

**SURAT PERNYATAAN  
PUBLIKASI KARYA ILMIAH**

Yang bertanda tangan di bawah ini, saya :

Nama : Yunita Rahman JMS  
NIM : B2A219058  
Fakultas/Jurusan : Fakultas Matematika dan Ilmu Pengetahuan Alam  
(FMIPA)/Statistika  
Jenis Penelitian : Skripsi  
Judul : Pemodelan Pengeluaran Per Kapita Di Provinsi Papua  
Menggunakan *Small Area Estimation* Dengan Pendekatan  
Semiparametrik *Penalized Spline*  
Email : [yunitarahman8@gmail.com](mailto:yunitarahman8@gmail.com)

Dengan ini menyatakan bahwa saya menyetujui untuk :

1. Memberikan hak bebas royalti kepada Perpustakaan Unimus atas penulisan karya ilmiah saya, demi pengembangan ilmu pengetahuan.
2. Memberikan hak menyimpan, mengalih mediakan/mengalih formatkan, mengelola dalam bentuk pengkalan data (*database*), mendistribusikannya, serta menampilkannya dalam bentuk *softcopy* untuk kepentingan akademik kepada Perpustakaan Unimus, tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta.
3. Bersedia dan menjamin untuk menanggung secara pribadi tanpa melibatkan pihak Perpustakaan Unimus, dari semua bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini.

Demikian pernyataan ini saya buat dengan sesungguhnya dan semoga dapat digunakan sebagaimana mestinya.

Semarang, 25 September 2020



Yunita Rahman JMS

# PEMODELAN PENGELUARAN PER KAPITA DI PROVINSI PAPUA MENGUNAKAN *SMALL AREA ESTIMATION* DENGAN PENDEKATAN SEMIPARAMETRIK *PENALIZED SPLINE*

Yunita Rahman JMS<sup>1)</sup>, Indah Manfaati Nur<sup>2)</sup>, Rocdhi Wasono<sup>3)</sup>

<sup>123)</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Muhammadiyah Semarang

Email: [yunitarahman9@gmail.com](mailto:yunitarahman9@gmail.com)

## Abstrak

JMS, Yunita Rahman 2020, Pemodelan Pengeluaran Per kapita di Provinsi Papua Menggunakan *Small Area Estimation* Menggunakan Pendekatan Semiparametrik *Penalized Spline*. Skripsi, Program Studi Statistika, Universitas Muhammadiyah Semarang. Pembimbing: I. Indah Manfaati Nur, M.Si, II. Dr. Rochdi Wasono, M.Si.

Kebutuhan informasi untuk domain dan wilayah yang lebih kecil semakin diminati. Keterbatasan sampel pada kegiatan survey menjadi kendala untuk menyediakan data untuk domain dan wilayah yang lebih kecil. *Small Area Estimation* (SAE) dapat dilakukan untuk mengatasi permasalahan ini memanfaatkan data dari domain besar untuk menduga parameter pada domain yang lebih kecil. Pada penelitian ini akan menerapkan pemodelan semiparametrik *penalized spline* untuk menduga area kecil level kabupaten/kota di Provinsi Papua. Berdasarkan hasil analisis nilai GCV minimum yang diperoleh adalah 5.309 dengan 7 lokasi titik knot. Berdasarkan metode RMSE *Jackknife*, Pendugaan pengeluaran per kapita dengan menggunakan estimasi tidak langsung mempunyai dugaan dengan tingkat akurasi dan presisi yang lebih baik dibandingkan dengan hasil pendugaan langsung.

Kata Kunci: Pengeluaran Per Kapita, *Small Area Estimation* (SAE), Semiparametrik *Penalized Spline*

## Abstract

JMS, Yunita Rahman 2020, *Modelling of Percapita expenditure in Papua province using Small Area Estimation using semiparametric approach of Penalized Spline*. Thesis, statistics study Program, University of Muhammadiyah Semarang. Supervisor: I. Indah Manfaati Nur, M.Si, II. Dr. Rochdi Wasono, M.Si.

Information needs for smaller domains and territories are increasingly in demand. The limited sample in survey activities becomes an obstacle to providing data for smaller domains and areas. *Small Area Estimation* (SAE) can be done to solve this problem by utilizing data from large domains to estimate parameters in smaller domains. This research will apply semiparametric *penalized spline* modeling to estimate small areas at district / city level in Papua Province. Based on the analysis, the minimum GCV value obtained was 5.309 with 7 knot point locations. Based on the *Jackknife* RMSE method, estimates of per capita expenditure using indirect estimates have predictions with a better degree of accuracy and precision than the results of direct estimates.

