

Small Area Estimation for Percentage of Out-of-School Children Aged 7-17 Years in Sumatera Island, 2023

Pendugaan Area Kecil untuk Persentase Anak Usia 7—17 Tahun yang Tidak Sekolah Level Kabupaten/Kota di Pulau Sumatera Tahun 2023

Wisly Ryan Eliezer¹, Aisyah ‘Azizah Nur Rahmah², Karina Himalaya³, Afidita Nabila Putri⁴, Aditya Prameswara Achmadi⁵, Azka Ubaidillah⁶, Shafiyah Asy Syahidah⁷

^{1,2,3,4,5,6} Program Studi D-IV Statistika Peminatan Statistika Ekonomi, ⁷ Badan Pusat Statistik RI

Email: wislyeliezer423@gmail.com, 212011312@stis.ac.id., 212011329@stis.ac.id, 212011398@stis.ac.id, 212011558@stis.ac.id, azka@stis.ac.id, asy.syahidah@bps.go.id

Abstract

Ensuring the quality of education is a fundamental commitment towards achieving sustainable development goals (SDGs). One effective strategy to enhance education quality is addressing the high number of children out of school. More precise district/city-level data on the percentage of out-of-school children needs to be provided. Estimation results from Susenas data show that Sumatera Island has the highest proportion of districts/cities with a Relative Standard Error (RSE) of over 25% compared to other islands in Indonesia. Therefore, this study applies Hierarchical Bayes (HB) Beta method by utilizing accompanying variables. The research reveals that the HB Beta estimator is the most effective in estimating the percentage of out-of-school children aged 7—17 years at the district/city level on Sumatera Island. The Small Area Estimation (SAE) model offers a more precise estimate than the direct estimator. Furthermore, there are 25 districts/cities with a high percentage of children aged 7—17 years who are not in school, with the majority located in the southern region of Sumatera Island.

Keywords: out-of-school children, small area estimation, hierarchical bayes beta

Abstrak

Menjamin kualitas pendidikan merupakan salah satu komitmen pada tujuan pembangunan berkelanjutan (SDGs). Salah satu strategi yang dapat dilakukan untuk meningkatkan kualitas pendidikan ada menekan besarnya angka anak yang tidak sekolah. Namun, belum terdapat data yang presisi untuk persentase anak yang tidak sekolah pada tingkat kabupaten/kota. Hasil estimasi pada data Susenas menunjukkan bahwa Pulau Sumatera memiliki persentase kabupaten/kota dengan nilai RSE lebih dari 25 persen yang paling tinggi dibandingkan pulau-pulau lainnya di Indonesia. Dalam mengatasi permasalahan tersebut, penelitian ini menerapkan



JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah 'Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

Hierarchical Bayes (HB) Beta dengan memanfaatkan bantuan variabel penyerta. Hasil penelitian menunjukkan bahwa penduga HB Beta menjadi penduga terbaik dalam mengestimasi persentase anak tidak sekolah 7—17 tahun level kabupaten/kota di Pulau Sumatera. Model SAE memberikan penduga yang lebih presisi dibandingkan penduga langsung. Selain itu, sebanyak 25 kabupaten/kota dengan nilai persentase anak usia 7—17 tahun yang tidak sekolah berkategori tinggi yang mayoritas berada pada kawasan selatan Pulau Sumatera.

Kata kunci: anak tidak sekolah, penduga area kecil, hierarchical bayes beta

1. PENDAHULUAN

Pendidikan merupakan kunci utama dalam pembangunan suatu negara [17]. Melalui pendidikan, seseorang akan memperoleh pengetahuan dan keterampilan untuk mengembangkan potensi diri dalam beradaptasi dengan kemajuan ilmu pengetahuan dan teknologi sehingga mampu bersaing dengan negara-negara di dunia [21]. Selain berdaya saing tinggi, sumber daya manusia yang berkualitas juga akan meningkatkan produktivitas dan efisiensi di berbagai sektor industri sehingga mampu mendorong pertumbuhan ekonomi [9].

Dengan demikian, menjamin kualitas pendidikan yang inklusif dan merata serta meningkatkan kesempatan belajar sepanjang hayat telah menjadi salah satu komitmen mendasar negara-negara di seluruh dunia di bawah program Tujuan Pembangunan Berkelanjutan atau *Sustainable Development Goals* [4]. Program ini mencakup 17 tujuan dengan 169 target yang telah disepakati oleh 193 negara termasuk Indonesia. Hal tersebut menggarisbawahi pengakuan universal terhadap pendidikan sebagai pondasi dalam pembangunan berkelanjutan.

Adapun salah satu indikator pada tujuan keempat SDGs adalah angka anak tidak sekolah [3]. UNESCO menyatakan telah terjadi peningkatan signifikan sebesar enam juta anak putus sekolah sehingga total terdapat 250 juta anak kehilangan akses terhadap pendidikan pada tahun 2023 [19]. Sementara itu, angka anak tidak sekolah di Indonesia tahun 2023 pada jenjang SD/ sederajat adalah 0,67%, SMP/ sederajat sebesar 6,93%, dan SMA/ sederajat mencapai 21,61% [2].

Strategi untuk meningkatkan kualitas pendidikan salah satunya adalah dengan menekan besarnya angka anak yang tidak sekolah. Ketersediaan data persentase anak yang tidak sekolah yang akurat diperlukan untuk merumuskan program dan kebijakan yang efektif dan tepat sasaran. Salah satu badan yang memiliki tugas untuk mengumpulkan data statistik dasar di Indonesia adalah Badan Pusat Statistik (BPS). Berkaitan dengan data terkait anak yang tidak sekolah maka dapat menggunakan data mikro yang berasal dari Survei Sosial Ekonomi Nasional (Susenas), yakni persentase anak usia 7—17 tahun yang tidak sekolah. Namun, estimasi pada level penyajian terendah kabupaten/kota masih memiliki nilai *Relative Standard Error* (RSE) yang lebih dari 25% bahkan 50 persen.

Berdasarkan data hasil pendugaan langsung oleh BPS, 99 dari 154 kabupaten/kota di Pulau Sumatera (atau sebesar 64,286 persen wilayah) memiliki nilai RSE yang lebih dari 25 persen. Bahkan, 13 kabupaten/kota diantaranya (atau sebesar 8,442 persen) memiliki nilai RSE lebih dari 50 persen. Selain itu, perlu diketahui bahwa persentase kabupaten/kota dengan nilai RSE lebih dari 25 persen bahkan 50 persen di Pulau Sumatera merupakan yang tertinggi dibandingkan pulau-pulau lainnya di Indonesia. Hal ini mengakibatkan analisis dari hasil pendugaan langsung tidak dapat dipercaya. Oleh karena itu, penggunaan pendugaan tidak langsung diperlukan untuk memperoleh penduga yang lebih presisi pada area kecil.

Data dengan presisi tinggi diperlukan untuk membuat kebijakan dalam rangka penurunan persentase anak berusia 7—17 tahun yang tidak sekolah. Rentang usia 7-17 tahun tercantum pada

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

data Survei Sosial Ekonomi Nasional (SUSENAS) Badan Pusat Statistik yang mencakup anak-anak yang menempuh pendidikan dasar, hal ini juga tertuang dalam *sustainable development goals (SDGs)* terkait dengan penyediaan data anak usia 7-17 tahun yang tidak sekolah. Anak yang berusia 7—17 tahun yang tidak sekolah menjadi subpopulasi dari anak yang tidak sekolah sehingga jumlah sampelnya dapat lebih kecil dari populasi asalnya. Sejalan dengan pendapat [23], ukuran sampel yang sudah cukup untuk memperoleh penduga langsung pada total populasi belum tentu cukup untuk memperoleh penduga langsung pada subpopulasinya. Presisi yang rendah dapat ditangani dengan meningkatkan ukuran sampel. Akan tetapi, peningkatan ukuran sampel dapat menimbulkan kendala berupa biaya yang semakin tinggi, tenaga yang semakin banyak, maupun waktu pelaksanaan survei yang makin lama. Alternatif penyelesaian dari masalah tersebut adalah penggunaan metode estimasi tidak langsung (*indirect estimation*). Metode estimasi tidak langsung memanfaatkan informasi sampel pada area lain yang dikaitkan dengan informasi tambahan (*auxiliary variables*) melalui persamaan model statistik atau yang disebut dengan metode *Small Area Estimation (SAE)*.

Small Area Estimation (SAE) merupakan metode untuk mengestimasi parameter di wilayah atau domain kecil, seperti kabupaten atau kecamatan. SAE membagi wilayah menjadi area-area kecil dan menggunakan informasi dari luar area untuk memperbaiki estimasi parameter. Area kecil didefinisikan sebagai wilayah yang tidak dapat diwakili oleh survei yang luas dan membutuhkan metode estimasi yang lebih efektif untuk mengestimasi parameter di wilayah tersebut [23]. Terdapat beberapa penelitian yang menggunakan metode SAE dalam menduga angka putus sekolah. Penelitian yang dilakukan oleh [16] mengestimasi angka anak putus sekolah di Kabupaten Semarang dengan metode EBLUP. Selain itu, [20] mengkaji penerapan metode Hierarchical Bayes angka anak putus sekolah yang berasal dari rumah tangga miskin. Namun, belum terdapat penelitian dengan metode SAE untuk mengestimasi persentase anak usia 7—17 tahun yang tidak sekolah khususnya di Pulau Sumatera.

Berdasarkan permasalahan yang telah diuraikan sebelumnya, diketahui bahwa terdapat kebutuhan data terkait persentase anak usia 7—17 tahun yang tidak sekolah di level kabupaten/kota yang lebih presisi. Oleh karena itu penelitian ini bertujuan untuk (i) menganalisis gambaran umum penduga langsung, (ii) memperoleh pendugaan area kecil menggunakan SAE HB Beta, (iii) membandingkan hasil penduga langsung dengan penduga tidak langsung SAE HB Beta, dan (iv) memetakan penduga persentase anak usia 7—17 tahun yang tidak sekolah berdasarkan metode estimasi terbaik. Dengan demikian, penelitian ini berkontribusi dalam mengembangkan data *official statistics* dengan menggunakan metode SAE untuk perencanaan pembangunan pendidikan di Indonesia, dengan fokus pada Pulau Sumatera.

