

# Pemodelan Pengangguran Terbuka di Jawa Timur dengan Menggunakan Pendekatan Regresi Spline Multivariabel

Ruli Sartika Sari dan I Nyoman Budiantara

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111

*E-mail:* i\_nyoman\_b@statistika.its.ac.id

**Abstrak**— Pengangguran merupakan salah satu masalah yang cukup serius yang dihadapi pemerintah Indonesia. Jawa Timur yang merupakan salah satu provinsi terpadat dan penyumbang jumlah penganggur terbesar ketiga di Indonesia. Pemodelan pengangguran terbuka di Jawa Timur dengan menggunakan regresi spline mampu mengestimasi data yang tidak memiliki pola tertentu. Model spline terbaik adalah model dengan nilai GCV minimum. Penelitian ini bertujuan untuk mengetahui faktor yang berpengaruh terhadap pengangguran terbuka di Jawa Timur dengan menggunakan regresi spline. Hasil pemodelan menunjukkan bahwa dengan regresi spline linier kombinasi tiga knot menghasilkan GCV minimum dengan persamaan

$$\hat{y} = 0,26t_1 - 17,91(t_2 - 117,38)_+^1 + 18(t_2 - 123,48)_+^1 + 0,16t_6 + \\ - 0,31(t_6 - 10,74)_+^1 + 0,20(t_6 - 32,52)_+^1$$

Variabel yang berpengaruh signifikan terhadap pengangguran terbuka di Jawa Timur adalah persentase penduduk usia kerja berumur 15 tahun ke atas berdasarkan pendidikan tertinggi yang ditamatkan adalah SMA/SMK, Angka Partisipasi Kasar, dan tingkat investasi menurut kabupaten/kota dengan nilai  $R^2$  sebesar 99,05 persen dan nilai MSE sebesar 0,3264.

**Kata Kunci**— GCV, Pengangguran Terbuka,  $R^2$ , Regresi Spline

## I. PENDAHULUAN

**T**ERCAPAINYA kesejahteraan masyarakat merupakan wujud dari pemerintah Indonesia dalam meningkatkan pembangunan nasional. Salah satu upaya pemerintah dalam meningkatkan kesejahteraan adalah meningkatkan stabilitas nasional, memacu pertumbuhan ekonomi, meningkatkan iklim investasi, dan menekan angka pengangguran. Pengangguran merupakan beban pekerjaan penting yang harus segera ditangani oleh pemerintah, khususnya Dinas Tenaga Kerja. Tingginya angka pengangguran di Indonesia disebabkan karena tidak ada kesesuaian antara penawaran tenaga kerja dengan kebutuhan di pasar tenaga kerja. Serta jumlah penduduk yang semakin meningkat yang tidak diimbangi dengan pertumbuhan lapangan usaha yang ada.

Jawa Timur adalah salah satu provinsi dengan kepadatan dan pertumbuhan penduduk yang cukup tinggi. Pada tahun 2011 Jawa Timur merupakan penyumbang pengangguran terbesar ketiga setelah Jawa Barat dan Jawa Tengah. Hal ini ditunjukkan dengan jumlah angkatan kerja sebesar 19,761 juta jiwa dan jumlah kesempatan kerja yang terserap sebesar 18,94 juta jiwa.

Jadi jumlah pengangguran pada tahun tersebut sebesar 821.546 jiwa. Masalah ketenagakerjaan dipengaruhi oleh tingkat ekonomi dari masing – masing daerah. Dimana masing-masing daerah memiliki kemampuan yang berbeda untuk mengembangkan potensi sektor penunjang perekonomian. Beberapa sektor penunjang yang menyerap tenaga kerja cukup besar adalah sektor per-dagangan, industri, dan per-tanian. Dalam penelitian ini memiliki kriteria dalam pembatas-an jumlah daerah yang akan dianalisis. Kriteria yang diguna-kan berdasarkan data PDRB menurut harga konstan dimana daerah yang memiliki peran di sektor perdagangan dan industri di atas 20 persen dan daerah yang memiliki peran di sektor pertanian kurang dari 20 persen merupakan daerah yang akan dianalisis lebih lanjut. Pemilihan ini ditujukan untuk menghindari kesalahan persepsi adanya pengangguran musiman.

