*unbiased*), dan terbaik (*best*) [4].

Berdasarkan uraian di atas, estimasi langsung pada *small area* dengan ukuran sampel kecil akan menghasilkan *standar error* estimasi yang besar. Oleh karena itu, untuk mengatasi permasalahan tersebut perlu dilakukan estimasi tidak langsung, yang dalam penelitian ini dilakukan dengan metode EBLUP pada kasus khusus model *small area* yaitu model Fay-Herriot.

2. TINJAUAN PUSTAKA

2.1 General Linear Mixed Model (GLMM)

Pandang bentuk model regresi linear berganda [3] berikut ini:

$$y = X\beta + \epsilon$$

dimana

y : vektor dari variabel respon, berukuran $nx1$

X : matriks dari variabel prediktor, berukuran $n \times p$, dimana $p = k + 1$

β : vektor dari parameter, berukuran $p \times 1$

ϵ : vektor dari *error*, berukuran $nx1$, dengan $\epsilon_i \sim NID(0, \sigma^2)$.

Pada model regresi linear tersebut, pengaruh tetap yang terkandung di dalam model telah diperhatikan, yaitu X . Namun, pengaruh acak yaitu v , tidak terkandung di dalam model. Sedangkan bentuk *General Linear Mixed Model* adalah sebagai berikut:

$$y = X\beta + Zv + e$$

dimana

y : vektor dari variabel respon, bersifat *random*, berukuran $nx1$

X : matriks *full rank* dari variabel prediktor yang diketahui, berukuran $n \times p$

β : vektor dari parameter, bersifat *fixed*, berukuran $p \times 1$

Z : matriks *full rank* dari variabel prediktor yang diketahui, berukuran $n \times h$

v : vektor dari parameter, bersifat *random*, berukuran $h \times 1$

e : vektor dari *error*, bersifat *random*, berukuran $nx1$

dengan asumsi:

$$\begin{aligned} E(v) &= \mathbf{0}, E(e) = \mathbf{0}, \\ v \text{ dan } e &\text{ saling independen,} \\ E(vv^T) &= G, E(ee^T) = R \end{aligned}$$

Karena v dan e saling independen, maka $cov(v, e) = cov(e, v) = \mathbf{0}$, sehingga $corr(v, e) = \mathbf{0}$. G dan R masing-masing adalah matriks varians-kovarian dari v dan e yang elemen-elemennya merupakan fungsi dalam $\delta = (\delta_1, \delta_2, \dots, \delta_q)^T$. δ adalah vektor parameter dari variansi v dan e . Bentuk dari matriks G dan R yaitu:

$$G(\delta) = \begin{pmatrix} g_{11}(\delta) & g_{12}(\delta) & \dots & g_{1h}(\delta) \\ g_{21}(\delta) & g_{22}(\delta) & \dots & g_{2h}(\delta) \\ \vdots & \vdots & \ddots & \vdots \\ g_{h1}(\delta) & g_{h2}(\delta) & \dots & g_{hh}(\delta) \end{pmatrix}$$

dan

$$R(\delta) = \begin{pmatrix} r_{11}(\delta) & r_{12}(\delta) & \dots & r_{1n}(\delta) \\ r_{21}(\delta) & r_{22}(\delta) & \dots & r_{2n}(\delta) \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1}(\delta) & r_{n2}(\delta) & \dots & r_{nn}(\delta) \end{pmatrix}$$

Sedangkan $E(y)$ dan $var(y)$ didapat dari:

$$\begin{aligned} E(y) &= X\beta \\ var(y) &= ZGZ^T + R = V(\delta) \end{aligned}$$

dengan $V(\delta)$ menyatakan matriks varians-kovarian dari y [4].

2.2 Model Area Level

Model area level adalah salah satu jenis model *small area*, dimana data pendukung yang tersedia hanya sampai pada tingkat area, yaitu $z_i^T = (z_{1i}, \dots, z_{pi})$. Parameter *small area* yang ingin diamati adalah θ_i . Parameter *small area* ini berhubungan linear dengan z_i^T mengikuti model linear berikut:

$$\theta_i = z_i^T \beta + b_i v_i \quad i = 1, \dots, m$$

dengan

$\beta = (\beta_1, \dots, \beta_p)^T$ adalah vektor parameter yang *fixed*, berukuran $p \times 1$

b_i konstanta positif yang diketahui

v_i pengaruh acak *small area*, diasumsikan $v_i \sim iid(0, \sigma_v^2)$

m jumlah *small area*.