*Key words: Percapita expenditure, Small Area Estimation (SAE), Penalized semiparametric Spline*

## PENDAHULUAN

Pertumbuhan ekonomi dan pendapatan per kapita merupakan tujuan dari proses pembangunan suatu negara. Suatu negara mengharapkan pertumbuhan ekonomi dan pendapatan per kapita tiap tahunnya berangsur-angsur meningkat. Pengeluaran per kapita merupakan salah satu indikator yang digunakan untuk melihat berhasil atau tidaknya pembangunan dan pertumbuhan ekonomi. Peningkatan pertumbuhan ekonomi berkaitan pula dengan peningkatan produksi barang dan jasa, dimana dalam hal ini dapat diukur dengan Produk Domestik Regional Bruto (PDRB). Pembangunan ekonomi merupakan suatu bentuk usaha untuk mengurangi kemiskinan, ketidakmerataan distribusi pendapatan serta pengangguran, yang merupakan suatu proses multidimensional dalam konteks pertumbuhan ekonomi secara menyeluruh (Todaro, 2003).

Keberhasilan pencapaian tujuan pembangunan secara umum dapat diukur dari tingkat kemakmuran dan kesejahteraan secara menyeluruh yang meliputi tingkat kemakmuran dan kesejahteraan wilayah maupun ketingkat yang lebih kecil yaitu individu atau rumah tangga. Salah satu indikator ekonomi yang banyak digunakan untuk mengukur tingkat kemakmuran dan kesejahteraan adalah pendapatan per kapita. Dalam kaitan itu, analisis tingkat pendapatan per kapita rumah tangga sangat diperlukan oleh pemerintah dalam rangka perumusan, pelaksanaan dan evaluasi kebijakan untuk pencapaian tujuan pembangunan.

Adanya era otonomi daerah pada saat sekarang, membuat pemerintahan daerah membutuhkan suatu perencanaan dan pengambilan kebijakan sampai level yang terkecil. Untuk itu diperlukan suatu informasi yang rinci mengenai indikator pembangunan, tidak hanya untuk level provinsi saja namun diharapkan sampai level kabupaten/kota bahkan untuk unit yang lebih kecil. Seiring dengan pembangunan otonomi

daerah tersebut, Badan Pusat Statistik (BPS) dituntut untuk memproduksi statistik dan menyajikannya sebagai informasi, sesuai kebutuhan daerah sebagai bentuk kontribusi positif BPS sebagai lembaga statistik nasional dalam mendukung perencanaan dan kebijakan pembangunan daerah. Salah satu indikator pembangunan daerah dapat dilihat pada pengeluaran per kapita rumah tangga.

Pengeluaran per kapita salah satu indikator yang bisa digunakan untuk mengukur dimensi kehidupan yang layak. Pengeluaran per kapita rumah tangga merupakan rata-rata dari besarnya pengeluaran setiap anggota keluarga. Pengeluaran per kapita rumah tangga terdiri dari pengeluaran makanan dan bukan makanan dapat menggambarkan bagaimana penduduk mengalokasikan kebutuhan rumah tangganya. Berikut rata-rata pengeluaran per kapita dilihat berdasarkan provinsi yang ada di Indonesia pada tahun 2019:



Gambar 1.1 Pengeluaran Per kapita Berdasarkan Provinsi di Indonesia Tahun 2019

Berdasarkan gambar 1.1 diatas dapat dilihat bahwa pengeluaran per kapita terendah di Indonesia terdapat pada Provinsi Papua dengan pengeluaran per kapita nya yaitu sebesar 7.336 juta rupiah per tahun. Rendahnya pengeluaran per kapita di Provinsi Papua saat ini menjadi perhatian bagi pemerintah pusat terutama pemerintah daerah untuk lebih meningkatkan lagi pengeluaran per kapita di Provinsi Papua.

*Small area estimation* (SAE) merupakan salah satu upaya untuk meningkatkan keakuratan, yaitu dengan menggunakan pendugaan secara tidak langsung (*indirect estimation*). Pendugaan parameter dalam area kecil dapat didekati

dengan dua metode, yaitu metode berbasis model (*model based estimator*) dan metode berbasis rancangan (*design based estimator*). Pendugaan menggunakan metode berbasis model yaitu menduga parameter suatu area yang didasarkan dengan informasi yang berhubungan dengan parameter. Pendugaan dengan metode berbasis rancangan dilakukan berdasarkan data dari contoh area tempat survey telah dilakukan. Salah satu metode yang termasuk adalah metode pendugaan langsung (*direct estimator*).

*Small area estimation* (SAE) merupakan salah satu upaya untuk meningkatkan keakuratan, yaitu dengan menggunakan pendugaan secara tidak langsung (*indirect estimation*). Pendugaan parameter dalam area kecil dapat didekati dengan dua metode, yaitu metode berbasis model (*model based estimator*) dan metode berbasis rancangan (*design based estimator*). Pendugaan menggunakan metode berbasis model yaitu menduga parameter suatu area yang didasarkan dengan informasi yang berhubungan dengan parameter. Pendugaan dengan metode berbasis rancangan dilakukan berdasarkan data dari contoh area tempat survey telah dilakukan. Salah satu metode yang termasuk adalah metode pendugaan langsung (*direct estimator*).

