ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 4, Nomor 1, Tahun 2015, Halaman 113 - 122
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

PEMODELAN TINGKAT PENGANGGURAN TERBUKA DI JAWA TENGAH MENGGUNAKAN REGRESI SPLINE

Seta Satria Utama[1], Suparti[2], Rita Rahmawati[3]
[1]Mahasiswa Jurusan Statistika FSM Universitas Diponegoro
[2,3]Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

ABSTRACK

Unemployment is one of the employment problems facing Indonesia. Central Java Province is one of the provinces with a high enough unemployment. The main indicators used to measure the unemployment rate in the labor force that is unemployed. Based on research Arianie (2012) labor force participation rate significantly affect the unemployment rate and based on research Sari (2012) the gross enrollment ratio significantly affects the rate of open unemployment. Therefore, in this study using the two predictor variables with the labor force participation rate as X_1 and gross enrollment rate as X_2. This study aimed to explore the model of open unemployment rate in the Province of Central Java. The method used is the method of spline regression. Spline regression has the ability to adapt more effectively to the data patterns up or down dramatically with the help of dots knots. Determination of the optimal point knots are very influential in determining the best spline models. The best spline models are models that have a minimum GCV (Generalized Cross Validation) Value. Best spline models for the analysis of the data rate of unemployment in Central Java Province is the spline regression model when order X_1 is 2 and order X_2 is 4 and large number of knots in the X_1 is 1 knot at the point 68.02394 and X_2 is 3 knots at the point 82.13, 87.19, and 87.65 with GCV value of 1.732746.

Keywords: Rate of Open Unemployment, Spline Regression, GCV

1. PENDAHULUAN

Salah satu tujuan dari pembangunan nasional adalah memajukan kesejahteraan umum dan mencerdaskan kehidupan bangsa. Sebagai negara yang berkembang, Indonesia berupaya untuk meningkatkan kesejahteraan demi mencapai tujuan nasional. Upaya pemerintah dalam meningkatkan kesejahteraan adalah meningkatkan stabilitas nasional, memacu pertumbuhan ekonomi, meningkatkan iklim investasi, dan menekan angka pengangguran. Salah satu masalah yang dihadapi Indonesia pada saat ini adalah masalah pengangguran. Jumlah pengangguran yang tinggi berdampak pada menurunnya tingkat kesejahteraan masyarakat.

Menurut BPS (2012), Indikator utama yang digunakan untuk mengukur angka pengangguran dalam angkatan kerja yaitu Tingkat Pengangguran Terbuka. Tingkat Pengangguran Terbuka (TPT) merupakan persentase jumlah pengangguran terhadap jumlah angkatan kerja. Upaya yang dapat dilakukan untuk mengurangi pengangguran salah satunya adalah dengan melakukan analisis terhadap faktor-faktor yang mempengaruhi TPT. Salah satu metode yang bisa digunakan adalah analisis regresi. Analisis regresi adalah metode statistik yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor.

Beberapa penelitian mengenai pengangguran pernah dilakukan oleh Sari (2012) dan Ariane (2012). Dari penelitian sebelumnya Sari (2012) menyatakan bahwa angka partisipasi kasar mempengaruhi TPT di Jawa Timur dan Ariane (2012) menyatakan bahwa tingkat partisipasi angkatan kerja mempengaruhi TPT di Jawa Timur dan Jawa Tengah. Dilihat dari angka korelasi, korelasi TPT terhadap tingkat partisipasi angkatan kerja dan angka partisipasi kasar SMP di Jawa Tengah lebih besar daripada korelasi TPT terhadap penduduk usia kerja, pertumbuhan ekonomi, dan angka partisipasi kasar SMA. Oleh karena itu, dalam penelitian ini penulis tertarik untuk menganalisis TPT di Provinsi Jawa Tengah dengan tingkat partisipasi angkatan kerja sebagai X_1 dan angka partisipasi kasar SMP sebagai X_2 menggunakan metode pendekatan regresi spline. Menurut Budiantara (2009), regresi spline merupakan analisis regresi yang mampu mengestimasi data yang tidak memiliki pola tertentu dan memiliki kecenderungan dalam mencari sendiri estimasi data dari pola yang terbentuk.

2. TINJAUAN PUSTAKA

2.1 Pengangguran Terbuka

Menurut BPS (2007), tenaga kerja adalah modal penting bagi bergeraknya roda pembangunan di suatu negara. Oleh karena itu, untuk mengetahui keberhasilan pembangunan di suatu negara dibutuhkan informasi mengenai statistik ketenagakerjaan, seperti Tingkat Partisipasi Angkatan Kerja (TPAK) dan Tingkat Pengangguran Terbuka (TPT).

