



**APLIKASI GAMLSS DALAM PEMODELAN *GENERALIZED*
GAMMA DENGAN ALGORITMA *MIXED* PADA
PEMULUSAN LOESS**

SKRIPSI

Oleh

**Ulfa Hasanah
NIM 111810101026**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015**



**APLIKASI GAMLSS DALAM PEMODELAN *GENERALIZED*
GAMMA DENGAN ALGORITMA *MIXED* PADA
PEMULUSAN LOESS**

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat untuk menyelesaikan Program Studi Matematika (S1) dan mencapai gelar Sarjana Sains

Oleh

Ulfa Hasanah
NIM 111810101026

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015**

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

1. Allah SWT yang telah memberikan kehidupan sempurna ini;
2. Ayahanda H. Bahruddin Ali dan Ibunda Hj. Umi Azizah yang telah memberikan doa, cinta dan kasih sayangnya kepada putri kecilnya;
3. Kakakku Hj. Nur Kamilah yang selalu memberikan semangat tanpa henti;
4. guru-guru sejak taman kanak-kanak sampai perguruan tinggi, yang telah memberikan ilmu dan membimbing dengan penuh kesabaran;
5. Almamater Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

MOTTO

Waktu itu bagaikan pedang, jika kamu tidak memanfaatkannya
menggunakan untuk memotong, ia akan memotongmu (menggilasmu)
(H.R. Muslim)

Mencoba untuk tidak menjadi orang yang sukses, tetapi cukup mencoba menjadi
orang yang berharga.*)

*) Einstein, A. Einstein quotes [on line]. <http://higherperspectives.com/einstein-quotes/> [8 September 2015]

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

nama : Ulfa Hasanah

NIM : 111810101026

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Aplikasi GAMLSS dalam Pemodelan *Generalized* Gamma dengan Algoritma *Mixed* pada Pemulusan Loess” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, September 2015

Yang menyatakan,

Ulfa Hasanah

111810101026

SKRIPSI

**APLIKASI GAMLSS DALAM PEMODELAN *GENERALIZED*
GAMMA DENGAN ALGORITMA *MIXED* PADA
PEMODELAN LOESS**

Oleh

Ulfa Hasanah
NIM 111810101026

Pembimbing:

Dosen Pembimbing Utama : Prof. Drs.I Made Tirta M.Sc., Ph.D

Dosen Pembimbing Anggota : Dian Anggraeni, S.Si., M.Si

RINGKASAN

Aplikasi GAMLSS dalam Pemodelan *Generalized Gamma* dengan Algoritma *Mixed* pada Pemulusan *Loess* ; Ulfa Hasanah,111810101026; 2015; 31 halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember

Generalized Additive Model for Location, Scale and Shape(GAMLSS) adalah metode pencocokan model bersifat semi-parametrik yang merupakan gabungan regresi parametrik dan non-parametrik. GAMLSS merupakan perluasan dari GLM (*Generalized Linear Model*) dan GAM (*Generalized Additive Model*) yang mencakup distribusi Keluarga Eksponensial dan distribusi non-Keluarga Eksponensial, salah satu distribusi yang tergolong Keluarga Eksponensial yaitu distribusi Gamma yang memiliki dua parameter, sedangkan distribusi yang tergolong non-Keluarga Eksponensial yaitu *Generalized Gamma* yang merupakan perluasan dari distribusi Gamma dengan tiga parameter. Metode GAMLSS ini dapat memodelkan 4 parameter yaitu parameter lokasi (μ) yang diperlukan untuk menggambarkan letak pemusatan data, parameter skala (σ) sebagai parameter dispersi/penyebaran data, parameter *skewness* (ν) dan parameter kurtosis(τ) digunakan untuk mendefinisikan bentuk kurva dari sebaran data.

Pada penelitian ini, langkah awal yang dilakukan yaitu memverifikasi distribusi Gamma merupakan bentuk khusus dari distribusi *Generalized Gamma*, dan langkah selanjutnya yaitu menganalisis distribusi Gamma dan *Generalized Gamma* untuk memperoleh distribusi yang paling baik, kemudian melakukan *fitting* model data dengan menggunakan prosedur GAMLSS. Data yang digunakan merupakan data sekunder yang berupa data Hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011. Setelah mendapatkan model data terbaik, kemudian dilihat hasil estimasi dari parameter lokasi, skala dan ukuran.

Berdasarkan kajian yang telah dilakukan, terverifikasi bahwa distribusi Gamma merupakan bentuk khusus dari *Generalized Gamma* dengan parameter

$v = 1$. Distribusi yang lebih baik untuk analisis data hasil Ujian Nasional adalah distribusi *Generalized Gamma*, dan formula model data terbaik yang didapatkan dari analisis *Generalized Gamma* yaitu:

UN~UJIAN.SEKOLAH+IQ+lo(~RAPOT, degree=1, span=.3, data=UAN, family=GG, method=mixed(60, 60)) dengan hasil *summary* dari model di atas menunjukkan bahwa nilai Ujian Sekolah dan hasil tes IQ berpengaruh secara signifikan terhadap hasil Ujian Nasional SMA Negeri 1 Situbondo. Berikut model terbaik untuk variabel hasil Ujian Nasional pada data UAN.

