



## Pemodelan Kemiskinan di Provinsi Bengkulu Menggunakan *Small Area Estimation* dengan Pendekatan Semiparametrik *Penalized Spline*

I Sriliana , E Sunandi, U Rafflesia

Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, Indonesia

### Info Artikel

*Sejarah Artikel:*

Diterima 11 Juli 2017

Disetujui 23 September 2017

Dipublikasikan 1 Oktober 2017

*Keywords:*

*Poverty, Semiparametric*

*Penalized Spline, Small Area*

*Estimation*

### Abstrak

Penelitian ini bertujuan untuk memodelkan tingkat kemiskinan di Provinsi Bengkulu. Metode yang digunakan dalam penelitian ini adalah *small area estimation* (SAE) dengan pendekatan semiparametrik *penalized spline* (P-Spline). SAE merupakan metode statistika yang sering digunakan untuk mendapatkan suatu informasi yang akurat tentang kemiskinan. Ketika asumsi linieritas pada model dasar SAE tidak terpenuhi, maka dilakukan pendekatan nonparametrik sebagai alternatif pilihan. Salah satunya adalah pendekatan semiparametrik *penalized spline*. Metode SAE dengan pendekatan semiparametrik mempunyai model yang lebih fleksibel karena mengakomodasi hubungan antara respon dengan prediktor yang bersifat linier dan nonlinier. Pada penelitian ini, dilakukan pemodelan kemiskinan di Provinsi Bengkulu berdasarkan rata-rata pengeluaran per kapita melalui pendugaan parameter model SAE menggunakan semiparametrik P-Spline sehingga diperoleh suatu persamaan regresi efek campuran sebagai model kemiskinan. Berdasarkan hasil analisis diperoleh model kemiskinan di Provinsi Bengkulu yaitu model P-Spline linier dengan 1 knot. Model ini mempunyai nilai GCV sebesar 148928361265,95, nilai AIC sebesar 13883,46, dan BIC sebesar 13904,38.

### Abstract

*This study aims to model the level of poverty in Bengkulu Province. The method used in this research is small area estimation (SAE) with semiparametric penalized spline (P-Spline) approach. SAE is a statistical method that is often used to obtain an accurate information about poverty. When the linearity assumptions on the basic SAE model are not met, a nonparametric approach is used as an alternative choice. One is the semiparametric approach of the penalized spline. The SAE method with semiparametric approach has a more flexible model because it accommodates the relationship between response with linear and nonlinear predictors. In this study, poverty modeling in Bengkulu Province was based on average per capita expenditure through the estimation of SAE model parameters using semiparametric P-Spline to obtain a mixed effect regression equation as a model of poverty. Based on the analysis result obtained poverty model in Bengkulu Province is model of P-Spline linear with 1 knot. This model has a GCV value of 148928361265,95, AIC value of 13883.46, and BIC of 13904.38.*

© 2017 Universitas Negeri Semarang

 Alamat korespondensi:

E-mail: [idhiasriliana@unib.ac.id](mailto:idhiasriliana@unib.ac.id)

ISSN 0215-9945

## PENDAHULUAN

Kemiskinan di Indonesia merupakan masalah sosial yang hingga sekarang belum bisa teratasi baik oleh pemerintah pusat maupun oleh pemerintah daerah. Pertumbuhan ekonomi yang terjadi di Indonesia tidak sejalan dengan pengurangan kemiskinan dan ketimpangan ekonomi. Berdasarkan data Badan Pusat Statistik (BPS) Maret 2017, jumlah penduduk miskin di Indonesia mengalami kenaikan 6.900 orang dari September 2016 ke Maret 2017. Salah satu provinsi di Indonesia yang mempunyai masalah kemiskinan kompleks dan perlu mendapat perhatian khusus dari pemerintah adalah Provinsi Bengkulu. Bengkulu merupakan provinsi di Indonesia dengan angka kemiskinan yang tinggi. Angka kemiskinan Provinsi Bengkulu hampir dua kali lipat angka kemiskinan nasional. Data Sosial Ekonomi BPS (Maret 2017) menunjukkan bahwa Bengkulu menduduki peringkat pertama provinsi termiskin di Sumatera dengan persentase penduduk miskin sebesar 16,45 %. Sedangkan di Indonesia, Bengkulu menempati urutan ke enam provinsi termiskin se-Indonesia. Jumlah penduduk miskin di Provinsi Bengkulu pada Maret 2017 mencapai 316.980 orang, berkurang sebesar 11.630 orang dibandingkan dengan penduduk miskin pada Maret 2016 yang sebesar 328.610 orang (17,23 %) (BPS 2017).