1. Tingkat Pengangguran Terbuka

   Pengangguran terbuka merupakan penduduk usia kerja yang tidak bekerja, sedang mencari pekerjaan, mempersiapkan suatu usaha, mereka yang tidak mencari pekerjaan karena merasa tidak mungkin mendapatkan pekerjaan, sudah punya pekerjaan tetapi belum mulai bekerja. Tingkat Pengangguran Terbuka dapat dihitung dengan cara membandingkan jumlah pengangguran terhadap jumlah angkatan kerja.

$$TPT = \frac{jumlah\ pengangguran}{jumlah\ angkatan\ kerja}\ x\ 100\%$$

2. Tingkat Partisipasi Angkatan Kerja

   Tingkat partisipasi angkatan kerja adalah perbandingan jumlah angkatan kerja terhadap penduduk usia kerja.

$$TPAK = \frac{jumlah\ angkatan\ kerja}{jumlah\ penduduk\ usia\ 15\ tahun\ ke\ atas}\ x\ 100\%$$

2.2 Pendidikan

Keberhasilan pembangunan suatu wilayah ditentukan oleh sumber daya manusia yang berkualitas. Pendidikan merupakan salah satu cara meningkatkan kualitas SDM tersebut. Oleh karena itu peningkatan mutu pendidikan harus terus diupayakan. Menurut Oey-Gardiner (2003) dalam BPS (2013), Pembangunan pendidikan di Indonesia telah menunjukkan keberhasilan yang cukup besar. Wajib belajar 6 tahun, yang didukung pembangunan infrastruktur sekolah dan diteruskan dengan wajib belajar 9 tahun adalah program sektor pendidikan yang diakui cukup sukses.

Untuk mengetahui daya serap pendidikan dapat dilihat dari persentase penduduk menurut Angka Partisipasi Sekolah (APS). Menurut BPS (2014), terdapat dua ukuran partisipasi sekolah yang utama, yaitu Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM).

1. Angka Partisipasi Sekolah

   Angka Partisipasi Sekolah adalah persentase anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jenjang pendidikan tersebut.

$$APS_i = \frac{P_i\ masih\ sekolah}{P_i}\ x\ 100\%$$

Dimana:

APS_i : APS pada kelompok umur ke-i

P_i : Jumlah penduduk pada kelompok umur ke-i

i : 1, 2, 3

i = 1 untuk kelompok umur 7 – 12 tahun

i = 2 untuk kelompok umur 13-15 tahun
i = 3 untuk kelompok umur 16-18 tahun

2. Angka Partisipasi Kasar

Angka Partisipasi Kasar adalah persentase jumlah murid di jenjang pendidikan tertentu, berapapun umurnya terhadap jumlah penduduk kelompok umur yang berkaitan dengan jenjang pendidikan tersebut.

$$APK_{SMP} = \frac{P_{SMP}}{P_{13-15\ tahun}}\ x\ 100\%$$

3. Angka Partisipasi Murni

Angka Partisipasi Murni adalah persentase murid dengan umur yang berkaitan dengan jenjang pendidikannya terhadap jumlah penduduk di umur yang sama.

$$APM_{SMP} = \frac{P_{13-15\ tahun}SMP}{P_{13-15\ tahun}}\ x\ 100\%$$

2.3 Regresi Spline

Regresi nonparametrik merupakan metode pendugaan model yang tidak terikat asumsi dan digunakan jika bentuk kurva regresinya tidak diketahui. Kurva regresi dalam regresi nonparametrik diasumsikan mulus (*smooth*) dan termuat dalam suatu ruang fungsi tertentu. Regresi nonparametrik memiliki fleksibilitas yang tinggi karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1999). Model regresi nonparametrik secara umum dapat dituliskan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i\ , i = 1, 2, \dots, n$$

dimana y_i : variabel respon pengamatan ke-i
x_i : variabel prediktor pengamatan ke-i
$f(x_i)$: fungsi regresi yang tidak diketahui
ε_i : error pengamatan ke-i

Menurut Wu dan Zang (2006), secara umum fungsi spline polinomial *truncated* berorde m dengan titik knot $(k_1, k_2, \dots, k_r)$ didefinisikan sebagai fungsi yang dirumuskan dalam bentuk persamaan berikut:

$$f(x) = \sum_{k=0}^{m-1} \beta_k x^k + \sum_{j=1}^{r} \beta_{j+m-1}\,(x - k_j)_+^{m-1} \quad (1)$$

dengan fungsi *truncated*

$$(x - k_j)_+^{m-1} = \begin{cases} (x - k_j)^{m-1} & ; x - k_j \geq 0 \\ 0 & ; x - k_j < 0 \end{cases}$$

