

EFISIENSI METODE EBLUP PADA *SMALL AREA ESTIMATION*

Studi Kasus: Estimasi Persentase Penduduk Miskin di Provinsi Nusa Tenggara Timur Tahun 2017

(*The Efficiency of EBLUP in Small Area Estimation*)

Case Study: Estimation of Poor Population Percentage in East Nusa Tenggara Province in 2017)

Easbi Ikhsan^{1*}, Chintia Arisandi Hidayat¹, Wirda Avie Nurizza¹

Badan Pusat Statistik¹, Jalan Otto Iskandardinata 64C, Jakarta Timur 13330

E-mail: *14.8096@stis.ac.id

Abstrak

Kemiskinan merupakan masalah multidimensional yang membutuhkan perhatian pemerintah dan segenap unsur masyarakat yang ada. Hal ini dikarenakan masalah kemiskinan mencakup berbagai aspek kehidupan. Ketersediaan data kemiskinan hingga wilayah terkecil sangat diperlukan agar program dan kebijakan pengentasan kemiskinan dapat dilaksanakan tepat sasaran. Seiring dengan berkembangnya pemerintahan, sistem pengambilan keputusan di setiap daerah membutuhkan ketersediaan data yang akurat dan representatif hingga ke wilayah terkecil. Permasalahan muncul ketika ingin diperoleh informasi kemiskinan untuk area kecil dengan desain survei Susenas Badan Pusat Statistik (BPS). Tingkat presisi hasil estimasi yang diperoleh menjadi rendah karena sampel yang digunakan terlalu sedikit. Hal ini dikarenakan *sampling design* BPS dirancang untuk estimasi langsung tingkat area yang luas seperti nasional, provinsi, dan kabupaten/kota. Salah satu upaya untuk dapat melakukan estimasi hingga area kecil tanpa harus menambah jumlah sampel adalah *Small Area Estimation* (SAE). Pada SAE dengan metode *Empirical Based Linear Unbiased Predictor* (EBLUP) berbasis level area serta menggunakan variabel penyerta mampu memberikan nilai estimasi yang lebih baik dibandingkan hasil estimasi langsung. Dari sampel yang tersedia, kecukupan presisi SAE dapat diketahui dengan melakukan simulasi *bootstrap resampling* pada ukuran sampel yang berbeda. Selanjutnya, dilakukan evaluasi dari simulasi yang telah dilakukan dengan melakukan penghitungan nilai efisiensi relatif sebagai perbandingan *Mean Squared Error* (MSE) dari SAE dengan estimasi langsung. Hasil penghitungan efisiensi relatif menunjukkan bahwa estimasi SAE metode EBLUP pada ukuran sampel 60 persen dari ukuran sampel Susenas sudah mampu menghasilkan presisi yang hampir sama dengan estimasi langsung dari sampel Susenas tersebut.

Kata kunci: EBLUP, kemiskinan, MSE, efisiensi relatif, SAE

Abstract

Poverty is a multidimensional problem that requires attention from the government and all elements of society. This is because the problem of poverty covers various aspects of life. The availability of poverty data to the smallest area is very necessary so that poverty alleviation programs and policies can be implemented on target. Along with the development of existing government systems, the decision-making system in each region requires the availability of accurate and representative data to the smallest areas. Susenas is one of survey that held by Indonesian Bureau of Statistics which provides data to get the poverty indicators, but it has limitations. The sampling design of this survey is only designed for direct estimation of large area levels such as national, provincial and district/city. One effort to be able to estimate up to a small area without having to increase the number of samples is Small Area Estimation (SAE). At SAE, the Empirical Based Linear Unbiased Prediction (EBLUP) method is based on area level and using the accompanying variables is able to provide a better estimation value than the direct estimation results. From the available samples, the adequacy of SAE precision can be determined by performing a bootstrap resampling simulation on different sample sizes. Furthermore, an evaluation of the simulation has been carried out by calculating the relative efficiency value as the Mean Squared Error (MSE) comparison of SAE with direct estimation. The results of relative efficiency calculations show that the SAE estimation of the EBLUP method at a sample size of 60 percent of the Susenas sample size has been able to produce precision that is almost the same as the direct estimation of the Susenas sample.

Keywords: EBLUP, poverty, MSE, relative efficiency, SAE

PENDAHULUAN

Kemiskinan merupakan suatu masalah yang tidak dapat dihindari oleh seluruh negara, baik negara maju maupun negara berkembang. Indonesia sebagai salah satu negara berkembang juga menghadapi masalah kemiskinan. Kemiskinan adalah suatu masalah multidimensional yang membutuhkan perhatian dari pemerintah dan segenap unsur masyarakat yang ada. Kemiskinan sering kali dikaitkan dengan aspek ekonomi, akan tetapi jika ditinjau lebih lanjut kemiskinan juga mencakup aspek kehidupan lainnya seperti pendidikan, kesehatan, dan sosial budaya.

Kemiskinan termasuk dalam salah satu indikator keberhasilan pembangunan nasional yang dilaksanakan oleh pemerintah. Berdasarkan data Susenas 2017 diketahui bahwa tingkat kemiskinan Indonesia saat ini berada pada angka 10,12 persen. Salah satu provinsi yang memiliki persentase penduduk miskin tertinggi di Indonesia adalah Provinsi Nusa Tenggara Timur (NTT). NTT menduduki peringkat persentase penduduk miskin tertinggi ketiga setelah Papua dan Papua Barat.