Untuk parameter lokasi (μ) dengan log sebagai *link-function* diperoleh :

$$\hat{\mu}_{\text{Ujian Nasional}} = \exp(1,182 + 8,143 \times 10^{-2} \text{ Ujian Sekolah} + 4,828 \times 10^{-3} \text{ IQ})$$

Untuk parameter *scale* (σ) dengan log sebagai *link-function* diperoleh :

$$\log(\hat{\sigma}_{\text{Ujian Nasional}}) = -3,56466$$

$$\hat{\sigma}_{\text{Ujian Nasional}} = \exp(-3,56466)$$

Dan untuk parameter *shape* (v) dengan Identity sebagai *link-function* diperoleh :

$$\hat{v}_{\text{Ujian Nasional}} = 62,963$$

PRAKATA

Alhamdulillah, puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga tugas akhir yang berjudul “Aplikasi GAMLSS dalam Pemodelan *Generalized* Gamma dengan Algoritma *Mixed* pada Pemulusan Loess” dapat terselesaikan dengan baik. Tugas akhir ini disusun untuk memenuhi syarat menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

- a. Prof. Drs.I Made Tirta M.Sc., Ph.D, selaku Dosen Pembimbing Utama dan Dian Anggraeni, S.Si, M.Si., selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan skripsi ini;
- b. Dr.Alfian Futuhul Hadi, S.Si., M.Si., dan Kosala Dwidja Purnomo, S.Si., M.Si., selaku dosen penguji atas saran-saran yang diberikan;
- c. seluruh staf pengajar Jurusan Matematika Fakultas MIPA Universitas jember yang telah memberikan ilmu serta bimbingannya sehingga penulis dapat menyelesaikan skripsi ini;
- d. seluruh keluarga dirumah, Ayahanda H. Bahruddin Ali, Ibunda Hj. Umi Azizah, dan kakak Hj. Nur Kamilah yang telah memberikan doa dan motivasi demi terselesaikannya skripsi ini;
- e. sahabat terbaik Lina Choiril O., Hijratul Ilahiyah dan Ivana Gabriella yang selalu memberikan dorongan dan semangat untuk menyelesaikan skripsi ini;
- f. teman-teman KRAMAT'11 yang telah memberikan motivasinya serta kakak dan adik angkatan Jurusan Matematika, terima kasih atas semangatnya;
- g. anak kost MAMA BELLA yang telah memberikan semangat, dukungan dan berbagi cerita;
- h. semua pihak yang tidak dapat disebutkan satu per satu.

Penulis juga menerima segala kritik dan saran dari semua pihak demi kesempurnaan skripsi ini. Akhirnya penulis berharap, semoga skripsi ini dapat bermanfaat.

Jember, September 2015

Penulis



DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN PEMBIMBING	v
HALAMAN PENGESAHAN.....	vi
RINGKASAN	vii
PRAKATA	ix
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN	xvi
BAB 1.PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat	3
BAB 2.TINJAUAN PUSTAKA.....	4
2.1 Model Linier	4
2.2 <i>Generalized Linear Model (GLM)</i>	4
2.3 <i>Generalized Additive Model (GAM)</i>	5
2.4 <i>Skewness dan Kurtosis</i>.....	5

2.4.1	<i>Skewness</i>	6
2.4.2	Kurtosis	6
2.5	Generalized Additive Model for Location, Scale and Shape (GAMLSS)	7
2.6	Bentuk GAMLSS	7
2.7	Distribusi Gamma	9
2.8	Distribusi <i>Generalized Gamma</i>	9
2.9	Metode Pemulusan Loess	10
2.10	Algoritma <i>Mixed</i>	12
2.11	Pemilihan Model melalui AIC	15
BAB 3. METODE PENELITIAN	16
3.1	Data	16
3.2	Metode Pengolahan Data & Analisis Data	16
BAB 4. HASIL DAN PEMBAHASAN	19
4.1	Perbandingan Bentuk Fungsi Kepadatan dari Gamma dan <i>Generalized Gamma</i>	19
4.2	Sebaran Data	21
4.3	<i>Fitting Model Data Menggunakan Prosedur GAMLSS dengan Pemulusan Loess dan Algoritma <i>Mixed</i></i>	24
4.3.1	Menentukan Distribusi Terbaik.....	24
4.3.2	Fitting Data dengan Pemulusan Loess	25
a.	Memilih <i>Degree</i> Polinomial Terbaik	25
b.	Memilih <i>Span</i> Terbaik	26
4.4	Menentukan Model Terbaik dengan Menggunakan stepGAIC ..	28
4.5	Menentukan Nilai Estimasi dari Model Terbaik	29
BAB 5. PENUTUP	31
5.1	Kesimpulan	31
5.2	Saran	31
DAFTAR SUMBER RUJUKAN	32

LAMPIRAN..... 33



DAFTAR TABEL

	Halaman
Tabel 3.1 Variabel Prediktor(x) yang Digunakan	6
Tabel 4.1 Nilai AIC Distribusi Model	25
Tabel 4.2 Nilai AIC Degree yang Terbaik	26
Tabel 4.3 Nilai AIC <i>Span</i> yang Terbaik.....	27
Tabel 4.4 <i>Coefficients</i> Parameter μ (mu) dengan Log sebagai <i>Link function</i>	29
Tabel 4.5 <i>Coefficients</i> Parameter σ (sigma) dengan Log sebagai <i>Link function</i>	29
Tabel 4.6 <i>Coefficients</i> Parameter ν (nu) dengan Identity sebagai <i>Link function</i>	30
Tabel 4.7 Nilai Estimasi, <i>p-value</i>	30