Persamaan (1) merupakan spline dengan derajat m dan r titik knot $k_1, k_2, \dots, k_r (a < k_1 < k_2 < \dots < k_r < b)$ dimana a merupakan nilai minimum dari x dan b nilai maksimum dari x. Jadi secara umum model spline *truncated* orde ke-m adalah sebagai berikut

$$y_i = \sum_{k=0}^{m-1} \beta_k x_i^k + \sum_{j=1}^{r} \beta_{j+m-1}\,(x_i - k_j)_+^{m-1} + \varepsilon_i \quad (2)$$

dimana m adalah derajat polinomial, k adalah titik knot pada fungsi truncated, dan ε_i adalah error random. Bentuk persamaan (2) dapat ditulis ke dalam bentuk matriks sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X_1\delta_1} + \boldsymbol{X_2\delta_2} + \boldsymbol{\varepsilon} \quad (3)$$

dengan

$$\mathbf{Y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix};\ \mathbf{X}_1=\begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} \\ 1 & x_2 & \cdots & x_2^{m-1} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{m-1} \end{bmatrix};\ \boldsymbol{\delta}_1=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m-1} \end{bmatrix};\ \boldsymbol{\delta}_2=\begin{bmatrix} \beta_{(m-1)+1} \\ \beta_{(m-1)+2} \\ \vdots \\ \beta_{(m-1)+r} \end{bmatrix}$$

$$\mathbf{X}_2=\begin{bmatrix} (x_1-k_1)_+^{m-1} & (x_1-k_2)_+^{m-1} & \cdots & (x_1-k_r)_+^{m-1} \\ (x_2-k_1)_+^{m-1} & (x_2-k_2)_+^{m-1} & \cdots & (x_2-k_r)_+^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ (x_n-k_1)_+^{m-1} & (x_n-k_2)_+^{m-1} & \cdots & (x_n-k_r)_+^{m-1} \end{bmatrix}$$

Bentuk matriks dari persamaan regresi spline (3) dapat disederhanakan menjadi

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon} \qquad (4)$$

dengan $\boldsymbol{X}=[\boldsymbol{X_1} \quad \boldsymbol{X_2}]$ dan $\boldsymbol{\beta}=\begin{bmatrix}\boldsymbol{\delta_1}\\ \boldsymbol{\delta_2}\end{bmatrix}$

Persamaan matriks (4) merupakan persamaan matriks regresi linier yang telah dimodifikasi dengan menambahkan beberapa kolom pada matriks $\mathbf{X}$ dan beberapa baris pada vektor $\boldsymbol{\beta}$. Pendugaan terhadap vektor parameter $\boldsymbol{\beta}$ dilakukan dengan menggunakan metode *least square* yaitu $\hat{\boldsymbol{\beta}}=(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{Y}$

Dengan sedikit modifikasi menyesuaikan persamaan regresi spline dengan titik-titik knot yang diberikan $\prod=\{k_1,k_2,\ldots,k_r\}$ maka estimasi vektor parameter $\boldsymbol{\beta}$ menjadi $\hat{\boldsymbol{\beta}}_\Pi=(\mathbf{X}_\Pi^{\mathbf{T}}\mathbf{X}_\Pi)^{-1}\mathbf{X}_\Pi^{\mathbf{T}}\mathbf{Y}$

Fungsi penduga dari $f(x)$ adalah sebagai berikut:

$$\begin{aligned}\hat{\boldsymbol{f}}_\Pi(x)&=\mathbf{X}_\Pi\hat{\boldsymbol{\beta}}_\Pi\\ &=\mathbf{X}_\Pi(\mathbf{X}_\Pi^{\mathbf{T}}\mathbf{X}_\Pi)^{-1}\mathbf{X}_\Pi^{\mathbf{T}}\mathbf{Y}\\ &=\mathbf{H}_\Pi\mathbf{Y}\end{aligned}$$

Dengan $\mathbf{H}_\Pi=\mathbf{X}_\Pi(\mathbf{X}_\Pi^{\mathbf{T}}\mathbf{X}_\Pi)^{-1}\mathbf{X}_\Pi^{\mathbf{T}}$ sedangkan $\mathbf{X}_\Pi$ adalah matriks desain berukuran n x (m+r) dari model yang membentuk penduga $\boldsymbol{f}_\Pi(x)$ dan bergantung pada titik knot.