Namun, dalam membuat kesimpulan tentang populasi di bawah model area level, diasumsikan estimator langsung $\hat{\theta}_i$ telah ada pada model dan dapat dituliskan sebagai:

$$\hat{\theta}_i = \theta_i + e_i \quad i = 1, \dots, m$$

dimana e_i adalah *sampling error* yang diasumsikan diketahui dengan $e_i \sim ind(0, \psi_i)$.

Kedua model linear di atas, jika digabungkan akan menjadi:

$$\hat{\theta}_i = \mathbf{z}_i^T \beta + b_i v_i + e_i \quad i = 1, \dots, m$$

dengan asumsi bahwa $v_i \sim iid(0, \sigma_v^2)$ dan $e_i \sim ind(0, \psi_i)$. Model ini merupakan salah satu kasus khusus dari *General Linear Mixed Model* (GLMM) dengan *block diagonal covariance structure*, yaitu GLMM dengan matriks varians-kovarians dari v_i , e_i , dan $\hat{\theta}_i$ masing-masing sebagai berikut:

$$\mathbf{G} = \begin{pmatrix} \sigma_v^2 & 0 & \dots & 0 \\ 0 & \sigma_v^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_v^2 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_n \end{pmatrix}$$

dan

$$\mathbf{V} = \mathbf{G} + \mathbf{R} = \begin{pmatrix} \sigma_v^2 + \psi_1 & 0 & \dots & 0 \\ 0 & \sigma_v^2 + \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_v^2 + \psi_n \end{pmatrix}$$

Pada model area level ini, yang akan diestimasi adalah parameter *small area*, yaitu $\theta_i = \mathbf{z}_i^T \beta + b_i v_i$, $i = 1, \dots, m$.

2.3 Model Fay-Herriot

Model Fay-Herriot diperkenalkan oleh Fay dan Herriot (1979), sebagai model dasar untuk mengestimasi pendapatan per kapita pada *small area - small area* (dengan populasi yang kurang dari 1.000 jiwa penduduk) di Amerika Serikat. Model Fay-Herriot ini merupakan kasus model area level pada *Small Area Estimation* (SAE) dengan $b_i = 1$. Berikut akan didefinisikan bentuk model Fay-Herriot yang memiliki estimator langsung $\hat{\theta}_i$ dan vektor variabel pendukung \mathbf{z}_i^T . Bentuk model Fay-Herriot [4] adalah sebagai berikut:

$$\begin{aligned} \hat{\theta}_i &= \mathbf{z}_i^T \beta + v_i + e_i & i = 1, 2, \dots, m \\ &= \theta_i + e_i \end{aligned}$$

dengan

$\hat{\theta}_i$: estimator langsung, berukuran 1×1

θ_i : parameter *small area*, berukuran 1×1

\mathbf{z}_i : vektor variabel pendukung, berukuran $p \times 1$

β : vektor parameter yang *fixed*, berukuran $p \times 1$

v_i : pengaruh acak *small area*, berukuran 1×1 , diasumsikan $v_i \sim NID(0, \sigma_v^2)$

e_i : *sampling error*, berukuran 1×1 , diasumsikan $e_i \sim NID(0, \psi_i)$, ψ_i diketahui,

dimana v_i dan e_i saling independen, sehingga $E(v_i e_i^T) = E(e_i v_i^T) = 0$ dan $E(v_i v_i^T) = \sigma_v^2$. Matriks varians-kovarians dari v_i dan e_i yaitu masing-masing \mathbf{G} dan \mathbf{R} merupakan matriks *block diagonal* dengan bentuk:

$$\mathbf{G} = \begin{pmatrix} \sigma_v^2 & 0 & \dots & 0 \\ 0 & \sigma_v^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_v^2 \end{pmatrix}$$

dan

$$\mathbf{R} = \begin{pmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_n \end{pmatrix}$$

Matriks \mathbf{G} merupakan matriks varians-kovarians, definit positif, berukuran $n \times n$, dari variansi antar *small area*, yang biasanya tidak diketahui dan harus diestimasi. Matriks \mathbf{R} adalah matriks varians-kovarians definit positif berukuran $n \times n$ dari *sampling error*. Sedangkan $\mathbf{V} = \mathbf{G} + \mathbf{R}$ adalah matriks varians-kovarians dari $\hat{\theta}_i$, yang berbentuk:

$$\mathbf{V} = \begin{pmatrix} \sigma_v^2 + \psi_1 & 0 & \dots & 0 \\ 0 & \sigma_v^2 + \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_v^2 + \psi_n \end{pmatrix}$$

Pada model Fay-Herriot ini, parameter yang akan diestimasi adalah $\theta_i = \mathbf{z}_i^T \boldsymbol{\beta} + v_i$, menggunakan metode EBLUP.