## TINJAUAN PUSTAKA

### 1. Pendugaan Area Kecil (*Small Area Estimation*)

SAE adalah salah satu teknik statistik yang digunakan untuk menduga parameter subpopulasi dengan ukuran sampel yang relatif kecil. Teknik ini mengembangkan data survei dan sensus untuk mengestimasi tingkat kesejahteraan atau indikator lainnya untuk unit geografis seperti kecamatan atau pedesaan (Davies, 2003).

#### 1.1 Model Area kecil

Terdapat dua ide utama yang digunakan untuk mengembangkan model pendugaan parameter area kecil yaitu:

a. model pengaruh tetap (*fixed effect model*) dimana asumsi bahwa keragaman di dalam area kecil, variabel respon dapat diterangkan

seluruhnya oleh hubungan keragaman yang bersesuaian pada informasi tambahan.

b. Pengaruh acak area kecil (*random effect*) dimana asumsi keragaman spesifik area kecil tidak dapat diterangkan oleh informasi tambahan.

Gabungan dari kedua asumsi tersebut membentuk suatu model pengaruh campuran (*mixed model*). Oleh karena variabel respon diasumsikan berdistribusi normal maka penduga area kecil yang dikembangkan merupakan bentuk khusus dari *General Linear Mixed Model* (GLMM).

Model *small area* biasanya menggunakan model linear campuran dalam bentuk:

$$y = Xb + Zu + e$$

### 2. Model Regresi Spline

Regresi nonparametrik merupakan suatu metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dan prediktor yang tidak diketahui bentuk fungsinya, hanya diasumsikan fungsi *smooth* (mulus) dalam arti termuat dalam suatu ruang fungsi tertentu, sehingga regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988). Model regresi nonparametrik secara umum dapat disajikan sebagai berikut:

$$y_i = m(x_i) + e_i, i = 1, 2, \dots, n$$

### 3. Regresi Penalized Spline

Regresi penalized spline yaitu regresi yang diperoleh berdasarkan kuadrat terkecil (least square) dengan penalty kekasaran. Penalized spline mempunyai banyak kesamaan dengan smoothing spline, tetapi jenis penalty yang digunakan pada penalized spline lebih umum dibandingkan pada smoothing spline (Ruppert, 2003).

Terdapat dua komponen penting dalam mengestimasi penalized spline, yang pertama adalah pemilihan karakter smoothing, sementara yang kedua adalah pemilihan jumlah knot dan lokasinya (Yao dan Lee, 2008). Pada persamaan (2.6) dapat dinyatakan ke dalam bentuk matriks yaitu:

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

Dimana

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & \cdots & x_1 & \cdots & x_k^p \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & x_n & \cdots & x_n^p \end{bmatrix},$$

$$Z = \begin{bmatrix} (x_1 - k_1)_+^p & \cdots & (x_1 - k_K)_+^p \\ \vdots & \ddots & \vdots \\ (x_n - k_1)_+^p & \cdots & (x_n - k_K)_+^p \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_o \\ \vdots \\ \beta_p \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}, \text{ dan } e = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

Estimator penalized spline diperoleh dengan meminimumkan fungsi penalized least square (PLS) sebagai berikut:

$$L = \|y - X\beta - Z\gamma\|^2 + \lambda \gamma^T \gamma$$

Dengan memisalkan  $C = [X, Z]$  dan  $\tilde{\theta} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$ , sehingga persamaan 9 dapat ditulis sebagai berikut:

$$L = \|y - C\tilde{\theta}\|^2 + \lambda \tilde{\theta}^T D \tilde{\theta}$$

Dimana diketahui D merupakan matrik penalty

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ \mathbf{0}_{(p+1) \times 2} & & & & & \mathbf{0}_{(p+1) \times K} \\ \mathbf{0}_{K \times (p+1) \times 1} & & & & & & \mathbf{I}_{K \times K} \end{bmatrix}$$

Dengan parameter  $\lambda$  parameter smoothing, dimana  $\lambda \geq 0$ . Suku pertama pada persamaan 10 adalah jumlah kuadrat error dan suku keduanya adalah penalty kekasaran. Menurut Djuraidah, et al (2006) Estimator penalized spline yang diperoleh adalah

$$\hat{\tilde{\theta}} = (C^T C + \lambda D)^{-1} C^T y$$

Dengan demikian didapatkan  $\hat{y} = C \hat{\tilde{\theta}}$

$$\hat{y} = (C^T C + \lambda D)^{-1} C^T y$$

Berdasarkan uraian di atas, nilai  $\hat{\theta}$  bergantung pada parameter smoothing  $\lambda$ . Jika nilai  $\lambda$  besar akan menghasilkan bentuk kurva regresi yang sangat halus. Sebaliknya, jika nilai  $\lambda$  kecil akan memberikan bentuk kurva regresi yang sangat kasar. Akibatnya pemilihan parameter smoothing optimal perlu dilakukan. Dengan menggunakan generalized cross-validation (GCV) yang didefinisikan sebagai berikut:

$$GCV(\lambda) = \frac{n^{-1}RSS(\lambda)}{[1 - n^{-1}df\lambda]^2} = \frac{MSE(\lambda)}{[n^{-1}tr(I - S_\lambda)]^2} \quad (2.12)$$