Tujuan dari penelitian ini adalah untuk mengetahui variabel yang berpengaruh signifikan/nyata terhadap pengangguran terbuka di Jawa Timur dengan menggunakan pendekatan regresi spline. Penelitian tentang pengangguran telah dilaku-kan dengan mengelompokkan faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Jawa Timur [1][2]. Rujukan [3] memodelkan faktor yang ber-pengaruh terhadap pengangguran terbuka di Jawa Timur dengan menggunakan regresi linier multivariate. Pada pe-nelitian ini melihat tidak adanya pola tertentu antara variabel respon dengan variabel prediktor – prediktornya sehingga didekati dengan menggunakan regresi spline. 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 [4]. Dari uraian di atas, maka dilakukan penelitian tentang pemodelan pengangguran terbuka dengan menggunakan regresi spline.

## II. METODE PENELITIAN

### A. Regresi Spline Multivariabel

Pada penelitian ini menggunakan metode regresi spline karena pola hubungan antara variabel respon dengan variabel prediktor-prediktornya tidak membentuk suatu pola tertentu. Oleh karena itu, metode yang sesuai untuk kasus ini adalah regresi spline. Spline adalah salah satu potongan polinomial yang memiliki sifat fleksibilitas yang lebih baik dari poli-nomial biasa. Spline juga memiliki kemampuan yang sangat

baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu. Secara umum, persamaan fungsi spline berorde  $p$  dengan titik knot  $K_1, K_2, \dots, K_r$  dapat ditulis sebagai berikut [4] :

$$f(t) = \sum_{k=0}^p \alpha_k t^k + \sum_{j=1}^r \beta_j (t - K_j)_+^p \quad (1)$$

dengan  $(t - K)_+^p = \begin{cases} (t - K)^p; & t \geq K \\ 0 & ; t < K \end{cases}$

dimana  $\alpha_k$  adalah parameter polinomial,  $\beta_j$  adalah parameter dari potongan polinomial, dan  $K_j$  adalah titik knot. Titik knot adalah titik perpaduan bersama yang menunjukkan perubahan pola perilaku data ([5] dan [4]).

Dalam analisis regresi spline, jika diberikan satu variabel respon dan variabel prediktornya lebih dari satu maka disebut regresi spline multivariabel. Model regresi spline multivariabel dapat ditulis sebagai berikut :

$$y_i = f(t_{i1}) + f(t_{i2}) + \dots + f(t_{ip}) + \varepsilon_i \\ = \sum_{l=1}^p f(t_{il}) + \varepsilon_i \quad (2)$$

dengan  $f(t_{il}) = \sum_{k=0}^p \alpha_{kl} t_{il}^k + \sum_{j=1}^r \beta_{lj} (t_{il} - K_{lj})_+^p$  dan

$$(t_{il} - K_{lj})_+^p = \begin{cases} (t_{il} - K_{lj})^p; & t_{il} \geq K_{lj} \\ 0 & ; t_{il} < K_{lj} \end{cases}$$

Dalam regresi nonparametrik sangat penting dalam menentukan titik knot optimal. Jika diperoleh titik knot optimal maka memberikan fungsi spline yang terbaik. Salah satu metode untuk pemilihan titik knot optimal adalah *Generalized Cross Validation* (GCV) ([4]). Fungsi GCV dapat ditulis menjadi :

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (3)$$

dengan  $MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ,  $(K_1, K_2, \dots, K_r)$

adalah titik knot, dan matrik  $A(K_1, K_2, \dots, K_r)$  yang diperoleh dari persamaan  $\hat{y} = A(K_1, K_2, \dots, K_r)y$ .

Setelah diperoleh model regresi spline terbaik, selanjutnya dilakukan pengujian kesesuaian parameter model regresi spline. Ada dua macam pengujian parameter model yaitu uji secara serentak (uji-F) dan uji secara parsial (uji-t).

### 1. Uji serentak

Uji serentak digunakan untuk mengetahui apakah parameter model regresi spline sudah signifikan atau belum. Pengujian ini dilakukan secara serentak dengan parameter yang ada dalam model. Hipotesis untuk uji serentak sebagai berikut :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Minimal ada satu } \beta_k \neq 0, k = 1, 2, \dots, p.$$

Statistik uji yang digunakan :

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \quad (4)$$

Keputusan : tolak  $H_0$  jika  $F_{hitung}$  lebih besar dari  $F_{tabel}$  ( $F_{\alpha; (k-1, n-k)}$ ).