Persentase kemiskinan NTT masih di atas angka 20 persen. Hal tersebut menunjukkan bahwa tingkat kemiskinan Provinsi NTT berada di atas tingkat kemiskinan nasional. Pada tahun 2017, pertumbuhan persentase kemiskinan NTT mengalami penurunan yakni dari 22,19 persen pada Maret 2017 menjadi 22,01 persen pada September 2017. Data hasil Susenas juga menunjukkan bahwa di Provinsi NTT persentase penduduk miskin di perdesaan lebih besar dari persentase penduduk miskin di perkotaan. Pada tahun 2017, persentase penduduk miskin di perdesaan berada pada angka 25,19 persen, sedangkan persentase penduduk miskin di perkotaan berada pada angka 10,17 persen.

Pemerintah memiliki tugas untuk menyelesaikan masalah kemiskinan. Pemerintah membutuhkan data kemiskinan yang akurat untuk dapat mengatasi masalah kemiskinan. Ketersediaan data kemiskinan hingga wilayah terkecil diperlukan agar program dan kebijakan pengentasan

kemiskinan dapat dilaksanakan dengan tepat sasaran.

Badan Pusat Statistik (BPS) merupakan instansi yang bertanggung jawab dalam menyediakan data yang dibutuhkan oleh pemerintah. Terdapat dua metode yang digunakan oleh BPS untuk mengumpulkan data, yaitu survei dan sensus. Survei memiliki kelebihan yakni hemat dalam biaya, waktu, dan sumber daya jika dibandingkan dengan sensus. Selain itu, pelaksanaan survei lebih efektif dan menghasilkan cakupan materi yang lebih luas. Namun, estimasi parameter yang dihasilkan survei hanya mampu untuk area yang memiliki kecukupan sampel seperti nasional, provinsi atau kabupaten.

Data kemiskinan yang digunakan sebagai dasar pembuatan kebijakan oleh pemerintah diperoleh dari hasil Survei Sosial Ekonomi Nasional (Susenas). Data yang digunakan untuk melakukan estimasi tingkat kemiskinan diperoleh dari data Susenas KOR dan Susenas modul konsumsi. Berdasarkan kedua data tersebut akan dilakukan estimasi terhadap besarnya pengeluaran konsumsi dan non konsumsi rumah tangga. Selanjutnya, nilai pengeluaran konsumsi dan non konsumsi tersebut akan dibandingkan dengan garis kemiskinan.

Seiring dengan berkembangnya sistem pemerintahan yang ada, sistem pengambilan keputusan di tiap-tiap daerah membutuhkan ketersediaan data yang akurat dan representatif di wilayah masing-masing. Permintaan pengukuran survei hingga ke wilayah terkecil (*small area statistics*) menjadi kebutuhan dalam pengaturan dan pembuatan kebijakan oleh pemerintah setempat. Kebutuhan data tidak hanya pada tingkat provinsi, namun juga pada area yang lebih kecil lagi baik level kabupaten, kecamatan, maupun desa.

Persoalan akan muncul ketika ingin diperoleh informasi untuk area kecil dengan desain survei BPS. Tingkat presisi hasil estimasinya menjadi rendah karena ukuran sampel yang digunakan terlalu sedikit. Hal ini dikarenakan desain sampel BPS hanya dirancang untuk *direct estimation* (estimasi langsung) pada tingkat area yang luas

seperti nasional, provinsi atau kabupaten. Pada pendekatan estimasi langsung (*simple random sampling, stratified sampling, cluster sampling* dan metode sampling lainnya), salah satu cara supaya dapat dilakukan estimasi hingga area kecil adalah dengan menambah ukuran sampel. Penambahan ukuran sampel akan meningkatkan biaya yang digunakan saat pelaksanaan survei.

Small Area Estimation (SAE) merupakan salah satu metode yang dapat diterapkan pada area dengan ukuran sampel kecil untuk meningkatkan presisi dari estimasi langsung. Hal ini dikarenakan pada SAE juga menggunakan estimasi tak langsung dengan memanfaatkan variabel penyerta. Oleh karena itu, pada penelitian ini akan digunakan SAE untuk mengestimasi persentase penduduk miskin dan melihat efisiensi model yang terbentuk.

Pada penelitian sebelumnya, Wijaya, dkk (2017) telah meneliti mengenai *Small Area Estimation* pada Tingkat Kemiskinan di Provinsi Jawa Tengah dengan Pendekatan *Empirical Best Linear Unbiased Prediction* (EBLUP). Hasil penelitian mereka menunjukkan estimasi area kecil dengan pendekatan EBLUP menghasilkan *Mean Squared Error* (MSE) dengan nilai yang kecil sehingga cocok digunakan sebagai penduga pada area yang memiliki ukuran sampel yang kecil. Pada penelitian ini akan dibahas mengenai SAE dengan pendekatan EBLUP pada kasus estimasi persentase penduduk miskin di Provinsi Nusa Tenggara Timur Tahun 2017. Selanjutnya, dilakukan simulasi *bootstrap resampling* pada ukuran sampel yang berbeda dari sampel yang tersedia untuk mengetahui kecukupan presisi SAE.

METODE

Small Area Estimation

Estimasi parameter populasi dapat dilakukan dengan estimasi langsung dan estimasi tak langsung. Estimasi langsung merupakan estimasi parameter yang dilakukan berdasarkan data sampel dari suatu domain. Oleh karena itu, estimasi langsung didasarkan pada teknik penarikan sampel. Akan tetapi, ketika digunakan

untuk estimasi area kecil akan menghasilkan varians yang lebih besar karena keterbatasan sampel (Sadik dan Notodiputro, 2006). Estimasi tak langsung merupakan estimasi berdasarkan model. Estimasi ini dapat menjadi salah satu solusi untuk estimasi area kecil dengan memanfaatkan kekuatan yang diperoleh dari variabel penyerta tambahan yang digunakan (*auxiliary variable*).