2. METODE PENELITIAN

2.1. Landasan Teori

Anak Usia 7-17 Tahun yang Tidak Sekolah

Mengacu pada Stranas ATS, definisi anak tidak sekolah yaitu anak berusia 7—18 tahun dengan kondisi tidak pernah sekolah, putus sekolah pada jenjang tertentu, dan menyelesaikan jenjang tertentu tapi tidak melanjutkan dan sebab-sebab lain seperti kesulitan akses geografi, bencana alam, dan daerah konflik [24]. Sementara menurut Badan Pusat Statistik, definisi operasional anak berusia 7—17 tahun yang tidak sekolah adalah ART berumur 7—17 tahun ke atas yang tidak terdaftar atau tidak aktif mengikuti pendidikan, baik di suatu jenjang pendidikan formal maupun non-formal (Paket A/B/C) hal ini digunakan untuk memfokuskan analisis pada tingkat pendidikan dasar. Selanjutnya variabel penyerta yang digunakan untuk mengestimasi indikator anak usia 7—17 tahun yang tidak sekolah pada level kabupaten/kota di Pulau Sumatera diantaranya adalah variabel-variabel yang memiliki korelasi terhadap indikator tersebut dan digolongkan ke dalam beberapa

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

faktor, yakni faktor ekonomi seperti pada penelitian [25], [29], faktor sosial pada penelitian [12], [28], faktor teknologi pada penelitian [8], [11], faktor lingkungan pada penelitian [22], dan faktor infrastruktur pada penelitian [10], [18].

Small Area Estimation Hierarchical Bayes (HB) Beta

Bayesian merupakan teori yang menggunakan dua distribusi dan satu fungsi, yaitu distribusi *posterior* yang merupakan perpaduan antara distribusi *prior* (informasi awal mengenai parameter yang biasanya didefinisikan di awal pemodelan) dan data observasi yang digunakan untuk menyusun fungsi *Likelihood* [5]. Menurut [23], dengan mengaplikasikan teorema Bayesian, didapatkan:

$$f(\theta, \lambda|y) = \frac{f(y, \theta|\lambda)f(\lambda)}{f_1(y)} \quad (2.1)$$

dengan $f_1(y)$ merupakan fungsi marginal dari y yang dapat dituliskan:

$$f_1(y) = \int f(y, \theta|\lambda)f(\lambda)d\theta d\lambda \quad (2.2)$$

Distribusi *posterior* $f(y, \theta|\lambda)$ pada persamaan 2.1 masih mengandung dua parameter. Ketika ingin melakukan estimasi pada parameter yang diamati θ , maka dapat dicari melalui:

$$f(\theta|y) = \int f(\theta, \lambda|y)d\lambda = \int f(\theta|y, \lambda)f(\lambda|y)d\lambda \quad (2.3)$$

dimana $f(\lambda)$ merupakan distribusi *prior* yang ditentukan secara subjektif oleh peneliti dengan parameter modelnya λ . Tujuannya adalah untuk memperoleh distribusi posterior ($f(\theta|y)$) dari parameter area kecil yang ingin diestimasi (θ) berdasarkan data yang tersedia (y).

Melalui persamaan 2.3, terlihat bahwa dalam memperoleh distribusi *posterior* dan mendapatkan besarannya membutuhkan integrasi yang bersifat multidimensi. Salah satu metode yang dapat mengatasi kesulitan komputasi dalam menyelesaikan integrasi berdimensi tinggi pada HB adalah metode *Markov Chain Monte Carlo* (MCMC). Metode ini membangkitkan sampel dari distribusi *posterior* dan menggunakan sampel yang disimulasikan untuk memperkirakan besaran distribusi *posterior* yang diinginkan [23].

Ketika data yang digunakan berupa proporsi dan distribusinya cenderung asimetris, maka pemodelan HB dapat menggunakan distribusi Beta. Model SAE HB Beta dapat dirumuskan sebagai berikut [14]:

1. *Sampling* model

$$\hat{\theta}_i | \theta_i \sim \text{Beta}(a_i, b_i), i = 1, \dots, m \quad (2.4)$$

dengan θ_i merupakan parameter proporsi, $\hat{\theta}_i$ penduga proporsi pada area kecil ke- i sedangkan $a_i = \theta_i k$ dan $b_i = (1 - \theta_i)k$ merupakan parameter distribusi beta. k yang muncul pada parameter beta merupakan konstanta yang diasumsikan berdistribusi $k \sim \text{Gamma}(g_1, g_2)$. Di mana g_1 dan g_2 ditentukan nilainya atau bersifat *fixed*.

2. *Linking* model

$$\text{logit} \left(\frac{a_i}{a_i + b_i} \right) | \beta, \sigma_v^2 \sim N(\mathbf{x}_i^T \beta, \sigma_v^2), i = 1, \dots, m \quad (2.5)$$

$$(2.6)$$

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

$$\text{logit}(\theta_i) | \beta, \sigma_v^2 \sim N(\mathbf{x}_i^T \beta, \sigma_v^2), i = 1, \dots, m$$

dengan $v_i \sim N(0, \sigma_v^2)$, $\sigma_v^2 \sim IG(c_1, c_2)$, dan inisial $\sigma_v^2 = 1$.

3. β dan σ_v^2 saling bebas dengan $\beta \sim N(\mu_\beta, \sigma_\beta^2)$, di mana inisial $\mu_\beta = 0$ dan $\sigma_\beta^2 = 1$ dengan β merupakan vektor pengaruh tetap atau vektor koefisien regresi berdistribusi dan σ_v^2 merupakan varians *area random effect*.

Ukuran kebaikan pada penduga sae HB Beta dapat dilihat melalui RSE yang dapat dihitung dengan formula berikut:

$$RSE(\hat{\theta}_i^{HB}) = \frac{\sqrt{\text{Var}(\phi_i^{HB})}}{\phi_i^{HB}} \times 100\% \quad (2.7)$$

dengan keterangan sebagai berikut :

$RSE(\hat{\theta}_i^{HB})$: *Relative standard error* penduga sae HB Beta

$\text{Var}(\phi_i^{HB})$: Varians penduga sae HB Beta

(ϕ_i^{HB}) : Penduga sae HB Beta

2.2. Sumber Data dan Variabel Penelitian

Penelitian ini menggunakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS). Data yang digunakan adalah data Survei Sosial Ekonomi Nasional (Susenas) KOR Maret 2023 untuk memperoleh *interest variable*, yakni persentase anak berusia 7—17 tahun yang tidak sekolah. Selain itu, digunakan data Potensi Desa (Podes) 2021 untuk, Daerah Dalam Angka 2023, dan Big Data (*The Payne Institute and its Research Areas*) untuk memperoleh *auxiliary variable* (variabel penyerta), dari data Potensi Desa diperoleh variabel-variabel yang terkait dengan infrastruktur pendidikan seperti jarak dari desa ke sekolah terdekat beserta jumlah sekolah untuk seluruh jenjang pendidikan, kemudian variabel terkait infrastruktur kesehatan seperti banyaknya fasilitas kesehatan seperti rumah sakit, puskesmas, tenaga kesehatan, dan lainnya beserta dengan jarak dari desa ke fasilitas kesehatan terdekat, kemudian infrastruktur wilayah seperti keberadaan kelompok pertokoan, hotel, restoran, industri-industri dan infrastruktur keamanan desa seperti jarak menuju pos polisi terdekat, dan jumlah anggota linmas. Kemudian diperoleh juga variabel-variabel yang menggambarkan teknologi seperti proporsi desa dengan sinyal internet yang kuat, proporsi desa dengan keberadaan warga yang menggunakan telepon seluler, serta proporsi desa dengan bangunan rumah dilalui SUTET, proporsi desa dengan penerangan jalan menggunakan tenaga surya, dan untuk faktor ekonomi diperoleh variabel proporsi desa dengan sumber penghasilan utama kepala keluarga berasal dari lapangan usaha pertanian dan jumlah warga desa yang berprofesi sebagai TKI. Dari publikasi Daerah Dalam Angka 2023 diperoleh variabel-variabel yang berhubungan dengan pendidikan seperti rasio guru terhadap murid dan rasio murid terhadap bangunan sekolah, sementara dari Big Data diperoleh variabel *nighttime light* untuk menggambarkan tingkat perekonomian suatu wilayah. Kemudian wilayah yang menjadi cakupan dalam penelitian ini adalah 154 kabupaten/kota di Pulau Sumatera yang meliputi Provinsi Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, dan Kepulauan Riau.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

2.3. Tahapan Penelitian

Penelitian ini menggunakan dua metode analisis, yaitu analisis deskriptif dan analisis inferensia. Analisis deskriptif akan menggambarkan secara umum mengenai karakteristik data persentase anak berusia 7—17 tahun yang tidak sekolah serta variabel-variabel pesertanya, yang digambarkan melalui tabel, grafik, dan peta tematik. Sedangkan analisis inferensia dilakukan untuk menghitung penduga tidak langsung persentase anak berusia 7—17 tahun yang tidak sekolah pada level kabupaten/ kota dengan menggunakan metode SAE HB Beta. Proses pengolahan dalam analisis ini menggunakan bantuan software *R-Studio*.