### 2. Uji parsial

Uji parsial digunakan untuk mengetahui parameter yang signifikan secara individu terhadap model. Hipotesis untuk uji parsial adalah sebagai berikut :

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, p$$

Statistik uji yang digunakan :

$$t_{hitung} = \frac{\hat{\beta}_k}{\text{st. error}(\hat{\beta}_k)} \quad (5)$$

Keputusan : tolak  $H_0$  jika  $|t_{hitung}|$  lebih besar dari  $t_{tabel}$  ( $t_{(\alpha/2; n-k)}$ ).

Setelah dilakukan uji kesesuaian parameter model regresi spline, selanjutnya memeriksa asumsi residual dari model regresi spline. Pemeriksaan asumsi residual dari model regresi spline bersifat sama seperti halnya pemeriksaan residual regresi parametrik yaitu asumsi IIDN.

### 1. Asumsi residual independen

Uji independen digunakan untuk mengetahui ada tidaknya korelasi antar residual. Cara mendeteksi residual bersifat independen atau tidak ada dua cara yaitu cara visual untuk mendeteksi korelasi antar residual dengan menggunakan plot *Autocorrelation Function* (ACF). Jika hasil plot menunjukkan tidak ada lag yang keluar dari batas maka residual bersifat independen. Selain itu, untuk lebih menyakinkan dilakukan uji *Durbin Watson* dengan hipotesis sebagai berikut [6] :

$$H_0 : \rho = 0 \text{ (residual independen)}$$

$$H_1 : \rho \neq 0 \text{ (residual tidak independen)}$$

Statistik uji yang digunakan adalah :

$$d_{hitung} = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (6)$$

Keputusan : tolak  $H_0$  jika  $d_{hitung} \leq d_{L, \alpha/2}$  atau

$$d_{L, \alpha/2} \leq (4 - d_{hitung}) \leq d_{U, \alpha/2}.$$

### 2. Asumsi residual identik

Uji identik digunakan untuk melihat homogenitas dari varians residual. Jika tidak homogen maka terjadi heterokedastisitas. Cara mendeteksi homogen tidaknya residual ada dua cara yaitu secara visual dengan membuat *scatter plot* antara residual dengan estimasi variabel respon ( $\hat{y}$ ). Jika hasil plot menunjukkan suatu pola yang membentuk corong, atau garis melengkung maka ada indikasi adanya heterokedastisitas atau varians tidak homogen. Yang kedua melakukan uji *gletser* [6]. Hipotesis yang digunakan adalah :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah :

$$F_{hitung} = \frac{\left[ \sum_{i=1}^n (\hat{e}_i - |\bar{e}|)^2 \right] / (k-1)}{\left[ \sum_{i=1}^n (e_i - |\hat{e}_i|)^2 \right] / (n-k)} \quad (7)$$

Keputusan : : tolak  $H_0$  jika  $F_{hitung}$  lebih besar daripada  $F_{tabel}$  ( $F_{\alpha; (s-1, n-s)}$ ). Nilai  $s$  adalah banyaknya parameter model gletser.

3. Asumsi residual berdistribusi normal

Uji normal digunakan untuk melihat apakah residual mengikuti distribusi normal atau tidak. Uji yang digunakan adalah uji *Kolmogorov Smirnov* dengan hipotesis sebagai berikut [7] :

$H_0$  : Residual mengikuti distribusi normal

$H_1$  : Residual tidak mengikuti distribusi normal

Statistik uji yang digunakan adalah :

$$Z_{hitung} = \sup_x |F_n(x) - F_0(x)| \quad (8)$$

Keputusan : tolak  $H_0$  jika  $Z_{hitung}$  lebih besar daripada  $Z_{\alpha}$

B. Sumber Data

Data yang digunakan dalam penelitian ini adalah data yang diambil dari hasil Survey Angkatan Kerja Nasional (Sakernas) BPS Provinsi Jawa Timur, data Publikasi Makro Sosial dan Ekonomi Jawa Timur, data Produk Domestik Regional Bruto (PDRB) Kabupaten/Kota Provinsi Jawa Timur, data Provinsi Jawa Timur dalam Angka Tahun 2011, dan data investasi dari Badan Penanaman Modal Provinsi Jawa Timur pada tahun 2010 yang mencakup tentang pengangguran terbuka dan faktor-faktor yang mempengaruhinya. Lokasi penelitian adalah 23 kabupaten / kota di Jawa Timur.