Dalam SAE terdapat dua model yang lazim digunakan yakni model implisit dan eksplisit. Model implisit digunakan berdasarkan pada *sampling design* yang digunakan dalam proses estimasi secara langsung. Selanjutnya, model eksplisit digunakan berdasarkan pada pengaruh acak (*random effect*) area kecil yang diperoleh keragaman variabel penyerta. Bentuk model eksplisit di antaranya adalah EBLUP dan *Empirical Bayes* (EB) (Rao, 2003).

Pendekatan *Empirical Best Linear Unbiased Prediction* (EBLUP)

Model campuran memiliki kelebihan dalam menduga kombinasi linear dari pengaruh tetap dan pengaruh acak. Teknik penyelesaian model pengaruh campuran untuk memperoleh Prediksi Tak Bias Linier Terbaik atau dikenal dengan *Best Linear Unbiased Prediction* (BLUP). Hasil estimasi parameter BLUP akan meminimumkan MSE di antara kelas-kelas estimasi parameter linier tak bias lainnya (Ghosh dan Rao, 1994). Asumsi yang digunakan dalam menghasilkan prediksi ini adalah komponen ragam diketahui. Faktanya, komponen ragam sulit bahkan tidak diketahui. Oleh karena itu, diperlukan estimasi terhadap komponen ragam tersebut melalui data sampel. Berikut merupakan bentuk *General Linear Mixed* (GLM) sebagai dasar pengembangan estimasi area kecil (Rao, 2003):

$$\theta = X\beta + Zv + e \quad (1)$$

dimana :

θ = vektor acak dari variabel respon yang terobservasi berukuran $n \times 1$

X = matriks dari variabel prediktor yang elemen-elemennya diketahui berukuran $n \times p$

- β = vektor parameter bersifat tetap yang tidak diketahui dan tidak terobservasi berukuran $p \times 1$
- Z = matriks dari variabel prediktor yang elemen-elemennya diketahui berukuran $n \times q$
- v = vektor acak parameter yang tidak diketahui dan tidak terobservasi berukuran $q \times 1$
- e = vektor *random error* yang tidak terobservasi berukuran $n \times 1$

EBLUP berbasis level area atau biasa disebut EBLUP *Fay Herriot* (EBLUP-FH) dituliskan sebagai berikut (Rao & Maolina, 2014):

$$\hat{\theta}_i = \mathbf{x}_i^T \beta + b_i v_i + e_i \quad (2)$$

dimana $\hat{\theta}_i$ merupakan variabel respon (yang diestimasi) dari area ke- i , $i = 1, \dots, m$, dengan m adalah jumlah area, \mathbf{x}_i^T merupakan vektor dari variabel prediktor, b_i merupakan konstanta yang bernilai 1, $v_i \sim N(0, \sigma_v^2)$ dengan σ_v^2 adalah varians dari pengaruh acak area, serta $e_i \sim N(0, \varphi_i^2)$ dengan φ_i^2 diketahui dari data sampel, dan v_i dan e_i saling independen. Adapun penduga BLUP dari θ_i dituliskan sebagai berikut (Rao & Maolina, 2015):

$$\tilde{\theta}_i^{BLUP} = \gamma_i \hat{\theta}_i + (1 - \gamma_i) \mathbf{z}_i^T \tilde{\beta} \quad (3)$$

dengan $\gamma_i = \frac{\sigma_v^2}{(\varphi_i + \sigma_v^2 b_i^2)}$.

Berikut merupakan koefisien regresi yang diduga dengan *Generalized Least Square (GLS)* (Rao, 2003):

$$\hat{\beta} = \hat{\beta}(\sigma_v^2) = \left[\sum_{i=1}^n \frac{\mathbf{x}_i \mathbf{x}_i^T}{(\varphi_i + \sigma_v^2 b_i^2)} \right]^{-1} \left[\sum_{i=1}^n \frac{\mathbf{x}_i \hat{\theta}_i}{(\varphi_i + \sigma_v^2 b_i^2)} \right] \quad (4)$$

Setelah dicari nilai penduga BLUP, untuk mengukur seberapa baik penduga BLUP maka dihitung nilai *Mean Square Error* (MSE) dengan rumus:

$$MSE(\tilde{\theta}_i^{BLUP}) = g_{1i}(\sigma_v^2) + g_{2i}(\sigma_v^2) \quad (5)$$

di mana $g_{1i}(\sigma_v^2) = \frac{\sigma_v^2 \varphi_i}{\varphi_i + \sigma_v^2} = \gamma_i \varphi_i$ dan

$$g_{2i}(\sigma_v^2) = (1 - \varphi_i)^2 \mathbf{x}_i^T \left[\sum_{i=1}^m \frac{\mathbf{x}_i \mathbf{x}_i^T}{(\varphi_i + \sigma_v^2 b_i^2)} \right]^{-1} \mathbf{x}_i.$$

Dalam praktiknya, varians pengaruh acak (σ_v^2) tidak diketahui, sehingga harus diduga terlebih dahulu. Metode yang dapat digunakan untuk estimasi varians pengaruh

acak (σ_v^2) adalah metode *Maximum Likelihood* (ML) atau bisa juga dengan *Residual Maximum Likelihood* (REML). Dalam penelitian ini menggunakan metode estimasi varians pengaruh acak dengan REML karena lebih robust dan sering digunakan dalam penelitian (Molina, Rao, dan Datta, 2015) sehingga diperoleh penduga baru sebagai berikut:

$$\hat{\theta}_i^{EBLUP} = \gamma_i \hat{\theta}_i + (1 - \gamma_i) \mathbf{z}_i^T \hat{\beta} \quad (6)$$