2.4 Estimator Langsung, Sintetik, dan Komposit

2.4.1 Estimator Langsung

Pada model Fay-Herriot parameter *small area* (θ_i) diestimasi secara langsung dengan $\hat{\theta}_i$. Nilai $\hat{\theta}_i$ ini berhubungan secara linear dengan θ_i mengikuti model linear:

$$\hat{\theta}_i = \underbrace{\mathbf{z}_i^T \boldsymbol{\beta} + v_i}_{\theta_i} + e_i = \theta_i + e_i \quad i = 1, \dots, m$$

dimana e_i adalah *sampling error* pada area ke- i yang diasumsikan diketahui dengan $e_i \sim \text{NID}(0, \psi_i)$. Estimasi langsung ini dilakukan hanya berdasarkan pada data dari sampel dalam *small area*. Namun, estimator langsung yang dihasilkan memiliki *standar error* yang besar karena ukuran sampel dari *small area* yang diamati terlalu kecil [4].

2.4.2 Estimator Sintetik

Estimator sintetik dengan informasi variabel-variabel pendukung (*auxiliary variables*) yang tersedia pada area ke- i , yaitu \mathbf{z}_i^T , dapat dinyatakan dalam bentuk:

$$\hat{\theta}_i^S = \mathbf{z}_i^T \hat{\boldsymbol{\beta}}; \quad i = 1, 2, \dots, m$$

dengan \mathbf{z}_i berukuran $p \times 1$ dan $\hat{\boldsymbol{\beta}}$ merupakan vektor parameter yang *fixed*, berukuran $p \times 1$.

Definisi estimator sintetik dijelaskan oleh Gonzales (1973) dalam [4], bahwa suatu estimator dikatakan estimator sintetik jika suatu estimator langsung yang *reliable* untuk area luas yang mencakup *small area - small area*, digunakan untuk menurunkan estimator tidak langsung untuk *small area*, di bawah asumsi *small area* tersebut memiliki karakteristik yang sama seperti area yang luas. Kelemahan dari estimator ini adalah memiliki bias yang besar dikarenakan adanya asumsi kesamaan karakteristik antara *small area* dengan area yang luas.

2.4.3 Estimator Komposit

Estimasi pada *small area* menggunakan dua jenis estimator (estimator langsung dan estimator sintetik) memiliki kelebihan dan kekurangan masing-masing. Estimator langsung bersifat *unbiased* karena hanya berdasarkan pada data sampel dari *small area* tersebut. Namun, estimator ini kurang stabil, karena memiliki *standar error* yang besar. Sedangkan estimator sintetik telah memiliki *standar error* yang lebih baik dari estimator langsung, tetapi memiliki bias yang besar. Oleh karena itu, untuk menyeimbangkan ketidakstabilan dari estimator langsung dan bias yang besar dari estimator sintetik, dibentuk estimator komposit yang merupakan rata-rata terboboti dari estimator langsung dan estimator sintetik. Bentuk estimator komposit yaitu:

$$\begin{aligned} \hat{\theta}_i^C &= \phi_i \hat{\theta}_i + (1 - \phi_i) \hat{\theta}_i^S; \quad i = 1, 2, \dots, m \\ &= \phi_i \hat{\theta}_i + (1 - \phi_i) \mathbf{z}_i^T \hat{\boldsymbol{\beta}} \end{aligned}$$

dengan ϕ_i merupakan bobot yang dipilih, $0 \leq \phi_i \leq 1$ dan m menunjukkan banyaknya *small area*. Bobot optimal untuk estimator komposit diperoleh dengan meminimalkan $\text{MSE}[\hat{\theta}_i^C]$ terhadap ϕ_i dengan asumsi $\text{cov}(\hat{\theta}_i, \hat{\theta}_i^S) = 0$. Bobot optimal tersebut dinyatakan sebagai [4]:

$$\phi_i^* = \frac{\text{MSE}[\hat{\theta}_i^S]}{\text{MSE}[\hat{\theta}_i] + \text{MSE}[\hat{\theta}_i^S]}$$

2.5 Mean Squared Error (MSE)

Misalkan θ merupakan suatu parameter dan $\hat{\theta}$ merupakan estimator θ . MSE dari $\hat{\theta}$ didefinisikan sebagai:

Misalkan: $E[\hat{\theta}] = a$, yang belum tentu θ ,

$$\begin{aligned} \text{MSE}[\hat{\theta}] &= E[(\hat{\theta} - \theta)^2] \\ &= E[(\hat{\theta} - a + a - \theta)^2] \\ &= E[(\hat{\theta} - a)^2 + 2(\hat{\theta} - a)(a - \theta) + (a - \theta)^2] \\ &= E[(\hat{\theta} - a)^2] + 2E[(\hat{\theta} - a)](a - \theta) + E[(a - \theta)^2] \end{aligned}$$

$$= E[(\hat{\theta} - a)^2] + 2 \underbrace{E[(\hat{\theta} - a)]}_{=0} (a - \theta) + E(a - \theta)^2$$

$$MSE[\hat{\theta}] = \text{var}[\hat{\theta}] + [\text{bias}(\hat{\theta})^2] ; \text{karena } 2E[(\hat{\theta} - a)] = 0$$

Berdasarkan definisi MSE, jika $\hat{\theta}$ yang diperoleh *unbiased*, maka MSE $\hat{\theta}$ akan sama dengan variansi $\hat{\theta}$. Sedangkan *standar error* dari $\hat{\theta}$ didefinisikan sebagai akar kuadrat positif dari $MSE[\hat{\theta}]$. Nilai MSE dari suatu estimator memiliki peranan penting untuk diketahui, salah satunya untuk mengukur seberapa baik estimator yang diperoleh [4].

3. METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini adalah data Susenas 2008 dan data Potensi Desa (PODES) 2008 yang diperoleh dari Badan Pusat Statistik (BPS). Berdasarkan Laporan BPS 2007, Jawa Timur merupakan propinsi paling miskin di Pulau Jawa dan Kabupaten Jember merupakan satu di antara kota/kabupaten yang memiliki tingkat kemiskinan paling tinggi. Kabupaten Jember terdiri dari 247 desa, dimana 14,17% nya (35 desa) terpilih sebagai sampel pada Susenas 2008. Setiap desa terpilih, 14 sampai dengan 16 rumah tangga terpilih sebagai sampel, sehingga jumlah keseluruhan rumah tangga sampel adalah 549 rumah tangga. Jumlah rumah tangga yang menjadi sampel pada setiap desa tersebut sangat kecil jika dibandingkan dengan jumlah rumah tangga di masing-masing desa, yaitu berkisar antara 0,1% sampai dengan 1,67% [2].

Dalam penelitian ini, estimasi pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember dilakukan dengan metode EBLUP. Setelah diperoleh estimator EBLUP pengeluaran rumah tangga per kapita, kemudian dicari *Mean Squared Error* (MSE) nya dan dibandingkan dengan MSE dari hasil estimasi langsung, untuk mengetahui metode penaksiran mana yang lebih baik.

Data variabel-variabel pendukung (*auxiliary variables*) diperoleh dari data PODES 2008. Berdasarkan ketersediaan data PODES, dipilih variabel-variabel pendukung yang berkaitan dengan faktor-faktor yang mempengaruhi pengeluaran rumah tangga [5]. Satuan masing-masing variabel berbeda-beda, sehingga dalam penelitian ini dilakukan standarisasi. Variabel-variabel pendukung tersebut antara lain:

- persentase keluarga yang menggantungkan hidupnya pada pertanian (\mathbf{z}_1^T),
- jumlah keluarga yang menerima askeskin dalam satu tahun terakhir (\mathbf{z}_2^T),
- jumlah keluarga pengguna listrik PLN (\mathbf{z}_3^T),
- jumlah keluarga yang mengenyam pendidikan SD, SMP, SMA, dan PT (\mathbf{z}_4^T),
- jumlah keluarga yang tinggal di pemukiman kumuh (\mathbf{z}_5^T),
- jumlah keluarga yang memiliki Surat Keterangan Tidak Mampu (SKTM) dan satu tahun terakhir (\mathbf{z}_6^T),
- jumlah keluarga yang pernah belajar di lembaga pendidikan dan ketrampilan (\mathbf{z}_7^T),
- jumlah keluarga yang memiliki anggota keluarga sebagai Tenaga Kerja Indonesia (TKI) (\mathbf{z}_8^T).

4. HASIL DAN ANALISIS

Dalam penelitian ini, estimasi parameter pada model Fay-Herriot dilakukan untuk memperoleh estimator EBLUP. Selanjutnya, estimasi langsung dan estimasi tidak langsung dengan metode EBLUP diterapkan untuk mengestimasi pengeluaran rumah tangga per kapita di desa di Kabupaten Jember. Pengeluaran rumah tangga per kapita merupakan indikator suatu keluarga dikatakan sebagai rumah tangga miskin. Untuk mengetahui seberapa baik estimator yang diperoleh, nilai MSE dari kedua estimator dibandingkan, nilai MSE yang lebih kecil menunjukkan estimator lebih baik.