Dimana  $(\lambda) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ,  $df\lambda = tr(S_\lambda)$

$S_\lambda = (C^T C + \lambda D)^{-1} C^T$  yang disebut dengan matriks smoothing (Ruppert, et al., 2003; Griggs, 2013)

#### 4. Pendugaan Area Kecil dengan Pendekatan Semiparametrik Penalized Spline

Pendugaan area kecil (SAE) adalah pendekatan yang digunakan untuk mengungkapkan hubungan antara variabel interest dengan variabel pendukung sebagai model linear dengan tambahan pengaruh acak area kecil. Dimisalkan  $\theta$  merupakan vektor dari parameter small area yang berukuran  $m \times 1$  dan diasumsikan vektor tersebut merupakan estimator langsung  $\theta$ . Jika dinyatakan  $m \times q$  adalah matriks dari variabel penyerta dari level area  $x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$  sehingga model SAE berbasis area dapat ditulis seperti persamaan (2.3) adalah sebagai berikut:

$$\theta_i = x_i^T \alpha + b_i v_i + ; i = 1, 2, \dots, m ; v_i \sim N(0, \sigma_v^2)$$

#### 5. Best Linear Unbiased Prediction (BLUP) dan Empirical Best Linear Unbiased Prediction (EBLUP)

Model small area terbagi menjadi model area level dan model unit level. Metode BLUP dan EBLUP salah satu metode yang digunakan untuk meminimumkan MSE. Pada metode BLUP, variansi pengaruh acak diasumsikan telah diketahui. Sedangkan pada metode EBLUP nilai variansi pengaruh acak small area tidak diketahui sehingga harus ditaksir dengan menggunakan metode Maximum Likelihood (ML). Misalkan data memenuhi model linear campuran berikut:

$$Y = X\beta + Z\gamma + e$$

#### 6. Pendugaan MSE dengan menggunakan Metode Model Jacknife dan Pendugaan MSE tidak langsung

Menurut Baillo dan Molina (2009), tujuan dari prosedur dan teknik yang digunakan dalam SAE adalah untuk memperoleh estimasi dengan tingkat presisi yang tinggi

pada area kecil tersebut. Tingkat presisi estimator ini dapat digambarkan oleh Mean Square Error (MSE). Penerapan jackknife pada SAE dilakukan untuk mengoreksi pendugaan MSE.

Fay dan Herriot (1979) mengembangkan model  $y_i = \mathbf{x}_i^T \beta + v_i + e_i$  sebagai dasar dalam pengembangan SAE. Untuk selanjutnya diasumsikan bahwa  $\beta$  dan  $\sigma_v^2$  tidak diketahui, akan tetapi  $\sigma_{ei}^2$  diketahui, dengan  $\beta_i = \sigma_{ei}^2 / (\sigma_v^2 + \sigma_{ei}^2)$  maka:

$$(\hat{\theta}_i^{EBLUP}) = (\hat{\theta}_i^{EBLUP} - \theta_i)^2 = (\hat{\theta}_i^{EBLUP}) + (\text{bias}(\hat{\theta}_i^{EBLUP}))^2$$

Persamaan tersebut dapat diuraikan menjadi:

$$(\hat{\theta}_i^{EBLUP}) = (\hat{\theta}_i^{EBLUP}) + (\hat{\theta}_i^{EBLUP} - \hat{\theta}_i^{EBLUP})^2$$

Metode jackknife pertama kali diperkenalkan oleh tukey pada tahun 1958 dan kemudian berkembang sebagai suatu metode untuk mengoreksi bias pada suatu estimator. Dengan melakukan penghapusan terhadap observasi ke- $i$  untuk  $i = 1, 2, \dots, m$  dan kemudian dilakukan pendugaan parameter misal  $\theta_{(i)}$ , maka penduga bias diduga dengan:

$$\text{bias}(\theta) = (m - 1)[\theta_{(i)} - \theta]$$

$$\theta_{(i)} = m^{-1} \sum_i^m \hat{\theta}(i)$$

Penduga jackknife diperoleh dari

$$\theta_{jack} = \theta - (\text{bias}(\theta)) \text{ dan } (\text{bias}(\theta)) = \left(\frac{n-1}{n}\right) \sum_i^m [\hat{\theta}_i - \theta]^2$$

### METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Survey Sosial Ekonomi Nasional (SUSENAS) (BPS.2020), Papua Dalam Angka tahun 2020 dan BPS Provinsi Papua. Objek dalam penelitian ini adalah Kabupaten/Kota di Provinsi Papua tahun 2019. Variabel yang digunakan dalam penelitian ini terbagi dalam dua variabel yaitu, variabel respon (Y) dan variabel penyerta/prediktor (X). Variabel yang menjadi perhatian dalam penelitian ini adalah rata-rata pengeluaran per kapita rumah tangga sebulan di Provinsi Papua.