Secara lengkap, langkah analisis dapat diuraikan sebagai berikut:

1. Melakukan estimasi langsung persentase anak berusia 7-17 tahun yang tidak sekolah pada level kabupaten/kota di 10 provinsi di Pulau Sumatera, dalam penelitian ini hasil estimasi langsung level kabupaten/kota telah disediakan oleh BPS. Menurut [23], estimasi langsung merupakan penghitungan pendugaan parameter populasi di suatu domain yang langsung menggunakan data sampel berdasarkan basis rancangan (*design based*). Teknik pendugaan design based merupakan teknik estimasi yang dilakukan sesuai dengan desain penarikan sampel (*sampling design*) dari pengumpulan data. *Sampling design* yang digunakan dalam survei-survei BPS bervariasi didasarkan pada kondisi lapangan dan permasalahan teknis lainnya. Adapun survei yang digunakan sebagai sumber data untuk memperoleh estimasi langsung dalam penelitian ini adalah Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2023 dengan *sampling scheme* sebagai berikut :

Tabel 2.1. *Sampling scheme* estimasi langsung

Tahap	Unit	Jumlah unit strata h		Metode pemilihan sampel	Peluang	Fraksi sampling	Keterangan
		Populasi	Sampel				
1	Blok Sensus	N_{gh}	n_{gh}	PPS, <i>size</i> M_{ghi}	$\frac{M_{ghi}}{M_{gh}}$	$n_{gh} \frac{M_{ghi}}{M_{gh}}$	Pembentukan <i>Master Sampling Frame</i> blok sensus
		n_{gh}	n'_{gh}	<i>Systematic</i>	$\frac{1}{n_{gh}}$	$\frac{n'_{gh}}{n_{gh}}$	Pemilihan sampel blok sensus Maret
2	Rumah Tangga	M_{ghi}^{up}	\bar{m}	<i>Systematic</i>	$\frac{1}{M_{ghi}^{up}}$	$\frac{\bar{m}}{M_{ghi}^{up}}$	Pemilihan sampel rumah tangga

Sumber : Susenas

Keterangan :

N_{gh} : jumlah populasi blok sensus di strata ke- h kabupaten/kota ke- g

n_{gh} : jumlah blok sensus yang digunakan sebagai *Master Sampling Frame* di strata ke- h kabupaten/kota ke- g ($h=1$ untuk daerah perkotaan, $h=2$ untuk daerah perdesaan)

n'_{gh} : jumlah sampel blok sensus Susenas Maret di strata ke- h kabupaten/kota ke- g

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

- M_{ghi} : jumlah populasi keluarga pada blok sensus ke-i strata ke-h kabupaten/kota ke-g pada master frame
- M_{gh} : jumlah populasi keluarga pada strata ke-h kabupaten/kota ke-g pada master frame
- M_{ghi}^{up} : jumlah muatan rumah tangga hasil pemutakhiran di blok sensus ke-i strata ke-h kabupaten/kota ke-g
- \bar{m} : jumlah sampel rumah tangga di setiap blok sensus

Persentase anak usia 7-17 tahun yang tidak sekolah diestimasi dengan menggunakan metode two stages one phase stratified sampling. Angka estimasi yang dihasilkan dari data susenas 2023 dihitung dengan menggunakan penimbang *sampling (sampling weight)* final yang sudah dilakukan berbagai *adjustment*. Berikut merupakan rumus yang digunakan dalam penghitungan sebagai berikut :

$$\hat{Y}_{gh} = \sum_{g=1}^G \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{ghij}^{(adj)} y_{ghij} \quad (2.8)$$

$$\hat{X}_{gh} = \sum_{g=1}^G \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{ghij}^{(adj)} x_{ghij} \quad (2.9)$$

Dengan $w_{ghij}^{(adj)}$ adalah pembobot pada kab/kota ke-g, strata ke-h, diblok Sensus (BS) ke-i dan rumah tangga ke-j. Kemudian, y_{ghij} adalah jumlah anak usia 7-17 tahun yang tidak sekolah pada kab/kota ke-g, strata ke-h, di Blok Sensus (BS) ke-i, dan rumah tangga ke-j serta x_{ghij} adalah jumlah anak usia 7-17 tahun pada kabupaten/kota ke-g, kelompok umur ke-h di blok sensus (BS) ke-i, dan rumah tangga ke-j. Persentase Anak usia 7-17 Tahun yang tidak sekolah adalah perbandingan jumlah anak usia 7-17 tahun yang tidak sekolah terhadap jumlah anak usia 7-17 tahun.

2. Melakukan pemodelan SAE HB Beta meliputi, penyeleksian variabel penyerta menggunakan *stepwise regression*, pengecekan asumsi non-multikolinieritas melalui nilai VIF yang kurang dari 10, selanjutnya menentukan banyaknya *iterasi*, *update*, *thin*, dan *burn-in* hingga mendapatkan model SAE yang konvergen.
3. Melakukan evaluasi model melalui RSE hasil penduga pemodelan SAE sebagaimana pada persamaan 2.7 dengan RSE penduga langsung.
4. Melakukan validasi hasil penduga pemodelan SAE level kabupaten/kota dengan angka hasil penduga langsung level provinsi yang telah diberikan oleh BPS.
5. Melakukan pemetaan persentase anak berusia 7—17 tahun yang tidak sekolah berdasarkan model terbaik menggunakan peta tematik.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila
Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

3. HASIL DAN PEMBAHASAN

3.1. Gambaran Umum Persentase Anak Usia 7-17 Tahun yang Tidak Sekolah di kabupaten/Kota di Pulau Sumatera

Penduga langsung yang telah diperoleh melalui proses penghitungan dengan memanfaatkan *sampling weight* pada Susenas 2023 disajikan melalui statistik deskriptif berikut:

Tabel 3.1. Statistik deskriptif pendugaan langsung dan RSE persentase anak usia 7—17 tahun yang tidak sekolah di kabupaten/kota di Pulau Sumatera

Statistik Deskriptif	Persentase Anak Usia 7—17 Tahun yang Tidak Sekolah (%)	RSE (%)
Minimum	0,28	16,04
Kuartil Pertama (Q1)	2,20	23,33
Median	3,83	29,24
Rata-rata	4,19	32,32
Kuartil Ketiga (Q3)	5,71	37,82
Maksimum	14,07	70,51

Sumber : Susenas 2023, data diolah

Tabel 3.1. menunjukkan gambaran umum penduga langsung persentase anak usia 7—17 tahun yang tidak sekolah level kabupaten/kota di pulau Sumatera beserta dengan RSEnya. Penduga langsung memiliki nilai kemencengan (*skewness*) yang lebih dari 0 yaitu 0,77 dan nilai keruncingan (*kurtosis*) yang lebih dari 3 yaitu 3,74. Artinya, penduga langsung cenderung asimetris dan menceng kanan. Sehingga digunakan model SAE HB dengan sampling model berdistribusi Beta dan fungsi penghubung logit. Selain itu, RSE penduga langsung memiliki rentang yang cukup besar, dengan rata – rata 32,32% dan nilai maksimum 70,51% (Banda Aceh), nilai rata-rata yang cukup tinggi tersebut menandakan bahwa hasil pendugaan langsung belum presisi.

3.2. Pemilihan Variabel Penyerta

Sebelum melakukan pemodelan SAE HB Beta dilakukan seleksi variabel penyerta melalui proses *stepwise regression* antara variabel penyerta dengan logit dari penduga langsung, kemudian dilakukan penyeleksian variabel dengan cara mereduksi variabel yang tidak signifikan dan memilih variabel dengan nilai *variance inflation factor* (VIF) yang kurang dari 10 agar mencegah adanya multikolinearitas antar variabel penyerta. Metode proses seleksi variabel penyerta dengan *stepwise regression* dilakukan dengan memasukan variabel independen yang mempunyai kontribusi terbesar terhadap variabel dependen, hal ini dilakukan secara terus menerus sampai seluruh variabel independen mempunyai kontribusi yang signifikan, adapapun proses yang dilalui menurut [27] adalah sebagai berikut :

1. Memasukkan variabel penyerta pertama dalam model dan menguji apakah variabel penyerta tersebut signifikan atau tidak terhadap variabel respon, kemudian memasukkan variabel penyerta kedua ke dalam model, apabila variabel penyerta kedua tidak signifikan maka proses berhenti.
2. Tahap selanjutnya ingin dilihat apakah variabel penyerta pertama akan tetap signifikan berada dalam model setelah variabel penyerta kedua masuk ke dalam model, apabila pengaruh variabel penyerta pertama setelah variabel penyerta kedua masuk dalam model tidak signifikan, maka variabel penyerta pertama dikeluarkan dari model, namun apabila

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah 'Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

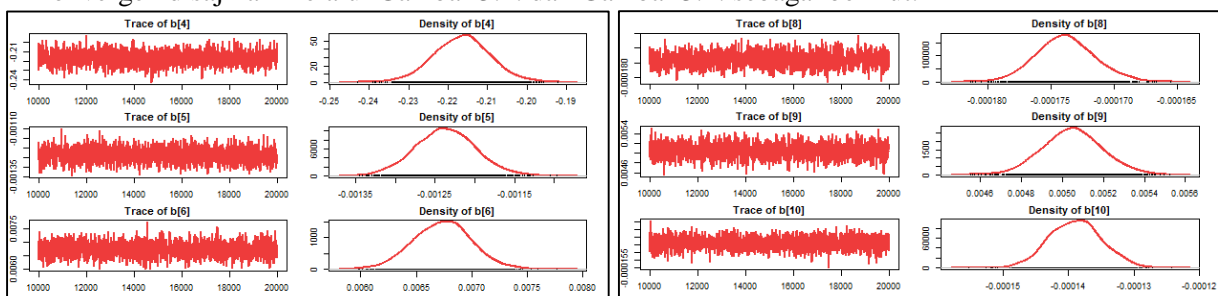
sebaliknya maka variabel penyerta pertama tetap masuk dalam model sampai pengujian tahap berikutnya.

3. Tahap selanjutnya menyeleksi variabel penyerta berikutnya yang akan dimasukkan ke dalam model, misal variabel penyerta ketiga. Ketika variabel penyerta ketiga signifikan dalam model, dilihat kembali apakah variabel penyerta pertama dan kedua masih boleh tetap berada di dalam model, atau harus dikeluarkan.
4. Begitu seterusnya sampai seluruh variabel diuj/seleksi.

