



COMPARISON OF GENERALIZED CROSS VALIDATION (GCV) METHODS WITH CROSS VALIDATION (CV) TO DETERMINE OPTIMAL KNOTS IN FOURIER SERIES NONPARAMETRIC REGRESSION (Case Study: Poverty Rate in North Sumatra Province)

M.Najib Amrullah¹, Siti Hariksa Amalia²

¹ Department of Mathematical Sciences, Universiti Teknologi Malaysia

^{2,3}Department of Statistics, Universitas Muhammadiyah ,Semarang, Indonesia

* E-mail: amrullah@utm.ac.id

ABSTRACTS

Poverty is multidimension problem that directly affected human life. North Sumatera Province is a province with the most poor population in Sumatera Island that is 1453.87 (thousand souls). A factor thought to affect the high poverty rate in the northern province of Sumatera is per capita income. Distribution of data on the number of poor people and per capita income of each regency/ city in North Sumatera Province tends to be repetitive and the pattern is unknown, then in this study used a nonparametric regression approach of fourier series. The purpose of this study is to model poverty in regency/ city in North Sumatera Province using nonparametric regression fourier series with GCV and CV methods to determine the optimal knots (K) and compare the results of modelling based on GCV and CV methods to determine the optimal K in nonparametric regression with fourier series. The population in this study were all regency/ cities in North Sumatera Province, amounting to 33 regency/ cities. The results showed that the GCV method in determining the optimal K in the nonparametric regression of the fourier series was better than the CV method for modelling poverty levels in North Sumatera Province. The GCV method produces an R^2 of 96% with an MSE value of 52.14 while the CV method produces an R^2 67% with an MSE value of 463.62. The results of the analysis of poverty level modelling in North Sumatera Province in 2017 with the GCV method for determining optimal K in nonparametric regression of fourier series obtained the highest poverty estimate of 134.52 (thousand inhabitants) found in Medan City and the lowest poverty estimate of 11.05 (thousand inhabitants) found in Pakpak Bharat Regency.

ARTICLE INFO

Article History:

Received 07 Okt 2020

Revised 03 Maret 2022

Accepted 04 April 2022

Available online 20 Mei 2022

Keyword:

Poverty,

GCV,

CV,

Nonparametric Regression,

Fourier Series

ABSTRAK

Kemiskinan merupakan masalah multidimensi yang langsung mempengaruhi kehidupan manusia. Provinsi Sumatera Utara merupakan provinsi yang memiliki jumlah penduduk miskin terbanyak di Pulau Sumatera yaitu berjumlah 1453.87 (ribu jiwa). Faktor yang diduga mempengaruhi tingginya angka kemiskinan di provinsi Sumatera Utara adalah pendapatan per kapita. Persebaran data jumlah penduduk miskin dan pendapatan per kapita setiap kabupaten/kota di provinsi Sumatera Utara cenderung berulang dan tidak diketahui polanya, maka dalam penelitian ini digunakan pendekatan regresi nonparametrik deret fourier. Tujuan dari penelitian ini adalah untuk memodelkan kemiskinan pada kabupaten/kota di Provinsi Sumatera Utara menggunakan regresi nonparametrik deret fourier dengan metode GCV dan CV untuk menentukan knot (K) optimalnya dan membandingkan hasil pemodelan berdasarkan metode GCV dan CV untuk menentukan K optimal dalam regresi nonparametrik deret fourier. Populasi dalam penelitian ini adalah seluruh kabupaten/kota di Provinsi Sumatera Utara yang berjumlah 33 kabupaten/kota. Hasil penelitian menunjukkan bahwa metode GCV dalam penentuan K optimal dalam regresi nonparametrik deret fourier lebih baik dibandingkan metode CV untuk memodelkan tingkat kemiskinan di Provinsi Sumatera Utara. Metode GCV menghasilkan R^2 sebesar 96% dengan nilai MSE 52.14 sedangkan metode CV menghasilkan R^2 sebesar 67% dengan nilai MSE 463.62. Hasil analisis pemodelan tingkat kemiskinan di Provinsi Sumatera Utara tahun 2017 dengan metode GCV

Keyword:

Kemiskinan,

GCV,

CV,

Regresi Nonparametrik, Deret

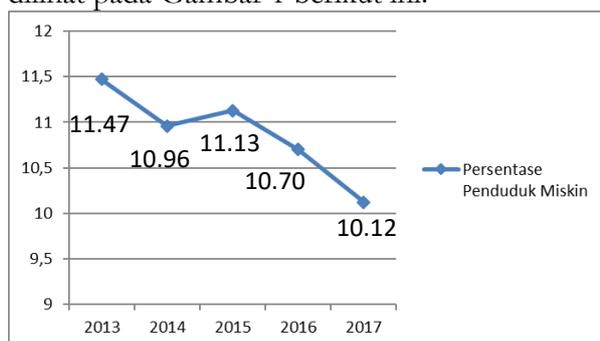
Fourier.

untuk penentuan K optimal dalam regresi nonparameterik deret fourier didapatkan estimasi kemiskinan tertinggi sebesar 134.52 (ribu jiwa) terdapat di Kota Medan dan estimasi kemiskinan terendah di sebesar 11.05 (ribu jiwa) terdapat di Kabupaten Pakpak Bharat.

