

PERBANDINGAN ESTIMATOR SPLINE DAN KERNEL DALAM MODEL REGRESI NONPARAMETRIK MENGGUNAKAN DATA PDRB DI INDONESIA

L. P. S. Pratiwi¹ dan I M. P. P. Wijaya²

^{1,2} Program Studi Sistem Informasi, ITB STIKOM Bali

¹putu_safitri@stikom-bali.ac.id, ²pasek_pradnyana@stikom-bali.ac.id

ABSTRACT

PRDP describes the ability of a region in managing its natural resources and human resources. Indonesia's economic growth in the third quarter of 2021 experienced a slowdown compared to the previous quarter. This was due to the spread of the Delta variant of the COVID-19 outbreak, so the government imposed the fourth level of emergency Community Activity Restrictions (PPKM). The spline and kernel estimators have several advantages, namely they have a flexible form and are mathematically relatively easy to work with and have a relatively fast convergence mean value. The research method in this study consisted of three stages of research, namely the preparation stage including problem identification, finding research reference sources, and formulating research hypotheses. Followed by the implementation stage of the research, the author collects data, the data used for both response variables and predictors comes from BPS secondary data for each province in Indonesia, and continues with PRDP data processing starting with descriptive analysis and making scatterplots so that it can be seen whether the data follows the pattern. from nonparametric regression or not. Followed by modeling the estimator and determining the optimal value of each estimator. Then proceed with analysis, data interpretation and drawing conclusions. Finally, the reporting stage and evaluation of research results. These three stages are very important to do so that researchers can carry out their research coherently and systematically.

Keywords : Kernel, PDRB, Spline

ABSTRAK

PDRB menggambarkan kemampuan suatu daerah dalam mengelola sumber daya alam dan sumber daya manusia yang dimiliki. Pertumbuhan ekonomi Indonesia tahun 2021 triwulan ketiga mengalami perlambatan dibanding triwulan sebelumnya. Hal ini disebabkan karena adanya penyebaran wabah COVID-19 varian Delta sehingga pemerintah memberlakukan Pembatasan Kegiatan Masyarakat (PPKM) darurat level ke-empat. Estimator spline dan kernel mempunyai beberapa kelebihan, yaitu mempunyai bentuk yang fleksibel dan secara matematik relatif mudah dikerjakan serta mempunyai nilai mean kekonvergenan yang relatif cepat. Metode penelitian pada penelitian ini terdiri dari tiga tahapan penelitian, yaitu tahap persiapan meliputi identifikasi masalah, mencari sumber acuan penelitian, dan merumuskan hipotesis penelitian. Dilanjutkan dengan tahap pelaksanaan penelitian penulis melakukan pengumpulan data, data yang digunakan baik untuk variabel respon dan prediktor berasal dari data sekunder BPS tiap provinsi di Indonesia, serta dilanjutkan dengan pengolahan data PDRB yang diawali dengan analisis deskriptif serta membuat scatterplot agar dapat diketahui apakah data mengikuti pola dari regresi

nonparametrik atau tidak. Dilanjutkan dengan memodelkan estimator serta ditentukan nilai optimal masing-masing estimator. Kemudian dilanjutkan dengan analisis, interpretasi data dan penarikan kesimpulan. Terakhir, tahap pelaporan serta evaluasi hasil penelitian. Ketiga tahapan ini sangat penting untuk dilakukan agar peneliti dapat melaksanakan penelitiannya dengan runtut dan sistematis.