Dalam penelitian ini variabel respon ( $y$ ) adalah persentase pengangguran terbuka dan variabel prediktor ( $t$ ) terdapat tujuh variabel yaitu

1. Persentase penduduk usia kerja berumur 15 tahun ke atas berdasarkan pendidikan tertinggi yang ditamatkan ( $T_1$ ).
2. Angka Partisipasi Kasar menurut kabupaten/kota ( $T_2$ ).
3. Laju pertumbuhan ekonomi daerah ( $T_3$ ).
4. Laju pertumbuhan penduduk menurut kabupaten/kota ( $T_4$ ).
5. Persentase perusahaan menurut kabupaten/kota ( $T_5$ ).
6. Tingkat investasi menurut kabupaten/kota ( $T_6$ ).
7. Tingkat upah minimum menurut kabupaten/kota ( $T_7$ ).

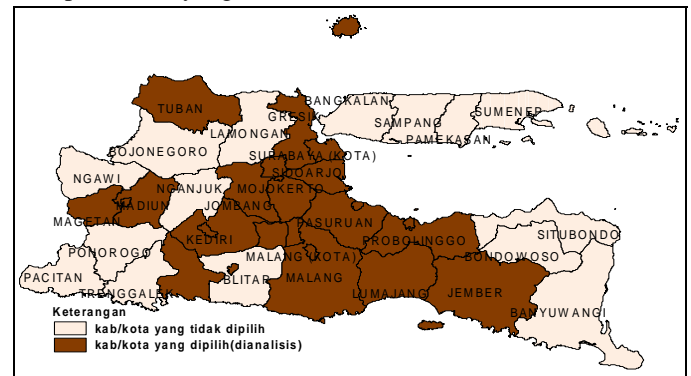
Penaksiran model yang digunakan dalam regresi spline adalah model regresi spline linier dengan menggunakan 3 knot dan kombinasinya. Model yang terbaik adalah model yang memiliki nilai GCV terkecil dari tiap-tiap kombinasi knot yang dihasilkan.

III. HASIL DAN PEMBAHASAN

A. Analisis Pengangguran Terbuka di Jawa Timur

Setiap kabupaten/kota di Jawa Timur memiliki wilayah demografis yang berbeda-beda. Sehingga menyebabkan

karakteristik perekonomian di masing-masing daerah berbeda pula. Akibat adanya perbedaan potensi dari masing-masing daerah mengakibatkan tingkat ekonomi dan tingkat ketenagakerjaan berbeda. Jika di daerah yang unggul di sektor pertanian, ada kecenderungan dari masyarakat bekerja pada saat musim tanam dan musim panen. Jika pada musim itu terlalui, biasanya mereka bekerja serabutan. Dan hal ini, juga berbeda untuk daerah memiliki potensi industri atau perdagangan, hotel, dan rumah makan. Untuk daerah disini, masyarakat yang tidak bekerja berarti masyarakat tersebut sedang mencari pekerjaan atau menganggur. Berikut ini, gambar dari kabupaten/kota yang akan dianalisis :



Gambar. 1. Pemilihan Kabupaten/Kota di Jawa Timur.

Gambar 1 menunjukkan bahwa tidak semua kabupaten / kota digunakan, sehingga hanya yang terpilih yang selanjutnya akan dianalisis. Beberapa kabupaten /kota yang dianalisis yaitu Kota Surabaya, Kota Pasuruan, Kota Batu, Kab. Jember, Kab. Gresik, Kab. Sidoarjo dan lainnya.

B. Model Regresi Spline Linier Multivariabel

Pemodelan dengan menggunakan regresi spline dilakukan dengan memodelkan 1, 2, 3, dan kombinasi 3 titik knot. Dari hasil pemodelan dihitung pula nilai GCV dari masing-masing model. Berikut ini ditampilkan nilai GCV dan pemodelan untuk masing-masing titik knot.

Tabel 1. Nilai GCV dari Model Regresi Spline

Model Regresi Spline	Nilai GCV
1 knot	1,59
2 knot	9,85
3 knot	35,93
Kombinasi 3 knot	0,81