1. PENDAHULUAN

Kemiskinan merupakan masalah multidimensi yang langsung mempengaruhi seluruh aspek kehidupan. Kemiskinan adalah suatu situasi dimana seseorang atau rumah tangga mengalami kesulitan untuk memenuhi kebutuhan dasar, sementara lingkungan pendukungnya kurang memberikan peluang untuk meningkatkan kesejahteraan secara berkesinambungan atau keluar dari kerentanan (Cahyat dkk, 2007).

Masalah kemiskinan selalu dihadapi oleh negara-negara di dunia, khususnya negara berkembang. Indonesia merupakan salah satu negara berkembang yang sedang berusaha mengurangi jumlah penduduk miskinnya setiap tahun. Pemerintah banyak mengeluarkan program-program penanggulangan kemiskinan seperti Program Keluarga Harapan (PKH), dana desa dan program subsidi lainnya, namun program-program tersebut belum bisa menekan angka persentase kemiskinan di Indonesia. Berdasarkan data Badan Pusat Statistik (BPS) persentase penduduk miskin di Indonesia pada tahun 2013 sampai 2017 dapat dilihat pada Gambar 1 berikut ini:



Gambar 1. Persentase Penduduk Miskin di Indonesia

Berdasarkan Gambar 1 dapat dilihat bahwa persentase penduduk miskin di Indonesia dari tahun 2013 sampai 2017 berfluktuasi. Persentase penduduk miskin pada tahun 2013 sampai tahun 2017 ini melebihi target pembangunan yang disepakati oleh pemerintah dalam Anggaran Perencanaan Belanja Negara (APBN) tahun 2018, sebesar 9.5% sampai 10% dari jumlah penduduk Indonesia. Oleh karena

itu perlu upaya yang konkret dari pemerintah agar Indonesia bebas dari masalah kemiskinan.

Provinsi Sumatera Utara merupakan provinsi di Pulau Sumatera yang memiliki jumlah penduduk miskin terbanyak, menurut Badan Pusat Statistik (BPS) berjumlah 1453.87 (ribu jiwa) pada tahun 2017. Menurut data BPS Sumatera Utara secara umum persentase penduduk miskin di beberapa kabupaten/kota masih tergolong tinggi dan masih terdapat beberapa kabupaten/kota yang memiliki persentase penduduk miskin di atas target pembangunan yang ditetapkan oleh pemerintah yaitu 22 kabupaten/kota dari 33 kabupaten/kota yang ada di Provinsi Sumatera Utara. Kondisi inilah yang dapat menjadi permasalahan bagi pemerintah Provinsi Sumatera Utara yang berdampak buruk terhadap perekonomian di masyarakat sekarang dan masa yang akan datang.

Untuk mengetahui faktor-faktor apa saja yang mempengaruhi kemiskinan di Provinsi Sumatera Utara, maka diperlukan sebuah analisis statistika. Salah satu analisis dalam statistika yang berguna menyelidiki dan memodelkan hubungan antara beberapa variabel adalah analisis regresi. Menurut persebaran pola datanya analisis regresi ada dua jenis yaitu regresi nonparametrik dan regresi parametrik. Regresi nonparametrik digunakan apabila pola data dari variabel penelitian tidak diketahui. Pendekatan regresi nonparametrik telah banyak dikembangkan antara lain menggunakan Spline, Kernel, Polinomial Lokal, Wavelet dan Fourier. Salah satu keunggulan pendekatan regresi nonparametrik dengan menggunakan deret fourier adalah mampu mengatasi data yang mempunyai sebaran trigonometri, dalam hal ini adalah sinus dan cosinus. Pola data yang sesuai dengan pendekatan fourier merupakan pola data yang berulang (fluktuatif), yaitu pengulangan terhadap nilai variabel dependen untuk variabel independen yang berbeda-beda (Prahutama, 2018).

Salah satu hal yang penting dalam pemodelan regresi nonparametrik

menggunakan deret Fourier adalah pemilihan parameter penghalus. Parameter penghalus atau Knot merupakan pengontrol keseimbangan fungsi. Jika λ besar, maka kurva estimator fungsi akan semakin smooth, tetapi kemampuan untuk mendekati pola data kurang baik. Sebaliknya jika λ kecil, maka estimator kurva akan semakin kasar. Oleh karena itu diperlukan nilai λ yang tidak terlalu besar ataupun kecil sehingga akan menghasilkan kurva estimator yang baik.