Nilai *Mean Square Error* (MSE) digunakan untuk mengukur seberapa baik penduga EBLUP, yaitu dengan rumus (Rao, 2003) :

$$MSE(\hat{\theta}_i^{EBLUP}) = g_{1i}(\hat{\sigma}_v^2) + g_{2i}(\hat{\sigma}_v^2) + 2g_{3i}(\hat{\sigma}_v^2) \quad (7)$$

dimana

$$g_{3i}(\hat{\sigma}_v^2) = \left[\frac{\psi_i^2 b_i^4}{(\psi_i + \hat{\sigma}_v^2 b_i^2)^4} \right] (\hat{\theta}_i - \mathbf{z}_i^T \tilde{\beta})^2 \times \bar{V}(\hat{\sigma}_v^2)$$

dan $\bar{V}(\hat{\sigma}_v^2) = 2m^{-2} \sum_{i=1}^m (\psi_i + \hat{\sigma}_v^2)^2$ merupakan ragam asimtot dari $\hat{\sigma}_v^2$.

Koefisien Variasi (CV)

CV merupakan ukuran kekonvergenan dari estimasi yang dihasilkan (Yang, 2008). Nilai CV estimasi distribusi *prior* salah satu model pada SAE dibandingkan dengan nilai CV estimasi langsung. Nilai CV diperoleh dengan membagi akar kuadrat varians dengan penduga dikali 100:

$$CV = \frac{\sqrt{MSE(\hat{\theta})}}{\hat{\theta}} \times 100 \quad (8)$$

Kebaikan Model

Dalam EBLUP-FH, alat untuk mengukur ketepatan model suatu set data digunakan beberapa ukuran. Dalam penelitian ini digunakan beberapa alat ukur yang mengacu pada beberapa penelitian sebelumnya dari Molina dan Marhuenda (2015) yaitu diantaranya adalah *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC):

$$AIC = -2\ln(\text{likelihood}) + 2(p + 1) \quad (9)$$

$$BIC = -2\ln(\text{likelihood}) + 2(p + 1)\log(i) \quad (10)$$

Selain itu, kebaikan model dari variabel penyerta di dalam model menurut

penelitian Lahiri dan Suntornchost (2015) seperti:

1. Simpangan Rata-rata (MD)

$$MD = \frac{\sum_{i=1}^n |y_i - \bar{y}|}{n} \quad (11)$$

2. Korelasi *Pearson*

$$r = \frac{(n \sum XY) - (\sum X \sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][(n \sum Y^2) - (\sum Y)^2]}} \quad (12)$$

3. *R-square* (R^2)

$$R^2 = \left(\frac{(n \sum XY) - (\sum X \sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][(n \sum Y^2) - (\sum Y)^2]}} \right)^2 \quad (13)$$

4. *R-square Adjusted* ($R\text{-adjusted}^2$)

$$R_{adj}^2 = 1 - \left[(1 - R^2) \left(\frac{n-1}{n-p-1} \right) \right] \quad (14)$$

Kemiskinan

BPS (2017) memandang kemiskinan sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Kebutuhan dasar meliputi kebutuhan sandang, pangan, dan papan. Menurut World Bank (2000), kemiskinan adalah kehilangan kesejahteraan (*deprivation of well being*). Ketidakmampuan memenuhi standar hidup rata-rata ditandai dengan kemampuan pendapatan yang rendah dalam memenuhi kebutuhan dasarnya. Hal ini akan berdampak kepada rendahnya kemampuan masyarakat dalam memenuhi kebutuhan selain kebutuhan dasar tersebut.

BPS (2017) mendefinisikan penduduk Miskin sebagai penduduk yang memiliki rata-rata pengeluaran perkapita per bulan di bawah garis kemiskinan. Penduduk yang memiliki rata-rata pengeluaran perkapita per bulan di bawah Garis Kemiskinan dikategorikan sebagai penduduk miskin. Garis kemiskinan (GK) adalah penjumlahan dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). GKM merupakan nilai pengeluaran kebutuhan minimum makanan yang disetarakan dengan 2100 kilokalori per kapita per hari. Paket komoditas kebutuhan dasar makanan diwakili oleh 52 jenis komoditas (padi-padian, umbi-umbian, ikan, daging, telur dan susu, sayuran, kacang-kacangan, buah-

buah, minyak dan lemak, dll). GKNM adalah kebutuhan minimum untuk perumahan, sandang, pendidikan dan kesehatan. Paket komoditas kebutuhan dasar non makanan diwakili oleh 51 jenis komoditas di perkotaan dan 47 jenis komoditas di perdesaan.

Rumus Penghitungan Garis Kemiskinan:

$$GK = GKM + GKNM \quad (15)$$

Efisiensi Relatif

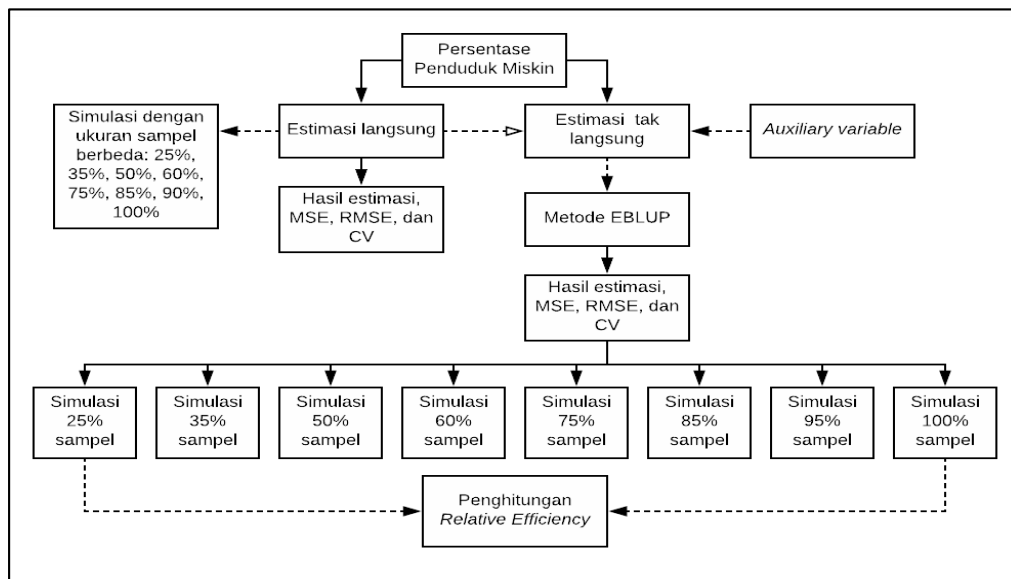
Suatu penduga dikatakan efisien jika memiliki varians yang lebih kecil di antara banyak penduga lainnya yang juga tak bias (Singh dan Chaudury, 1986). Pada penelitian ini dilakukan pengukuran efisiensi relatif antara penduga EBLUP di setiap alternatif ukuran sampel dengan penduga SRS dari seluruh sampel Survei Sosial Ekonomi Nasional (Susenas) di setiap kabupaten/kota.

$$RE = \frac{MSE(\hat{\theta}^{EBLUP})}{MSE(\hat{\theta}^{Langsung})} \quad (16)$$

Variabel Penyerta

Menurut Rao (2003) variabel penyerta yang baik adalah peubah yang berhubungan erat dengan peubah yang menjadi perhatian dan berasal dari data sensus atau data administratif. Salah satu contoh data sensus yang sering digunakan sebagai penyerta adalah Potensi Desa (Podes). Pemilihan variabel penyerta menggunakan ukuran korelasi antara variabel penyerta dengan variabel respons (Darsyah, 2013). Beberapa variabel penyerta yang memiliki pengaruh dan korelasi kuat dengan persentase kemiskinan di provinsi Nusa Tenggara Timur adalah Koperasi Industri Kecil dan Kerajinan Rakyat (X_1), Jumlah Koperasi Simpan Pinjam (X_2), Jumlah Kantor Pelayanan Nasabah Bank Perkreditan Rakyat (X_3), Jumlah Keluarga Tanpa Listrik PLN (X_4), Bahan Bakar Berupa Kayu untuk Memasak (X_5), Sumber Air Minum (X_6), dan Pemukiman Kumuh (X_7).

Tahapan Analisis



Gambar 1. Tahapan analisis

Tahapan analisis penelitian ini dapat dilihat pada Gambar 1 dengan uraian sebagai berikut:

1. Melakukan estimasi langsung dengan SRS beserta MSE-nya terhadap persentase penduduk miskin dengan cara melakukan pengkodean terhadap pengeluaran perkapita di setiap rumah tangga. Apabila nilainya lebih kecil atau sama dengan GK, maka diberikan kode 1. Apabila lebih besar diberikan kode 0. Kemudian dilakukan agregasi rata-rata penduduk miskin di setiap kabupaten/kota yang kemudian menjadi persentase penduduk miskin.
2. Memilih variabel penyerta yang berkorelasi dan signifikan terhadap variabel yang diteliti yang diperoleh dari Podes 2014 serta variabel penyerta berdasarkan penelitian terkait.
3. Penerapan pada pendekatan EBLUP
 - i. Melakukan uji normalitas terhadap hasil estimasi langsung dengan menggunakan Uji Lilliefors karena jumlah area yang diobservasi relatif kecil. Jika data tidak normal maka akan dilakukan transformasi ArcSin (Rao, 2003).
 - ii. Melakukan pendugaan ($\hat{\beta}$) dan varian pengaruh acak dengan menggunakan metode REML pada *software R*.
- iii. Menduga persentase kemiskinan MSE dengan metode estimasi tidak langsung (EBLUP-FH) dengan memasukkan komponen-komponen pada rumus.
- iv. Memeriksa asumsi-asumsi klasik seperti: normalitas residual, homoskedastis, dan non-multikolinearitas.
- v. Melihat kebaikan (*goodness*) model EBLUP-FH dari kriteria seperti R^2 , R^2 *adjusted*, simpangan rata-rata, AIC, dan BIC.
4. Melakukan estimasi parameter persentase penduduk miskin di setiap kabupaten/kota dengan melakukan *resampling* yakni menarik ulang sampel, yaitu:
 - i. Mengambil sampel dengan alternatif ukuran sampel 25, 35, 50, 60, 75, 85, dan 90 persen dari data sampel Susenas pada setiap kabupaten/kota dan diiterasi sebanyak B kali.
 - ii. Setiap iterasi pada ukuran sampel alternatif dilakukan estimasi persentase penduduk miskin setiap kabupaten/kota beserta MSE-nya dengan estimasi langsung menggunakan SRS.
 - iii. Persentase penduduk miskin di tiap

- kabupaten/kota dipasangkan dengan variabel penyerta yang sudah ditetapkan sebelumnya untuk diestimasi tidak langsung dengan metode EBLUP.
- iv. Mencari rata-rata estimasi persentase penduduk miskin di tiap kabupaten/kota beserta MSE dengan metode estimasi langsung SRS dan estimasi tidak langsung dengan metode EBLUP.
 - v. Membandingkan Hasil *resampling*

untuk 7 ukuran alternatif 25, 35, 50, 60, 75, 85 dan 90 persen dengan membandingkan MSE, RMSE, RRMSE antar penduga SRS dengan penduga SAE metode EBLUP.