DAFTAR GAMBAR

	Halaman
Gambar 3.2 Diagram Metode Penelitian.....	18
Gambar 4.1 Kurva Fungsi Kepadatan Peluang Distribusi (a) Gamma (b)	21
Gambar 4.2 Sebaran Data Hasil Ujian Nasional.....	22
Gambar 4.3 Hubungan Kolinieran antara Variabel Respon dan Masing-masing Variabel Eksplanatori	23
Gambar 4.4 Histogram Distribusi (a) Gamma (b) <i>Generalized</i> Gamma	24
Gambar 4.5 Hasil Visualisasi <i>Smoothing</i> dengan (a) <i>Span</i> 0,3 (b) <i>Span</i> 0,6 (c) <i>Span</i> 0,8	27

DAFTAR LAMPIRAN

	Halaman
A. Data Hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/201117	33
B. Pencocokan <i>Generalized Additive Model for Location Scale And Shape</i> (GAMLSS) Secara Umum.....	35
C. Tabel Daftar Distribusi Kontinu yang Dapat Diimplementasikan dalam Paket GAMLSS	37
D. Program.....	39
D1. Ilustrasi Fungsi Kepadatan Peluang dari distribusi Gamma dan <i>Generalized</i> Gamma	39
D2. Sebaran Data	39
D3. Analisis Distribusi Terbaik	40
D4. Analisis Derajat Polinomial (<i>degree</i>) terbaik.....	42
D5. Analisis <i>Span</i> Terbaik	44
D6. Visualisasi Model Setelah Dilakukan Smoothing dengan Beberapa Pilihan <i>Span</i>	56
D7. Analisis Model Terbaik dengan Menggunakan stepGAIC	56

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan model statistika yang mempelajari hubungan fungsional dari dua variabel yaitu variabel respon dan variabel explanatori/penjelas, yang mana dapat menganalisis data cacahan maupun kontinu. Perkembangan analisis regresi saat ini semakin berkembang seiring dengan jumlah data yang dibutuhkan semakin meningkat, untuk itu diharapkan model regresi lebih fleksibel dan algoritma yang digunakan memungkinkan untuk mengolah data dengan cepat dan lebih spesifik.

Perkembangan pemodelan stokastik, terutama model linier dapat dikatakan dimulai pada abad ke 19 yang ditandai adanya kajian oleh Galton (1822-1911) tentang hubungan tinggi badan ayah dengan anaknya. Model linier normal atau *Normal Linear Models* (NLM) merupakan model linier yang paling sederhana yang diasumsikan variabel respon berdistribusi normal. Pada kenyataannya variabel respon tidak hanya berdistribusi normal saja sehingga para pelopor statistik mengembangkan model linier yang diasumsikan variabel respon tidak hanya berdistribusi normal melainkan juga berdistribusi eksponensial yang dikenal dengan *Generalized Linier Model* (GLM) dan *Generalized Additive Model* (GAM). Namun kedua model linier tersebut tidak memodelkan *skewness* dan kurtosis yang sangat penting untuk melihat keruncingan dan kemencengan pada data. Sehingga dengan mengetahui hal tersebut maka pemodelan statistik yang dilakukan bisa lebih valid.

Beberapa kekurangan dari setiap model linier yang dijelaskan di atas, Rigby & Stasinopoulos (2001, 2005) dan Akantziliotou *et al.* (2002) pada (Rigby & Stasinopoulos, 2007) mengenalkan metode pencocokan model yaitu *Generalized Additive Model for Location Scale & Shape* (GAMLSS), GAMLSS merupakan perluasan dari GLM dan GAM yang mencakup distribusi keluarga Eksponensial dan non-Eksponensial dan dapat memodelkan *skewness* dan kurtosis. GAMLSS bersifat semi-parametrik karena mengasumsikan fungsi

parametrik untuk variabel respon dan memodelkan fungsi parameter dari variabel eksplanatori yang melibatkan fungsi *smooth* non-parametrik. Metode ini juga menyediakan algoritma yang membantu dalam proses *fitting* data, salah satunya yaitu algoritma *mixed*. Algoritma *mixed* merupakan gabungan dari algoritma Rigby&Stasinopoulus (RS) dan algoritma Cole&Green (CG) dengan iterasi j dari algoritma RS kemudian diteruskan iterasi k dari algoritma CG. Distribusi yang tercakup keluarga Eksponensial salah satunya adalah distribusi Gamma (GA), distribusi ini memiliki dua parameter yaitu parameter mu (μ) dan sigma (σ). Dan distribusi yang tercakup keluarga non-Eksponensial salah satunya adalah distribusi *Generalized* Gamma (GG) yang merupakan perluasan dari distribusi Gamma yang memiliki tiga parameter yaitu parameter mu (μ), sigma (σ) dan nu (ν).

Pada era ini banyak prosedur pemulusan pada kurva dan permukaan regresi dari sebuah data, salah satunya yaitu dengan menggunakan pemulusan loess. Loess merupakan suatu strategi untuk pemulusan kurva dari data empiris dan menyediakan suatu rangkuman grafis hubungan antara variabel respon dan variabel eksplanator (Jacoby, 2000).