$$\mathbf{X}_\Pi=\begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} & (x_1-k_1)_+^{m-1} & \cdots & (x_1-k_r)_+^{m-1} \\ 1 & x_2 & \cdots & x_2^{m-1} & (x_2-k_1)_+^{m-1} & \cdots & (x_2-k_r)_+^{m-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{m-1} & (x_n-k_1)_+^{m-1} & \cdots & (x_n-k_r)_+^{m-1} \end{bmatrix}$$

2.4 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan yang menunjukkan perubahan perilaku fungsi spline pada selang yang berbeda. Salah satu metode pemilihan titik knot optimal adalah *Generalized Cross Validation* (GCV). Menurut Eubank (1999), metode GCV dapat dituliskan sebagai berikut:

$$GCV(\Pi)=\frac{MSE\ (\Pi)}{(n^{-1}\ trace[I-H(\Pi)])^2}$$

dengan I merupakan matriks identitas n x n dan $MSE(\Pi)=n^{-1}\sum_{i=1}^{n}(y_i-\hat{f}_{\Pi\, i})^2$.

2.5 *Mean Absolute Precentage Error* (MAPE)

Menurut Juanda dan Junaidi (2012) kriteria *Mean Absolute Percentage Error* (MAPE) dapat digunakan untuk evaluasi kesalahan dalam peramalan. MAPE mengukur kesalahan nilai dugaan model yang dinyatakan dalam bentuk rata-rata persentase absolut residual. Menurut Makridakis (1999) MAPE dapat ditulis sebagai berikut:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \text{ x } 100\%$$

dimana y_i = Data aktual pada pengamatan ke-i
$\hat{y}_i$ = Data hasil prediksi pada pengamatan ke-i
n = banyak pengamatan yang diprediksi

Suatu model mempunyai kinerja sangat bagus jika nilai MAPE berada di bawah 10%, dan mempunyai kinerja bagus jika nilai MAPE berada di antara 10% dan 20% (Sobri Harun dalam Zainun dan Eftekhari (2010)).

2.6 Koefisien Determinasi (R^2)

Koefisien determinasi merupakan besaran yang digunakan untuk mengukur seberapa jauh kecocokan suatu model regresi. Nilai dari koefisien determinasi berada pada kisaran $0 \le R^2 \le 1$. Semakin besar nilai koefisien determinasi berarti model semakin cocok (Gujarati, 2003).

$$R^2 = \frac{JKR}{JKT} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

dengan JKR : Jumlah Kuadrat Regresi
JKT : Jumlah Kuadrat Total

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari buku Jawa Tengah dalam Angka 2013, Keadaan Angkatan Kerja di Jawa Tengah Agustus 2012 hasil Sakernas 2012, dan Statistik Pendidikan Jawa Tengah 2012 hasil Susenas 2012. Unit pengamatan yang digunakan adalah 35 kabupaten/kota di Provinsi Jawa Tengah.

Variabel respon dan variabel prediktor yang digunakan dalam penelitian ini adalah sebagai berikut:

Y : Tingkat Pengangguran Terbuka
X_1 : Tingkat Partisipasi Angkatan Kerja
X_2 : Angka Partisipasi Kasar SMP

Software yang digunakan dalam penelitian ini adalah Minitab 14, R 2.14.2, dan Microsoft Excel 2010.

Langkah-langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Memasukkan data variabel respon dan variabel prediktor.
2. Membuat *scatterplot* antara variabel respon dengan masing-masing variabel prediktor.
3. Menentukan orde dan banyaknya titik knot.
4. Mengkombinasikan beberapa orde dan titik knot.
5. Menentukan model spline untuk setiap kombinasi orde dan titik knot.
6. Menghitung GCV
7. Menentukan titik knot optimal dengan GCV (*Generalized Cross Validation*) minimum.
8. Menentukan model spline terbaik.

4. HASIL DAN PEMBAHASAN

4.1 Statistika Deskriptif

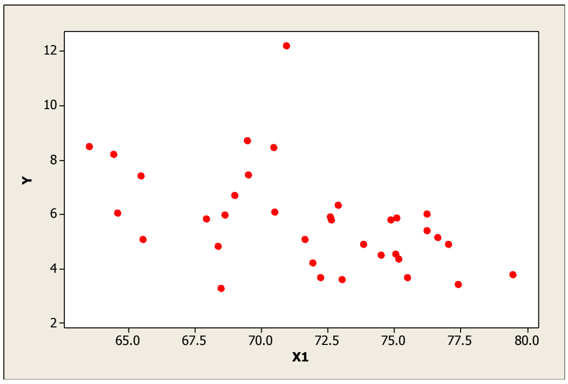
Statistika Deskriptif dari setiap variabel penelitian dapat dilihat pada Tabel 1.