Kata kunci : Kernel, PDRB, Spline

I. PENDAHULUAN

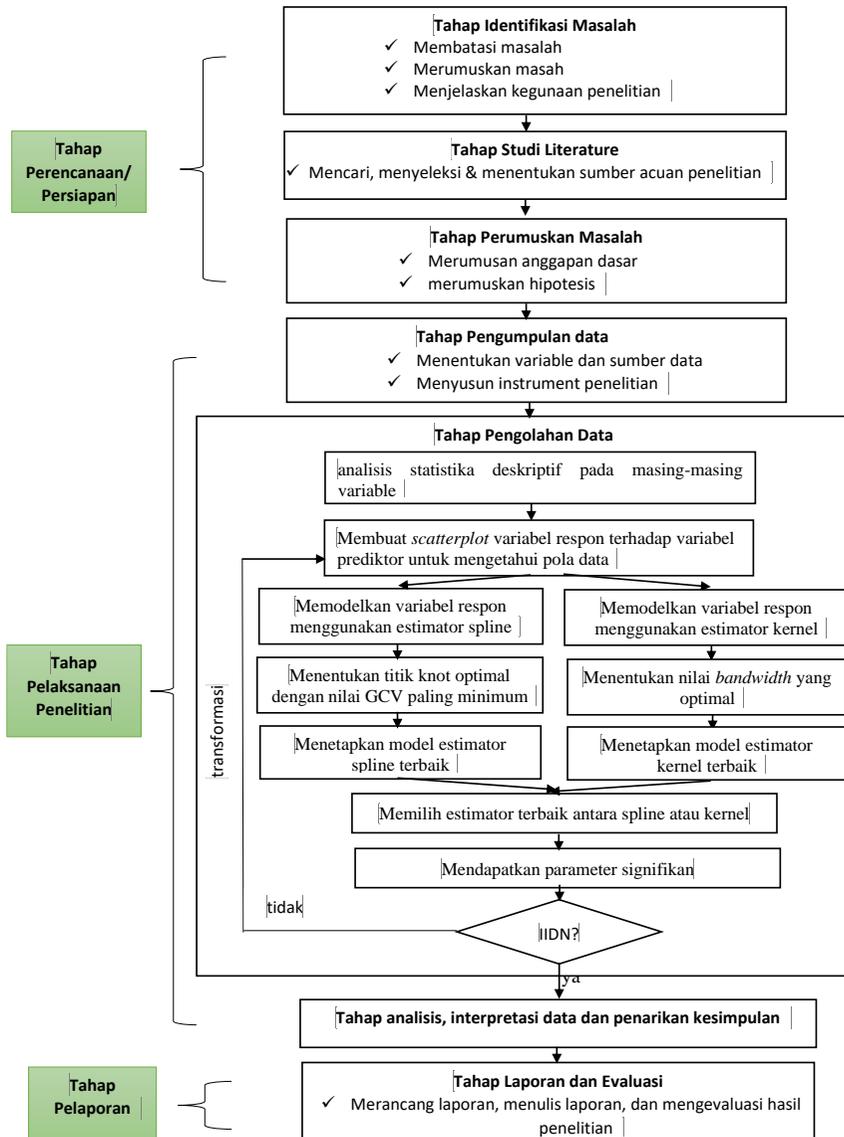
Pembangunan ekonomi yang dilakukan oleh pemerintah merupakan serangkaian kebijaksanaan yang bertujuan untuk meningkatkan taraf hidup masyarakat, memperluas lapangan kerja, pemerataan pendapatan masyarakat, dan meningkatkan hubungan antardaerah. Pada tahun 2021 triwulan ketiga pertumbuhan ekonomi di Indonesia mengalami perlambatan dibanding triwulan sebelumnya yang mencapai 3,51 persen namun masih tumbuh positif, sedangkan triwulan dua naik hingga mencapai 7,07 persen dari negative 0,74 persen pada triwulan satu [1]. Hal ini disebabkan karena adanya penyebaran wabah COVID-19 varian Delta sehingga pemerintah memberlakukan Pembatasan Kegiatan Masyarakat (PPKM) darurat level ke-empat. Penerapan PPKM ini, berdampak besar pada segala sector terutama sektor ekonomi. Salah satu cara untuk mengetahui aktivitas ekonomi secara keseluruhan adalah dengan melihat hasil perhitungan tingkat kenaikan Produk Domestik Regional Bruto (PDRB).

Persoalan utama analisis regresi nonparametrik adalah menentukan bentuk estimator. Estimator dalam regresi nonparametrik ada beberapa macam yaitu deret fourier, kernel, spline, polinomial lokal. Untuk menentukan estimator, dalam penelitian ini diusulkan suatu estimasi yaitu spline dan kernel yang masing-masing merupakan metode yang populer dikalangan statistik digunakan untuk pemodelan nonparametrik. Menurut Khusniawati [2] regresi kernel merupakan teknik statistik nonparametrik untuk menaksir nilai ekspektasi bersyarat misalnya nilai ekspektasi bersyarat dari Y diberikan ke X dari suatu variabel random. Sedangkan, estimator spline terdapat titik knot dimana titik knot ini berfungsi dalam mencari nilai estimasi data dan akan bergerak sendiri berdasarkan pola data [3]. Penelitian mengenai perbandingan estimator spline dan kernel pernah diteliti oleh Sukarsa dkk [4]; Yuliati & Sihombing [5], mereka berpendapat bahwa hal terpenting dari kedua estimator tersebut adalah pemilihan nilai optimal, untuk kernel sendiri pemilihan nilai yang optimalnya terletak pada nilai bandwidth dengan menggunakan kriteria Generalized Cross Validation (GCV) dilakukan dengan cara memperkecil nilai errornya. Sedangkan spline memilih nilai optimal berdasarkan pada kriteria GCV juga dimana untuk memperoleh titik knot optimal dapat dilihat dari nilai GCV yang paling minimum.