**Tabel 3.2 Variabel Penelitian**

Variabel	N Variabel	Tipe
Y	Rata-Rata Pengeluaran per kapita Tahun 2019	Kontinu
X1	Rata-Rata Banyaknya Anggota Keluarga	Kontinu
X2	Jumlah Penggunaan Listrik PLN	Kontinu
X3	Jumlah Penduduk yang Sedang Sekolah	Kontinu
X4	Kepadatan Penduduk	Kontinu
X5	Keluarga Petani	Kontinu

Sebelum melakukan tahap penelitian, terlebih dahulu dilakukan tahap *Pre-Processing* data yang akan diolah sebagai berikut: (2.22)

1. Menyiapkan data pengeluaran per kapita dari data BPS Provinsi Sumatera Barat.
2. Menyimpan data untuk variabel prediktor ( $X_1, X_2, X_3, X_4$  dan  $X_5$ ) per kabupaten/kota di Provinsi Papua tahun 2019.

Selanjutnya metode dan tahapan penelitian yang akan dilakukan untuk mencapai tujuan penelitian adalah sebagai berikut:

- a. eksplorasi data variabel penelitian:
  1. Eksplorasi data dan membuat statistik deskriptif untuk variabel respon dan variabel penyerta/prediktor
  2. Membuat plot data antara variabel respon dengan semua variabel prediktor secara parsial kemudian hasil tersebut ditentukan komponen parametrik dan variabel komponen nonparametrik menentukan variabel prediktor yang menggunakan kurva parametrik linier dan kurva nonparametrik *spline*
- Menduga rata-rata pengeluaran per kapita dengan secara tidak langsung dengan menggunakan metode semiparametrik *penalized spline* dilakukan sebagai berikut:
  - a. Proses penentuan model menggunakan metode *Generalized Additive*  $y = f(x_1, x_2, x_3, x_4) + f(x_5) + v$ , dimana mencari nilai parametrik untuk  $f$  (*parametrik*), nilai dugaan parametrik digunakan untuk tahapan pencarian nilai nonparametrik *penalized*

spline untuk  $f(x_5)$ .

b. Memodelkan komponen nonparametrik dengan pendekatan *penalized spline*, kemudian menentukan jumlah knot yang digunakan dengan menggunakan *fixed selection method*.

c. Menghitung nilai GCV untuk masing-masing jumlah knot

d. Memilih jumlah knot optimum berdasarkan nilai GCV minimum

e. Menghitung nilai lambda ( $\lambda$ ) menggunakan kriteria GCV

f. Menyatakan bentuk area kecil dengan mengikuti suatu fungsi *penalized spline* berbasis polynomial terputus

$$\theta^{\square} = \Xi\beta + Z\gamma + bv$$

1. Menghitung  $\beta$ ,  $\gamma$  dan  $v$

2. Menduga jumlah per kapita untuk masing-masing kabupaten/kota di Provinsi Papua.

3. Menghitung RMSE hasil penduga pengeluaran per kapita untuk level kabupaten/kota dengan pendekatan *Jackknife* dan juga menggunakan RMSE langsung.

## HASIL DAN PEMBAHASAN

### 1. Gambaran Umum Wilayah Studi



Gambar 4.1 Peta Provinsi Papua

Provinsi Papua merupakan Provinsi yang paling luas wilayahnya dari seluruh Provinsi di Indonesia. Provinsi Papua terletak antara 2°25'-9° Lintang Selatan dan 130°-141° Bujur Timur. Provinsi Papua yang memiliki luas 317.062 km<sup>2</sup> atau 17,04 persen dari luas Indonesia, merupakan provinsi dengan wilayah terluas di Indonesia. Pada tahun 2019, Papua dibagi menjadi 29 kabupaten/kota dan 1 kota dimana Merauke merupakan kabupaten/kota terluas (56,84%) dan Kota Jayapura merupakan kabupaten/kota terkecil di Papua (0,1%).

Papua di bagian utara dibatasi Samudra Pasifik, sebelah selatan berbatasan dengan Laut Arafuru, sebelah barat berbatasan dengan Laut Seram, Laut Banda, Provinsi Papua Barat, Provinsi Maluku dan sebelah timur berbatasan dengan Papua New Guinea (BPS, 2010).