Tabel 1. Statistika Deskriptif

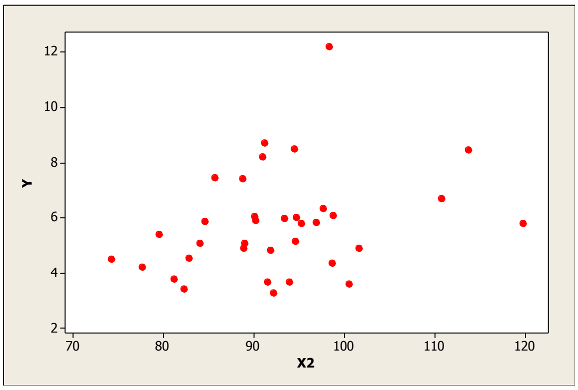
Variabel	Minimal	Maksimal	Rata-Rata	Varian
Y	3,28	12,20	5,75	3,39
X_1	63,51	79,47	71,74	16,52
X_2	74,31	119,85	92,60	89,89

4.2 *Scatterplot*

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menyelidiki pola hubungan antara varaiabel prediktor dengan variabel respon. Bentuk pola hubungan fungsional antara variabel prediktor dengan variabel respon dapat diperkirakan dengan membuat *scatterplot* yang memuat informasi tentang hubungan kedua variabel tersebut.



Gambar 1



Gambar 2

Ket: *Scatterplot* antara variabel X_1 terhadap variabel respon (Gambar 1) dan *Scatterplot* antara variabel X_2 terhadap variabel respon (Gambar 2).

Berdasarkan Gambar 1 dan Gambar 2 plot antara Tingkat Pengangguran Terbuka (TPT) dengan Tingkat Partisipasi Angkatan Kerja (TPAK) dan Angka Partisipasi Kasar (APK) SMP tidak diketahui bentuk kurva regresinya. Kurva regresi yang diidentifikasi melalui plot-plot yang tersebar menunjukkan pola yang tidak beraturan atau dapat dikatakan tidak mengikuti suatu pola tertentu.

4.3 Pemilihan Titik Knot Optimal

Estimasi kurva regresi menggunakan regresi spline dilakukan dengan menentukan banyak dan letak titik knot dalam beberapa orde. Banyak titik knot yang akan digunakan pada setiap variabel prediktor beragam yaitu satu knot, dua knot, tiga knot, dan juga kombinasi banyaknya knot pada setiap variabel prediktor sehingga memungkinkan setiap variabel prediktor mempunyai banyak knot yang tidak sama. Letak titik knot merupakan hal yang sangat penting untuk mendapatkan model terbaik. Pendekatan dimulai dari satu titik knot dengan orde 2 sampai dengan tiga titik knot dengan orde 4. Salah satu metode yang dapat digunakan untuk pemilihan titik knot optimal adalah metode *Generalized Cross Validation* (GCV). Penambahan orde dan banyaknya titik knot dilakukan untuk menghasilkan estimasi yang lebih optimal. Namun, jika model spline dengan orde yang lebih besar ataupun dengan titik knot lebih banyak menghasilkan nilai GCV yang lebih besar, maka penambahan orde dan banyaknya titik knot memberikan estimasi yang tidak efisien.

4.4 Model Spline Terbaik

Nilai GCV minimum untuk masing-masing kombinasi banyak knot pada setiap variabel terangkum dalam Tabel 2 sebagai berikut:

Tabel 2. Perbandingan Nilai GCV Minimum

Banyaknya Knot		Orde		Titik Knot		GCV
X_1	X_2	X_1	X_2	X_1	X_2	
1	1	2	4	68,34636	78,45	2,442249
1	2	2	2	64,63848	77,07;77,53	3,274236
1	**3**	**2**	**4**	**68,02394**	**82,13;87,19;87,65**	**1,732746**
2	1	2	2	68,50758;68,83	79,83	2,769046
2	2	2	4	68,66879;68,83	79,37;89,49	2,498986
2	3	2	2	68,50758;68,83	75,23;75,69;76,15	2,795415
3	1	2	2	68,34636;68,50758;68,66879	74,77	2,90187
3	2	2	2	68,34636;68,50758;68,66879	74,77;76,61	2,927942
3	3	2	2	68,34636;68,50758;68,66879	74,77;75,69;77,53	2,927942

Pemilihan estimasi model spline terbaik diperoleh dengan membandingkan nilai GCV minimum pada masing-masing kombinasi orde dan banyak titik knot dalam setiap variabel prediktor. Titik knot optimal adalah titik knot yang mempunyai nilai GCV minimum. Berdasarkan Tabel 2 nilai GCV minimum diperoleh pada saat X_1 berorde 2 dan X_2 berorde 4 dan banyaknya knot X_1 sebanyak 1 knot yaitu pada titik 68,02394 dan banyaknya knot X_2 sebanyak 3 knot yaitu pada titik 82,13; 87,19 dan 87,65.