Peneliti sebelumnya, (Wandi, 2012) telah mengaplikasikan GAMLSS untuk meneliti data agroklimatologi pada proyek Pekalen-Sampean Kabupaten Jember dengan algoritma Rigby-Stasinopoulus (RS) dan pemulusan *cubic spline*. Hasil penelitiannya menunjukkan distribusi yang sesuai untuk data tersebut adalah *Generalized* Gamma. Berdasarkan penelitian diatas, penelitian kali ini akan mengaplikasikan GAMLSS dalam memodelkan *Generalized* Gamma dengan algoritma *mixed* pada pemulusan loess. Data yang digunakan merupakan data sekunder dari (Permono, 2013) yang berupa data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011.

1.2 Rumusan Masalah

Setelah mengetahui uraian diatas dapat dirumuskan suatu permasalahan yang akan dibahas dalam tugas akhir ini yaitu sebagai berikut :

- a. bagaimana perbedaan distribusi Gamma dan distribusi *Generalized* Gamma didalam regresi GAMLSS?
- b. bagaimana memodelkan data dengan analisis *Generalized* Gamma dengan algoritma *mixed* pada pemulusan loess ?

1.3 Tujuan Masalah

Adapun tujuan dari penyusunan tugas akhir ini adalah sebagai berikut :

- a. mengetahui perbedaan distribusi Gamma dan *Generalized* Gamma didalam regresi GAMLSS.
- b. mendapatkan model data dari analisis *Generalized* Gamma dengan algoritma *mixed* pada pemulusan loess

1.4 Manfaat

Manfaat dari penyusunan tugas akhir ini adalah untuk menambah wawasan kepada pembaca, terutama bagi pihak yang mempelajari pemodelan statistika tentang distribusi Gamma, *Generalized* Gamma di dalam regresi GAMLSS dengan algoritma *mixed* pada pemulusan loess.

BAB 2. TINJAUAN PUSTAKA

2.1 Model Linier

Pemodelan stokastik, terutama model linier mulai berkembang pada abad ke-19 yang ditandai dengan adanya kajian tentang hubungan tinggi badan ayah dan anaknya yang dikaji oleh Galton (1822-1911). Pada abad yang sama perkembangan model linier juga dimulai dengan perkembangan analisis regresi oleh Pearson yang mengembangkan korelasi setelah itu. Teori regresi ini yang menjadi dasar perkembangan teori model linier yang tidak bisa dilepaskan dengan perkembangan teori matriks atau aljabar linier. Melalui teori matriks yaitu determinan, invers dan perkalian matriks pembahasan model linier dapat didekati secara umum.

Bentuk umum model linier sederhana yang dapat dituliskan sebagai berikut (Tirta, 2009) :

$$Y = X\beta + \epsilon \quad (2.1)$$

dengan X adalah peubah tetap yang bukan acak, β merupakan parameter yang menentukan peubah tetap tadi, dan ϵ merupakan kesalahan atau galat yang diasumsikan merupakan peubah acak yang berasal dari suatu distribusi tertentu. Jika galat yang diasumsikan peubah acak berdistribusi Normal maka model linier (2.1) disebut model linier normal (Normal Linier Model). Pada model linier normal variabel respon hanya diasumsikan berdistribusi normal, namun dalam kenyataan variabel respon tidak hanya berdistribusi normal saja melainkan juga berdistribusi eksponensial yang dikenal dengan *Generalized Linear Model* (GLM).

2.2 *Generalized Linier Model* (GLM)

Generalized Linear Model (GLM) merupakan perluasan dari model linier normal yang dapat menangani kondisi yang mana variabel respon tidak hanya berdistribusi Normal, tetapi masih *independent* (saling bebas). Model linier ini menggunakan asumsi variabel respon berdistribusi keluarga Eksponensial.

Adapun distribusi keluarga Eksponensial yang banyak dikenal yaitu distribusi Normal, Gamma, dan Poisson.

Ada tiga hal yang terkandung dalam *Generalized Linier Model* (GLM) (Tirta, 2009) yaitu:

1. komponen tetap yang disebut prediktor linier $\eta = \sum_{j=0}^p x_{ij}\beta_j$
2. respon y_i berdistribusi secara independen dalam keluarga eksponensial
3. hubungan antara mean dengan prediktor linier ditunjukkan fungsi $g(\cdot)$ yang disebut fungsi 'link' sedemikian hingga $g(\mu) = \eta$. Fungsi $g(\cdot)$ disebut fungsi hubungan (*link-function*).

Pada kenyataannya, GLM juga memiliki keterbatasan yaitu hanya mampu mencocokkan model dengan variabel respon berdistribusi Eksponensial dengan prediktor linier. Dari keterbatasan tersebut diperlukan model linier dengan prediktor non linier seperti model aditif.

2.3 *Generalized Additive Model* (GAM)

Generalized Additive Model merupakan perluasan dari model aditif yang mengasumsikan bahwa mean dari variabel dependen tergantung pada prediktor aditif. Model linier ini juga mengasumsikan variabel respon berdistribusi Eksponensial. *Generalized Additive Model* (GAM) juga merupakan perluasan *Generalized Linear Model* (GLM) dengan mengadopsi prediktor linier $\sum_{j=0}^p x_j\beta_j$ menjadi $\sum_{j=0}^p f_j x_j$ yang merupakan prediktor aditif.