4.1. Metode EBLUP pada Model Fay-Herriot

Berbeda dengan metode BLUP yang variansi pengaruh acak *small area* sudah diketahui, pada estimasi parameter model Fay-Herriot dengan metode EBLUP, variansi pengaruh acak (σ_v^2) *small area* tidak diketahui nilainya, sehingga harus diestimasi dari data empiris. Salah satu metode yang dapat digunakan adalah metode *Maximum Likelihood* (ML), dengan *mean* dan variansi dari $\hat{\theta}_i$ yaitu:

$$E(\hat{\theta}_i) = \mathbf{z}_i^T \boldsymbol{\beta}$$

$$\text{var}(\hat{\theta}_i) = \sigma_v^2 + \psi_i.$$

Oleh karena v_i dan e_i berdistribusi Normal, maka $\hat{\theta}_i$ juga berdistribusi Normal. Dengan demikian, $\hat{\theta}_i \sim N(\mathbf{z}_i^T \boldsymbol{\beta}, (\sigma_v^2 + \psi_i))$. Sedangkan fungsi *likelihood* dan *log-likelihood* dari $\hat{\theta}_i$ yaitu:

$$L(\boldsymbol{\beta}, \sigma_v^2; \hat{\theta}_i) = f(\hat{\theta}_i) = \frac{1}{(2\pi)^{1/2} (\sigma_v^2 + \psi_i)^{1/2}} \exp \left[-\frac{1}{2} \left((\hat{\theta}_i - \mathbf{z}_i^T \boldsymbol{\beta})^T (\sigma_v^2 + \psi_i)^{-1} (\hat{\theta}_i - \mathbf{z}_i^T \boldsymbol{\beta}) \right) \right]$$

$$\ln L(\hat{\beta}, \sigma_v^2; \hat{\theta}_i) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \left[\ln(\sigma_v^2 + \psi_i) + (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})^T (\sigma_v^2 + \psi_i)^{-1} (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta}) \right]$$

Fungsi *log-likelihood* dari $\hat{\theta}_i$ kemudian diturunkan terhadap σ_v^2 , dinotasikan dengan $\mathbf{s}(\hat{\beta}, \sigma_v^2)$, sehingga untuk differensial terhadap elemen ke- j diperoleh formula:

$$\mathbf{s}_j(\hat{\beta}, \sigma_v^2) = \frac{\partial \ln L(\hat{\beta}, \sigma_v^2; \hat{\theta}_i)}{\partial \sigma_v^2} = -\frac{1}{2} \frac{1}{(\sigma_v^2 + \psi_i)} + \frac{1}{2} \frac{(\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})^T (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})}{(\sigma_v^2 + \psi_i)^2}$$

Berdasarkan prosedur metode *Maximum Likelihood*, akan dicari solusi dari persamaan:

$$\frac{1}{2} \frac{1}{(\sigma_v^2 + \psi_i)} = \frac{1}{2} \frac{(\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})^T (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})}{(\sigma_v^2 + \psi_i)^2}$$

dimana σ_v^2 tidak dapat diselesaikan secara analitik. Oleh karena itu, estimator σ_v^2 diselesaikan secara numerik, menggunakan *Scoring Algorithm*, dengan iterasi ke- $(a + 1)$ [4]:

$$\sigma_v^{2(a+1)} = \sigma_v^{2(a)} + [\mathcal{L}(\sigma_v^{2(a)})]^{-1} \mathbf{s}(\hat{\beta}(\sigma_v^{2(a)}), \sigma_v^{2(a)})$$

dimana

$$\mathcal{L}(\sigma_v^2) = \frac{\partial^2 \ln L(\hat{\beta}, \sigma_v^2; \hat{\theta}_i)}{\partial \sigma_v^2 \partial \sigma_v^2} = \frac{1}{2} \frac{1}{(\sigma_v^2 + \psi_i)^2}$$

dan

$$\mathbf{s}(\hat{\beta}, \sigma_v^2) = -\frac{1}{2} \frac{1}{\sigma_v^2 + \psi_i} + \frac{1}{2} \frac{(\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta})^2}{(\sigma_v^2 + \psi_i)^2}$$