Secara umum kemiskinan didefinisikan sebagai kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Kemiskinan merupakan masalah multi dimensional, sehingga tidak mudah untuk mengukur kemiskinan dan perlu kesepakatan pendekatan pengukuran yang dipakai. Untuk mengukur kemiskinan, BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar (*basic needs approach*). Dengan pendekatan ini, kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Jadi penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita perbulan dibawah garis kemiskinan (BPS 2012).

Metode yang bisa digunakan menduga rata-rata pengeluaran per kapita sebagai indikator pengukuran kemiskinan adalah *Small Area Estimation (SAE)*. SAE merupakan suatu metode statistika untuk menduga parameter pada suatu subpopulasi dimana jumlah contohnya berukuran kecil atau bahkan tidak ada. Teknik pendugaan ini memanfaatkan data dari domain besar untuk menduga parameter pada domain yang lebih kecil yang dapat berupa desa/kelurahan, kecamatan, kabupaten, kelompok suku, maupun kelompok umur. Metode SAE mempunyai konsep dalam pendugaan parameter secara tidak langsung di suatu area yang relatif kecil dalam percontohan survei (*survey sampling*) dimana pendugaan langsung tidak mampu memberikan ketelitian yang cukup bila ukuran sampel dalam *small area* berukuran kecil/sedikit, sehingga statistik yang dihasilkan akan memiliki variansi yang besar atau bahkan pendugaan tidak dapat dilakukan karena tidak terwakili dalam survei (Prasad & Rao 1990).

Pada umumnya, SAE menggunakan pemodelan parametrik untuk menghubungkan statistik area kecil dengan variabel-variabel pendukungnya. Pendugaan parameter model dasar SAE umumnya menggunakan metode EBLUP (*Empirical Best Linear Unbiased Prediction*) yang membangun suatu model linier campuran. Pemodelan ini kurang fleksibel dalam menyesuaikan dengan pola data hasil survei yang mungkin saja tidak mirip sama sekali dengan distribusi formal yang ada. Sehingga pendekatan nonparametrik menjadi alternatif pilihan, salah satunya dengan menggunakan pendekatan semiparametrik *penalized spline* mempunyai model yang lebih fleksibel karena keberadaan dua komponen dalam model yang mengakomodasi hubungan antara respon dengan prediktor yang bersifat linier, dan hubungan antar respon dengan prediktor yang bersifat nonlinier.

Berbagai penelitian yang telah dilakukan menggunakan *small area estimation* dengan pendekatan nonparametrik antara lain: (Apriani 2017) menggunakan SAE dengan pendekatan Semiparametrik *Penalized Spline* untuk memodelkan pengeluaran per kapita di Kabupaten Sleman, Sriliana *et al.* (2016) melakukan pemetaan kemiskinan di Kabupaten Mukumuko menggunakan SAE dengan pendekatan regresi

*Penalized Spline*, (Baskara 2014) SAE dengan pendekatan *P-spline* untuk menduga pengeluaran per kapita di Kabupaten Sumenep, (Iriyanto & Darsyah 2014) SAE Kernel-*Bootstrap* untuk menduga tingkat kemiskinan di Indonesia, Salvati *et al.* (2008) menggunakan model nonparametrik *based direct estimator*, dan Opsomer *et al.* (2008) Pengembangan SAE dengan pendekatan *penalized spline regression*.