Bentuk umum dari model linier ini didefinisikan sebagai berikut :

$$Y = \sum_{j=0}^p f_j x_j + \epsilon \quad (2.2)$$

Namun GLM dan GAM tidak memodelkan *skewness* dan kurtosis secara eksplisit, tetapi keduanya hanya memodelkan *skewness* dan kurtosis secara implisit yang tergantung terhadap mean.

2.4 *Skewness dan Kurtosis*

Karakteristik dari model regresi GAMLSS salah satunya terdapat *skewness* dan kurtosis. Nilai dari *skewness* dan kurtosis tersebut diilustrasikan

pada sebuah kurva yang menggambarkan kemencengan dan keruncingan pada sebuah data. Dengan melihat arah kemencengan dan bentuk keruncingan tersebut dapat melakukan pemodelan statistik dengan valid.

2.4.1 Skewness

Skewness atau kemencengan dari suatu kurva merupakan derajat ketidaksimetrisan dari distribusi. Ketidaksimetrisan tersebut dapat dilihat dari letak penyebaran datanya disekitar rata-rata yang tidak simetris. Dikatakan simetris jika ketiga ukuran pemusatan data (*mean*, median, modus) berada pada titik yang sama. Jika distribusi itu menceng ke kanan maka memiliki kemencengan positif (*positive skewness*) yang mana dapat dilihat dengan ekor dari kurva lebih menjulur ke kanan atau ke arah positif, namun sebaliknya jika distribusi tersebut menceng ke kiri maka memiliki kemencengan negatif (*negative skewness*) yang ditandai dengan menjulurnya ekor kurva ke arah kiri.

Menurut (Brown, 2011) nilai *skewness* dapat dicari dengan rumus dibawah ini

$$Skewness (S) = \frac{m_3}{m_2^{3/2}} \quad (2.3)$$

dengan

$$m_3 = \frac{\sum_{j=0}^p (x_j - \bar{x})^3}{n} \text{ dan } m_2 = \frac{\sum_{j=0}^p (x_j - \bar{x})^2}{n} \quad (2.4)$$

dimana :

\bar{x} = mean

n = jumlah sampel

m_3 = moment ketiga

m_2 = varian

2.4.2 Kurtosis

Ketinggian dan keruncingan (*peakedness*) kurva dari suatu distribusi dinamakan kurtosis. Tingkat keruncingan diukur dengan membandingkan bentuk keruncingan kurva distribusi data dengan kurva normal atau distribusi Normal. Distribusi Normal memiliki nilai kurtosis sama dengan 3 dan dinamakan mesokurtik, sedangkan kurva yang memiliki nilai kurtosis lebih dari 3 dinamakan leptokurtik dengan ciri puncak kurva lebih runcing dibanding dengan distribusi

normal, sementara kurva yang memiliki nilai kurtosis kurang dari 3 dinamakan platikurtik dengan ciri puncak kurva lebih datar.

Menurut (Brown, 2011) nilai skewness dapat dicari dengan rumus dibawah ini:

$$Kurtosis (K) = \frac{m_4}{m_2^2} \quad (2.5)$$

dengan

$$m_4 = \frac{\sum_{j=0}^p (x_j - \bar{x})^4}{n} \quad \text{dan} \quad m_2 = \frac{\sum_{j=0}^p (x_j - \bar{x})^2}{n} \quad (2.6)$$

dimana :

\bar{x} = mean

n = jumlah sampel

m_4 = moment keempat

m_2 = varian

2.5 Generalized Additive Model For Location Scale & Shape (GAMLSS)

Generalized Additive Model For Location Scale & Shape (GAMLSS) adalah regresi yang bersifat semi- parametrik, pada model ini bersifat parametrik karena mengasumsikan parametrik untuk variabel respon, dan memodelkan fungsi parameter dari variabel eksplanatori yang melibatkan penggunaan fungsi *smooth* non-parametrik. Pada model regresi GAMLSS ini variabel respon (y) dapat berasal dari distribusi Keluarga Eksponensial (distribusi Gamma) maupun dari non Keluarga Eksponensial (distribusi *Generalized* Gamma). GAMLSS juga baik dalam memodelkan parameter *skewness* dan kurtosis pada distribusi diskrit dan distribusi kontinu.

2.6 Bentuk GAMLSS

Model GAMLSS mengasumsikan variabel respon y_i untuk $i = 1, 2, \dots, n$ dengan fungsi kepadatan peluang $f(y_i | \theta^i)$ yang mana $\theta^i = (\theta_{i1}, \theta_{i2}, \dots, \theta_{ip})$ yang merupakan vektor dengan p parameter, p parameter didefinisikan sebagai jumlah parameter yang ada pada observasi. Jumlah maksimal parameter yang ada pada observasi adalah 4 parameter yaitu μ_i, σ_i, ν_i , dan τ_i . Dua parameter pertama yaitu

μ_i dan σ_i dikarakteristikan sebagai parameter lokasi dan parameter skala, untuk yang lainnya yaitu v_i dan τ_i dikarakteristikan sebagai parameter ukuran, yang mana parameter ukuran dibagi dua jenis yaitu parameter skewness(v_i) dan kurtosis (τ_i). Misal $y^T = (y_1, y_2, \dots, y_n)$ vektor dari variabel respon dengan panjang n , dan $g_k(\cdot)$ dimisalkan sebagai fungsi link monotonik yang menghubungkan θ_k dengan parameter distribusi dan variabel eksplanator, untuk $k = 1, 2, \dots, p$.