Proses iterasi tersebut berhenti jika $(\widehat{\sigma}_v^2)^{(a+1)} \approx (\widehat{\sigma}_v^2)^{(a)}$, yang dalam penelitian ini, ditetapkan $|(\widehat{\sigma}_v^2)^{(a+1)} - (\widehat{\sigma}_v^2)^{(a)}| < 10^{-5}$. Kemudian $(\widehat{\sigma}_v^2)^{(a+1)}$ yang diperoleh merupakan estimator σ_v^2 . Berdasarkan bentuk estimator komposit, maka estimator pada model Fay-Herriot yang bersifat linear, *unbiased*, dan terbaik (memiliki *standar error* kecil), berbentuk [4]:

$$\begin{aligned} (T(\hat{\theta}_i))(\widehat{\sigma}_v^2) &= \phi_i \hat{\theta}_i + (1 - \phi_i) \mathbf{z}_i^T \hat{\beta} \\ &= \mathbf{z}_i^T \hat{\beta} + \frac{\widehat{\sigma}_v^2}{(\widehat{\sigma}_v^2 + \psi_i)} (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta}) \\ &= \mathbf{z}_i^T \hat{\beta} + \hat{v}_i \end{aligned}$$

dengan

$$\begin{aligned} \hat{\beta}(\widehat{\sigma}_v^2) &= (\mathbf{z}_i(\widehat{\sigma}_v^2 + \psi_i)^{-1} \mathbf{z}_i^T)^{-1} \mathbf{z}_i(\widehat{\sigma}_v^2 + \psi_i)^{-1} \hat{\theta}_i \\ \hat{v}_i(\widehat{\sigma}_v^2) &= \frac{\widehat{\sigma}_v^2}{(\widehat{\sigma}_v^2 + \psi_i)} (\hat{\theta}_i - \mathbf{z}_i^T \hat{\beta}(\widehat{\sigma}_v^2)). \end{aligned}$$

4.2. Penerapan Metode EBLUP pada Kasus Kemiskinan di Kabupaten Jember

Dalam penelitian ini, data yang digunakan sebagai estimator langsung adalah rata-rata pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember. Estimator langsung dan variansi sampel dihitung dari data pengeluaran rumah tangga per kapita yang terpilih menjadi sampel pada Susenas 2008, sejumlah 35 desa. Berdasarkan variansi sampel tersebut dapat dihitung nilai MSE dari estimator langsung. Estimasi langsung pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember dilakukan menggunakan *software Ms. Excel*. Estimator langsung pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember ($\hat{\theta}_i$) didapat dengan menghitung rata-rata pengeluaran rumah tangga per kapita di setiap desa, yaitu:

$$\hat{\theta}_i = \frac{\text{Jumlah pengeluaran rumah tangga per kapita di desa-i}}{\text{Jumlah rumah tangga di desa-i}}$$

Sedangkan MSE dan *standar error* (SE) yang merupakan akar kuadrat positif dari MSE estimator langsung, dihitung dengan:

$$\begin{aligned} \text{MSE}(\hat{\theta}_i) &= \frac{s_i^2}{n_i}, \text{ dengan } s_i^2 = \frac{1}{n_i - 1} \sum_i^{n_i} (\hat{\theta}_i - \bar{\theta}_i)^2 \\ \text{SE}_{\hat{\theta}_i} &= \sqrt{\text{MSE}(\hat{\theta}_i)} \end{aligned}$$

Hasil estimasi langsung pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember ($\hat{\theta}_i$) disajikan pada Tabel 1. Pengujian asumsi normalitas dengan Uji Kolmogorov-Smirnov menunjukkan estimator langsung pengeluaran rumah tangga per kapita memenuhi asumsi normalitas.

Tabel 1. Estimator Langsung Pengeluaran Rumah Tangga per Kapita per Bulan di Desa di Kabupaten Jember (dalam Rupiah)

No. Desa	Estimator Langsung	No. Desa	Estimator Langsung	No. Desa	Estimator Langsung	No. Desa	Estimator Langsung	No. Desa	Estimator Langsung	No. Desa	Estimator Langsung
1	669.589	7	681.211	13	671.319	19	634.862	25	605.607	31	569.281
2	655.184	8	656.333	14	624.151	20	588.754	26	661.572	32	708.618
3	578.315	9	630.343	15	658.970	21	610.272	27	731.584	33	519.447
4	581.532	10	622.306	16	615.928	22	577.459	28	586.261	34	539.347
5	596.978	11	627.596	17	633.036	23	608.950	29	598.145	35	487.680
6	689.277	12	609.002	18	798.599	24	566.675	30	585.499		

Tabel 1 menunjukkan tiga desa dengan pengeluaran rumah tangga per kapita yang cukup tinggi dibanding desa yang lain, yaitu desa ke-18, desa ke-27, dan desa ke-32, dimana desa yang memiliki pengeluaran rumah tangga per kapita tertinggi adalah desa ke-18, sebesar Rp 798.599,-.