Pada penelitian ini, dilakukan analisis SAE dengan pendekatan semiparametrik *P-spline*. Pendugaan parameter model dengan menggunakan pendekatan *P-spline* ini kemudian digunakan untuk memodelkan pengeluaran per kapita level desa di Provinsi Bengkulu berdasarkan beberapa variabel indikator kemiskinan. Evaluasi hasil pendugaan dilakukan dengan melihat nilai GCV pada model.

**METODE**

**Small Area Estimation**

*Small Area Estimation* (SAE) merupakan suatu teknik statistika untuk menduga parameter-parameter subpopulasi yang ukuran sampelnya kecil atau bahkan area yang tidak tersampling. Dalam SAE terdapat dua jenis model dasar yang digunakan, yaitu model berbasis area dan model berbasis unit (Rao 2003). Pada model SAE berbasis area, data pendukung yang tersedia hanya sampai level area. Model level area menghubungkan penduga langsung area kecil dengan data pendukung dari domain lain untuk setiap area.

Parameter area kecil yang ingin diamati adalah  $\theta_i$ . Model linier yang menjelaskan hubungan tersebut adalah :

$$\theta_i = x_i^T \beta + z_i v_i \tag{1}$$

dengan  $\beta = (\beta_1, \dots, \beta_p)^T$  adalah koefisien regresi berukuran  $p \times l$ ,  $z_i =$  konstanta positif yang diketahui,  $v_i =$  pengaruh acak area kecil, diasumsikan  $v_i \sim iid N(0, \sigma^2)$  dimana  $i = 1, 2, \dots, m$  dan  $x_i^T$  adalah data pendukung area ke- $i$ .

Dalam membuat kesimpulan tentang populasi diasumsikan bahwa nilai estimasi langsung  $\hat{\theta}_i$  diketahui maka dapat dinyatakan sebagai berikut :

$$\hat{\theta}_i = \theta_i + e_i \tag{2}$$

di mana  $e_i$  adalah sampling error, diasumsikan  $e_i \sim iid N(0, \psi_i)$  dan  $i = 1, 2, \dots, m$ .

Model SAE untuk level area terdiri dari dua tingkat komponen model yaitu komponen model estimasi tidak langsung sesuai dengan persamaan (1) dan komponen model estimasi langsung sesuai persamaan (2). Model pada persamaan (1) dan (2) jika digabungkan membentuk persamaan sebagai berikut :

$$\hat{\theta}_i = x_i^T \beta + z_i v_i + e_i \tag{3}$$

di mana  $i = 1, 2, \dots, m$ .

**Penalized Spline Regression**

Regresi *Penalized Spline* atau *P-spline* adalah suatu metode *smoothing* yang sangat menarik karena mempunyai sifat sederhana (Eubank 1988). Diberikan model:

$$y_i = m(x_i) + \varepsilon_i \tag{4}$$

di mana  $\varepsilon_i$  adalah peubah acak yang saling bebas dengan rata-rata  $npl$  dan varian  $\sigma_\varepsilon^2$ , Fungsi  $m(x_i)$  adalah fungsi yang tidak diketahui dan diasumsikan dapat didekati dengan *P-spline*:

$$m(x_i) = \beta_0 + \beta_1 x + \dots + \beta_p x^p + \sum_{j=1}^K \gamma_j (x_i - k_j)_+^p \tag{5}$$

di mana  $p$  adalah derajat spline (*fixed*),  $(x_i - k_j)_+ = maks\{0, (x_i - k_j)\}$ ,  $k_j, j = 1, \dots, K$ ,

adalah himpunan titik knot.  $\beta = (\beta_0, \dots, \beta_p)^T$  merupakan vektor koefisien parametrik dari parameter yang tidak diketahui,

$\gamma = (\gamma_1, \dots, \gamma_k)^T$  adalah vektor koefisien *spline*.