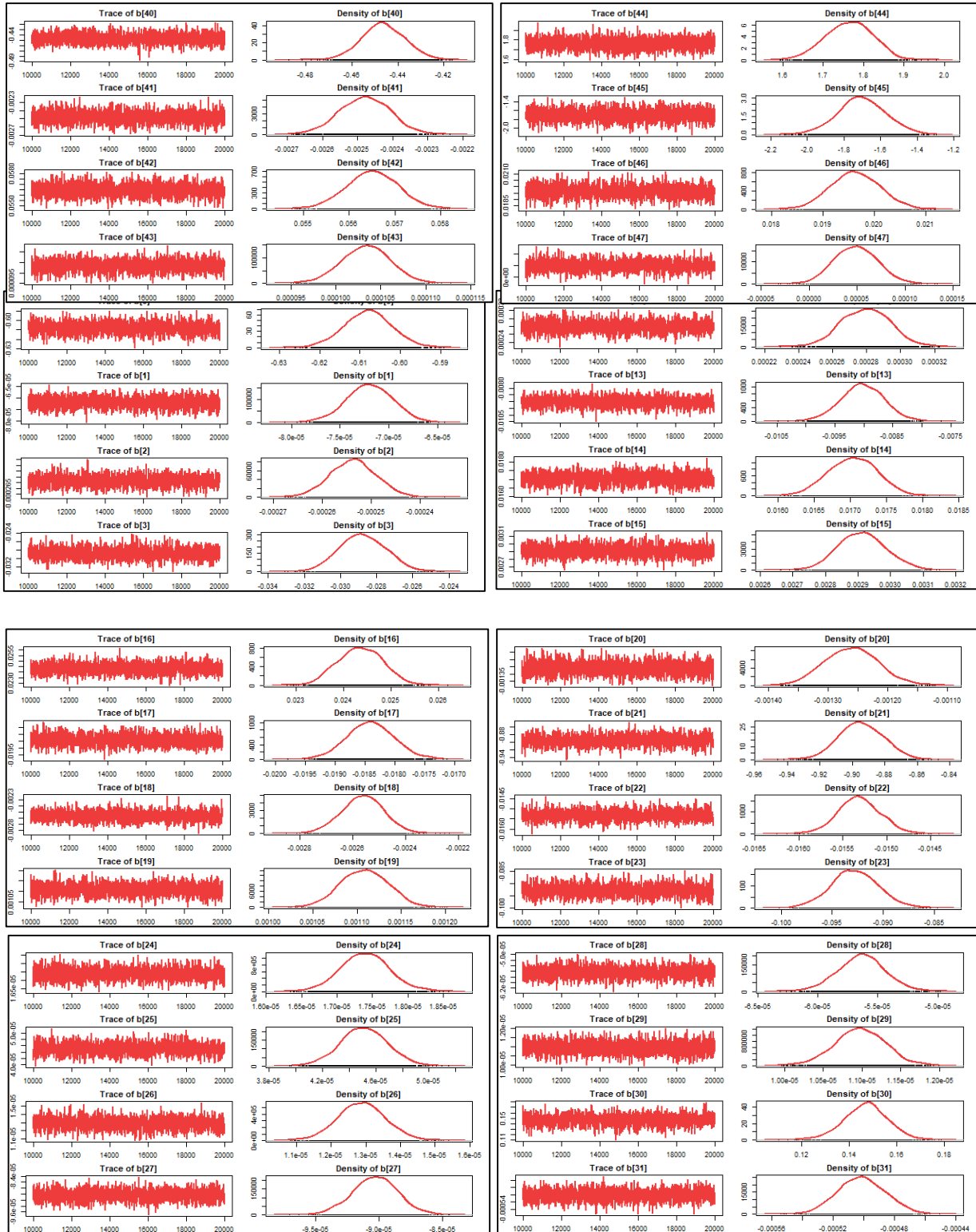
Dari 109 variabel penyerta yang diajukan, pada model akhir diperoleh 47 variabel yang secara bersamaan signifikan memengaruhi logit penduga langsung dan memiliki nilai VIF di bawah 10, diantaranya ialah jumlah warga desa yang bekerja sebagai TKI (x_1), jumlah tenaga kesehatan (x_3), jumlah pabrik gula (x_5), proporsi desa dengan sumber penghasilan utama penduduk dari lapangan usaha pertanian (x_6), jumlah kelompok pertokoan (x_9), jumlah pasar dengan bangunan permanen (x_{10}), jumlah pasar tanpa bangunan (x_{12}), jumlah restoran (x_{14}), jumlah kedai (x_{15}), jumlah hotel (x_{16}), jumlah industri tekstil (x_{21}), jumlah industri pakaian (x_{22}), jumlah industri minuman (x_{26}), jumlah industri tembakau (x_{27}), jumlah industri kertas (x_{28}), jumlah industri reparasi (x_{32}), jumlah rumah sakit (x_{33}), jumlah puskesmas dengan rawat inap (x_{35}), jumlah puskesmas tanpa rawat inap (x_{36}), jumlah puskesmas pembantu (x_{37}), jumlah tempat praktik bidan (x_{41}), jumlah pos kesehatan desa (x_{42}), jarak menuju sd terdekat (x_{44}), jarak menuju mi terdekat (x_{45}), jarak menuju smp terdekat (x_{46}), jarak menuju rumah sakit bersalin terdekat (x_{53}), jarak menuju puskesmas dengan rawat inap terdekat (x_{54}), jarak menuju poliklinik terdekat (x_{57}), jarak menuju tempat praktik dokter terdekat (x_{58}), jarak menuju pos kesehatan desa terdekat (x_{61}), jarak menuju pondok bersalin desa terdekat (x_{62}), proporsi desa dengan akses air minum aman (x_{65}), jumlah bangunan rumah dibawah saluran udara tegangan ekstra tinggi (SUTET) (x_{68}), proporsi desa dengan adanya keberadaan warga yang menggunakan telepon seluler (x_{75}), proporsi desa dengan sinyal internet telepon seluler 4G/LTE (x_{79}), proporsi desa yang mengalami kejadian tanah longsor (x_{80}), proporsi desa yang mengalami kejadian gempa bumi (x_{83}), proporsi desa yang mengalami kejadian angin puyuh (x_{86}), proporsi desa yang mengalami kejadian gunung meletus (x_{87}), proporsi desa dengan keberadaan tindak kejahatan pemerkosaan dalam 1 tahun terakhir (x_{92}), proporsi desa dengan keberadaan tindak kejahatan pembunuhan dalam 1 tahun terakhir (x_{93}), *nighttime light* (x_{97}), proporsi desa dengan penerangan jalan menggunakan tenaga surya (x_{98}), jumlah embung (x_{102}), jarak menuju pos polisi terdekat (x_{103}), jumlah anggota linmas (x_{104}), rasio guru terhadap murid SMA (x_{109}).

3.3. Pemodelan SAE HB Beta

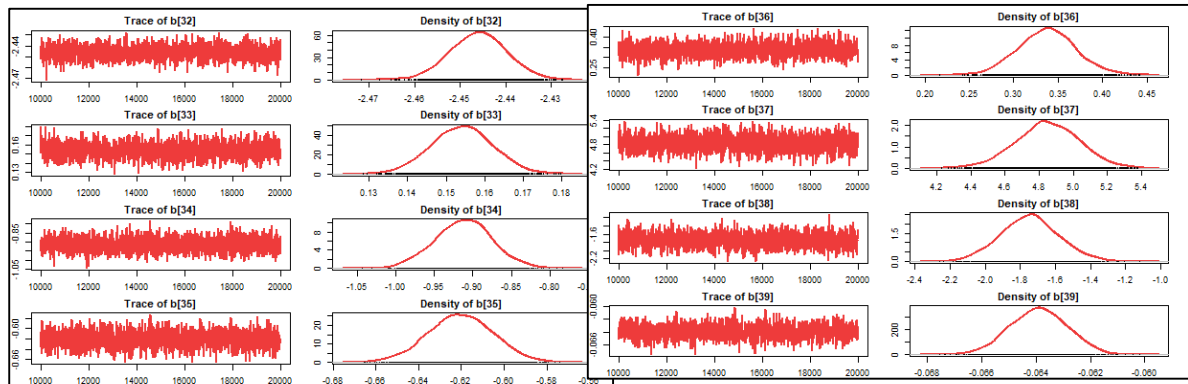
Pemodelan SAE HB Beta dilakukan dengan memanfaatkan proses simulasi *Markov Chain Monte Carlo* (MCMC). Proses tersebut dapat menyelesaikan proses integrasi yang kompleks dalam pemodelan SAE HB Beta. Distribusi posterior yang akurat akan diperoleh ketika kondisi *equilibrium* tercapai, yakni saat algoritma telah konvergen. Setelah dilakukan iterasi sebanyak 20.000, *burn-in* sebanyak 10.000, *thin* sebanyak 5, dan *update* sebanyak 75 maka kondisi konvergen dapat tercapai. Kondisi tersebut ditandai dengan *trace plot* yang telah stasioner atau tidak ada pola yang *periodic*, *density plot* yang bentuknya cenderung *smooth* dan menyerupai kurva normal, serta *autocorrelation plot* yang berbentuk *cut off* setelah lag pertama. Visualisasi untuk memperlihatkan kondisi konvergen disajikan melalui Gambar 3.1. dan Gambar 3.2. sebagai berikut.