Berdasarkan Tabel 1 diketahui bahwa model regresi spline yang menghasilkan nilai GCV paling minimum adalah model regresi spline dengan menggunakan kombinasi tiga knot. Nilai GCV paling minimum sebesar 0,81 dimana letak titik knot berada di  $T_1=86,10$ ;  $T_2=119,38$  dan  $119,48$ ;  $T_3=6,83$  dan  $6,90$ ;  $T_4=0,33$ ;  $T_5=0,14$ ;  $T_6=10,74$  dan  $32,52$ ; dan  $T_7=0,18$ . Sehingga model regresi spline terbaik adalah model dengan menggunakan kombinasi 3titik knot, yang dapat dituliskan sebagai berikut :

$$\hat{y} = 3,92 - 0,26 t_1 - 0,03(t_1 - 86,10)_+^1 + 0,19 t_2 - 17,91(t_2 - 119,38)_+^1 + 18(t_2 - 119,48)_+^1 - 0,46 t_3 + 9,61(t_3 - 6,83)_+^1 - 11,35(t_3 - 6,90)_+^1 + 0,73 t_4 + 2,41(t_4 - 0,33)_+^1 - 29,37 t_5 + 56,60(t_5 - 0,14)_+^1 + 0,16 t_6 + - 0,31(t_6 - 10,74)_+^1 + 0,20(t_6 - 32,52)_+^1 + 51,33 t_7 - 75,08(t_7 - 0,18)_+^1$$

dari model yang telah diperoleh selanjutnya dilakukan uji kesesuaian parameter model dan asumsi residual.

C. Pengujian Parameter Model secara Serentak

Hipotesis yang digunakan untuk mengetahui pengaruh parameter secara serentak terhadap model yang telah diperoleh sebagai berikut :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{17} = 0$$

$$H_1 : \text{Minimal ada satu } \beta_k \neq 0, k = 1,2,\dots,17.$$

Tabel 2  
ANOVA Model Regresi Spline Linier

Source of Variation	df	Sum of Square	Mean Square	F <sub>hitung</sub>
Regression	17	134,7709	7,9277	24,288
Error	6	1,9584	0,3264	
Total	23	136,7293		

Berdasarkan hasil pengujian ANOVA pada Tabel 2 diperoleh nilai F<sub>hitung</sub> sebesar 24,288, sedangkan nilai F<sub>(0,15;17;6)</sub> sebesar 2,4477. Jadi dari hipotesis di atas dapat diputuskan tolak H<sub>0</sub> yang berarti dengan uji serentak paling tidak ada satu parameter model berpengaruh signifikan terhadap model regresi spline linier multivariabel.

D. Pengujian Parameter Model secara Parsial

Setelah dilakukan uji parameter secara serentak maka selanjutnya adalah dilakukan pengujian parameter secara parsial dengan hipotesis sebagai berikut.

Hipotesis

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 1,2,\dots,17$$

Tabel 3.  
Pengujian Parameter Model secara Parsial

Parameter	Coef	t <sub>hitung</sub>	t <sub>(0,075;6)</sub>	Keputusan
$\beta_1$	-0,2663	1,7641	1,6994	signifikan
$\beta_2$	-0,0301	0,14403	1,6994	tidak signifikan
$\beta_3$	0,1932	1,0363	1,6994	tidak signifikan
$\beta_4$	-17,9198	1,8789	1,6994	signifikan
$\beta_5$	18,0033	1,8865	1,6994	signifikan
$\beta_6$	-0,4695	0,60931	1,6994	tidak signifikan
$\beta_7$	9,6105	0,67524	1,6994	tidak signifikan
$\beta_8$	-11,3570	0,71934	1,6994	tidak signifikan
$\beta_9$	0,7387	0,23375	1,6994	tidak signifikan
$\beta_{10}$	2,4178	0,77865	1,6994	tidak signifikan
$\beta_{11}$	-29,3745	1,10 × 10 <sup>-6</sup>	1,6994	tidak signifikan
$\beta_{12}$	56,6061	0,72906	1,6994	tidak signifikan
$\beta_{13}$	0,1677	2,4015	1,6994	signifikan
$\beta_{14}$	-0,3062	2,6406	1,6994	signifikan
$\beta_{15}$	0,2002	3,3528	1,6994	signifikan

Tabel 3.  
Pengujian Parameter Model secara Parsial (Lanjutan)

Parameter	Coef	t <sub>hitung</sub>	t <sub>(0,075;6)</sub>	Keputusan
$\beta_{16}$	51,3334	1,92 × 10 <sup>-6</sup>	1,6994	tidak signifikan
$\beta_{17}$	-75,0851	0,81463	1,6994	tidak signifikan