Misal diberikan  $X = \begin{bmatrix} 1 & x_i & \dots & x_i^p \end{bmatrix}_{1 \leq i \leq n}$ ,

$$Z = \begin{bmatrix} (x_i - k_1) & \dots & (x_i - k_K)_+^p \end{bmatrix}_{1 \leq i \leq n}$$

dengan

$$(x_i - k_j)_+^p = \begin{cases} (x_i - k_j)_+^p & \text{untuk } x_i \geq k_j \\ 0 & \text{untuk } x_i < k_j \end{cases}$$

sehingga model pada persamaan (4) dapat ditulis dalam bentuk:

$$y_i = \beta_0 + \beta_1 x + \dots + \beta_p x^p + \sum_{j=1}^K \gamma_j (x_j - k_j)_+^p + e_i \tag{6}$$

Atau dapat ditulis dalam bentuk:

$$Y = X\beta + Z\gamma + e \text{ di mana } Y = (y_1 \dots y_n)^T \tag{7}$$

Persamaan (6) disebut sebagai model regresi *spline smoothing*. Fungsi spline pada model menunjukkan bahwa *spline* merupakan model polinomial terputus, tapi masih bersifat kontinu pada knot-knotnya (Eubank 1988). Salah satu metode pemilihan titik knot optimal adalah dengan menggunakan *Generalized Cross Validation* (GCV). Definisi GCV dapat ditulis sebagai berikut :

$$GCV(\mathbf{K}) = \frac{MSE(\mathbf{K})}{[n^{-1}trace(\mathbf{I} - A(\mathbf{K}))]^2} \quad (8)$$

di mana  $MSE(\mathbf{K}) = n^{-1}\mathbf{Y}^T(\mathbf{I} - A(\mathbf{K}))^T(\mathbf{I} - A(\mathbf{K}))\mathbf{Y}$ ,  $K = (K_1, K_2, \dots, K_N)$  adalah titik knot dan matriks  $A(\mathbf{K})$  diperoleh dari persamaan  $\hat{\mathbf{Y}} = A(\mathbf{K})\mathbf{Y}$ .

**Small Area Estimation Model using Semiparametric Penalized Spline**

Pada model small area estimation dengan pendekatan semiparametrik *P-spline*, model penalized spline merupakan pengaruh acak yang dapat dikombinasikan dengan model SAE berbasis area untuk mendapatkan estimasi area kecil secara semiparametrik berdasarkan model linier campuran. Berdasarkan persamaan (3) dan persamaan (7), model semiparametrik Fay-Herriot dapat ditulis sebagai berikut:

$$\hat{\theta} = \begin{bmatrix} \mathbf{X} \\ \mathbf{X}_1 \end{bmatrix} [\boldsymbol{\alpha} \quad \boldsymbol{\beta}] + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{D}\mathbf{u} + e \quad (9)$$

Menurut Giusti *et al.* (2012), jika terdapat peubah lain yang perlu disertakan dalam model, maka peubah tersebut ditambahkan ke dalam  $\mathbf{X}$  sebagai matriks pengaruh tetap.

Misalkan terdapat  $T$  area kecil,  $U_1, U_2, \dots, U_T$  adalah parameter yang akan diestimasi. Definisikan  $d_{it} = I_{\{i \in U_t\}}$  dan untuk setiap pengamatan  $\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{iT})$ ,

$$\mathbf{Y} = (y_1, y_2, \dots, y_n)^T,$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & L & x_1^p \\ M & M & O & M \\ 1 & x_n & L & x_n^p \end{bmatrix},$$

$$\mathbf{Z} = \begin{bmatrix} (x_1 - k_1)_+^p & L & (x_1 - k_K)_+^p \\ M & O & M \\ (x_n - k_1)_+^p & L & (x_n - k_K)_+^p \end{bmatrix},$$

di mana

$$(x_i - k_j)_+^p = \begin{cases} (x_i - k_j)_+^p & \text{untuk } x_i \geq k_j \\ 0 & \text{untuk } x_i < k_j \end{cases}$$