Menurut (Rigby & Stasinopoulos, 2007) model GAMLSS diberikan sebagai berikut :

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} Z_{jk} \gamma_{jk} \quad (2.7)$$

Parameter θ_k dikarakteristikan sebagai parameter lokasi, skala dan ukuran, sehingga didapat model

$$g_1(\mu) = \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} Z_{j1} \gamma_{j1}$$

$$g_1(\sigma) = \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} Z_{j1} \gamma_{j1}$$

$$g_1(v) = \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} Z_{j1} \gamma_{j1}$$

$$g_1(\tau) = \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} Z_{j1} \gamma_{j1}$$

Jika $J_k = 0$ maka persamaan (2.7) menjadi

$$g_k(\theta_k) = \eta_k = X_k \beta_k \quad (2.8)$$

Sedangkan jika pada persamaan (2.7) $Z_{jk} = I_n$, dengan I_n merupakan matrik identitas $n \times n$ dan $\gamma_k = h_{jk} = h_{jk}(X_{jk})$ untuk semua kombinasi j dan k , sehingga diperoleh model GAMLSS semi-parametrik sebagai berikut

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} h_{jk} X_k \quad (2.9)$$

Keterangan :

μ, σ, ν, τ : parameter lokasi, skala, ukuran (skewness & kurtosis)

β_k : vektor parameter dengan panjang J'_k

X_k : matrik dengan ukuran $n \times J'_k$

Z_{jk} : variabel random

γ_{jk} : parameter random

Fungsi h_{jk} merupakan fungsi yang tidak diketahui dari variabel eksplanatory X_{jk} dan $h_{jk} = h_{jk}(x_{jk})$ yang merupakan vektor dari hasil taksiran dari fungsi h_{jk} dan x_{jk} .

2.7 Distribusi Gamma (GA)

Distribusi Gamma(GA) tergolong dalam distribusi keluarga Eksponensial yang kontinu dan memiliki dua parameter . Distribusi Gamma memiliki ruang rentang $0 < y < \infty$ dengan hubungan mean dan varian berbentuk kuadratik, dan sebaran data semakin kekanan semakin lebar (Tirta, 2009). Fungsi gamma didefinisikan sebagai berikut :

$$\Gamma(\sigma) = \int_0^{\infty} y^{\frac{1}{\sigma^2}-1} e^{-y} dy \quad (2.10)$$

Fungsi kepadatan peluang dari Gamma dinotasikan $GA(\sigma, \mu)$, dengan σ paramater skala dan μ merupakan parameter lokasi. Menurut (Rigby & Stasinopoulos, 2014) fungsi kepadatan peluang distribusi Gamma didefinisikan sebagai berikut:

$$f_y(y|\sigma, \mu) = \frac{1}{(\sigma^2\mu)^{1/\sigma^2}} \frac{y^{1/\sigma^2-1} e^{-y/(\sigma^2\mu)}}{\Gamma(\frac{1}{\sigma^2})}, y > 0, \mu > 0, \sigma > 0 \quad (2.11)$$

Untuk mean dan varian didefinisikan

$$E(Y) = \mu \text{ dan } Var(Y) = \sigma^2 E(Y)^2$$

2.8 Distribusi Generalized Gamma (GG)

Distribusi *Generalized* Gamma(GG) tergolong dalam distribusi keluarga non Eksponensial yang memiliki tiga parameter yaitu σ, μ, ν dan distribusi ini merupakan perluasan dari distribusi Gamma. Fungsi kepadatan peluang dari

Generalized Gamma dinotasikan GG (σ, μ, ν) , dengan σ merupakan parameter skala, μ merupakan parameter lokasi, dan ν parameter ukuran. Fungsi kepadatan peluang distribusi *Generalized* Gamma sebagai berikut (Rigby & Stasinopoulos, 2010):

$$f_y(y|\sigma, \mu, \nu) = \frac{\nu \theta^\theta z^\theta \exp\{-\theta z\}}{\Gamma(\theta) y}, y > 0, -\infty < \nu < \infty, \mu > 0, \sigma > 0 \quad (2.12)$$

untuk

$$z = (y/\mu)^\nu \text{ dan}$$

$$\theta = \frac{1}{(\sigma^2 \nu^2)}$$

Mean dan varian dari distribusi *Generalized* Gamma diberikan sebagai berikut (Rigby & Stasinopoulos, 2010) :

$$E(Y) = \frac{\mu \Gamma(\theta + \frac{1}{\nu})}{\left[\frac{1}{\theta \nu} \Gamma(\theta) \right]} \text{ dan}$$

$$Var(Y) = \frac{\mu^2 \left\{ \Gamma(\theta) \Gamma\left(\theta + \frac{2}{\nu}\right) - \left[\Gamma\left(\theta + \frac{1}{\nu}\right) \right]^2 \right\}}{\left[\theta^2 / \nu \Gamma(\theta) \right]^2}$$

2.9 Metode Pemulusan Loess

Loess merupakan suatu strategi untuk pemulusan kurva dari data empiris dan menyediakan bentuk rangkuman grafis hubungan antara variabel respon dan variabel eksplanator (Jacoby, 2000). Loess menambahkan *scatterplot* agar dengan mudah melihat hubungan prosedur model data statistik. Proses pemulusan dianggap lokal karena setiap nilai yang dimuluskan ditentukan oleh titik data (x_i, y_i) yang berdekatan pada suatu *span*, dengan i didefinisikan $i = 1, \dots, n$ dan n merupakan banyaknya data.