Berdasarkan penjabaran di atas maka diperlukan pemahaman mengenai permasalahan tersebut sehingga penelitian ini akan digunakan suatu estimator spline dan kernel untuk mendapatkan model PDRB di Indonesia, yang bertujuan untuk memodelkan serta membandingkan kedua estimator sehingga diketahui model mana yang lebih baik digunakan pada studi kasus PDRB di Indonesia.

II. METODE PENELITIAN

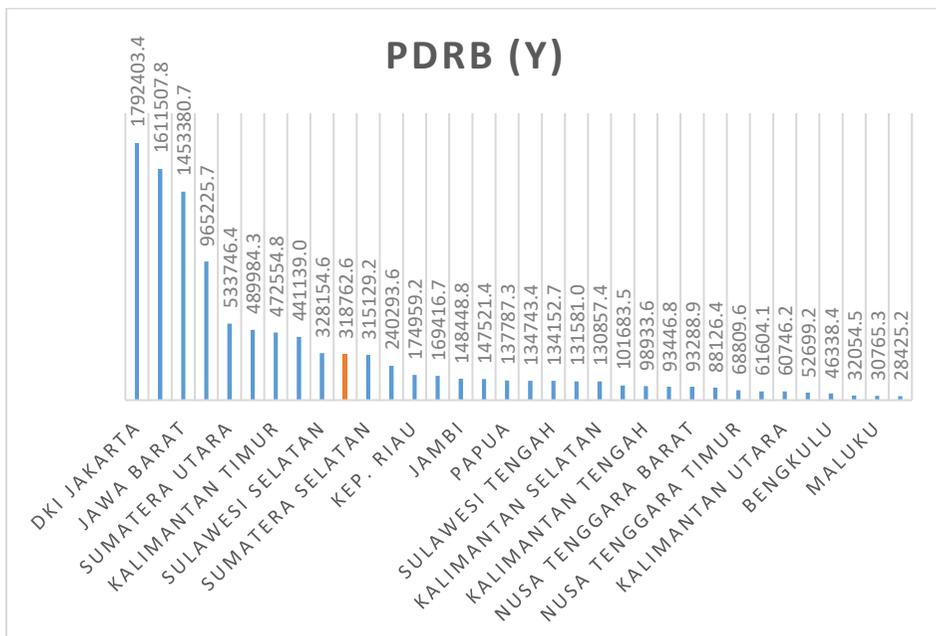
Data berasal dari BKPM (Badan Koordinasi Penanaman Modal) dan Badan Pusat Statistik (BPS). Objek penelitian terdiri dari 34 Provinsi di Indonesia. Variabel yang digunakan adalah, untuk variable respon yaitu Produk Domestik Regional Bruto (PDRB), sedangkan untuk variable predictor terdiri dari Tingkat Partisipasi Angkatan Kerja, Persentase Realisasi Investasi Penanaman Modal Asing, dan Realisasi Investasi Penanaman Dalam negeri. Adapun tahapan analisis data yang akan dilakukan dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 : Tahapan Analisis Data

III. HASIL DAN PEMBAHASAN

PDRB di 34 provinsi di Indonesia disajikan dalam diagram batang untuk memudahkan melihat Provinsi mana saja yang memiliki PDRB tertinggi hingga terendah yang ditunjukkan pada Gambar 2. Berdasarkan Gambar 2 Provinsi yang memiliki PDRB tertinggi pada tahun 2021 yaitu DKI Jakarta 1792403,4 Miliar Rupiah. Sementara itu, Provinsi Gorontalo memiliki PDRB terendah sebesar 28425,2 Miliar Rupiah selama tahun 2021. Terdapat 9 Provinsi di Indonesia yang memiliki PDRB di atas rata-rata keseluruhan PDRB di Indonesia. Provinsi tersebut antara lain DKI Jakarta, Jawa Timur, Jawa Barat, Jawa Tengah, Sumatera Utara, Riau, Kalimantan Timur, Banten, Sulawesi selatan. Sementara itu 25 provinsi lain PDRB-nya dibawah rata-rata keseluruhan PDRB di Indonesia.



Gambar 2 : Diagram Batang Variabel Respon y

3.1 Pemilihan Titik Knot Optimal Spline Linier Satu Titik Knot

Pada bagian ini dibahas pemilihan titik knot optimal pada regresi Spline linier satu titik knot pada PDRB di Indonesia dengan tiga variabel prediktornya. Berikut ini adalah estimasi model regresi nonparametrik Spline linier dengan satu titik knot pada kasus PDRB.