Opsomer *et al.* (2008) menggunakan *P-spline* untuk mengestimasi area kecil dengan menambahkan pengaruh acak area kecil pada persamaan (7), sehingga diperoleh:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{D}\mathbf{u} + e \quad (10)$$

di mana fungsi semiparametrik  $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$  adalah fungsi *spline* yang memuat komponen linier dan nonlinier,  $\mathbf{D}\mathbf{u}$  adalah pengaruh acak area kecil,  $\mathbf{D} = (d_1, \dots, d_n)^T$  adalah matriks covarian, dan  $\mathbf{u}$  adalah vektor pengaruh area kecil, setiap komponen acak diasumsikan independen satu sama lain, dan

$$\begin{aligned} \boldsymbol{\gamma} &= (0, \Sigma_\gamma), \Sigma_\gamma = \sigma_\gamma^2 I_K \\ \boldsymbol{\gamma} &= (0, \Sigma_u), \Sigma_u = \sigma_u^2 I_T \\ \boldsymbol{\gamma} &= (0, \Sigma_\epsilon), \Sigma_\epsilon = \sigma_\epsilon^2 I_n \end{aligned}$$

Jika komponen ragam diketahui, pendugaan pengaruh tetap  $\boldsymbol{\beta}$  dapat dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) dengan menganggap  $\boldsymbol{\gamma}$  dan  $\mathbf{u}$  sebagai pengaruh acak. Persamaan (10) dapat ditulis sebagai (Baskara 2014):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}^* \text{ di mana } \boldsymbol{\epsilon}^* = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{D}\mathbf{u} + e \quad (11)$$

Penduga parameter  $\boldsymbol{\beta}$  dapat diperoleh dengan memaksimumkan *fungsi likelihoodnya* sehingga diperoleh:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{Y} \quad (12)$$

di mana  $\mathbf{V} = \mathbf{Z}\Sigma_\gamma\mathbf{Z}^T + \mathbf{D}\Sigma_u\mathbf{D}^T + \Sigma_e$  adalah matriks varian kovarian dari  $\mathbf{Y}$ .

Penduga terbaik untuk variabel prediktor  $\gamma$  dan  $u$  diperoleh dengan meminimumkan MSE dari  $\gamma$  dan  $u$ . Sehingga diperoleh prediktor linier tak bias terbaik (BLUP) untuk  $\gamma$  dan  $u$  sebagai berikut:

$$\hat{\gamma} = \sum_{\gamma} \mathbf{Z}^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\beta}) \quad (13)$$

$$\hat{u} = \sum_u \mathbf{Z}^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\beta}) \quad (14)$$

Untuk area kecil  $U_T$  yang diberikan, maka akan dilakukan pendugaan terhadap:

$$\bar{y}_t = \bar{x}_t\beta + \bar{z}_t\gamma + u_t \quad (15)$$

di mana  $\bar{x}_t$  adalah nilai rata-rata dari  $x_i$ ,  $\bar{z}_t$  adalah basis fungsi spline, dan  $u_t$  adalah pengaruh acak area kecil dengan  $u_t = \mathbf{d}_t\mathbf{u} = \mathbf{e}_t\mathbf{u}$ , dan  $\mathbf{e}_t$  adalah vektor dengan nilai 1 saat ke- $t$ , dan bernilai 0 untuk  $t$  lainnya. Sehingga penduga untuk  $\bar{y}_t$  yaitu:

$$\hat{y}_t = \bar{x}_t\hat{\beta} + \bar{z}_t\hat{\gamma} + \mathbf{e}_t\hat{u}_t \quad (16)$$

yang merupakan kombinasi linier penduga GLS pada (10) dan BLUP pada (11) dan (12), sehingga  $\bar{y}_t$  merupakan BLUP untuk  $\bar{y}_t$ .