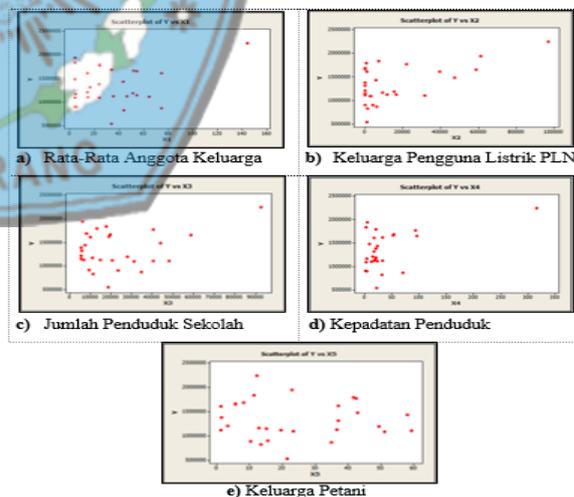
### 2. Eksplorasi

Tabel 4.1 Nilai Ringkasan Pendugaan Statistik Rata-Rata Pengeluaran Per Kapita

Statistik	Pendugaan Pengeluaran Per kapita
Mean	1341523
Std. Deviasi	389209
Variansi	1,515E+11
Minimum	541094
Maximum	2244874

Berdasarkan pada Tabel 4.1 rata-rata pengeluaran per kapita tertinggi di Provinsi Papua pada tahun 2019 terdapat pada Kota Jayapura yaitu sebesar 2244874, untuk rata-rata pengeluaran per kapita terendah pada tahun 2019 di Provinsi Papua terdapat pada kabupaten Dogiyai yaitu sebesar 541094.

### 3. Eksplorasi Data Hubungan Antar Variabel



Gambar 4.2 Scatterplot Lima Variabel

Berdasarkan hasil eksplorasi data hubungan antara variabel respon dan variabel prediktor, terdapat lima variabel prediktor yang digunakan dalam pembentukan model dimana empat variabel prediktor yaitu rata-rata anggota keluarga ( $X_1$ ), keluarga pengguna listrik PLN ( $X_2$ ), jumlah penduduk

yang sedang sekolah ( $X_3$ ), kepadatan penduduk ( $X_4$ ), di Provinsi Papua diasumsikan adanya pola hubungan linear yang mengikuti garis lurus, dengan kelinearan yang menunjukkan arah kanan, artinya hubungan kelinearannya tersebut bersifat positif, keempat variabel tersebut merupakan variabel yang bersifat parametrik. Dan ada satu variabel prediktor yaitu keluarga petani ( $X_5$ ) bersifat non linear atau mempunyai pola data tertentu dimana perilaku pola data ini dapat berubah-ubah, sehingga variabel keluarga petani ( $X_5$ ) merupakan variabel nonparametrik, berikut merupakan output hasil uji statistik linearitas antara variabel respon dengan variabel prediktor pada Tabel 4.3 :

**Tabel 4.3 Uji Linearitas Secara Parsial Antara Variabel Y dan X**

Predictor	Coef	SE Coef	T	Ket
Constant	48822	18062	2,70	Linear
X1	-165,3	236,5	-0,70	Linear
X2	-0,7764	0,6339	-1,22	Linear
X3	2,9870	0,7097	4,21	Linear
X4	234,0	215,2	1,09	Linear
X5	6,3	441,2	0,01	TL

Sehingga pada penelitian ini dilakukan pembentukan model *Small Area Estimation (SAE)* dengan pendekatan Semiparametrik *Panelized Spline*.

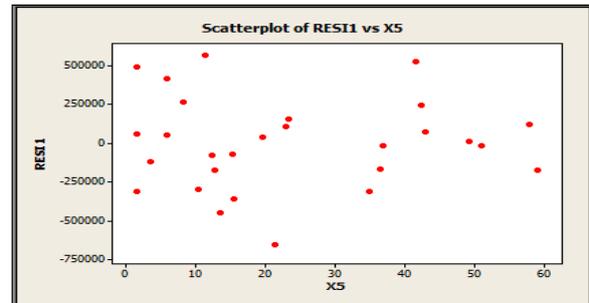
#### 4. Pemodelan Pengeluaran Per Kapita dengan Menggunakan Semiparametrik Panelized Spline

Proses pembentukan model dengan *Generalized Additive Model* untuk mencari nilai nilai parametrik  $f$  (*parametrik*) menghasilkan nilai parametrik adalah sebagai berikut:

$$y = 34.060 + (-62.805) x_1 + 1.121 x_2 + 1.596 x_3 + (-45.583) x_4 + f(x_5) + v \quad (2.26)$$

Selanjutnya dilakukan tahapan kedua yaitu pencarian nilai nonparametrik dengan pembentukan model SAE menggunakan

pendekatan *penalized spline*, proses ini dilakukan dengan memodelkan nilai residual pada model parametrik dengan variabel prediktor yang nonlinier ( $X_5$ ). Pola hubungan antara residual parametrik dengan variabel nonparametrik disajikan pada Gambar 4.3.