5. Membuat kesimpulan.

HASIL DAN PEMBAHASAN

Estimasi Langsung

Tabel 1. Persentase Penduduk Miskin, MSE, dan Jumlah Rumah Tangga Sampel

Kabupaten/Kota	Penduduk Miskin (%)	MSE	Jumlah Rumah Tangga Sampel
(1)	(2)	(3)	(4)
Sumba Barat	28,18	0,0063	440
Sumba Timur	12,62	0,0054	515
Kupang	13,69	0,0050	555
Timor Tengah Selatan	16,13	0,0044	626
Timor Tengah Utara	23,85	0,0058	478
Belu	24,42	0,0054	516
Alor	18,78	0,0060	458
Lembata	25,98	0,0064	435
Flores Timur	7,35	0,0054	517
Sikka	18,25	0,0050	559
Ngada	2,05	0,0063	438
Manggarai	6,28	0,0050	557
Rote Ndao	9,11	0,0063	439
Manggarai Barat	12,11	0,0058	479
Sumba Tengah	15,25	0,0069	400
Sumba Barat Daya	38,46	0,0053	520
Nagekeo	20,71	0,0058	478
Manggarai Timur	2,08	0,0058	480
Sabu Raijua	13,78	0,0069	399
Malaka	26,25	0,0069	400

Estimasi langsung persentase penduduk miskin dilakukan pada 21 kabupaten dan 1 kota yang ada di Provinsi Nusa Tenggara Timur. Jumlah sampel yang diambil di setiap kabupaten menggunakan data Susenas Maret 2017. Proses estimasi langsung ini menggunakan metode SRS. Adapun hasil estimasi SRS dapat dilihat pada Tabel 1. Berdasarkan hasil estimasi langsung dengan metode SRS diperoleh kabupaten dengan persentase penduduk miskin tertinggi berada ialah Kabupaten Sumba Barat Daya dengan persentase 38,46 persen. Sedangkan kabupaten yang

memiliki persentase penduduk miskin terendah ialah Kabupaten Ngada dengan persentase sebesar 2,05 persen. Dapat dilihat juga nilai MSE setiap kabupaten/kota sehingga diperoleh rata-rata MSE kabupaten/kota di Provinsi NTT sebesar 0,0057.

Uji kenormalan data persentase penduduk miskin menunjukkan data berdistribusi normal sehingga dalam hal ini SAE dengan metode EBLUP dapat digunakan. Selanjutnya, pada tahapan estimasi persentase penduduk miskin area kecil dengan EBLUP membutuhkan

beberapa variabel penyerta yang diperoleh dari Podes. Pemilihan variabel penyerta dilakukan dengan cara mencari variabel yang berkorelasi signifikan dan variabel yang telah dibangun pada penelitian terdahulu terhadap persentase penduduk miskin. Variabel yang terpilih adalah X_1 , X_2 , X_3 , X_4 , X_5 , X_6 , X_7 . Berikut hasil estimasi koefisien dengan menggunakan metode EBLUP dapat dilihat pada Tabel 2.

Tabel 2. Hasil estimasi dengan metode EBLUP

Variabel	Estimasi beta	Standar error	t-value	p-value
(1)	(2)	(3)	(4)	(5)
X_0	0,2006	0,0451	4,4530	0,0000
X_1	0,0037	0,0025	1,4968	0,1344
X_2	-0,0012	0,0007	-1,8149	0,0695
X_3	-0,0155	0,0166	-0,9363	0,3491
X_4	0,0000	0,0000	0,6171	0,5371
X_5	-0,0003	0,0004	-0,7979	0,4249
X_6	0,0036	0,0017	2,0530	0,0401
X_7	0,0054	0,0101	0,5358	0,5921

Terlihat pada Tabel 3 bahwa hasil pengujian bahwa kriteria AIC, BIC, dan simpangan rata-rata menunjukkan nilai yang kecil, kriteria *R-squared Adjusted*, *R-squared*, dan korelasi menunjukkan nilai yang cukup kuat dikarenakan lebih dari 0,5 serta kriteria normalitas dan non-heteroskedastisitas tidak terjadi pelanggaran asumsi. Oleh karena itu, model yang terbentuk sudah tepat untuk digunakan dalam proses estimasi.

Tabel 3. Nilai Keباikan Suai model

Kriteria	Nilai/p-value	Keterangan
(1)	(2)	(3)
AIC	-45,78704	Nilai AIC kecil
BIC	-35,96766	Nilai BIC kecil
<i>R-squared Adjusted</i>	0,535992	Nilai R^2 -adj lebih dari 0,5
<i>R-squared</i>	0,6906613	Nilai R^2 lebih dari 0,5
Korelasi	0,831075	Nilai Korelasi kuat
Simpangan rata-rata	0,05580736	Nilai Simpangan rata-rata relatif kecil

Tabel 4 menunjukkan hasil estimasi EBLUP pada data Susenas untuk estimasi

persentase penduduk miskin masing-masing kabupaten dan kota di NTT:

Tabel 4. Hasil Estimasi Langsung dan Estimasi EBLUP

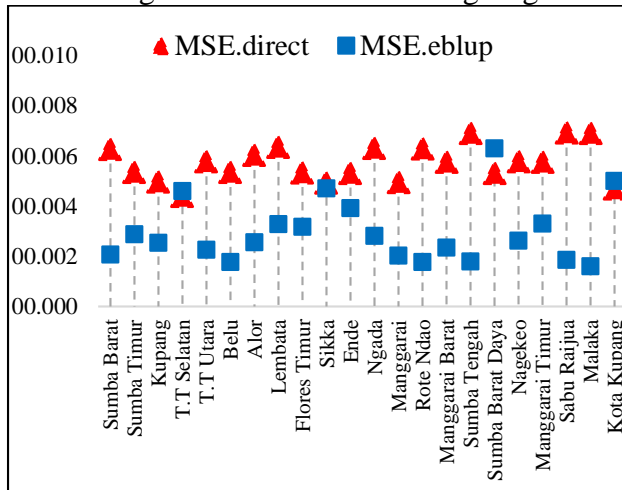
Kabupaten/Kota	Y direct	Y eblup
(1)	(2)	(3)
Sumba Barat	28,18	17,22
Sumba Timur	12,62	16,71
Kupang	13,69	15,42
Timor Tengah Selatan	16,13	17,22
Timor Tengah Utara	23,85	19,80
Belu	24,42	15,95
Alor	18,78	23,75
Lembata	25,98	22,64
Flores Timur	7,35	1,35
Sikka	18,25	20,50
Ende	9,25	12,20
Ngada	2,05	7,63
Manggarai	6,28	6,42
Rote Ndao	9,11	14,79
Manggarai Barat	12,11	5,62
Sumba Tengah	15,25	19,92
Sumba Barat Daya	38,46	37,40
Nagekeo	20,71	17,35
Manggarai Timur	2,08	8,05
Sabu Raijua	13,78	18,69
Malaka	26,25	20,84
Kota Kupang	5,62	10,06

Berdasarkan Tabel 4 dan Gambar 2, dapat dilihat nilai estimasi penduga langsung dan EBLUP serta nilai MSE-nya di setiap kabupaten/kota. Untuk dapat membandingkan besaran hasil estimasi dilakukan perbandingan terhadap nilai MSE yang diperoleh untuk setiap hasil estimasi. Terlihat dalam grafik bahwa MSE EBLUP memiliki nilai MSE yang lebih kecil dibandingkan dengan nilai MSE estimasi langsung. Selanjutnya, dilakukan pengujian simulasi pengurangan jumlah sampel rumah tangga untuk menguji tingkat efisiensi estimasi EBLUP.

Simulasi Estimasi Persentase Penduduk Miskin dengan SAE pada Level Kabupaten/Kota

Pemodelan yang telah dilakukan sebelumnya (di atas) dilakukan pada 100 persen ukuran sampel yang terpilih pada desain sampel Susenas yang diestimasi langsung dengan SRS. Pemodelan ini dilakukan untuk mengukur sejauh mana

kebaikan atau presisi estimasi SAE EBLUP dibandingkan metode estimasi langsung.

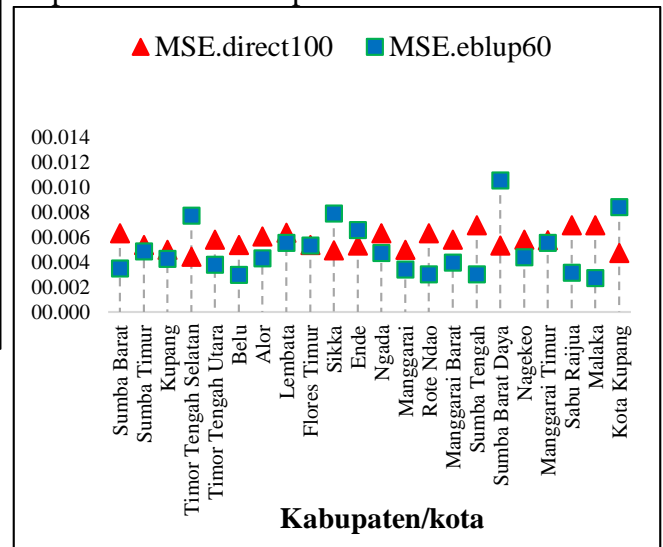


Gambar 2. MSE hasil estimasi langsung dan EBLUP (seluruh sampel)

Pada bagian sebelumnya, terlihat bahwa model SAE digunakan untuk estimasi suatu indikator pada area kecil. Untuk mengetahui tingkat performa model SAE terutama pada area yang memiliki ukuran sampel yang kurang dari yang ditetapkan pada desain sampel, maka pada penelitian ini dilakukan simulasi. Simulasi dilakukan dengan estimasi persentase penduduk miskin pada level kabupaten/kota dengan menggunakan ukuran sampel alternatif yaitu 25, 35, 50, 60, 75, 85, dan 90 persen. Pada setiap ukuran sampel alternatif, dilakukan penarikan sampel secara SRS di tiap-tiap kabupaten/kota sesuai dengan ukuran sampel alternatifnya sebanyak 1000 kali. Hal ini dilakukan menyerupai proses *bootstrapping* agar bisa memberikan keputusan mengenai presisi dari masing-masing ukuran sampel alternatif tersebut.

Pada Gambar 3 ditampilkan perbandingan hasil estimasi dan MSE simulasi dengan 60 dan 100 persen ukuran sampel. Terlihat bahwa grafik estimasi dari hasil simulasi 60 persen ukuran sampel mendekati hasil estimasi dari simulasi 100 persen ukuran sampel. Begitu pula dengan grafik MSE yang terbentuk, kedua simulasi

tersebut sudah memberikan hasil yang mendekati. Oleh karena itu, efisiensi SAE dengan 60 persen ukuran sampel sudah tercermin karena diperoleh presisi yang hampir sama dengan hasil simulasi 100 persen ukuran sampel.



Gambar 3. MSE hasil simulasi estimasi EBLUP (60 persen ukuran sampel) dan estimasi SRS (100 persen ukuran sampel)

Efisiensi Relatif

Hasil pengukuran dari efisiensi relatif bertujuan mengetahui ukuran sampel alternatif yang menghasilkan dugaan yang sama efisiennya antara SAE EBLUP-FH dengan penduga SRS dari seluruh ukuran sampel (100 persen). Efisiensi relatif antara tiap-tiap ukuran sampel alternatif dengan penduga SRS seluruh ukuran sampel di setiap kabupaten/kota menunjukkan bahwa pada ukuran sampel alternatif 60 persen memiliki efisiensi relatif 0,9102. Hal ini berarti penduga SAE dengan 60 persen ukuran sampel sudah mampu menghasilkan presisi yang hampir sama dengan penduga SRS untuk seluruh ukuran sampel Susenas tersebut. Berikut pada Tabel 5 diberikan hasil pengukuran efisiensi relatif dan hasil estimasi rata-rata beserta MSE untuk provinsi NTT.