Penduga BLUP tergantung pada komponen ragam yang biasanya tidak diketahui. Pendugaan komponen ragam menggunakan MLE akan menghasilkan penduga bias, sehingga pendugaan dilakukan dengan menggunakan REML (*Restricted Maksimum Likelihood*) yang didasarkan pada residual yang dihitung setelah  $\beta$  dihitung. Untuk mengestimasi komponen ragam dengan REML dapat digunakan beberapa metode, yaitu metode *Newton Rhapsod* dan algoritma EM (*expectation and maximization*).

## HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistika Provinsi Bengkulu (Data Susenas dan Podes 2014). Objek penelitian adalah desa yang menjadi sampel pada Susenas 2014 di Provinsi Bengkulu.

Berdasarkan survei SUSENAS 2014, terdapat 502 desa di 113 kecamatan yang menjadi sampel dalam pendugaan rata-rata pengeluaran perkapita level desa di Provinsi Bengkulu. Adapun variabel yang digunakan dalam penelitian meliputi: rata-rata pengeluaran perkapita sebagai variabel respon, sedangkan variabel prediktor: Jumlah Keluarga Tanpa Listrik ( $X_1$ ), Jumlah Sarana Pendidikan ( $X_2$ ), Jumlah Sarana Kesehatan ( $X_3$ ), Jumlah Penerima Jamkesmas ( $X_4$ ), dan Jumlah SKTM ( $X_5$ ). Penelitian dimulai dengan melakukan eksplorasi data untuk memperoleh gambaran secara umum faktor-faktor yang mempengaruhi kemiskinan di Provinsi Bengkulu. Selanjutnya dilakukan analisis data yaitu pembentukan model kemiskinan menggunakan SAE dengan pendekatan semiparametrik *P-spline*. Model dibangun berdasarkan data sampel hingga diperoleh suatu persamaan model regresi linier efek campuran sebagai model SAE.

Berdasarkan survei SUSENAS 2014, terdapat 502 desa di 113 kecamatan yang menjadi sampel dalam pendugaan rata-rata pengeluaran per kapita level desa di Provinsi Bengkulu. Dari sebanyak 502 sampel terdapat 15 data sampel yang tidak valid, sehingga pada penelitian ini dilakukan pendugaan area kecil untuk 487 desa dengan menggunakan pendekatan regresi *penalized spline*. Eksplorasi data pengeluaran per kapita pada level desa di Provinsi Bengkulu dapat dilihat pada Tabel 1.

**Tabel 1.** Rata-rata Pengeluaran Per kapita di Provinsi Bengkulu Tahun 2014

Statistik	Pengeluaran Per kapita (Rp)
Rata-rata	773041,14
Kuartil ke-1	536709,08
Kuartil ke-3	902426,89
Minimum	3115622,06
Maksimum	248376,69

Tabel 1 menunjukkan nilai rata-rata dari rata-rata pengeluaran per kapita di Provinsi Bengkulu tahun 2014 sebesar Rp. 773.041,14. Sekitar 75% desa sampel yang ada di Provinsi Bengkulu memiliki rata-rata pengeluaran per kapita sebesar Rp 902.426,06 dan 25% sebesar Rp 536.709,08. Rata-rata pengeluaran per kapita terbesar berada di Kelurahan Kebun Dahri Kecamatan Ratu Samban Kota Bengkulu sebesar Rp. 3.115.622,06 dan terendah berasal dari Desa

Sekunyit Kecamatan Kaur Selatan Kabupaten Kaur sebesar Rp 248.376,69.

Berdasarkan hasil eksplorasi pola hubungan antara variabel respon dengan variabel prediktor, terdapat lima variabel prediktor yang digunakan dalam pembentukan model dimana empat variabel prediktor ( $X_1, X_2, X_3$ , dan  $X_4$ ) diasumsikan bersifat linier dan satu variabel prediktor yaitu Jumlah SKTM ( $X_5$ ) bersifat non linier. Sehingga pada penelitian ini dilakukan pembentukan model SAE dengan pendekatan semiparametrik *P-spline*. Pembentukan model SAE dengan pendekatan semiparametrik *P-spline* dilakukan melalui dua tahapan dengan menggunakan *Generalized Additive Model*, yaitu mencari nilai parametrik untuk  $f(x_1, x_2, x_3, x_4)$  dan nilai nonparametrik *P-spline* untuk  $f(x_5)$ . Proses pembentukan model pada tahapan pertama untuk mencari nilai parametrik dengan menggunakan *Generalized Additive Model* dapat menghasilkan model berikut:

$$Y = -93479,34 - 309,89X_1 - 14697,81X_2 + 45049,73X_3 - 46,89X_4 + f(x_5) + u \quad (17)$$