Estimasi tidak langsung pengeluaran rumah tangga per kapita di desa di Kabupaten Jember dilakukan menggunakan *software R*, yaitu program EBLUP ML.r. Berdasarkan *output* program EBLUP ML.r, diperoleh estimator pengaruh tetap, $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_8)^T$ dan pengaruh acak, \hat{v}_i , yang masing-masing diberikan pada Tabel 2 dan Tabel 3.

Tabel 2. Estimator Pengaruh Tetap $\hat{\beta}$

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$
-63.875	-47.021	59.037	-6.977	3.049	-51.072	-17.284	-59.079

Tabel 3. Estimator Pengaruh Acak \hat{v}_i

No. Desa	\hat{v}_i	No. Desa	\hat{v}_i	No. Desa	\hat{v}_i	No. Desa	\hat{v}_i	No. Desa	\hat{v}_i	No. Desa	\hat{v}_i
1	166.389	7	504.496	13	843.733	19	674.222	25	33.579	31	314.672
2	-109.469	8	96.544	14	288.033	20	581.511	26	344.855	32	492.838
3	367.437	9	790.676	15	579.096	21	472.122	27	256.939	33	622.253
4	763.553	10	854.161	16	560.107	22	420.615	28	-66.699	34	283.873
5	346.565	11	746.070	17	1.091.283	23	464.819	29	489.578	35	361.510
6	426.916	12	343.443	18	809.113	24	547.013	30	-154.404		

Pengujian asumsi normalitas menggunakan Uji Kolmogorov-Smirnov juga menunjukkan estimator pengaruh acak *small area* berdistribusi Normal. Dengan mengambil nilai awal $\hat{\sigma}_v^{2(0)}=5.000.000$, diperoleh estimator variansi pengaruh acak menggunakan *Scoring Algorithm* yaitu $\hat{\sigma}_v^2 = 280.108.000.000$. Estimator EBLUP pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember ($\hat{\theta}_i$) disajikan pada Tabel 4.

Tabel 4. Estimator EBLUP Pengeluaran Rumah Tangga per Kapita per Bulan di Desa di Kabupaten Jember (dalam Rupiah)

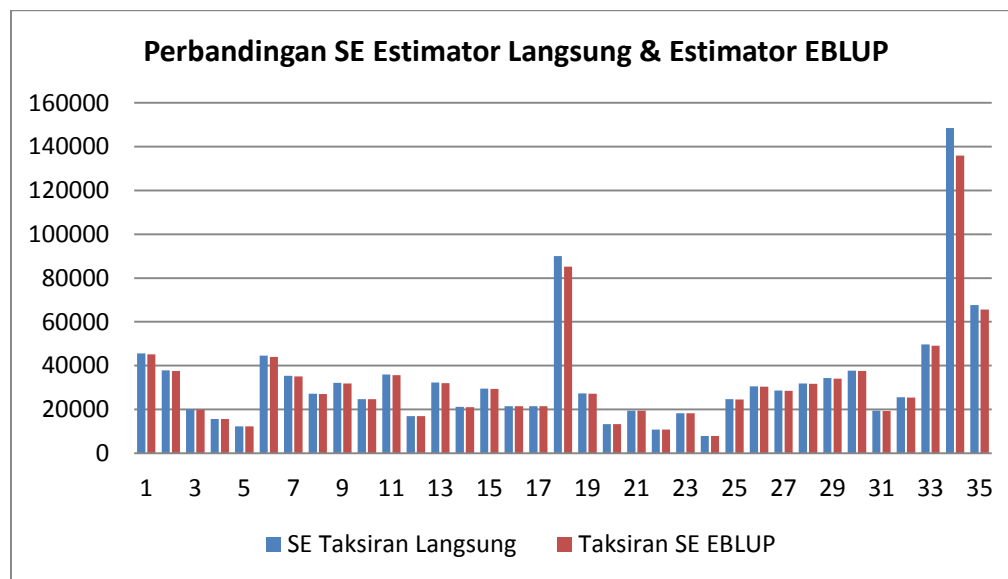
No. Desa	Estimator EBLUP	No. Desa	Estimator EBLUP	No. Desa	Estimator EBLUP	No. Desa	Estimator EBLUP	No. Desa	Estimator EBLUP	No. Desa	Estimator EBLUP
1	668.350	7	678.950	13	668.187	19	633.065	25	605.534	31	568.858
2	655.744	8	656.078	14	623.691	20	588.384	26	660.426	32	707.471
3	577.793	9	627.437	15	657.172	21	609.637	27	730.834	33	513.974
4	580.859	10	620.439	16	615.006	22	577.281	28	586.503	34	516.977
5	596.793	11	624.152	17	631.232	23	608.394	29	596.075	35	481.776
6	686.255	12	608.647	18	775.181	24	566.552	30	586.284		

Berdasarkan Tabel 4, desa-desa yang memiliki pengeluaran rumah tangga per kapita cukup tinggi adalah desa ke-18, desa ke-27, dan desa ke-32, dimana desa yang memiliki pengeluaran rumah tangga per kapita tertinggi adalah desa ke-18, yaitu sebesar Rp 775.181,-. Perbandingan nilai *standar error* kedua estimator (langsung dan EBLUP) dirangkum dalam Tabel 5 dan disajikan dalam bentuk diagram batang pada Gambar 1.