Model regresi nonparametrik Spline linier yang terbaik diperoleh dari titik-titik knot yang optimum. Titik knot optimum diperoleh dari nilai GCV yang paling kecil. Berikut adalah hasil analisis perhitungan GCV pada regresi nonparametrik dengan satu knot.

Tabel 1 : Nilai GCV untuk Satu Titik Knot

Knot			GCV
x1	x2	x3	
62.43	112.26	1610.67	19294411611
64.74	856.81	10118.27	15124373896
66.4	1388.62	16195.13	13432013950
67.72	1814.08	21056.61	12075116876
73.01	3515.89	40502.56	9330595629
76.32	4579.52	52656.27	14920277138
76.65	4685.88	53871.64	14909108004
76.98	4792.25	55087.01	17624980262
77.64	5004.97	57517.76	17427128369
77.97	5111.34	58733.13	17427126488

Berdasarkan Tabel 1, terlihat bahwa nilai GCV paling kecil adalah sebesar 9330595629 dengan titik knot optimal adalah sebagai berikut:

K1= 73.01, : K2= 3515.89, : K3= 40502.56

3.2. Pemilihan Titik Knot Optimal Spline Linier Dua Titik Knot

Setelah diperoleh GCV minimum pada Spline linier satu titik knot kemudian dilanjutkan menjadi dua titik knot pada setiap variabel. Berikut ini adalah estimasi model regresi nonparametrik Spline linier dengan dua titik knot pada kasus PDRB.

Tabel 2 : Nilai GCV untuk Dua Titik Knot

Knot			GCV
x1	x2	x3	
63.75	537.72	6472.16	18048570695
64.74	856.81	10118.27	
64.08	644.08	7687.53	12510159277
77.31	4898.61	56302.39	
70.03	2558.62	29564.21	10598794632
75.66	4366.79	50225.53	
70.03	2558.62	29564.21	9314544993
77.31	4898.61	56302.39	
70.37	2664.98	30779.59	5916587562
71.36	2984.07	34425.7	
70.37	2664.98	30779.59	8211797988

73.34	3622.25	41717.93	
70.37	2664.98	30779.59	9458310283
74	3834.98	44148.67	
70.37	2664.98	30779.59	11163846193
75.32	4260.43	49010.16	
76.32	4579.52	52656.27	14928028823
76.98	4792.25	55087.01	
77.64	5004.97	57517.76	17427128998
77.97	5111.34	58733.13	

Berdasarkan Tabel 2, terlihat bahwa nilai GCV minimum adalah 5916587562, dengan titik knot optimal adalah:

: (K1=70.37; K2=71.36), : (K3=2664.98; K4=2984.07),

: (K5=30779.59; K6=34425.7),

3.3. Pemilihan Titik Knot Optimal Spline Linier Tiga Titik Knot

Setelah diperoleh dua titik knot, kemudian dilanjutkan dengan tiga titik knot dengan estimasi model regresi nonparametrik Spline linier tiga knot ditunjukkan pada Tabel 3.

Tabel 3 : Nilai GCV untuk Tiga Titik Knot

Knot			GCV
x1	x2	x3	
62.43	112.26	1610.67	23932703263
62.76	218.63	2826.04	
63.09	324.99	4041.41	
63.42	431.35	5256.79	15336906986
66.73	1494.99	17410.5	
70.37	2664.98	30779.59	
64.08	644.08	7687.53	12459407176
70.37	2664.98	30779.59	
75.32	4260.43	49010.16	
66.07	1282.26	14979.76	11200073364
66.73	1494.99	17410.5	
75.32	4260.43	49010.16	
69.37	2345.89	27133.47	9587419636
76.98	4792.25	55087.01	
77.31	4898.61	56302.39	
70,37	2664,98	30779,59	5199284230

71,03	2877,71	33210,33	
71,36	2984,07	34425,70	
72.68	3409.52	39287.19	9332221895
77.31	4898.61	56302.39	
77.97	5111.34	58733.13	
74	3834.98	44148.67	10041111622
74.33	3941.34	45364.04	
74.66	4047.7	46579.41	
75.99	4473.16	51440.9	11608471381
76.98	4792.25	55087.01	
77.97	5111.34	58733.13	
77.31	4898.61	56302.39	17619309287
77.64	5004.97	57517.76	
77.97	5111.34	58733.13	