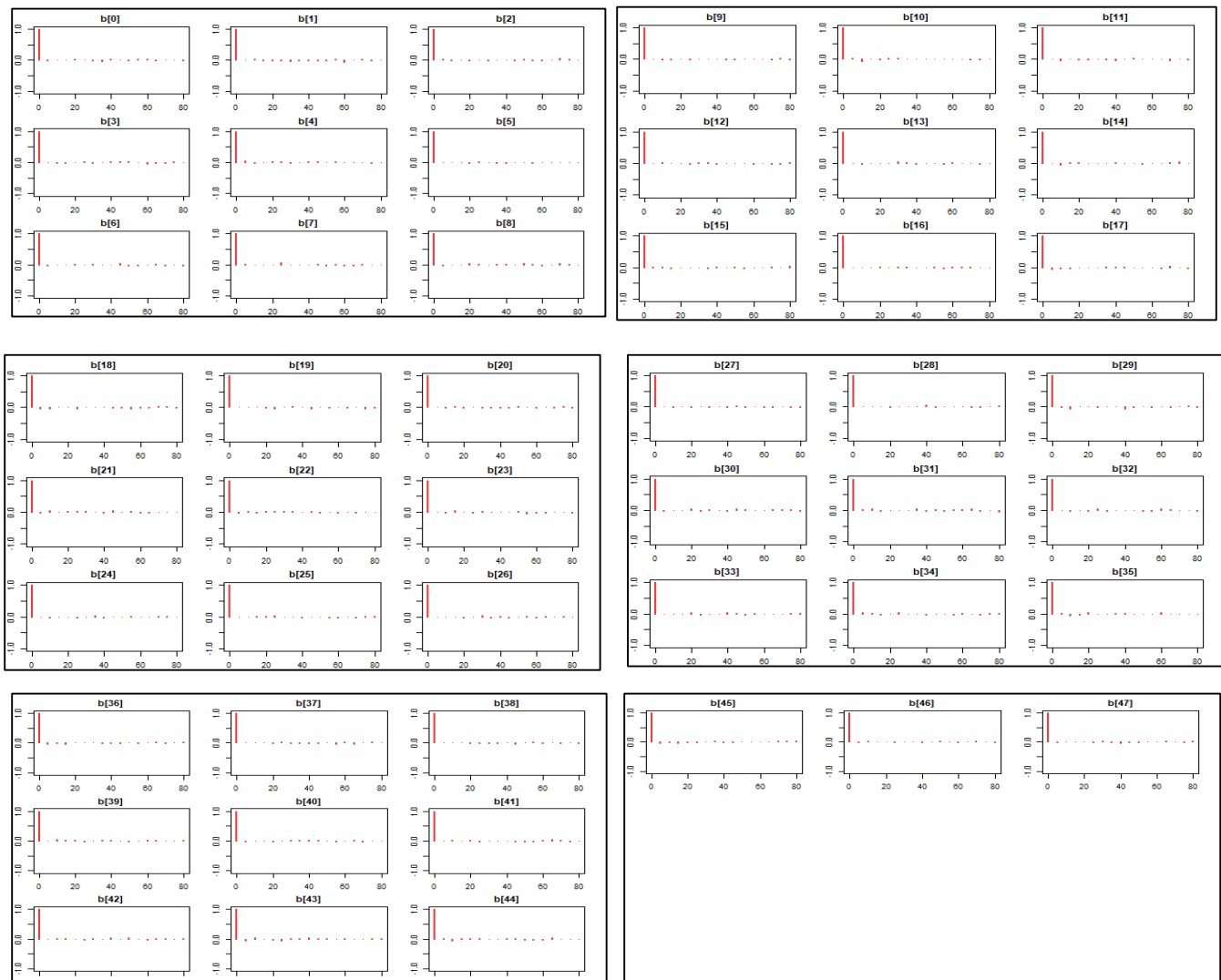
JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
 Wisly Ryan Eliezer, Aisyah 'Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah



JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
 Wisly Ryan Eliezer, Aisyah 'Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila
 Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah



Gambar 3.1. Trace plot dan Density plot untuk seluruh parameter



Gambar 3.2. Autocorrelation plot untuk seluruh parameter

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
**Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila
 Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah**

Berikutnya akan dilihat signifikansi variabel penyerta terhadap pendugaan SAE HB Beta, dalam model HB Beta signifikansi pada model dapat dilihat melalui *credible interval* yang berada di rentang 2,5% hingga 97,5%. Variabel penyerta dikatakan signifikan jika nilai *credible interval* pada rentang tersebut tidak memuat nilai nol.

Estimasi *random effect variance* yang dihasilkan sebesar 0.0794 dan estimasi nilai parameternya disajikan melalui Tabel 3.2. berikut.

Tabel 3.2. Estimasi parameter penduga SAE HB Beta

Variabel	Mean	2,50%	97,50%	Variabel	Mean	2,50%	97,50%
(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
β_0	$-6,080 \times 10^{-1}$	$-6,196 \times 10^{-1}$	$-5,967 \times 10^{-1}$	x_{53}	$1,740 \times 10^{-5}$	$1,672 \times 10^{-5}$	$1,808 \times 10^{-5}$
x_1	$-7,198 \times 10^{-5}$	$-7,661 \times 10^{-5}$	$-6,760 \times 10^{-5}$	x_{54}	$4,513 \times 10^{-5}$	$4,171 \times 10^{-5}$	$4,848 \times 10^{-5}$
x_3	$-2,536 \times 10^{-4}$	$-2,625 \times 10^{-4}$	$-2,446 \times 10^{-4}$	x_{57}	$1,289 \times 10^{-5}$	$1,157 \times 10^{-5}$	$1,423 \times 10^{-5}$
x_5	$-2,878 \times 10^{-2}$	$3,117 \times 10^{-2}$	$-2,624 \times 10^{-2}$	x_{58}	$-9,029 \times 10^{-5}$	$-9,382 \times 10^{-5}$	$-8,677 \times 10^{-5}$
x_6	$-2,165 \times 10^{-1}$	$-2,302 \times 10^{-1}$	$-2,025 \times 10^{-1}$	x_{61}	$-5,635 \times 10^{-5}$	$-6,024 \times 10^{-5}$	$-5,263 \times 10^{-5}$
x_9	$-1,234 \times 10^{-3}$	$-1,309 \times 10^{-3}$	$-1,163 \times 10^{-3}$	x_{62}	$1,097 \times 10^{-5}$	$1,033 \times 10^{-5}$	$1,162 \times 10^{-5}$
x_{10}	$6,735 \times 10^{-3}$	$6,234 \times 10^{-3}$	$7,237 \times 10^{-3}$	x_{65}	$1,470 \times 10^{-1}$	$1,284 \times 10^{-1}$	$1,646 \times 10^{-1}$
x_{12}	$6,583 \times 10^{-3}$	$6,236 \times 10^{-3}$	$6,914 \times 10^{-3}$	x_{68}	$5,013 \times 10^{-4}$	$-5,305 \times 10^{-4}$	$-4,713 \times 10^{-4}$
x_{15}	$-1,738 \times 10^{-4}$	$-1,782 \times 10^{-4}$	$-1,6912 \times 10^{-4}$	x_{75}	$-24,460 \times 10^{-1}$	$-24,580 \times 10^{-1}$	$-24,340 \times 10^{-1}$
x_{16}	$5,049 \times 10^{-3}$	$4,764 \times 10^{-3}$	$5,341 \times 10^{-3}$	x_{79}	$1,541 \times 10^{-1}$	$1,390 \times 10^{-1}$	$1,695 \times 10^{-1}$
x_{21}	$-1,390 \times 10^{-4}$	$-1,467 \times 10^{-4}$	$-1,313 \times 10^{-4}$	x_{80}	$-9,121 \times 10^{-1}$	$-9,844 \times 10^{-1}$	$-8,396 \times 10^{-1}$
x_{22}	$5,443 \times 10^{-4}$	$5,236 \times 10^{-4}$	$5,655 \times 10^{-4}$	x_{83}	$-6,204 \times 10^{-1}$	$-6,501 \times 10^{-1}$	$-5,913 \times 10^{-1}$
x_{26}	$2,796 \times 10^{-4}$	$2,525 \times 10^{-4}$	$3,070 \times 10^{-4}$	x_{86}	$3,375 \times 10^{-1}$	$2,770 \times 10^{-1}$	$4,015 \times 10^{-1}$
x_{27}	$-9,026 \times 10^{-3}$	$-9,750 \times 10^{-3}$	$-8,311 \times 10^{-3}$	x_{87}	$48,488 \times 10^{-1}$	$44,774 \times 10^{-1}$	$52,011 \times 10^{-1}$
x_{28}	$1,703 \times 10^{-2}$	$1,640 \times 10^{-2}$	$1,766 \times 10^{-2}$	x_{93}	$-17,463 \times 10^{-1}$	$-20,597 \times 10^{-1}$	$-14,132 \times 10^{-1}$
x_{32}	$2,913 \times 10^{-3}$	$2,776 \times 10^{-3}$	$3,060 \times 10^{-3}$	x_{97}	$-6,391 \times 10^{-2}$	$-6,596 \times 10^{-2}$	$-6,192 \times 10^{-2}$
x_{35}	$2,436 \times 10^{-2}$	$2,347 \times 10^{-2}$	$2,522 \times 10^{-2}$	x_{98}	$-4,463 \times 10^{-1}$	$-4,643 \times 10^{-1}$	$-4,280 \times 10^{-1}$
x_{36}	$-1,841 \times 10^{-2}$	$-1,916 \times 10^{-2}$	$-1,762 \times 10^{-2}$	x_{102}	$-2,479 \times 10^{-3}$	$-2,620 \times 10^{-3}$	$-2,341 \times 10^{-3}$
x_{37}	$-2,562 \times 10^{-3}$	$-2,717 \times 10^{-3}$	$-2,408 \times 10^{-3}$	x_{103}	$5,650 \times 10^{-2}$	$5,539 \times 10^{-2}$	$5,759 \times 10^{-2}$
x_{41}	$1,111 \times 10^{-3}$	$1,056 \times 10^{-3}$	$1,160 \times 10^{-3}$	x_{104}	$1,303 \times 10^{-4}$	$9,813 \times 10^{-5}$	$1,085 \times 10^{-4}$
x_{42}	$-1,261 \times 10^{-3}$	$-1,343 \times 10^{-3}$	$-1,173 \times 10^{-3}$	x_{109}	$17,642 \times 10^{-1}$	$16,508 \times 10^{-1}$	$18,700 \times 10^{-1}$
x_{44}	$-8,946 \times 10^{-1}$	$-9,206 \times 10^{-1}$	$-8,682 \times 10^{-1}$	x_{92}	$-17,022 \times 10^{-1}$	$-19,670 \times 10^{-1}$	$-14,371 \times 10^{-1}$
x_{45}	$-1,531 \times 10^{-2}$	$-1,578 \times 10^{-2}$	$-1,483 \times 10^{-2}$	x_{33}	$1,963 \times 10^{-2}$	$1,871 \times 10^{-2}$	$2,059 \times 10^{-2}$
x_{46}	$-9,276 \times 10^{-2}$	$-9,732 \times 10^{-2}$	$-8,803 \times 10^{-2}$	x_{14}	$4,870 \times 10^{-5}$	$5,907 \times 10^{-6}$	$9,378 \times 10^{-5}$