Tabel 5 dan Gambar 1 menunjukkan bahwa nilai *standar error* dari estimator EBLUP lebih kecil dibanding *standar error* dari estimasi langsung. Oleh karena itu, estimasi tidak langsung (EBLUP) dikatakan lebih baik daripada estimasi langsung dalam kasus kemiskinan di Kabupaten Jember.

Tabel 5. Perbandingan *Standar Error* Estimator Langsung & EBLUP Pengeluaran Rumah Tangga per Kapita per Bulan di Desa di Kabupaten Jember (dalam Rupiah)

No. Desa	SE Langsung	SE EBLUP	No. Desa	SE Langsung	SE EBLUP	No. Desa	SE Langsung	SE EBLUP	No. Desa	SE Langsung	SE EBLUP
1	45.675	45.230	10	24.746	24.678	19	27.326	27.182	28	31.889	31.687
2	37.841	37.637	11	35.957	35.666	20	13.356	13.339	29	34.412	34.120
3	19.957	19.904	12	17.015	17.001	21	19.412	19.378	30	37.726	37.495
4	15.709	15.688	13	32.245	32.011	22	10.888	10.879	31	19.398	19.356
5	12.229	12.216	14	21.157	21.102	23	18.304	18.281	32	25.535	25.475
6	44.527	43.924	15	29.492	29.321	24	7.952	7.948	33	49.634	49.121
7	35.432	35.122	16	21.478	21.435	25	24.633	24.552	34	148.570	135.871
8	27.204	27.090	17	21.519	21.466	26	30.511	30.362	35	67.634	65.597
9	32.088	31.866	18	90.039	85.201	27	28.588	28.450			



Gambar 1. Grafik Perbandingan *Standar Error* Estimator Langsung dan EBLUP Pengeluaran Rumah Tangga per Kapita per Bulan di Desa di Kabupaten Jember

5. KESIMPULAN

Small area merupakan gambaran area geografis kecil atau subpopulasi kecil. Karena areanya yang kecil, estimasi langsung pada *small area* menghasilkan *standar error* yang besar, sehingga perlu dilakukan estimasi tidak langsung yang dalam penelitian ini menggunakan metode EBLUP. Bentuk estimator EBLUP merupakan rata-rata terbobot untuk menyeimbangkan bias yang besar dari estimator langsung dan ketidakstabilan estimator sintetik, dengan bobot yang optimal. Nilai MSE digunakan untuk mengetahui seberapa baik kedua estimator yang diperoleh. Estimator EBLUP menghasilkan nilai MSE yang lebih kecil dibanding MSE estimator langsung. Oleh karena itu, estimator EBLUP pengeluaran rumah tangga per kapita per bulan di desa di Kabupaten Jember pada model Fay-Herriot lebih baik dibanding estimator langsung. Untuk penelitian berikutnya, pengaruh spasial pada *small area* dapat diperhitungkan dengan menggunakan metode Spasial EBLUP.

REFERENSI

- [1] Caroline, L. C. Penaksiran Pengeluaran per Kapita di Kabupaten Lumajang dengan Menggunakan Metode Empirical Best Linear Unbiased Prediction (EBLUP) pada Small Area Estimation (SAE). Depok: Dept. Matematika, FMIPA, UI; 2010.
- [2] Matualage, D. Metode Prediksi Tak Bias Linear Terbaik Empiris Spasial Pada Area Kecil Untuk Pendugaan Pengeluaran per Kapita. Bogor: Sekolah Pasca Sarjana, IPB; 2012.
- [3] Montgomery, D. C., Peck, E. A., & Vining, G. G. Introduction to Linear Regression Analysis. New York: John Wiley and Sons, Inc; 2001.
- [4] Rao, J. N. Small Area Estimaion. New Jersey: John Wiley & Sons, Inc; 2003.
- [5] Sunandi, E. Model Spasial Bayes dalam Pendugaan Area Kecil dengan Peubah Respon Biner (Kasus: Pendugaan Proporsi Keluarga Miskin di Kabupaten Jember, Jawa Timur). Bogor: Pascasarjana, IPB; 2011.