Terdapat beberapa metode yang dapat digunakan untuk menentukan parameter penghalus yang optimal dalam regresi nonparametrik deret Fourier antara lain metode *Generalized Cross Validation* (GCV) (Wahba, 1990 dalam Devi, 2015), *Cross Validation* (CV) (Craven dan Wahba, 1979 dalam Devi, 2015), *Unbiased Risk* (UBR) (Wahba, 1990; Wang, 1998 dalam Devi, 2015) dan *Generalized Maximum Likelihood* (GML) (Eubank, 1999 dalam Devi 2015). Salah satu metode yang banyak digunakan dalam menentukan parameter penghalus yang optimal adalah GCV. Adapun kelebihan yang dimiliki metode GCV antara lain sederhana dan efisien dalam perhitungan, optimal secara asimtotik, invarian terhadap transformasi dan dalam perhitungannya varians populasi tidak perlu diketahui (Wahba, 1990 dalam Devi, 2015). Metode GCV dan CV merupakan metode yang terkenal karena sifat optimalnya. Metode GCV merupakan metode CV yang terboboti, yaitu dengan memberikan bobot yang sama pada setiap pengamatan, sedangkan CV merupakan regresi terboboti dengan bobot yang berbeda-beda pada masing-masing pengamatan sesuai dengan kontribusinya.

Penelitian-penelitian mengenai regresi nonparametrik deret Fourier menggunakan GCV untuk penentuan K optimal yang telah dilakukan sebelumnya diantaranya, Penelitian Prahutama (2013) Model Regresi Nonparametrik dengan Pendekatan Deret Fourier Pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur. Penelitian Nurjanah (2015) Model Regresi Nonparametrik dengan pendekatan deret Fourier pada Pola Data Curah Hujan Di Kota Semarang.

Penelitian – penelitian terdahulu mengenai variabel pendapatan per kapita yang

berpengaruh signifikan terhadap tingkat kemiskinan diantaranya, Penelitian Pulungan (2016) Analisis Faktor-faktor yang Mempengaruhi Tingkat Kemiskinan di Provinsi Sumatera Utara. Penelitian Puspita (2015) Analisis Determinan Kemiskinan di Provinsi Jawa Tengah. Penelitian Hudaya (2009) Faktor – faktor yang Mempengaruhi Tingkat Kemiskinan di Indonesia.

Berdasarkan uraian di atas, variabel respon (Y) yang digunakan adalah jumlah penduduk miskin di Provinsi Sumatera Utara dan variabel prediktor (X) adalah pendapatan per kapita di Provinsi Sumatera Utara. Penelitian ini akan membahas tentang perbandingan hasil pemilihan parameter penghalus optimal dengan metode GCV dan CV dalam regresi nonparametrik menggunakan deret Fourier untuk memodelkan kemiskinan pada Kabupaten/Kota di Provinsi Sumatera Utara. Hal ini dikarenakan pola sebaran data tingkat kemiskinan setiap Kabupaten/Kota di Provinsi Sumatera Utara dan faktor pengaruhnya cenderung berulang dan pola datanya tidak diketahui.

TINJAUAN PUSTAKA

1. Analisis Regresi

Analisis regresi adalah suatu metode statistika yang umum digunakan untuk melihat pengaruh antara variabel independen dengan variabel dependen. Jika kurva regresi merupakan model parametrik maka disebut sebagai regresi parametrik dan apabila model yang diasumsikan ini benar, maka pendugaan parametrik sangat efisien, tetapi jika tidak, menyebabkan interpretasi data yang menyesatkan (Hardle, 1994). Misalnya Y adalah variabel respon dan X adalah variabel prediktor, secara umum hubungan antara Y dan X dapat ditulis sebagai berikut:

$$Y_i = \beta_0 + \beta_1 x_i + \dots + \beta_n x_n + \varepsilon_i \\ i = 1, 2, \dots, n$$

Dimana,

Y_i = Variabel respon

β_i = Parameter

x_i = Variabel predictor

ε_i = Error

Apabila tidak terdapat informasi apapun tentang bentuk fungsi, maka digunakan

pendekatan nonparametrik (Hardle,1994).

2. Regresi Nonparametrik

Regresi nonparametrik adalah regresi yang diasumsikan tidak diketahui bentuk kurva regresi atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, 1999). Adapun model regresi nonparametrik secara umum disajikan pada persamaan berikut:

$$Y_i = f(x_i) + \epsilon_i, i = 1,2, \dots, n$$

Dimana,

Y_i = variabel respon

x_i = variabel prediktor

$f(x_i)$ = fungsi regresi

ϵ_i = error yang berdistribusi independen dengan mean nol dan varians σ^2

$f(x_i)$ merupakan kurva regresi yang tidak diketahui bentuknya. Kurva $f(x_i)$ diasumsikan smooth pada ruang fungsi tertentu.