Tabel 3 menunjukkan bahwa dari keseluruhan parameter model yang diperoleh ternyata  $\beta_1, \beta_4, \beta_5, \beta_{13}, \beta_{14}, \beta_{15}$  yang berpengaruh signifikan secara parsial terhadap model regresi spline linier. Hal ini diketahui karena nilai |t<sub>hitung</sub>| > t<sub>(0,075;6)</sub>. Sehingga dari uji parameter secara serentak maupun parsial dapat dituliskan model regresi spline linier yang signifikan sebagai berikut :

$$\hat{y} = 0,26 t_1 - 17,91(t_2 - 117,38)_+^1 + 18(t_2 - 123,48)_+^1 + 0,16 t_6 + - 0,31(t_6 - 10,74)_+^1 + 0,20(t_6 - 32,52)_+^1$$

Berdasarkan ketujuh variabel prediktor yang diduga berpengaruh terhadap pengangguran terbuka terdapat tiga variabel prediktor yang signifikan. Jika diasumsikan (t<sub>2</sub>) dan (t<sub>6</sub>) adalah konstan maka model spline linier untuk t<sub>1</sub> adalah  $\hat{y} = 0,26 t_1$ . Dari model tersebut menjelaskan bahwa pengangguran terbuka akan bertambah sebesar 0,26 persen jika variabel penduduk usia kerja berumur 15 tahun ke atas berdasarkan pendidikan tertinggi yang ditamatkan adalah SMA/SMK (t<sub>1</sub>) bertambah satu persen. Dari model tersebut menginformasikan bahwa pengaruh pendidikan cukup besar terhadap pengangguran terbuka. Hal ini dikarenakan, jika seseorang memiliki tingkat pendidikan yang tinggi dan ketrampilan yang memadai maka semakin produktif seseorang dalam bekerja. Bila hal ini seseorang membuka lapangan usaha, maka jumlah pengangguran terbuka akan turun. Sesuai dengan penelitian [8] yang menyatakan bahwa peranan pendidikan mempengaruhi pengangguran, sehingga perlu peningkatan kualitas dan kuantitas.

Selanjutnya model spline linier untuk (t<sub>2</sub>) dengan asumsi (t<sub>1</sub>) dan (t<sub>6</sub>) konstan adalah

$$\hat{y} = -17,91(t_2 - 117,38)_+^1 + 18(t_2 - 123,48)_+^1 = \begin{cases} 2102,2 - 17,91 t_2 & ; 117,38 \leq t_2 < 123,48 \\ -120,36 + 0,09 t_2 & ; t_2 \geq 123,48 \end{cases}$$

Dari model di atas terdapat dua segmen yang menunjukkan pada Angka Partisipasi Kasar (t<sub>2</sub>) antara 117,38 sampai dengan 123,48 maka pengangguran terbuka akan berkurang sebesar 17,91 persen jika Angka Partisipasi Kasar bertambah satu persen. Sedangkan, pada Angka Partisipasi Kasar lebih dari 123,48 maka pengangguran terbuka akan bertambah sebesar 0,09 persen jika Angka Partisipasi Kasar bertambah satu persen. Angka Partisipasi Kasar digunakan untuk melihat kesetaraan gender di bidang pendidikan yang terserap di pasar kerja. Dari model spine untuk (t<sub>2</sub>) menginformasikan bahwa semakin tinggi Angka Partisipasi Kasar maka jumlah pengangguran semakin meningkat. Pada pemodelan ini kurang sesuai dengan teori sosial, karena seharusnya semakin tinggi Angka Partisipasi Kasar maka pengangguran akan turun. Ketidaksesuaian model ini disebabkan kemungkinan adanya otonomi daerah dalam meningkatkan kualitas pendidikan sehingga terjadi perbedaan fasilitas dan kualitas yang di-dapatkan untuk daerah perkotaan

dengan pedesaan. Untuk itu, perlu ada upaya dari pemerintah untuk meningkatkan fasilitas dan kualitas di pedesaan agar tidak tertinggal jauh dengan daerah perkotaan.