**Gambar 4.3 Scatterplot antara Residual  $f_1$  dan  $X_5$**

Gambar 4.3 merupakan scatterplot yang dihasilkan antara residual nilai parametrik dengan variabel Keluarga Petani ( $X_5$ ) terlihat bahwa pola data yang terbentuk tidak membentuk hubungan tertentu, sehingga Gambar 4.3 bersifat nonparametrik, dengan yang tidak mempunyai pola tertentu saja dapat didekati dengan menggunakan *spline* mempunyai model *polynomial tersegmen* yang dibangun sedemikian rupa sehingga kurva yang dibentuk mulus pada titik-titik yang disebut *knot*, dimana dua polinomial disatukan, akan tetapi *spline* mempunyai kekurangan yaitu jika *knot* yang dihasilkan terlalu banyak akan mengakibatkan model tersebut menjadi *overfitting*.

Karena nilai  $x_i$  yang digunakan *unique*, artinya setiap data  $x_i$  tidak ada yang bernilai sama. Dengan demikian jumlah *knot* yang dapat dihitung sebagai berikut:

$$\begin{aligned} K &= \min \left( \frac{1}{4} \times n, 35 \right) \\ &= \min \left( \frac{1}{4} \times 29, 35 \right) \\ &= \min (7,25, 35) \\ &= 7,25 \end{aligned}$$

Hal ini berarti jumlah *knot* yang akan digunakan dalam penelitian ini adalah sebanyak 7 *knot*. Kemudian menggunakan nilai GCV untuk masing-masing *knot* yang

telah ditentukan berdasarkan *fixed selection method*. Nilai GCV minimum yang diperoleh adalah 5.309. Berdasarkan kriteria GCV minimum diperoleh knot optimum sebanyak 7 buat knot yang masing-masing terletak pada titik 5.883, 11.429, 13.581, 19.870, 25.130, 37.770, 44.306.

Model semiparametrik *penalized spline* yang terbentuk dengan proses *generalized additive model* adalah sebagai berikut:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \gamma_1(x_5-5.883) + \gamma_2(x_5-11.429) + \gamma_3(x_5-13.581) + \gamma_4(x_5-19.870) + \gamma_5(x_5-25.130) + \gamma_6(x_5-37.770) + \gamma_7(x_5-44.306) + v$$

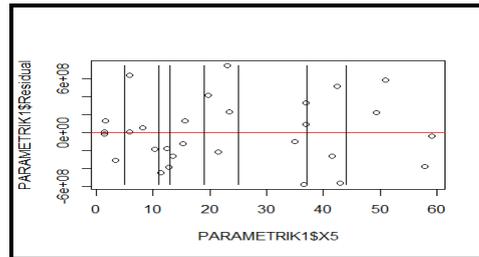
Pada model persamaan (2.27) di atas merupakan hasil knot yang didapatkan dengan menggunakan *penalized spline*, kemudian juga model yang akan digunakan untuk menduga pengeluaran per kapita di Provinsi Papua Barat dengan memaksimalkan fungsi *likelihood* atau *loglikelihood* sehingga dapat nilai estimasi  $\beta$  pada Tabel 4.4

**Tabel 4.4** Nilai estimasi  $\beta$

Parameter	Estimasi
$\beta_0$	34.060
$\beta_1$	-62.805
$\beta_2$	1.121
$\beta_3$	1.596
$\beta_4$	-45.583

Berdasarkan Tabel 4.4 di atas menunjukkan hasil nilai estimasi, sehingga nilai Interpretasi pada model persamaan (2.28) di atas menunjukkan bahwa terdapat perbedaan hasil untuk setiap kabupaten/kota, kemudian titik lokasi yang dihasilkan dengan menggunakan *penalized spline*.

Kemudian lokasi titik knot yang dihasilkan dengan menggunakan *penalized spline* disajikan pada Gambar 4 berikut.



Gambar 4.4 Lokasi titik Knot

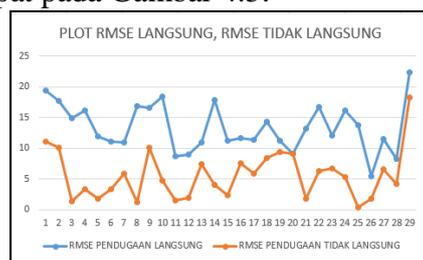
### 5. Model yang Terbentuk Untuk Small Area Estimation dengan Pendekatan Semiparametrik Penalized Spline untuk Data Pengeluaran Per Kapita

Berdasarkan hasil penelitian didapatkan model yang terbentuk untuk Small Area Estimation dengan pendekatan semiparametrik *penalized spline* untuk data pengeluaran per kapita tertinggi di Provinsi Papua adalah Kota jayapura dengan model yang terbentuk sebagai berikut :

$$y = 34.060 + (-62.805)x_1 + 1.121x_2 + 1.596x_3 + (-45.583)x_4 + 6.457y_1 + 0.911y_2 + (-1.24y_3) + (-7.529y_4) + (12.79y_5) + (-25.43y_6) + (-31.966y_7) + 608977.2$$

### 6. Kebaikan Model Pendugaan

Kebaikan model merupakan salah satu hal yang dilakukan untuk melihat apakah model yang telah terbentuk baik atau tidak, untuk mengetahui model yang telah terbentuk baik atau tidak akan dilakukan pendugaan nilai RMSE dengan menggunakan pendugaan tidak langsung (*indirect*) dan langsung (*direct*). Pendugaan tidak langsung pada subbab ini merupakan pendugaan yang menggunakan model yang mempunyai efek *random* untuk mengetahui *small area estimation* yang dihasilkan, sedangkan pada pendugaan langsung untuk menduga *Plot* perbandingan nilai RMSE dari hasil pendugaan langsung dan nilai RMSE pendugaan tidak langsung terdapat pada Gambar 4.5.