Selanjutnya dilakukan tahapan kedua yaitu pencarian nilai nonparametrik dengan pembentukan model SAE menggunakan pendekatan *P-spline*, proses ini dilakukan dengan memodelkan nilai residual pada model parametrik dengan variabel prediktor yang tidak linier ( $X_5$ ). Berdasarkan hasil analisis data menggunakan program R diperoleh nilai GCV minimum untuk ketiga model *P-spline*, yaitu *P-spline* linier, kuadrat, dan kubik adalah GCV dengan satu titik knot. Tabel 1 menyajikan nilai GCV untuk model *P-spline* linier, kuadrat, dan kubik dengan satu titik knot. Berdasarkan Tabel 2 dapat dilihat bahwa GCV optimum diperoleh dari model *P-spline* linier dengan satu titik knot. Sehingga dapat disimpulkan bahwa model SAE dengan pendekatan semiparametrik *P-spline* yang digunakan untuk memodelkan kemiskinan berdasarkan rata-rata pengeluaran perkapita pada desa sampel di Provinsi Bengkulu diperoleh dari model *P-spline* linier dengan satu titik knot.

**Tabel 2.** Nilai GCV Values untuk Model *P-spline* Linier, Kuadrat, dan Kubik

Jumlah Titik Knot	Nilai GCV		
	<i>P-spline</i> Linier	<i>P-spline</i> Kuadrat	<i>P-spline</i> Kubik
1	14892836	14954630	Singular
	1265,95*	1175,71	
2	15079346	15143110	Singular
	4357,71	5071,33	
3	15153244	15269231	Singular
	4408,51	5419,77	
4	15242622	15436424	Singular
	1524,96	4917,04	
5	15373464	15513524	15734737
	5823,07	5424,28	5809,58

Setelah diketahui lokasi titik knot dan model *P-spline* dengan GCV Optimum, langkah selanjutnya adalah mengestimasi pengaruh tetap dan pengaruh acaknya. Estimasi pengaruh tetap  $\beta$  dengan memaksimumkan fungsi *likelihood* atau *log likelihoodnya*, dan mencari  $\hat{\gamma}$  dan  $\hat{u}$  yang merupakan EBLUP (*Empirical Best Linear Unbiased Predictors*) dari  $\gamma$  dan  $u$  sebagai pengaruh acak. Nilai penduga  $\beta$  untuk model *P-spline* terbaik (*Spline* linier dengan satu knot) dapat dilihat pada Tabel 3.

**Tabel 3.** Penduga Pengaruh Tetap

Parameter	Penduga
$\beta_0$	-93479,34
$\beta_1$	-309,89
$\beta_2$	-14697,81
$\beta_3$	45049,73
$\beta_4$	-46,89
$\beta_5$	45,25

sehingga diperoleh model penduga sebagai berikut:

$$Y = -93479,34 - 309,89X_1 - 14697,81X_2 + 45049,73X_3 - 46,89X_4 + 45,25X_5 + \gamma_1(X_5 - 756,19)_+^1 + u \quad (18)$$

Dengan  $\gamma_i$  dan  $u$  adalah faktor pengaruh acak dengan nilai penduga  $\hat{\gamma}_i$  bergantung pada titik knot

dan  $\hat{u}$  bergantung pada masing-masing area. Model pada persamaan (18) merupakan model semiparametrik *P-spline* linier dengan satu titik knot yaitu pada titik 756,19 dan mempunyai nilai penalty atau parameter penghalus ( $\lambda$ ) sebesar 288504,70.