Model spline X_1 berorde 2 dengan 1 knot dan X_2 berorde 4 dengan 3 knot adalah $\hat{f}(x) = \hat{\alpha}_{00} + \hat{\alpha}_{11}x_1 + \hat{\beta}_{11}(x_1 - k_{11})_+ + \hat{\alpha}_{12}x_2 + \hat{\alpha}_{22}x_2^2 + \hat{\alpha}_{32}x_2^3 + \hat{\beta}_{21}(x_2 - k_{21})_+^3$ $+\hat{\beta}_{22}(x_2 - k_{22})_+^3 + \hat{\beta}_{23}(x_2 - k_{23})_+^3$. Berdasarkan *running* program menggunakan *software* R 2.14.2, diperoleh estimasi parameter model spline terbaik. Estimasi parameter model spline terbaik dapat dilihat pada Tabel 3.

Tabel 3. Estimasi Parameter Model Spline Terbaik

Variabel	Parameter	Estimasi Parameter
Intersep	α_{00}	14927,99
X_1	α_{11}	0,31
	β_{11}	0,09
X_2	α_{12}	562,11
	α_{22}	7,06
	α_{32}	0,03
	β_{21}	0,05
	β_{22}	0,15
	β_{23}	0,12

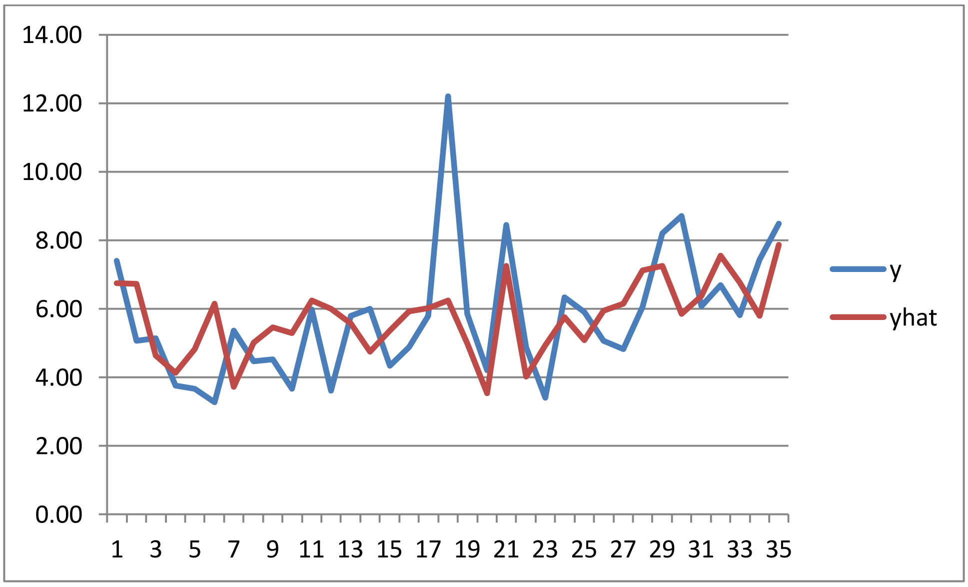
Berdasarkan Tabel 3 estimasi model spline terbaik dapat dituliskan sebagai berikut:

$$\hat{f}(x) = 14927{,}99 - 0{,}31x_1 + 0{,}09\,(x_1 - 68{,}02394)_+ - 562{,}11x_2 + 7{,}06x_2^2 - 0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)_+^3 - 0{,}15\,(x_2 - 87{,}19)_+^3 + 0{,}12\,(x_2 - 87{,}65)_+^3$$