Berdasarkan Tabel 3, terlihat bahwa GCV minimum sebesar **5199284230**. Berikut ini adalah nilai titik knot optimum untuk tiga titik knot pada masing-masing variabel komponen nonparametrik, sebagai berikut:

: (K1=70,37; K2=2664,98; K3=30779,59),

: (K4=71,03; K5=2877,71; K6=33210,33),

: (K7=71,36; K8=2984,07; K9=34425,70)

3.4. Berdasarkan Estimator Kernel

Dalam penggunaan Estimator Kernel, hal terpenting adalah menentukan nilai bandwidth. Pemilihan bandwidth terbaik dilihat dari nilai bandwidth yang optimum yang bersesuaian dengan fungsi kernel yang digunakan. Fungsi kernel yang digunakan adalah kernel Gaussian. Pemilihan bandwidth optimum dapat dilakukan dengan menggunakan dua cara yaitu secara subyektif dengan metode coba-coba dan secara obyektif dengan metode validasi silang. Pemilihan bandwidth dengan metode coba-coba memiliki kelemahan dimana peneliti harus memiliki sense yang baik dalam memilih bandwidth yang sesuai digunakan pada fungsi kernel. Untuk itu pada penelitian ini dilakukan pemilihan bandwidth optimum secara obyektif dengan metode validasi silang yaitu GCV, dimana bandwidth optimum ditentukan dari nilai GCV yang terkecil.

Tabel 4 : Nilai Bandwith dan GCV Minimum

Keterangan	h1	h2	h3
Bandwidth Optimal	1540	15400	23050
GCV Minimum	228158049.8		
MSE	2.97646E+11		

Setelah diperoleh nilai bandwidth optimal maka diperoleh estimator Nadaraya Watson sebagai berikut, dengan nilai h_j tertera pada Tabel 4.

$$m(x_i) = \frac{\sum_{i=1}^n \prod_{j=1}^3 K\left(\frac{x_{ij} - x_j}{h_j}\right) y_i}{\sum_{i=1}^n \prod_{j=1}^3 K\left(\frac{x_{ij} - x_j}{h_j}\right)} \quad (1)$$

Berdasarkan hasil Estimator Spline dengan GCV satu knot, dua knot, tiga knot dan Estimator kernel Nadaraya Watson fungsi kernel Gaussian akan dibandingkan masing-masing GCV minimum dapat dilihat dari nilai GCV terkecil maka estimator tersebut yang terbaik dan dapat memodelkan PDRB di Indonesia, dapat dilihat pada Tabel 5.

Tabel 5 : Perbandingan Nilai GCV minimum dari titik knot

Estimator	Nilai GCV Minimum
Spline Satu Titik Knot	9330595629
Spline Dua Titik Knot	5916587562
Spline Tiga Titik Knot	5199284230
Kernel	228158049.8

Pada Tabel 5 diperoleh nilai GCV terkecil terletak pada estimator Kernel yaitu sebesar 228158049.8.

IV. KESIMPULAN

Berdasarkan hasil analisis maka kesimpulan yang dapat diambil dari penelitian ini adalah estimator terbaik untuk memodelkan angka PDRB di Indonesia yaitu estimator kernel dalam model regresi nonparametrik, dengan nilai GCV minimum yaitu sebesar 228158049.8.

DAFTAR PUSTAKA

- [1]. BPS (Badan Pusat Statistik), *Berita Resmi Statistik: Pertumbuhan Ekonomi Indonesia Triwulan III-2021. No. 83/11/Th. XXIV*, 5 November 2021, Jakarta:BPS.
- [2]. Khusniawati, F., Pengujian Hipotesis Parameter Komponen Spline Dalam Model Regresi Nonparametrik Campuran Spline Dan Kernel, Tesis, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya, 2017.
- [3]. Budiantara, I.N., Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang, *Jurnal Sains Dan Seni Pomits*, 3 No. 1, 2009, (2014) 2337-3520 (2301-928X Print), Institut Teknologi Sepuluh November, Surabaya.
- [4]. Sukarsa, I.K.G., Srinadi, I. G. A. M., & Lestari, N L A P., Perbandingan Estimator Kernel Dan Estimator Spline Dalam Model Regresi Nonparametrik, *Jurnal KNM XVI 3-6 Juli 2012 UNPAD*, 2012, hal. 581-590.
- [5]. Yuliati, I.F & Sihombing, P.R., Pemodelan Fertilitas di Indonesia Tahun 2017 Menggunakan Pendekatan Regresi Nonparametrik Kernel dan Spline, *Jurnal Statistika dan Aplikasinya (JSA)*, 4 No. 1, Juni 2020.