3. Deret Fourier

Fungsi Deret Fourier adalah sebagai berikut :

$$f(t) = \frac{1}{2} a_0 + \gamma t + \sum_{k=1}^K a_k \cos\left(\frac{2\pi kt}{2L}\right)$$

Dengan a_0, a_k , dan b_k adalah koefisien Fourier (Asrini, 2012).

Tingkat kemulusan estimator deret Fourier ditentukan oleh pemilihan parameter pemulus K. Semakin kecil parameter pemulus K, semakin mulus estimasinya dan semakin besar parameter pemulus K, semakin kurang mulus estimasi dari f . Oleh karena itu, perlu dipilih K yang optimal.

4. Generalized Cross Validation (GCV)

Penentuan nilai K optimal akan menghasilkan nilai koefisien R2 yang tinggi. Generalized Cross Validation (GCV) didefinisikan sebagai berikut :

$$GCV(K) = \frac{MSE(K)}{\left[1 - \left(\frac{trace(A_K)}{n}\right)\right]^2}$$

(Wu dan Zhang ,1988)

Dimana,

$$MSE(K) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$(A_K) = A(A^T A)^{-1} A^T$$

$$\hat{y}_i = (A_K)y$$

n = banyak pengamatan

Nilai GCV terkecil akan menghasilkan K yang optimal (Asrini, 2012 dalam Prahutama, 2013).

5. Cross Validation (CV)

$$CV(K) = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - f_{K_i}}{1 - s_i}\right)^2$$

Pemilihan K optimal dengan metode CV didapatkan dengan meminimalkan fungsi CV (Devi, 2015).

6. Estimasi Parameter Regresi Nonparametrik Deret Fourier

$f(t)$ merupakan kurva yang tidak diketahui bentuknya maka $f(t)$ didekati dengan menggunakan Deret Fourier yaitu sebagai berikut :

$$f(t) = \frac{1}{2} \alpha_0 + \gamma t + \sum_{k=1}^K \alpha_k \cos kt$$

Persamaan di atas didapatkan elemen diagonalnya yaitu :

$$f(t) = \frac{1}{2} \alpha_0 + \gamma t + (\alpha_{i1} \cos t_i + \alpha_{i2} \cos t_i + \dots + \alpha_{ik} t_i)$$

Jika $f(t) = B\hat{\theta}$, maka :

$$B = \begin{bmatrix} 1 & t_1 & \cos t_1 & \dots & \cos kt_1 \\ 1 & t_2 & \cos t_2 & \dots & \cos kt_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cos t_n & \dots & \cos kt_n \end{bmatrix}_{n(k+2)} \quad \hat{\theta} = \begin{bmatrix} \phi \\ \gamma \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}_{(k+2) \times 1}$$

Dengan $\phi = \frac{n}{2} \alpha_0$

7. Kemiskinan Penduduk

Kemiskinan adalah kondisi kehidupan yang serba kekurangan yang dialami seseorang yang pengeluaran per kapitanya selama sebulan tidak cukup untuk memenuhi kebutuhan standar hidup minimum. Kebutuhan standar hidup minimum digambarkan dengan garis kemiskinan (GK), yaitu batas minimum pengeluaran per kapita per bulan untuk memenuhi kebutuhan minimum makanan dan nonmakanan. Batas kecukupan minimum makanan mengacu pada Widya Karya Nasional Pangan dan

Gizi pada tahun 1978, yaitu besarnya rupiah yang dikeluarkan untuk makanan yang memenuhi kebutuhan minimum energy 2100 kalori per kapita per hari, sedangkan kebutuhan minimum nonmakanan mencakup pendapatan untuk perumahan, penerangan, bahan bakar, pakaian, pendidikan, kesehatan, transportasi, barang-barang tahan lama serta barang dan jasa esensial lainnya. Jumlah orang miskin dapat dilihat melalui jumlah orang yang berada di bawah atau sama dengan garis kemiskinan (Badan Pusat Statistik, 2002).

8. Pendapatan Per kapita

Pendapatan per kapita atau produk domestik bruto per kapita adalah pendapatan rata-rata penduduk suatu negara (Untoro, 2010:13). Persamaan untuk mencari pendapatan per kapita adalah :

$$\text{Pendapatan per kapita} = \frac{\text{Produk Nasional Bruto (GNP)}}{\text{Jumlah Penduduk}}$$

2. METODE

1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui publikasi lembaga pemerintahan resmi yaitu Badan Pusat Statistik Provinsi Sumatera Utara periode tahun 2017. Unit observasi dalam penelitian ini adalah Kabupaten/Kota di Provinsi Sumatera Utara.

2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini ada dua variabel yaitu variabel respon dan variabel prediktor. Variabel respon dalam penelitian ini adalah jumlah penduduk miskin dari 33 kabupaten/kota di Provinsi Sumatera Utara, sedangkan variabel prediktor dalam penelitian ini adalah pendapatan per kapita dari 33 kabupaten/kota di Provinsi Sumatera Utara.