Berdasarkan Tabel 3.2., dapat dilihat bahwa seluruh variabel penyerta memiliki nilai estimasi koefisien parameter yang berada pada range 2,5% hingga 97,5% yang tertera masing-masing pada kolom (3) dan kolom (4) tidak ada yang melewati angka nol untuk semua parameternya. Hal ini memiliki arti bahwa seluruh variabel penyerta signifikan memengaruhi pemodelan SAE HB Beta untuk persentase anak usia 7-17 tahun yang tidak sekolah dengan arah pengaruh yang tetap. Berdasarkan hasil estimasi parameter penduga SAE HB Beta yang diperoleh pada Tabel 3.2., diatas terlihat bahwa variabel-variabel yang mewakili faktor ekonomi seperti jumlah warga desa yang

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah 'Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

bekerja sebagai TKI dan proporsi desa dengan sumber penghasilan utama penduduk dari lapangan usaha pertanian memiliki pengaruh negatif terhadap persentase anak usia 7-17 tahun yang tidak sekolah seperti tertera pada kolom (2) nilai *mean* estimasi koefisien variabel tersebut memiliki nilai negatif, hal ini berarti peningkatan variabel tersebut akan menurunkan persentase anak usia 7-17 tahun yang tidak sekolah, hasil ini sejalan oleh penelitian yang dilakukan oleh [15], [25], faktor ekonomi disini dijelaskan bahwa pada rumah tangga miskin yang memiliki tabungan rendah dan pendapatan yang rendah, mungkin akan sangat terkendala dalam menyekolahkan anak-anak mereka, sehingga anak mereka terpaksa tidak mengikuti kegiatan persekolahan, sehingga semakin baik kondisi perekonomian maka dapat menurunkan angka anak tidak sekolah, variabel-variabel yang menggambarkan tingkat perekonomian suatu wilayah seperti *nighttime light* (cahaya malam), jumlah pabrik gula, jumlah kedai, jumlah industri tekstil dan jumlah industri tembakau juga memiliki pengaruh negatif terhadap persentase anak tidak sekolah seperti tertera pada kolom (2) nilai *mean* estimasi koefisien variabel tersebut memiliki nilai negatif. Namun, variabel-variabel lain yang menggambarkan tingkat perekonomian suatu wilayah pada Tabel 3.2 seperti jumlah pasar, jumlah industri pakaian, jumlah industri minuman, dan jumlah industri reparasi memiliki pengaruh positif seperti tertera pada kolom (2) nilai *mean* estimasi koefisien variabel tersebut memiliki nilai positif, hal ini berarti peningkatan variabel tersebut akan meningkatkan persentase anak usia 7-17 tahun yang tidak sekolah, hasil ini sejalan dengan penelitian yang dilakukan oleh [26] yang menyatakan bahwa anak-anak dari keluarga miskin yang tinggal disekitar kawasan industri cenderung memilih untuk tidak melanjutkan sekolah dan bekerja di pasar atau di pabrik untuk membantu meningkatkan perekonomian. Variabel-variabel yang mewakili faktor teknologi seperti proporsi desa dengan adanya keberadaan warga desa yang menggunakan telepon seluler memiliki pengaruh negatif terhadap persentase anak tidak sekolah seperti tertera pada kolom (2) Tabel 3.2 nilai *mean* estimasi koefisien variabel tersebut memiliki nilai negatif, hal ini berarti peningkatan variabel tersebut akan menurunkan persentase anak usia 7-17 tahun yang tidak sekolah, hasil ini sejalan dengan penelitian yang dilakukan oleh [8], peningkatan akses terhadap teknologi dapat meningkatkan partisipasi anak dalam mengikuti sekolah dengan sistem pembelajaran yang lebih menarik dan efisien. Variabel-variabel yang mewakili faktor infrastruktur dalam hal ini infrastruktur kesehatan seperti jumlah tenaga kesehatan, jumlah puskesmas, jumlah pos kesehatan desa memiliki pengaruh negatif terhadap persentase anak tidak sekolah seperti tertera pada kolom (2) Tabel 3.2 nilai *mean* estimasi koefisien variabel tersebut memiliki nilai negatif, hal ini berarti peningkatan variabel tersebut akan menurunkan persentase anak usia 7-17 tahun yang tidak sekolah, hasil tersebut sejalan penelitian yang dilakukan oleh [7]. Semakin banyak dan mudahnya akses terhadap infrastruktur pendidikan akan meningkatkan angka partisipasi sekolah siswa. Variabel yang menggambarkan kondisi lingkungan di suatu wilayah seperti kejadian bencana alam dalam hal ini angin puyuh dan gunung meletus memiliki pengaruh positif terhadap persentase anak tidak sekolah seperti tertera pada kolom (2) Tabel 3.2 nilai *mean* estimasi koefisien variabel tersebut memiliki nilai positif, hal ini berarti peningkatan variabel tersebut akan meningkatkan persentase anak usia 7-17 tahun yang tidak sekolah, hasil ini sejalan oleh penelitian yang dilakukan oleh [22] bahwa peristiwa bencana yang dialami anak berpengaruh terhadap pendidikannya. Bencana alam dan dampak kerusakan rumah akibat bencana alam berpengaruh signifikan terhadap putus sekolah.

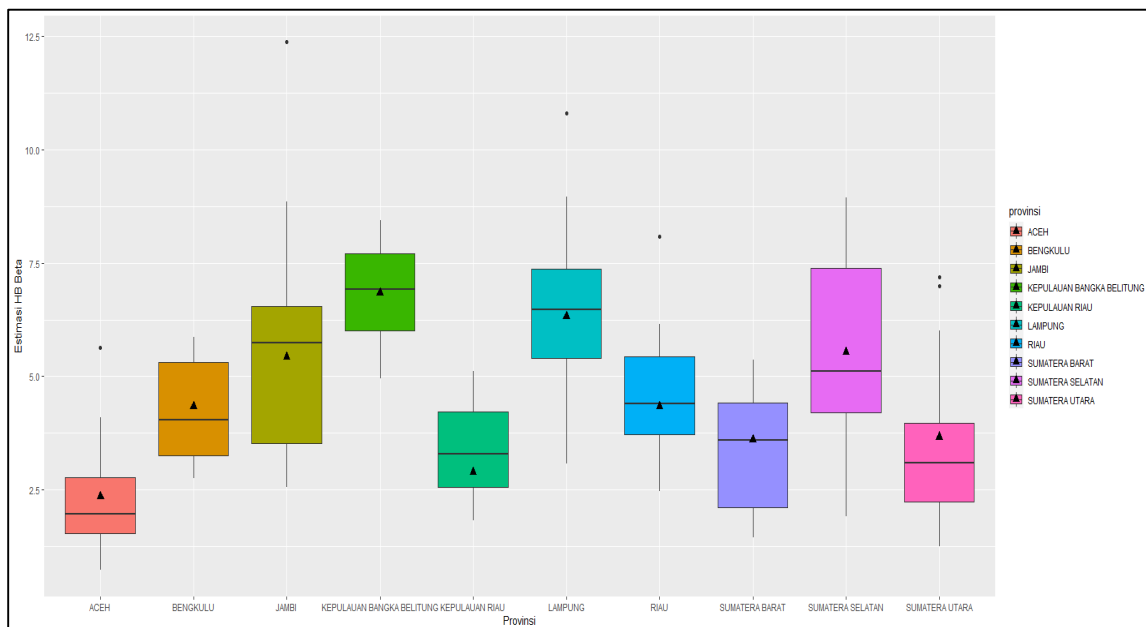
Di samping itu, untuk melihat gambaran terkait hasil penduga dari penduga HB Beta secara keseluruhan maka disajikan statistik deskriptif sebagai berikut.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

Tabel 3.3. Statistik deskriptif pendugaan dan RSE persentase anak usia 7—17 tahun yang tidak sekolah dengan metode HB Beta

Statistik Deskriptif	Persentase Anak Usia 7—17 Tahun yang Tidak Sekolah (%)	RSE (%)
Minimum	0,73	11,77
Kuartil Pertama (Q1)	2,57	15,74
Median	3,85	17,72
Rata-rata	4,19	17,95
Kuartil Ketiga (Q3)	5,58	19,80
Maksimum	12,37	24,75

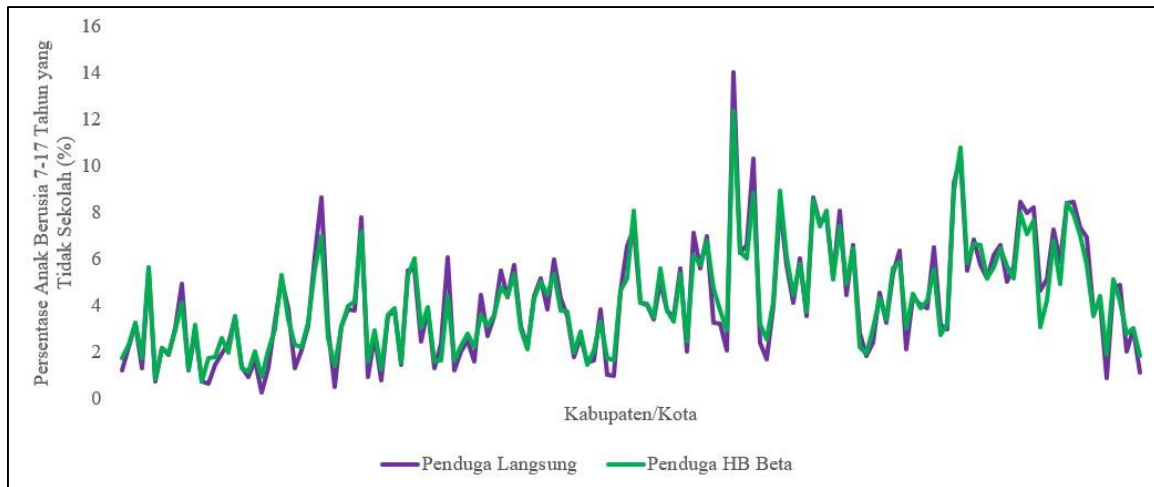
Melalui Tabel 3.3., terlihat bahwa nilai RSE yang dihasilkan dari model HB Beta berada pada rentang 11,77% hingga 24,75% dengan rata-rata sebesar 17,95%. Rendahnya nilai RSE tersebut mengindikasikan bahwa hasil pendugaan dengan metode HB Beta memberikan presisi yang tinggi dibandingkan pendugaan langsung, tidak adanya nilai RSE yang diatas 25% juga menandakan bahwa penduga HB Beta reliabel untuk digunakan [1].