(Cleveland, 1979) menyatakan ada empat pokok yang menjadi dasar untuk melakukan pemulusan loess yaitu pemilihan $f(\text{span})$, W (fungsi bobot), d (derajat polinomial) dan t (iterasi). Parameter f merupakan *span* (rentang) untuk menentukan jumlah pemulusan yang memberikan takaran dari sebuah observasi yang digunakan setiap daerah regresi. Nilai *span* di spesifikasi antara 0 sampai 1, yang mana nilai f tersebut mempengaruhi pemulusan terhadap kurva, semakin

besar nilai f maka *fitting* kurva semakin mulus, dan berlaku untuk sebaliknya. W merupakan fungsi bobot yang terkandung didalam *span*, berikut fungsi bobot *tricube*:

$$W(x) \begin{cases} (1 - |x^3|)^3, & |x| < 0 \\ 0, & |x| \geq 0 \end{cases} \quad (2.13)$$

untuk memperoleh *local regression* yaitu dengan menghitung estimasi $\hat{\beta}_m(x_i)$ yang merupakan nilai dari $\hat{\beta}_m$ dengan meminimalkan persamaan berikut :

$$\hat{\beta} = \sum_{k=1}^n w_k (y_k - \beta_0 - \beta_1 x_k - \dots - \beta_d x_k^d)^2 \quad (2.14)$$

dengan $\hat{\beta}_m$ merupakan parameter pada regresi polinomial berderajat d dari y_k pada x_k , yang mana pencocokan dengan menggunakan bobot kuadrat terkecil dengan bobot $w_k(x_i)$ untuk titik (x_k, y_k) dan d merupakan parameter *degree* polinomial untuk pencocokan lokal disetiap titik pada *scatterplot*, jika yang dimasukkan $d = 1$ maka polinomial yang digunakan berbentuk persamaan linier, namun jika yang dimasukkan $d = 2$ polinomial yang digunakan berbentuk persamaan kuadrat. Pemulusan titik pada x_i dengan menggunakan *locally weighted regression* berderajat d adalah (x_i, \hat{y}_i) , dengan \hat{y}_i merupakan *fitted value* dari regresi pada x_i . Sehingga dapat ditulis persamaannya sebagai berikut:

$$\hat{y}_i = \sum_{m=0}^d \hat{\beta}_m(x_i) x_i^m = \sum_{k=1}^n r_k(x_i) y_k \quad (2.15)$$

dengan r_k tidak bergantung pada y_j . Misalkan $e_i = y_i - \hat{y}_i$ dan smedian dari $|e_i|$ dengan e_i merupakan *residuals* dari *fitted value*, sehingga didefinisikan *robustness weight* $\delta_k = B\left(\frac{e_k}{6s}\right)$. Perhitungan \hat{y}_i baru untuk setiap i diperoleh dari *fitting* polinomial berderajat d menggunakan bobot kuadrat terkecil dengan bobot $\delta_k w_k(x_i)$ pada titik (x_k, y_k) .

Proses pemulusan akan berhenti dengan beberapa iterasi t jika proses tersebut mendapatkan hasil yang maksimal. Jumlah iterasi yang diindikasikan memadai pada sebuah percobaan yaitu dengan dua iterasi untuk semua situasi.

Fungsi `loess()` adalah fungsi *loess* yang digunakan dalam implementasi *GAMLSS* didalam program R yang merupakan satu-satunya fungsi yang menyediakan pemulusan lebih dari satu variabel eksplanator. Implementasi fungsi

$\log(\cdot)$ pada GAMLSS sama dengan implementasi fungsi $\log(\cdot)$ pada GAM (Rigby & Stasinopoulos, 2007).

2.10 Algoritma *Mixed*

Algoritma *mixed* merupakan gabungan algoritma Rigby & Stasinopoulos (RS) dan algoritma Cole & Green (CG), dengan iterasi j dari algoritma RS kemudian diteruskan iterasi k dari algoritma CG. Untuk memulai proses analisis, terlebih dahulu menetapkan batasan maksimum iterasi untuk algoritma RS dan algoritma CG.

Misalkan $\gamma_{jk}^{r,i,m}$ merupakan nilai terbaru yang didapat dari γ_{jk} pada saat ke- r , ke- i dan algoritma backfitting ke- m dengan r merupakan indeks iterasi dari *outer cycle*, k merupakan parameter indeks, i merupakan indeks iterasi dari *inner cycle*, m indeks algoritma *backfitting*, dan j *random effect* (non parametrik). Dan misalkan $\gamma_{jk}^{r,i}$ menyatakan nilai dari γ_{jk} pada saat nilai *backfitting* konvergen saat ke- i dan ke- r dengan $j = 1, 2, \dots, j_k$ dan $k = 1, 2, \dots, p$.