Y : Jumlah penduduk miskin

X : Pendapatan per kapita

3. Analisis Data

Tahapan analisis yang digunakan dalam penelitian ini adalah :

1. Membuat plot data tingkat kemiskinan kabupaten/kota di provinsi Sumatera Utara (x,y).
2. Memodelkan tingkat kemiskinan kabupaten/kota di Provinsi Sumatera Utara dengan pendekatan deret fourier, dengan langkah-langkah sebagai berikut :
 - a. Diberikan model regresi nonparametrik : $y = f(x) + \varepsilon$
 - b. Membuat kurva regresi nonparametrik dengan pendekatan deret fourier :

$$\hat{f}(t) = \frac{1}{2}a_0 + \sum_{k=1}^K \cos\left(\frac{2\pi kt}{2L}\right)$$

Menentukan K optimal dengan metode GCV pada data tingkat kemiskinan kabupaten/kota di provinsi Sumatera Utara. Dengan persamaan sebagai berikut :

$$GCV(K) = \frac{MSE(K)}{\left[1 - \left(\frac{\text{trace}(A_k)}{n}\right)\right]^2}$$

Menentukan K optimal dengan metode CV pada data tingkat kemiskinan kabupaten/kota di provinsi Sumatera Utara. Dengan persamaan sebagai berikut :

$$CV(K) = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - f_{K_i}}{1 - s_i}\right)^2$$

- c. Setelah didapat K optimal, langkah selanjutnya adalah menjalankan program estimasi model regresi nonparametric deret fourier pada data tingkat kemiskinan kabupaten/kota di Provinsi Sumatera Utara.
3. Pemilihan model terbaik berdasarkan nilai R2 dan MSE dari metode GCV dan CV untuk penentuan K optimal dalam regresi nonparametrik deret fourier.

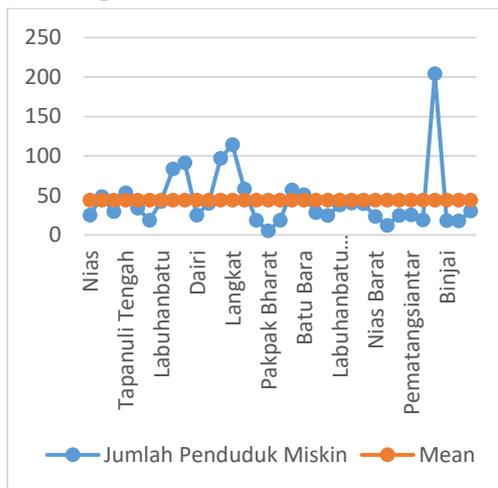
Analisis hasil pemodelan terbaik tingkat kemiskinan dengan pendekatan Deret Fourier.

3. HASIL DAN PEMBAHASAN

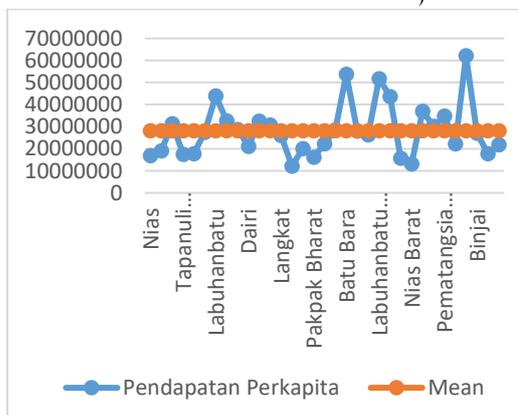
1. Pemodelan Tingkat Kemiskinan Kabupaten/kota di Provinsi Sumatera Utara dengan Pendekatan Regresi Nonparametrik Deret Fourier

Pemodelan tingkat kemiskinan kabupaten/kota di Provinsi Sumatera Utara dengan pendekatan deret fourier dengan variabel respon yaitu pendapatan perkapita dan variabel prediktornya yaitu jumlah penduduk miskin. Dalam membuat model tingkat kemiskinan kabupaten/kota di Provinsi Sumatera Utara dengan pendekatan regresi nonparametrik deret fourier, langkah-langkahnya yaitu:

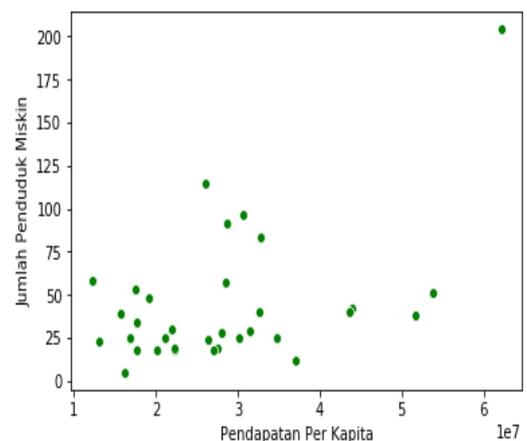
1. Plot data pendapatan per kapita dan jumlah penduduk miskin (x, y)