Tabel 5. Hasil pengukuran Efisiensi Relatif (RE), rata-rata estimasi pada level provinsi beserta MSE

Ukuran sampel	RE	Rata-rata estimasi (%)	MSE
(1)	(2)	(3)	(4)
90 persen	0,6062	15,92	0,0057
85 persen	0,6514	15,88	0,0033
75 persen	0,7279	15,89	0,0035
60 persen	0,9102	15,89	0,0040
50 persen	1,0915	15,90	0,0050
35 persen	1,5615	15,88	0,0060
25 persen	2,1889	15,89	0,0120

KESIMPULAN

Pemodelan SAE yang tepat untuk mengestimasi persentase penduduk miskin pada level kabupaten/kota di Provinsi Nusa Tenggara Timur Tahun 2017 adalah: berdasarkan beberapa kriteria kebaikan model, model EBLUP yang terbangun menunjukkan hasil yang baik dengan variabel penyerta yang digunakan untuk mengestimasi ialah Jumlah Koperasi Industri Kecil dan Kerajinan Rakyat (Kopinkra) yang masih aktif/beroperasi (X_1), Jumlah Koperasi Simpan Pinjam (Kospin) yang masih aktif/beroperasi (X_2), Jumlah kantor pelayanan nasabah Bank Perkreditan Rakyat (BPR) di wilayah desa/kelurahan (X_3), Jumlah keluarga tanpa listrik (X_4), Bahan bakar berupa kayu bakar untuk memasak yang digunakan oleh sebagian besar keluarga (X_5), Sumber air untuk minum sebagian besar keluarga yang tidak layak (kolam, sungai, air hujan, dan lainnya) (X_6), Jumlah lokasi permukiman kumuh (X_7).

Pada ukuran sampel alternatif 60 persen, efisiensi relatif bernilai 0,9102 yang berarti bahwa estimasi SAE dengan ukuran sampel alternatif 60 persen dari sampel Susenas saja sudah mampu menghasilkan presisi yang hampir sama dengan estimasi

SRS dengan seluruh sampel Susenas tersebut.

Saran yang dapat diberikan untuk penelitian berikutnya agar dapat diterapkan pada survei kompleks sehingga dapat diketahui efisiensi dari tiap-tiap metode estimasi langsung. Selain itu, lebih selektif dalam pemilihan variabel penyerta yang digunakan pada SAE. Variabel penyerta yang digunakan diharapkan berkorelasi kuat dan signifikan di daerah yang diteliti agar dapat menggambarkan kemiskinan dengan lebih baik.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2017). Berita Resmi Statistik Profil Kemiskinan Provinsi Nusa Tenggara Timur Tahun 2014-2016. Nusa Tenggara Timur: BPS.
- Badan Pusat Statistik. (2018). Laporan Bulanan Data Sosial Ekonomi Edisi 93. Jakarta: BPS
- Badan Pusat Statistik. (2017). Potensi Ekonomi Provinsi Nusa Tenggara Timur. Nusa Tenggara Timur: BPS
- Badan Pusat Statistik. (2018). Definisi kemiskinan dan Penduduk Miskin. www.bps.go.id. Diakses pada bulan Maret 2018.
- Darsyah, M. Y. (2013). Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep Dengan

- Pendekatan Nonparametrik. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 1(2).
- Ghosh and Rao. (1994). *Small Area Estimation: An Appraisal*. Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to Statistical Science.
- Molina, I., & Marhuenda, Y. (2015). sae: An R package for small area estimation. *The R Journal*, 7(1), 81-98.
- Molina, I., Rao, J. N. K., & Datta, G. S. (2015). Small area estimation under a Fay–Herriot model with preliminary testing for the presence of random area effects. *Survey Methodology*, 41(1), 1-19.
- Lahiri, P., & Suntornchost, J. (2015). Variable selection for linear mixed models with applications in small area estimation. *Sankhya B*, 77(2), 312-320.
- Rao. J. (2003). *Small Area Estimation*. United States of America: John Wiley & Sons. Inc.
- Rao. J. (2014). *A Comparison of Small Area Estimation Methods for Poverty Mapping*. Calle Madrid: Universidad Carlos III de Madrid.
- Rao. JNK dan Molina (2015). *Small Area Estimation 2nd Edition*. New York: John Wiley and Sons. Inc.
- Sadik. K. dan Notodiputro. K.A. 2006. Small Area Estimation with Time and Area Effects Using Two Stage Estimation. Proceeding at the First International Conference on Mathematics and Statistics. MSMSSEA. 19-21 June 2006. Bandung. 32
- Singh, Daroga dan Chaudhary, Fauran S. (1986). *Theory and Analysis of Sample Survey Designs*, Wiley Eastern Limited
- Wijaya, A., Darsyah, M. Y., & Suprayitno, I. J. (2017). *Small Area Estimation Pada Tingkat Kemiskinan Di Provinsi Jawa Tengah Dengan Pendekatan Empirical Best Linier Unbiased Prediction*. In *Prosiding Seminar Nasional & Internasional*.
- Yang. Yuhong. (2007). *Consistency of Cross Validation for Comparing Regression Procedures*. Institute of
- Mathematical Statistics. University of Minnesota.