Model pada persamaan (18) menunjukkan bahwa terdapat perbedaan hasil untuk setiap desa, kemudian lokasi titik knot dengan nilai 756.19 yang artinya jika nilai  $x_i \leq 756,19$  dan variabel prediktor lainnya diasumsikan bernilai konstan, maka setiap kenaikan satu satuan akan berpengaruh sebesar  $(42,25X_5 + \gamma_1)$  satuan terhadap variabel respon ( $Y$ ).

## SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa *small area estimation* dengan pendekatan regresi *penalized spline* dapat digunakan untuk memodelkan kemiskinan pada level desa di Provinsi Bengkulu. Hasil analisis data menunjukkan model semiparametrik *P-splines* terbaik untuk pendugaan area kecil adalah model *P-spline* linier dengan 1 knot. Model ini mempunyai nilai GCV sebesar 148928361265,95.

## DAFTAR PUSTAKA

- Apriani F. 2017. *Pemodelan Pengeluaran Per Kapita Menggunakan Small Area Estimation dengan Pendekatan Semiparametrik Penalized Spline*. (Thesis). Institut Teknologi Sepuluh Nopember. Surabaya.  
[[http://repository.its.ac.id/3155/1/1315201026-Master\\_Theses.pdf](http://repository.its.ac.id/3155/1/1315201026-Master_Theses.pdf)]
- Badan Pusat Statistika (BPS) Provinsi Bengkulu. 2017. *Berita Resmi Statistik -Tingkat Kemiskinan di Provinsi Bengkulu Maret 2017*. No. 42/07/17/XI, 17 Juli 2017.
- Badan Pusat Statistika (BPS). 2012. *Data Strategis BPS*. Katalog BPS 1103003: No. 03220.1202. ISSN, 2087-2011.
- Baskara ZW. 2014. *Pendugaan Area Kecil Menggunakan Pendekatan Penalized Spline*. (Thesis). Institut Teknologi Sepuluh Nopember. Surabaya.  
[<http://digilib.its.ac.id/ITS-paper-13121140006367/34319>]
- Iriyanto S & Darsyah MY. 2014. Analysis of Poverty in Indonesia with Small Area Estimation: case in

- Demak District. *South East Asia Journal of Contemporary Business, Economics and Law* 5(3): 18–23. [<http://seajbel.com/previous-issues/december-2014/vol-5-issue-3-june-2014-economic/>]
- Eubank RL. 1998. *Spline Smoothing and Nonparametric Regression*. New York: Marcel Decker.
- Giusti C, Marchetti S, Pratesi M & Salvati N. 2012. Semiparametric Fay-Herriot Model using Penalized Spline. *JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS (JISAS)* 66(1): 1-14. [[www.isas.org.in/jisas](http://www.isas.org.in/jisas)]
- Opsomer DJ, Claeskens G, Ranalli MG, Kauermann G & Breidt FJ. 2008. Non-Parametric Small Area Estimation using Penalized Spline Regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70(1): 265–286. DOI: 10.1111/j.1467-9868.2007.00635.x
- Prasad NGN & Rao JNK. 1990. The Estimation of The Mean Squared Error of The Small Area Estimators. *Journal of American Statistical Association* 85(409): 163-171. DOI: 10.1080/01621459.1990.10475320
- Rao JNK. 2003. *Small Area Estimation*. London: Wiley.
- Salvati N, Chandra H, Ranalli MG & Chambers R. 2008. *Small Area Estimation Using a Nonparametric Model Based Direct Estimator*. Wollongong NSW: Centre for Statistical & Survey Methodology, University of Wollongong.
- Sriliana I, Agustina D & Sunandi E. 2016. Pemetaan Kemiskinan di Kabupaten Mukomuko Menggunakan Small Area Estimation dengan Pendekatan Regresi Penalized Spline. *Jurnal Matematika Integratif* 12(2): 125-133. DOI:10.24198/jmi.v12.n2.11929.125-133