Berdasarkan gambar diatas dapat dijelaskan bahwa pada tahun 2017 kabupaten/kota di Provinsi Sumatera Utara yang memiliki jumlah penduduk miskin terbanyak yaitu Kota Medan yaitu berjumlah 204.22 ribu jiwa sedangkan kabupaten/kota yang memiliki jumlah penduduk miskin terendah adalah Kabupaten Pakpak Bharat yang berjumlah 4.95 ribu jiwa, dengan rata-rata sebesar 44.06 ribu jiwa dan varians sebesar 1480.63 ribu jiwa.



Berdasarkan gambar diatas dapat dilihat dijelaskan bahwa pada tahun 2017 kabupaten/kota di Provinsi Sumatera Utara yang memiliki pendapatan per kapita terbanyak yaitu Kota Medan yaitu 62173471.6 ribu rupiah per tahun, sedangkan kabupaten/kota yang memiliki pendapatan per kapita terendah adalah Kabupaten Nias Selatan yaitu 12183709.28 ribu rupiah per tahun, dengan rata-rata sebesar 28211904.9 ribu rupiah per tahun dan varians sebesar 1.438E+14 ribu rupiah per tahun.



Pada gambar diatas terlihat bahwa plot variabel respon yaitu jumlah penduduk miskin dengan variabel prediktor yaitu pendapatan per kapita mengindikasikan kurva tidak mempunyai pola tertentu dan plot data menunjukkan adanya perubahan perilaku kurva pada selang waktu tertentu.

1. Pemodelan Regresi dengan Menggunakan Metode GCV sebagai Pemilihan K Optimal Pada Data Tingkat Kemiskinan Kab/kota di Provinsi Sumatera Utara
 - a. tingkat kemiskinan kabupaten/kota di Provinsi Sumatera Utara dengan Deret Fourier

Nilai GCV Data Tingkat Kemiskinan			
Nilai K	GCV	Nilai K	GCV
2	127171.17	34	639.32
4	33886.11	36	287.03
6	15570.63	38	2601.55
8	8509.96	40	639.02
10	5448.72	42	971.27
12	3496.23	44	137.96
14	2498.11	46	83.25
16	1863.1	48	83.79
18	1419.48	50	131.98
20	1112.01	52	139.68
22	297.79	54	18.93
23	267.45	56	17.76
26	1236.67	58	23.23
28	1065.34	60	23.19
30	1047.49	62	86.74
32	869.17	64	11.91

Berdasarkan table diatas didapatkan nilai GCV pada K=64 merupakan K yang paling optimal karena nilai GCVnya yang paling kecil diantara nilai GCV yang lain. Jika nilai K=64 yang dipakai maka banyaknya parameter yang harus ditaksir sebanyak 66 parameter, hal ini diperoleh berdasarkan persamaan 4.6 yaitu dengan mengetahui jumlah parameter yang ditaksir dari persamaan tersebut, maka akan diketahui model yang dihasilkan baik atau memenuhi, dilihat dari nilai R2 nya untuk K=1 sampai dengan K=64.

Nilai R ² dan MSE Setiap K Optimal Metode GCV		
Nilai K	R ²	MSE
23	0.90 (90%)	141.44
46	0.96 (96%)	52.14
64	0.99 (99%)	9.22

Berdasarkan table diatas terlihat bahwa untuk nilai K=23 sudah menghasilkan R2 yang cukup tinggi, sedangkan untuk nilai K=46 menghasilkan R2 sebesar 96%. Jika nilai K yang dipilih adalah K=64 maka estimasi parameter yang

harus dicari adalah sebanyak 66 parameter, hal ini dapat dilihat pada persamaan 4.6 dengan melihat jumlah parameter yang ditaksir. Metode yang dipilih adalah metode dengan R2 yang besar, nilai MSE yang kecil dan juga model yang parsimony (sederhana) maka model yang dipilih adalah K=46 dengan jumlah parameter yang ditaksir sebanyak 46 parameter.