Gambar 3.3. Deteksi validasi model HB Beta

Selain memiliki RSE yang rendah, hasil estimasi model SAE juga harus valid pada level penyajian di atasnya atau penduga langsung area besarnya. Dalam konteks penelitian ini, penduga SAE persentase anak usia 7—17 tahun yang tidak sekolah disuatu kabupaten harus memuat angka proporsi penduduk miskin level provinsi yang telah dipublikasikan oleh BPS. Terlihat pada Gambar 3.3. bahwa model HB Beta sudah valid dikarenakan nilai penduga SAE HB Beta kabupaten/kota sudah mencakup nilai estimasi pada level provinsi untuk seluruh provinsi di Pulau Sumatera.

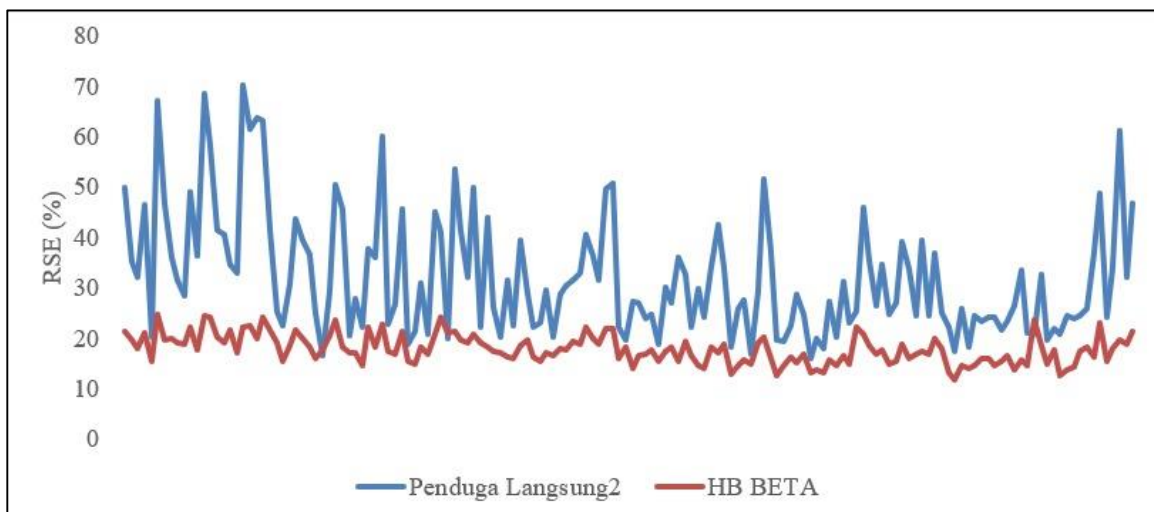
JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah



Gambar 3.4. Perbandingan penduga langsung dan HB Beta persentase anak usia 7—17 tahun yang tidak sekolah

Gambar 3.4. menunjukkan bahwa pemodelan SAE HB Beta menghasilkan penduga dengan pola yang serupa dengan pendugaan langsung. Persamaan pola ini menunjukkan penduga model SAE yang dihasilkan bersifat tidak bias dari nilai sebenarnya [6].

3.4. Perbandingan Pendugaan Langsung dan Pendugaan SAE HB Beta dalam memodelkan Persentase Anak Usia 7—17 Tahun yang Tidak Sekolah Kabupaten/Kota di Pulau Sumatera.



Gambar 3.5. Perbandingan RSE/RRMSE penduga langsung dan HB Beta persentase anak usia 7—17 tahun yang tidak sekolah

Evaluasi model dilakukan dengan membandingkan nilai RSE/RRMSE dari masing-masing penduga. Supaya lebih terlihat perbedaannya maka perbandingannya disajikan melalui grafik garis. Berdasarkan Gambar 3.5. dapat dilihat bahwa metode SAE HB Beta memberikan penurunan RSE

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
**Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila
 Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah**

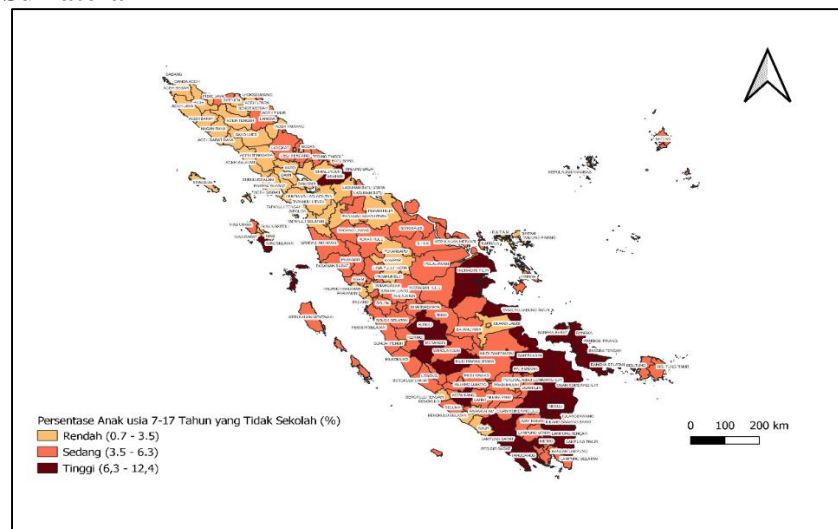
dibandingkan dengan penduga langsung. Secara umum RRMSE dari penduga HB Beta lebih rendah dibandingkan dengan RSE penduga langsung, hasil ini sejalan dengan penelitian yang dilakukan oleh [13] yang menemukan bahwa penduga SAE dalam hal ini HB Beta menghasilkan estimasi yang lebih presisi dengan standar error yang lebih kecil dibandingkan dengan pendugaan langsung. Ringkasan jumlah kabupaten/kota berdasarkan kelompok RSE dan metode pendugaan yang digunakan disajikan melalui Tabel 3.4.

Tabel 3.4. Jumlah kabupaten/kota berdasarkan kelompok RSE dan metode estimasi persentase anak usia 7—17 tahun yang tidak sekolah di Pulau Sumatera tahun 2023

Kelompok RSE	Jumlah Kabupaten/Kota	
	Penduga Langsung	Penduga SAE
RSE < 25%	55	154
25% < RSE ≤ 50%	86	0
RSE > 50%	13	0

Berdasarkan Tabel 3.4., terlihat bahwa penduga HB Beta mampu menghasilkan nilai estimasi yang lebih unggul sebagaimana ditandai dengan nilai RSE untuk seluruh wilayah kabupaten dan kota di Pulau Sumatera telah berada di bawah 25 persen. Berdasarkan hasil evaluasi model yang telah dilakukan, dapat disimpulkan bahwa penduga HB Beta merupakan penduga yang terbaik jika dibandingkan dengan penduga langsung. Oleh karena itu, penduga HB Beta akan digunakan sebagai dasar pemetaan untuk melihat gambaran persentase anak usia 7—17 tahun yang tidak sekolah di Pulau Sumatera berdasarkan aspek kewilayahan.

3.5. Pemetaan Persentase Anak Usia 7—17 Tahun yang Tidak Sekolah Kabupaten/Kota di Pulau Sumatera



Gambar 3.6. Peta tematik (*choropleth map*) persentase anak usia 7—17 tahun yang tidak sekolah kabupaten/kota di pulau sumatera 2023 berdasarkan penduga HB Beta

Untuk melihat kondisi persentase penyandang disabilitas miskin pada setiap kabupaten dan kota di Pulau Sumatera, maka disajikan peta tematik dalam bentuk *choropleth map*. Nilai persentase dikelompokkan menjadi tiga kategori dengan menggunakan metode *Natural Breaks*.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

Berdasarkan Gambar 3.6., dapat dilihat bahwa sebagian besar kabupaten atau kota di Pulau Sumatera memiliki persentase anak usia 7—17 tahun yang tidak sekolah di rentang 0,7 persen hingga 3,5 persen, dalam penelitian ini *range* persentase tersebut dikelompokkan sebagai kategori rendah. Lebih rinci lagi, dari hasil pengelompokkan nilai persentase menggunakan metode *Natural Breaks*, dihasilkan sebanyak 66 wilayah berada di rentang 0,7 hingga 3,5 persen (kategori rendah), 63 wilayah berada di rentang 3,5 hingga 6,30 persen (kategori sedang), dan 25 wilayah berada di rentang 6,3 persen hingga 12,4 persen (kategori tinggi). Yang termasuk dalam 25 wilayah dengan persentase anak 7—17 tidak sekolah berkategori tinggi termasuk dalam cakupan administrasi kabupaten. Wilayah tersebut tersebar di Sumatera Utara, Riau, Jambi, Sumatera Selatan, Lampung, dan Kepulauan Bangka Belitung. Kabupaten di Sumatera Utara yang masuk dalam kategori tinggi adalah Asahan dan Nias Selatan, kabupaten di Riau yang masuk dalam kategori tinggi adalah kabupaten Indragiri Lahir, kabupaten di jambi yang masuk dalam kategori tinggi diantaranya Merangin, Tanjung Jabur Timur, Tanjung Jabung Barat, dan Bungo. Kabupaten di Sumatera Selatan yang masuk dalam kategori tinggi diantaranya Ogan Komering Ilir, Banyu Asin, Ogan Komering Ulu Selatan, Ogan Komering Ulu Timur, Empat Lawang, dan Musi Rawas Utara. Kabupaten di Lampung yang masuk dalam kategori tinggi diantaranya Lampung Barat, Tanggamus, Lampung Timur, Lampung Tengah, Tulangbawang, Mesuji, Tulang Bawang barat, dan Pesisir Barat. Sementara untuk kabupaten di Kepulauan Bangka Belitung yang masuk dalam kategori tinggi diantaranya Bangka, Bangka Barat, Bangka Tengah, dan Bangka Selatan.