Persamaan model spline terbaik beserta *truncated* yang membentuknya adalah sebagai berikut:

$$\hat{f}(x) = \begin{cases} 14927{,}99 - 0{,}31x_1 - 562{,}11x_2 + 7{,}06x_2^2 - 0{,}03\,x_2^3 \\ x_1 < 68{,}02394\,;\ x_2 < 82{,}13 \\ 14927{,}99 - 0{,}31x_1 - 562{,}11x_2 + 7{,}06x_2^2 - 0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3 \\ x_1 < 68{,}02394\,;\ 82{,}13 \le x_2 < 87{,}19 \\ 14927{,}99 - 0{,}31x_1 - 562{,}11x_2 + 7{,}06x_2^2 - 0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3 \\ -0{,}15\,(x_2 - 87{,}19)^3\ ;x_1 < 68{,}02394\,;\ 87{,}19 \le x_2 < 87{,}65 \\ 14927{,}99 - 0{,}31x_1 - 562{,}11x_2 + 7{,}06x_2^2 - 0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3 \\ -0{,}15\,(x_2 - 87{,}19)^3\ + 0{,}12\,(x_2 - 87{,}65)^3\ ;x_1 < 68{,}02394\,;x_2 \ge 87{,}65 \\ 14927{,}99 - 0{,}31x_1 + 0{,}09\,(x_1 - 68{,}02394)\ - 562{,}11x_2 + 7{,}06x_2^2 \\ -0{,}03\,x_2^3\,;x_1 \ge 68{,}02394\,;\ x_2 < 82{,}13 \\ 14927{,}99 - 0{,}31x_1 + 0{,}09\,(x_1 - 68{,}02394)\ - 562{,}11x_2 + 7{,}06x_2^2 \\ -0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3\ ;x_1 \ge 68{,}02394\,;\ 82{,}13 \le x_2 < 87{,}19 \\ 14927{,}99 - 0{,}31x_1 + 0{,}09\,(x_1 - 68{,}02394)\ - 562{,}11x_2 + 7{,}06x_2^2 \\ -0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3\ - 0{,}15\,(x_2 - 87{,}19)^3 \\ x_1 \ge 68{,}02394\,;\ 87{,}19 \le x_2 < 87{,}65 \\ 14927{,}99 - 0{,}31x_1 + 0{,}09\,(x_1 - 68{,}02394)\ - 562{,}11x_2 + 7{,}06x_2^2 \\ -0{,}03\,x_2^3 + 0{,}05\,(x_2 - 82{,}13)^3\ - 0{,}15\,(x_2 - 87{,}19)^3\ + 0{,}12\,(x_2 - 87{,}65)^3 \\ x_1 \ge 68{,}02394\,;\ x_2 \ge 87{,}65 \end{cases}$$

Setelah mendapatkan model spline terbaik, maka dapat diperoleh nilai estimasi yang disajikan pada Gambar 3.



Gambar 3. Grafik Estimasi Model Spline Terbaik

Pada Gambar 3 dapat dilihat bahwa estimasi yang dihasilkan mendekati sebaran data aktualnya. Hal tersebut menunjukkan bahwa regresi spline memiliki kemampuan menyesuaikan diri lebih efektif terhadap pola data yang naik atau turun secara tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus.

4.5 *Mean Absolute Precentage Error* (MAPE)

MAPE merupakan kuantitas error yang membandingkan antara data aktual tingkat pengangguran terbuka dengan data peramalan dari hasil estimasi model spline terbaik. Dengan y_i = TPT Aktual dan $\hat{y}_i$ = TPT Prediksi, diperoleh nilai MAPE sebagai berikut:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \text{ x } 100\% = \frac{1}{35}\left(\left|\frac{6{,}76-6{,}33}{6{,}76}\right| + \ldots + \left|\frac{9{,}25-5{,}66}{9{,}25}\right|\right) \text{ x } 100\% = 19{,}03\%$$

Dilihat dari nilai MAPE yang kurang dari 20% maka dapat disimpulkan bahwa estimasi model spline terbaik mempunyai kemampuan peramalan yang baik.

4.6 Koefisien Determinasi (R^2)

Koefisien determinasi merupakan besaran yang digunakan untuk mengukur seberapa jauh kecocokan suatu model regresi. Dengan nilai y_i dan $\hat{y}_i$ berdasarkan Lampiran 2 dan $\bar{y}$ = 5,75 maka dapat dihitung nilai koefisien determinasi sebagai berikut:

$$R^2 = \frac{JKR}{JKT} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\{(6{,}74 - 5{,}75)^2 + \cdots + (7{,}86 - 5{,}75)^2\}}{\{(7{,}40 - 5{,}75)^2 + \cdots + (8{,}49 - 5{,}75)^2\}} = \frac{38{,}93}{118{,}62} = 0{,}328$$

Nilai R^2 tersebut menunjukkan bahwa pengaruh tingkat partisipasi angkatan kerja dan angka partisipasi kasar SMP terhadap tingkat pengangguran terbuka sebesar 32,80%.