Model yang didapatkan dari K=46 untuk data kemiskinan terendah adalah sebagai berikut :

$$\hat{y} = 36.99 + 0.01t - 3.85 \cos t - 8.25 \cos 2t - 6.79 \cos 3t - 0.06 \cos 4t - \dots - 20.63 \cos 46t$$

- b. Pemodelan Regresi dengan Menggunakan Metode CV sebagai Pemilihan K Optimal Pada Data Tingkat Kemiskinan Kab/kota Provinsi Sumatera Utara

Nilai CV Data Tingkat Kemiskinan			
Nilai K	CV	Nilai K	CV
1	11.77	34	699.22
4	51.98	36	22985.88
6	35.35	38	141922400
8	0.37	40	3329151
10	513.83	42	3850756
12	235.22	44	328970100
14	3040.98	46	638857
16	3501.45	48	344722
18	736416.2	50	5635817
20	64433780	52	29012940
22	279385900	54	1346.8
24	133.81	56	993.95
26	0.41	58	560719.3
28	43075.41	60	357590.9
30	73228.85	62	2423971
32	3.26	64	50489.96

Berdasarkan tabel diatas didapatkan nilai CV pada K=8 merupakan K yang paling optimal karena nilai CVnya yang paling kecil diantara nilai CV yang lain. Jika nilai

K=8 yang dipakai maka banyaknya parameter yang harus ditaksir sebanyak 10 parameter, hal ini diperoleh berdasarkan persamaan 4.6 yaitu dengan mengetahui jumlah parameter yang ditaksir dari persamaan tersebut, maka akan diketahui model yang dihasilkan baik atau memenuhi, dilihat dari nilai R² nya untuk K=1 sampai dengan K yang memiliki nilai CV terendah setelah K=8.

Nilai R² dan MSE Setiap K Optimal Metode CV

Nilai K	R ²	MSE
8	0.55 (55%)	632.97
26	0.67 (67%)	463.42
32	0.70 (70%)	418.09

Berdasarkan tabel 7 terlihat bahwa untuk nilai K=32 menghasilkan R² sebesar 70%. Sedangkan untuk K=26 menghasilkan R² sebesar 67%. Jika nilai K yang dipilih adalah K=8 maka R² yang dihasilkan cukup rendah yaitu sebesar 55%. Metode yang dipilih adalah metode dengan R² yang besar, nilai MSE yang kecil, dan juga model yang parsimony (sederhana) maka model yang dipilih adalah K=26 dengan R² sebesar 67% dan jumlah parameter yang ditaksir sebanyak 28 parameter.

Model yang didapatkan dari K=46 untuk data kemiskinan terendah adalah sebagai berikut :

$$\hat{y} = -35.46 + 2.08 \times 10^{-6}t + 5.37 \cos t + 2.26 \cos 2t + 23.32 \cos 3t - 4.81 \cos 4t - \dots + 4.83 \cos 26t$$

2. Analisis Hasil Pemodelan Tingkat Kemiskinan di Provinsi Sumatera Utara Menggunakan Pendekatan Regresi Nonparametrik Deret Fourier

Nilai R² dan MSE
Metode GCV dan CV

Metode	MSE	R ²
GCV	52.14	96%
CV	463.42	67%

Berdasarkan hasil pemodelan yang telah dilakukan di dapatkan nilai R² dan MSE seperti pada tabel diatas Penaksiran model

tingkat kemiskinan di provinsi Sumatera Utara menggunakan regresi nonparametrik deret fourier dengan metode GCV untuk menentukan K optimalnya menghasilkan R² sebesar 96% dan nilai MSE sebesar 52,14. Sedangkan untuk penaksiran model tingkat kemiskinan di provinsi Sumatera Utara menggunakan regresi nonparametrik deret fourier dengan metode CV untuk menentukan K optimalnya menghasilkan R² sebesar 67% dan nilai MSE sebesar 463.42. Jadi dapat disimpulkan bahwa pemodelan tingkat kemiskinan di provinsi Sumatera Utara pada tahun 2017 menggunakan regresi nonparametrik deret fourier dengan metode GCV untuk menentukan K optimalnya menghasilkan model yang lebih baik dengan R² sebesar 96% dan nilai MSE sebesar 52,14.

Model yang didapatkan dari penentuan K dengan metode GCV adalah :

$$\hat{y} = 4.19 + 0.01t - 13.33 \cos t - 17.82 \cos 2t + 40.43 \cos 3t + 9.04 \cos 4t - \dots + 9.19 \cos 8t$$

Dari model diatas diatas dapat dijelaskan bahwa hasil estimasi jumlah penduduk miskin (\hat{y}) tertinggi berjumlah 209.29 (ribu jiwa), hal tersebut tidak jauh berbeda dengan data aktual penduduk miskin tertinggi yaitu 204.22 (ribu jiwa) terdapat di Kota Medan, dengan error yang didapatkan sebesar -5.07. Hasil estimasi Jumlah penduduk miskin (\hat{y}) terendah berjumlah 11.05 (ribu jiwa), hal tersebut tidak jauh berbeda dengan data aktual penduduk miskin terendah yaitu 4.95 (ribu jiwa) terdapat di Kabupaten Pakpat Bharat dengan error yang didapatkan sebesar - 6.10.