Selain itu, untuk model spline linier ( $t_6$ ) dengan asumsi ( $t_1$ ) dan ( $t_2$ ) konstan adalah

$$\hat{y} = 0,16 t_6 - 0,31(t_6 - 10,74)_+ + 0,20(t_6 - 32,52)_+^2$$

$$= \begin{cases} 0,16 t_6 & ; t_6 < 10,74 \\ 3,33 - 0,15 t_6 & ; 10,74 \leq t_6 < 32,52 \\ -3,17 + 0,05 t_6 & ; t_6 \geq 32,52 \end{cases}$$

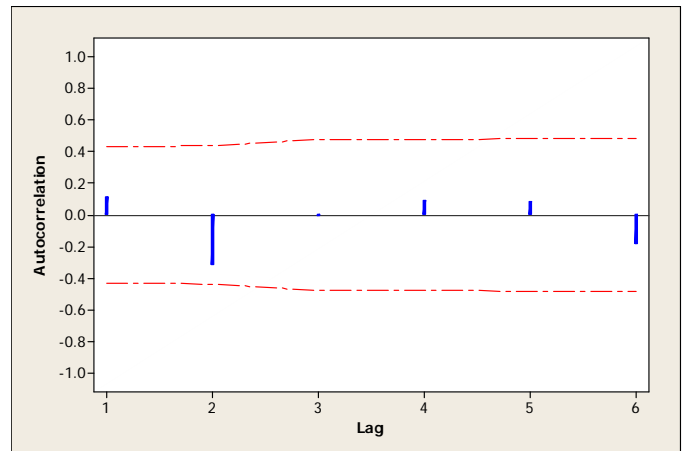
Untuk model ketiga ini terbagi menjadi tiga segmen yaitu untuk segmen pertama tingkat investasi ( $t_6$ ) kurang dari 10,74 persen maka pengangguran terbuka akan bertambah sebesar 0,16 persen jika variabel tingkat investasi ( $t_6$ ) bertambah satu persen. Untuk segmen kedua jika tingkat investasi ( $t_6$ ) antara 10,74 sampai kurang dari 32,52 persen maka pengangguran terbuka akan berkurang sebesar 0,15 persen apabila tingkat investasi bertambah satu persen. Sedangkan tingkat investasi lebih dari 32,52 persen, maka pengangguran terbuka akan bertambah sebesar 0,05 persen jika tingkat investasi dinaikkan sebesar satu persen.

Investasi merupakan salah satu upaya yang dilakukan para pelaku pemilik usaha dalam meningkatkan pengembangan usaha mereka, sehingga dengan semakin tinggi nilai investasi di suatu perusahaan maka akan menyerap jumlah tenaga kerja yang cukup besar. Menurut rujukan [9] menyatakan bahwa peningkatan investasi baik dari dalam negeri maupun luar negeri berpengaruh terhadap pertumbuhan sektor ekonomi dan jumlah tenaga kerja yang terserap. Dari model yang dihasilkan di atas diperoleh nilai  $R^2$  sebesar 99,05 persen dan nilai MSE sebesar 0,3264 sehingga model regresi spline linier dapat dikatakan baik dalam memodelkan pengangguran terbuka di Jawa Timur.

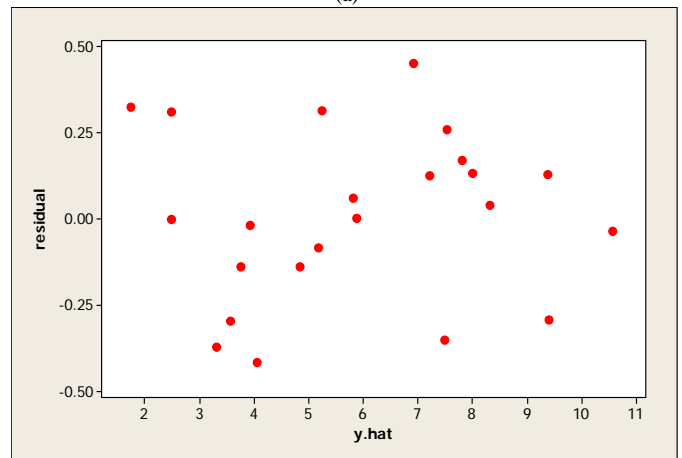
**E. Pemeriksaan Asumsi Residual**

Pengujian asumsi residual meliputi uji independen, identik, dan residual berdistribusi normal. Berikut ini akan ditampilkan plot dari masing-masing uji asumsi residual.