Gambar 4.5 Plot nilai RMSE

Berdasarkan Gambar 4.5 menunjukkan perbandingan RMSE yang dihasilkan pada metode *small area estimation* dengan menggunakan pendugaan langsung dan metode pendugaan tidak langsung, terlihat bahwa RMSE yang dihasilkan dengan menggunakan metode pendugaan tidak langsung nilai yang dihasilkan lebih kecil dibandingkan dengan hasil pendugaan langsung. Perbedaan RMSE yang dihasilkan disebabkan nilai pembagi pada pendugaan langsung dan tidak langsung berbeda. Secara umum dapat dikatakan bahwa pendugaan pengeluaran per kapita dengan menggunakan metode tidak langsung menghasilkan dugaan dengan tingkat akurasi dan presisi yang lebih baik dibandingkan dengan hasil pendugaan langsung.

Berdasarkan Gambar 4.5 terlihat bahwa RMSE menggunakan pendugaan langsung dan tidak langsung menghasilkan perbedaan untuk setiap kabupaten/kota. Ada beberapa pendugaan tidak langsung yang menghasilkan nilai RMSE yang cukup besar dibandingkan kabupaten lain yaitu pada kota Jayapura dan kabupaten Marauke. Secara visual hasil perbandingan antar plot RMSE penduga langsung dan penduga tidak langsung disajikan pada Gambar 4.5 suatu area tidak memerlukan model.

#### KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa *small area estimation* dengan pendekatan semiparametrik *penalized spline* dapat digunakan untuk memodelkan Pengeluaran per kapita pada level kabupaten/kota di Provinsi Papua. Hasil analisis data menunjukkan model semiparametrik *penalized spline* terbaik untuk pendugaan area kecil adalah model *penalized spline* dengan 7 lokasi titik knot. Model yang didapatkan:

$$y = 34.060 + (-62.805)x_1 + 1.121x_2 + 1.596x_3 + (-45.583)x_4 + \gamma_1(x_5 - 5.883) \pm \gamma_2(x_5 - 11.429) + \gamma_3(x_5 - 13.581) + \gamma_4(x_5 - 19.870) + \gamma_5(x_5 - 25.130) \pm \gamma_6(x_5 - 37.770) + \gamma_7(x_5 - 44.306) + v \quad (2.27)$$

Pengeluaran per kapita terbesar di Provinsi Papua pada tahun 2019 yaitu terdapat pada Kota Jayapura dan terkecil

pada kabupaten Dogiyai. Penggunaan metode RMSE *jackknife* untuk kebaikan model didapatkan hasil bahwa menggunakan pendugaan tidak langsung (pemodelan menggunakan *Small Area Estimation* (SAE)) menghasilkan dugaan dengan tingkat akurasi dan presisi yang lebih baik dibandingkan dengan hasil dugaan langsung.

#### DAFTAR PUSTAKA

- Badan Pusat Statistik (BPS). (2019). Papua dalam Angka 2020. Provinsi Papua.
- Baskara, (2014). Pendugaan Area Kecil Menggunakan Pendekatan Penalized Spline pada Pendugaan Pengeluaran Per kapita Tingkat Kecamatan Kabupaten Sumenep. Tesis. Institut Teknologi Sepuluh November. Surabaya.
- Chandra, H., Chamber, R., dan Salvati, N. (2009). Small Area Estimation of Proportion in Business Surveys. Australia: The University of Wollongong
- Darsyah, M.Y. (2013). *Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep dengan pendekatan Kernel-Bootstrap*. Jurnal. Intitut Teknologi Sepuluh Nopember, Surabaya.
- Darsyah, M.Y., Wasono, R. (2013). *Pendugaan Tingkat Kemiskinan di Kabupaten Sumenep dengan Pendekatan SAE*. Prosiding Semnas Satatistika UII, Yogyakarta.
- Darsyah, M.Y., Wasono, R. (2013). *Small Area Estimation terhadap Pengeluaran Per kapita di Kota Semarang ppendekatan Kernel*. Prosiding Semnas MIPA UNDIP, Semarang.
- Djuraidah, A. Dan Aunuddin. (2006). *Pendugaan Regresi Spline Terpenalti degan Pendekatan Model Linear Campuran*. Statistika. Vol.6. No.1, hal. 47-54.
- Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Deker
- Efron, B. Tibshirani, R. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.
- Wang, Y. (2011). *Smoothing Spline Methods and Applications*, CRC Press, Taylor & Francis Group, Boca Raton, London, New York.