4. KESIMPULAN

1. Model tingkat kemiskinan menggunakan pendekatan regresi nonparametrik deret fourier dengan metode GCV yang paling minimum adalah sebagai berikut :

$$\hat{y} = 36.99 + 0.01t - 3.85 \cos t - 8.25 \cos 2t - 6.79 \cos 3t - 0.06 \cos 4t - \dots - 20.63 \cos 46t$$
2. Model tingkat kemiskinan menggunakan pendekatan regresi nonparametrik deret fourier dengan metode CV yang paling minimum adalah sebagai berikut :

$$\hat{y} = -35.46 + 2.08 \times 10^{-6}t + 5.37 \cos t + 2.26 \cos 2t + 23.32 \cos 3t - 4.81 \cos 4t - \dots + 4.83 \cos 26t$$

Berdasarkan hasil analisis data menunjukkan bahwa model regresi nonparametrik deret fourier dengan pemilihan K optimal menggunakan metode GCV memberikan hasil yang lebih baik dibandingkan model regresi nonparametrik deret fourier dengan pemilihan K optimal menggunakan CV. Kesimpulan tersebut dengan memperhatikan kriteria kebaikan model, yaitu MSE yang relative lebih kecil dan R2 yang lebih besar, dimana metode GCV menghasilkan R2 sebesar 96% dan nilai MSE 52.14 sedangkan metode CV menghasilkan R2 sebesar 55% dan nilai MSE 632.97.

3. Hasil analisis pemodelan terbaik tingkat kemiskinan di Provinsi Sumatera Utara tahun 2017 sebagai berikut :

6. DAFTAR PUSTAKA

- Asrini, Luh Juni. 2012, Regresi Semiparametrik Deret Fourier, Prosiding Seminar Nasional FMIPA Universitas Negeri Surabaya, hal.77-80, 24 November 2012.
- Badan Pusat Statistik Sumatera Utara. 2002. Sumatera Utara dalam Angka. BPS: Medan.
- Badan Pusat Statistik Sumatera Utara. 2018. Sumatera Utara dalam Angka. BPS: Medan.
- Cahyat Ade, Gonner Christian, Haug Michaela. 2007. Mengkaji Kemiskinan dan Kesejahteraan Keluarga. Bogor: Cifor.
- Devi, Alvita Rachama, 2015, Metode Unbiased Risk (UBR) dan Cross-Validation (CV) untuk Pemilihan Titik Knot Optimal dalam Regresi Nonparametrik Multivariabel Spline Truncated, Tesis, Statistika, Institut Teknologi Sepuluh November, Surabaya.
- Eubank, R. 1999. Nonparametric Regression and Spline Smoothing 2nd Edition. Marcel Dekker. New York.
- Hardle, W. 1994. Applied Nonparametric Regression. New York: Cambridge University Press.
- Nurjanah, F., 2015, Model Regresi Nonparametrik Dengan Pendekatan Deret Fourier Pada Pola Data Curah Hujan Di Kota Semarang, Skripsi, Statistika Universitas Muhammadiyah Semarang, Semarang.
- Prahotama, A., 2013, Model Regresi Nonparametrik dengan Pendekatan deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur, In Prosiding Seminar Nasional Statistika Universitas Diponegoro 2013.
- Prahotama, A., Devi, A.R., Santoso, R., Suparti. 2018. Regresi Nonparametrik. Ponorogo : Wade Group.
- Pulungan, M. H., 2016, Analisis Faktor-faktor yang Mempengaruhi Tingkat Kemiskinan di Provinsi Sumatera Utara, Skripsi, Ilmu Ekonomi Institut Pertanian Bogor.
- Puspita, D. W. 2015. Analisis Determinan Kemiskinan di Provinsi Jawa Tengah. *JEJAK Journal of economics and policy*, 8 (1):100-107. doi: <http://dx.doi.org/10.15294/jejak.v8i1.3858>.
- Untoro, Joko. 2010. *Ekonomi*. Jakarta : Kawahmedia.

4. Hasil estimasi jumlah penduduk miskin () tertinggi berjumlah 209.29 (ribu jiwa), hasil estimasi penduduk miskin tersebut tidak jauh berbeda dengan data aktual penduduk miskin tertinggi yaitu 204.22 (ribu jiwa) terdapat di Kota Medan dengan error yang didapatkan sebesar -5.07. Hasil estimasi jumlah penduduk miskin () terendah berjumlah 11.05 (ribu jiwa), hasil estimasi penduduk miskin tersebut tidak jauh berbeda dengan data aktual penduduk miskin yaitu 4.95 (ribu jiwa) terdapat di Kabupaten Pakpat Bharat dengan error yang didapatkan sebesar - 6.10.

5. UCAPAN TERIMA KASIH

Ucapan terima kasih kepada Badan Pusat Statistik Provinsi Sumatera Utara atas kerjasamanya dalam menyediakan data penelitian.

Wu, H. dan Zhang, J.T. 2006. Nonparametric Regression Methods for Longitudinal Data Analysis, A John-Wiley and Son Inc. Publication, New Jersey.