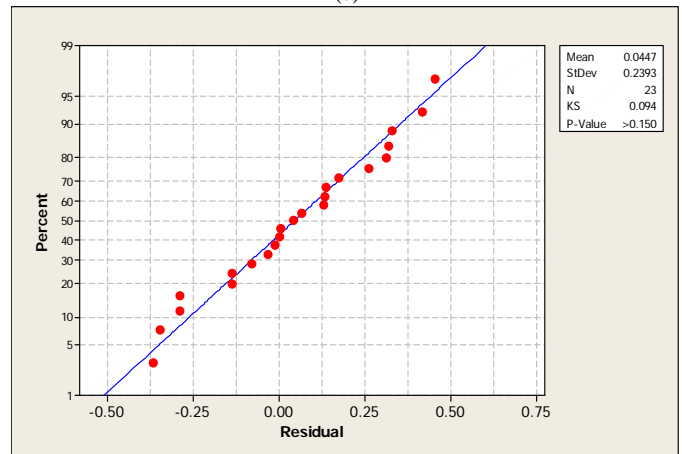
Hasil pengujian asumsi residual menunjukkan untuk uji residual independen dilihat dari hasil plot ACF residual terlihat tidak ada lag yang keluar dari garis sehingga dapat disimpulkan bahwa asumsi residual telah independen terpenuhi. Untuk uji residual identik terlihat hasil plot antara  $\hat{y}_t$  dengan  $e_t$  menunjukkan pola yang menyebar dan tidak terlihat suatu pola corong, garis lurus atau garis melengkung, sehingga dari plot di atas dapat disimpulkan bahwa asumsi residual bersifat identik telah terpenuhi. Sedangkan untuk uji normalitas Gambar 2 (c) diperoleh nilai KS sebesar 0,094 dan  $p$ -value lebih dari 0,15 berarti bahwa residual mengikuti distribusi normal.



(a)



(b)



(c)

Gambar. 2. Uji Pemeriksaan Asumsi Residual (a) Plot ACF dari Residual (b) Plot antara  $\hat{y}$  dengan Residual (c) Probability Plot Residual.

#### IV. KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini adalah pemodelan persentase pengangguran terbuka di Jawa Timur dengan menggunakan regresi spline linier multivariabel adalah sebagai berikut :

$$\hat{y} = 0,26t_1 - 17,91(t_2 - 117,38)_+^1 + 18(t_2 - 123,48)_+^1 + 0,16t_6 + \\ - 0,31(t_6 - 10,74)_+^1 + 0,20(t_6 - 32,52)_+^1$$

Model regresi spline ini menghasilkan  $R^2$  sebesar 99,05 persen dan nilai MSE sebesar 0,3264. Variabel yang berpengaruh signifikan terhadap model persentase pengangguran terbuka adalah persentase penduduk usia kerja berumur 15 tahun ke atas berdasarkan pendidikan tertinggi yang ditamatkan adalah SMA/SMK, Angka Partisipasi Kasar, dan tingkat investasi menurut kabupaten/kota di Jawa Timur.

#### DAFTAR PUSTAKA

- [1] Santoso, N. (2009). *Klasifikasi Kabupaten/kota di Jawa Timur Berdasarkan Tingkat Pengangguran Terbuka dengan Pendekatan Multivariate Adaptive Regression Spline (MARS)*. Surabaya: Institut Teknologi Sepuluh Nopember
- [2] Lailiya, A. R. (2011). *Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Kesamaan Nilai Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka dengan Metode Hirarki dan Nonhirarki*. Surabaya: Institut Teknologi Sepuluh Nopember.
- [3] Asti, F. (2011). *Regresi Multivariate pada Faktor yang Berpengaruh terhadap Pengangguran Terbuka di Jawa Timur*. Surabaya: Institut Teknologi Sepuluh Nopember.
- [4] Budiantara, I. N. (2009). "Spline dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang", *Pidato Pengukuhan untuk Jabatan Guru Besar dalam Bidang Ilmu Matematika Statistika dan Probabilitas, pada Jurusan Statistika, Fakultas MIPA*. Surabaya: ITS Press.
- [5] Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing 2nd Edition*. New York: Marcel Dekker.
- [6] Gujarati, D. N. (2006). *Dasar-Dasar Ekonometrika Edis Pertama*. Jakarta: Erlangga.
- [7] Kvam, P. H., dan Vidakovic, B. (2007). *Nonparametric Statistics with Applications to Science and Engineering*. New Jersey: John Wiley & Sons.
- [8] Merizal, Y. (2008). *Analisis Pengaruh Pendidikan, Tingkat Upah Minimum Kabupaten, dan Kesempatan Kerja terhadap Pengangguran Terdidik di Kabupaten Semarang*. Semarang: Universitas Diponegoro.
- [9] Sitompul, N. L. (2008). *Analisis Pengaruh Investasi dan Tenaga Kerja terhadap PDRB Sumatrer Utara*. Medan: Universitas Sumatera Utara.