5. PENUTUP

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan, maka dapat diambil beberapa kesimpulan antara lain:

1. Model spline terbaik untuk analisis data tingkat pengangguran terbuka adalah model regresi spline X_1 berorde 2 dan X_2 berorde 4 dan banyaknya titik knot pada X_1 adalah 1 knot yaitu pada titik 68,02394 dan X_2 adalah 3 knot yaitu pada titik 82,13; 87,19 dan 87,65 dengan nilai GCV sebesar 1,732746.
2. Persamaan model spline terbaik adalah sebagai berikut:

$$\begin{aligned}\hat{f}(x) = {} & 14927{,}99 - 0{,}31x_1 + 0{,}09\ (x_1 - 68{,}02394)_+ - 562{,}11x_2 + 7{,}06x_2^2 \\ & - 0{,}03\ x_2^3 + 0{,}05\ (x_2 - 82{,}13)_+^3 - 0{,}15\ (x_2 - 87{,}19)_+^3 \\ & + 0{,}12\ (x_2 - 87{,}65)_+^3\end{aligned}$$

5.2 Saran

Pada penelitian ini, pemodelan tingkat pengangguran terbuka terbatas menggunakan 2 variabel prediktor. Penelitian selanjutnya disarankan untuk menambah variabel prediktor yang berhubungan dengan tingkat pengangguran terbuka. Selain itu, dalam penelitian ini hanya mencobakan orde 2, 3, dan 4 dengan banyak titik knot maksimal 3 titik dalam setiap variabel prediktornya. Oleh karena itu, perlu dicobakan pula orde lebih dari 4 dan banyaknya titik knot dalam setiap variabel prediktor lebih dari 3 titik menggunakan *package software* terbaru.

6. DAFTAR PUSTAKA

Ariane, S. 2012. *Pendekatan Regresi Ridge untuk Memodelkan Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Jawa Timur dan Jawa Tengah*. Surabaya: Tugas Akhir, Institut Teknologi Sepuluh November.

BPS. 2007. *Analisis Perkembangan Statistik Ketenagakerjaan (Laporan Sosial Indonesia 2007)*. BPS, Jakarta.

BPS. 2012. *Berita Resmi Statistik Keadaan Ketenagakerjaan Jawa Tengah*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2013. *Berita Resmi Statistik Keadaan Ketenagakerjaan*. BPS, Jakarta.

BPS. 2013. *Jawa Tengah Dalam Angka 2013*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2013. *Keadaan Angkatan Kerja di Jawa Tengah 2012*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2013. *Statistik Pendidikan Jawa Tengah 2012*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2014. *Jawa Tengah Dalam Angka 2014*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2014. *Keadaan Angkatan Kerja di Jawa Tengah 2013*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2014. *Statistik Sosial dan Kependudukan Jawa Tengah 2013*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2014. http://sirusa.bps.go.id/index.php?r=indikator/view&id=44 (diakses tanggal 18 Juni 2014).

BPS. 2014. http://www.datastatistik-indonesia.com/portal/index.php?option=com_content&ask=view&id=800&Itemid=800 (diakses tanggal 18 Juni 2014).

Budiantara, I. N. 2009. *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.

Eubank, R. L. 1999. *Spline Smoothing and Nonparametric Regression Second Edition*. Texas: Department of Statistics Southern Methodist Dallas University.

Gujarati, D. N. 2003. *Basic Econometrics Fourth Edition*. McGraw-Hill, Inc. New York.

Hardle, W. 1994. *Applied Nonparametric Regression*. Berlin: Humboldt University.

Juanda, Bambang dan Junaidi. 2012. *Ekonometrika Deret Waktu Teori dan Aplikasi*. Bogor: IPB Press

Makridakis, Wheelwright, McGee. 1999. *Metode dan Aplikasi Peramalan*. Jakarta: Binarupa Aksara.

Sari, R. S. 2012. *Pemodelan Pengangguran Terbuka di Jawa Timur dengan Menggunakan Pendekatan Regresi Spline Multivariabel*. Surabaya: Tugas Akhir, Institut Teknologi Sepuluh November.

Wu, H. dan Zang, J. T. 2006. *Nonparametric Regression Methods for Longitudinal Data Analysis*. New Jersey: John Wiley and Sons.

Zainun, N. Yasmin dan Eftekhari, M. 2010. *Forecasting Low-Cost Housing Demand in Urban Area in Malaysia using ANN*. UK: Loughborough University.