4. KESIMPULAN

Berdasarkan hasil analisis yang dilakukan, dengan menggunakan metode pendugaan langsung diketahui hanya 55 dari 154 kabupaten/kota di Pulau Sumatera yang memiliki RSE kurang dari 25 persen. Dengan kata lain, sebagian besar kabupaten/kota di Pulau Sumatera memiliki RSE yang tinggi sehingga hasil estimasinya kurang presisi. Metode SAE HB Beta sebagai bagian dari metode pendugaan tidak langsung menjadi metode terbaik dalam mengestimasi persentase anak usia 7-17 tahun yang tidak sekolah level kabupaten/kota di Pulau Sumatera tahun 2023. Hal ini ditandai dengan RSE dari penduga HB Beta yang lebih rendah secara keseluruhan dibandingkan dengan RSE penduga langsung. Selain itu, penduga HB Beta yang diajukan telah reliabel, valid, dan *unbiased*. Variabel-variabel yang menggambarkan keberhasilan perekonomian wilayah, kemajuan teknologi, dan infrastruktur memiliki peran dalam menurunkan persentase anak tidak sekolah, sementara faktor lingkungan dalam hal ini bencana alam menyebabkan semakin banyaknya anak yang tidak sekolah. Sebagian besar kabupaten/kota di Pulau Sumatera memiliki persentase anak usia 7-17 tahun yang tidak sekolah di rentang 0,7 persen hingga 3,5 persen. Dalam penelitian ini, rentang persentase tersebut dikelompokkan sebagai kategori rendah. Namun, masih terdapat sebanyak 25 kabupaten/kota dengan nilai persentase anak usia 7—17 tahun yang tidak sekolah yang tinggi (berada pada rentang 6,3 persen hingga 12,4 persen) yang mayoritas berada pada kawasan selatan pulau Sumatera seperti provinsi Jambi, Sumatera Selatan, Lampung, dan Kepulauan Bangka Belitung.

ACKNOWLEDGEMENT

Penulis mengucapkan terimakasih kepada Direktorat Kesejahteraan Rakyat Badan Pusat Statistik RI yang telah menyediakan data estimasi langsung beserta RSE-nya dalam penelitian ini. Penulis juga mengucapkan terimakasih kepada Prodi D-IV Statistika Politeknik Statistika STIS yang telah membantu penyelesaian *project small area estimation* kami.

REFERENCES

- [1] Australian Bureau of Statistics, 2021. *Confidentiality and Relative Standard Error*. <https://www.abs.gov.au/statistics/microdata-tablebuilder/tablebuilder/confidentiality-and-relative-standard-error#cite-window2>
- [2] Badan Pusat Statistik, 2024. *Angka Anak Tidak Sekolah Menurut Jenjang Pendidikan dan Kelompok Pengeluaran, 2022-2023*. <https://www.bps.go.id/id/statistics-table/2/MTk4OCMy/angka-anak-tidak-sekolah-menurut-jenjang-pendidikan-dan-kelompok-pengeluaran.html>
- [3] Bappenas. 2024a. *Metadata Indikator SDGs*. <https://sdgs.bappenas.go.id/metadata-indikator-sdgs/>
- [4] Bappenas, 2024b. *Sustainable Development Goals 4: Pendidikan Berkualitas*. <https://sdgs.bappenas.go.id/17-goals/goal-4/>
- [5] Box, G., & Tiao, G. 1992. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons Ltd. <https://doi.org/10.1002/9781118033197>
- [6] Eurostat, 2019. *Guidelines on Small Area Estimation For City Statistics and other functional geographies*. European Union.
- [7] Firmando, G., & Ubaidillah, A., 2021. Pendugaan Area Kecil Untuk Angka Partisipasi Kasar Pendidikan Dasar Dan Menengah Tingkat Kabupaten/Kota Di Provinsi Jawa Tengah Tahun 2018. *Seminar Nasional Official Statistics, 2020(1)*, 651–661. <https://doi.org/https://doi.org/10.34123/semnasoffstat.v2020i1.466>
- [8] Gu, H., & Ozturk, U., 2016. Technology, enrollment gender gap and the impact of social norms in rural China. *China Agricultural Economic Review, 8(2)*, 335–344. <http://dx.doi.org/10.1108/CAER-01-2016-0016>
- [9] Hakim, A., 2020. Faktor Penyebab Anak Putus Sekolah. *Jurnal Pendidikan, 21(2)*, 122–132. <https://doi.org/10.33830/jp.v21i2.907.2020>
- [10] Ikhsan, E., Ratu, N. Y., & Nurizza, W. A. 2019. Estimasi Small Area Estimation Angka Partisipasi Kasar Di Perguruan Tinggi Pulau Kalimantan Tahun 2018. *Prosiding Seminar Nasional Matematika, Statistika Dan Aplikasinya, 1*, 82–88. <https://jurnal.fmipa.unmul.ac.id/index.php/SNMSA/article/view/530>
- [11] Kiong, J. F., 2023. The Impact of Technology on Education: A Case Study of Schools. *Journal of Education Review Provision, 2(2)*, 43–47. <https://doi.org/10.55885/jerp.v2i2.153>
- [12] Lanyasanya, R., Ngala, F. B. J. A., & Tikoko, B., 2023. Relationship between ecological, economic, conflict, socio-cultural factor and enrolment of girls in rural public primary schools in Samburu County, Kenya. *Journal of Education and Learning, 2(1)*, 154–160. <https://doi.org/10.51317/jel.v2i1.454>
- [13] Liang, X., Kamata, A., & Li, J., 2020. Hierarchical Bayes Approach to Estimate the Treatment Effect for Randomized Controlled Trials. *Educational and Psychological Measurement, 80(6)*, 1090–1114. <https://doi.org/10.1177/0013164420909885>
- [14] Liu, B., 2009. *Hierarchical Bayes Estimation And Empirical Best Prediction Of Small Area Proportions* [University of Maryland]. https://drum.lib.umd.edu/bitstream/handle/1903/9149/liu_umd_0117e_10245.pdf?sequence=1
- [15] Maghfirah, D. A., 2019. The Determinant Factors of Drop Out Students at High School/Vocational School Level in Mataram City. *Jurnal Kebijakan Pendidikan, 8(3)*, 215–222. <https://doi.org/10.21831/sakp.v8i3.15862>
- [16] Malik, N. F. J., Hoyyi, A., & Ispriyanti, D., 2014. Pendugaan Angka Putus Sekolah Di Kabupaten Semarang Dengan Metode Prediksi Tak Bias Linier Terbaik Empirik Pada Model Pendugaan Area Kecil. *Jurnal Gaussian, 3(2)*, 101–110.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Wisly Ryan Eliezer, Aisyah ‘Azizah Nur Rahmah, Karina Himalaya, Afidita Nabila Putri, Aditya Prameswara Achmadi, Azka Ubaidillah, Shafiyah Asy Syahidah

- [https://doi.org/10.14710/j.gauss.3.1.101 - 110](https://doi.org/10.14710/j.gauss.3.1.101-110)
- [17] Muhardi, 2004. Kontribusi Pendidikan Dalam Meningkatkan Kualitas Bangsa Indonesia. *MIMBAR: Jurnal Sosial Dan Pembangunan*, XX(4), 478–492. <https://doi.org/10.29313/mimbar.v20i4.153>
- [18] Mukherjee, M., 2012. Do Better Roads Increase School Enrollment? Evidence from a Unique Road Policy in India. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2207761>
- [19] Nations, U. 2023. *UNESCO: 250 million children now out of school*. <https://news.un.org/en/story/2023/09/1140882>
- [20] Noviani, A., 2016. Small Area Estimation dengan Pendekatan Hierarchical Bayesian Neural Network Untuk Kasus Anak Putus Sekolah dari Rumah Tangga Miskin di Provinsi Jawa Timur. *Thesis*. <https://repository.its.ac.id/1470/>
- [21] Nurjanah, S., 2024. Factors Affecting Gross Enrollment Rates in Higher Education in Indonesia. *International Journal of Applied and Advanced Multidisciplinary Research*, 2(3), 243–258. <https://doi.org/10.59890/ijaamr.v2i3.1566>
- [22] Putri, S. P. O., 2023. Pengaruh Bencana Alam Terhadap Putus Sekolah di Indonesia. *EDUCATIONIST: Journal of Educational and Cultural Studies*, 2(2), 79–93. <https://jurnal.litnuspublisher.com/index.php/jecs/article/view/170>
- [23] Rao, J. N. ., & Molina, I., 2015. *Small Area Estimation*. John Wiley & Sons Ltd.
- [24] Raya, U., 2021. *Akses Pendidikan Dasar Bagi Semua Anak*. https://sikompak.bappenas.go.id/pembelajaran/view/40/id/materi_belajar/Strategi Tata Kelola Penanganan Anak Tidak Sekolah di Wilayah Dukungan KOMPAK/download.pdf
- [25] Rudji, H., 2023. Analisis Kondisi Demografi, Ekonomi dan Lingkungan pada Anak Putus Sekolah di Kecamatan Banawa Kabupaten Donggala (Studi Kasus Jenjang Pendidikan Sekolah Menengah Pertama). *Jurnal Kolaboratif Sains*, 6(6), 519–526. <https://doi.org/10.56338/jks.v6i6.3672>
- [26] Soetrisnaadisendjaja, D., & Sari, N., 2019. Fenomena Anak Putus Sekolah di Kawasan Industri Kota Cilegon. *Hermeneutika: Jurnal Hermeneutika*, 5(2), 89. <https://doi.org/10.30870/hermeneutika.v5i2.7383>
- [27] Yamashita, T., Yamashita, K., & Kamimura, R., 2007. A stepwise AIC method for variable selection in linear regression. *Communications in Statistics - Theory and Methods*, 36(13), 2395–2403. <https://doi.org/10.1080/03610920701215639>
- [28] Zainuri, M., Matsum, J. H., & Thomas, Y., 2020. Tingkat Pendapatan, sosial, Budaya dan Jarak Rumah dengan Sekolah sebagai Faktor Penyebab Anak Putus Sekolah. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa*, 3(10), 1–15. <https://jurnal.untan.ac.id/index.php/jpdpb/article/download/6812/7556>
- [29] Zimmerman, F. J., 2001. Determinants of school enrollment and performance in Bulgaria: The role of income among the poor and rich. *Contemporary Economic Policy*, 19(1), 87–98. <https://doi.org/10.1111/j.1465-7287.2001.tb00052.x>