Langkah-langkah pada Algoritma RS adalah sebagai berikut:

1. Mulai-> pemberian nilai awal *fitted value* $\theta_k^{(1,1)}$ dan *random effect* $\gamma_{jk}^{1,1,1}$ untuk $j = 1, 2, \dots, j_k$ dan $k = 1, 2, \dots, p$. Lalu dievaluasi nilai prediktor aditif $\eta_k^{1,1} = \mathbf{g}_k(\boldsymbol{\eta}_k^{1,1})$ untuk $k = 1, 2, \dots, p$.
2. Mulai dengan memasukkan nilai r (indeks *outer cycle*) dengan $r = 1, 2, \dots$ hingga konvergen untuk $k = 1, 2, \dots, p$.
 - a. Mulai pemberian nilai awal *inner cycle* $i = 1, 2, \dots$ hingga konvergen.
 - i. Evaluasi nilai terbaru $\mathbf{u}_k^{(r,i)}$, $\mathbf{W}_{kk}^{r,i}$ dan $\mathbf{Z}_k^{r,i}$,
 - ii. Mulai pemberian nilai awal algoritma *backfitting* dengan $m = 1, 2, \dots$ hingga konvergen
 - iii. Meregresi nilai residual terbaru secara partial dari $\boldsymbol{\varepsilon}_{0k}^{r,i,m} = \mathbf{Z}_k^{r,i} - \sum_{j=1}^{j_k} \mathbf{Z}_{jk} \gamma_{jk}^{r,i,m}$ yang berupa matriks \mathbf{X}_k , menggunakan iterasi berbobot $\mathbf{W}_{kk}^{r,i}$ untuk mendapatkan parameter estimasi terbaru $\boldsymbol{\beta}_k^{r,i,m+1}$

- iv. Untuk $j = 1, 2, \dots, j_k$ pemulusan parsial residual $\boldsymbol{\varepsilon}_{0k}^{r,i,m} = \mathbf{Z}_k^{r,i} - \mathbf{X}_k \boldsymbol{\beta}_k^{r,i,m+1} - \sum_{t=1, t \neq j}^{j_k} \mathbf{Z}_{tk} \boldsymbol{\gamma}_{tk}^{r,i,c}$ menggunakan *shrinking* (pemulusan) matriks \mathbf{S}_{jk} diberikan oleh persamaan $\mathbf{S}_{jk} = \mathbf{Z}_{jk}^T (\mathbf{W}_{kk} \mathbf{Z}_{jk} + \mathbf{G}_{jk})^{-1} \mathbf{Z}_{jk}^T \mathbf{W}_{kk}$ untuk mendapatkan prediktor aditif terbaru $\mathbf{Z}_{jk} \boldsymbol{\gamma}_{jk}^{r,i,m}$
- v. *Backfitting* berakhir ketika didapat nilai yang konvergen dari $\boldsymbol{\beta}_k^{r,i}$ dan $\mathbf{Z}_{jk} \boldsymbol{\gamma}_{jk}^{r,i}$ dengan $\boldsymbol{\beta}_k^{r,i+1} = \boldsymbol{\beta}_k^{r,i}$ dan $\boldsymbol{\gamma}_{jk}^{r,i+1} = \boldsymbol{\gamma}_{jk}^{r,i}$ untuk $j = 1, 2, \dots, j_k$. Jika tidak, *update m* dan kembali mengulang *backfitting*.
- vi. Kalkulasi $\boldsymbol{\gamma}_{jk}^{r,i+1}$ dan $\boldsymbol{\theta}_k^{r,i+1}$ terkini.

- b. *Inner cycle* berakhir dengan didapat $\boldsymbol{\beta}_k^r$ yang konvergen dan prediktor aditif $\mathbf{Z}_{jk} \boldsymbol{\gamma}_{jk}^r$ dengan $\boldsymbol{\beta}_k^{r+1,i} = \boldsymbol{\beta}_k^{r,i}$, $\boldsymbol{\gamma}_{jk}^{r+1,i} = \boldsymbol{\gamma}_{jk}^{r,i}$ dan $\boldsymbol{\theta}_{jk}^{r+1,i} = \boldsymbol{\theta}_k^r$ untuk $j = 1, 2, \dots, j_k$.

3. *Update* nilai k

4. *Outer cycle* berakhir jika didapatkan (*penalized*) likelihood yang cukup kecil. Jika tidak *update r* dan ulangi kembali ke *outer cycle*

langkah-langkah Algoritma CG adalah sebagai berikut:

1. Mulai -> pemberian nilai awal *fitted value* $\boldsymbol{\theta}_k^{(1,1)}$ dan *random effect* $\boldsymbol{\gamma}_{jk}^{1,1,1}$ untuk $j = 1, 2, \dots, j_k$ dan $k = 1, 2, \dots, p$. Lalu dievaluasi nilai prediktor aditif $\boldsymbol{\eta}_k^1 = \boldsymbol{\eta}_k^{1,1} = \mathbf{g}_k(\boldsymbol{\eta}_k^{1,1})$ untuk $k = 1, 2, \dots, p$.
2. Mulai dengan memasukkan nilai r (indeks *outer cycle*) dengan $r = 1, 2, \dots$
3. Melakukan evaluasi dan memperbaiki nilai terbaru dari $\mathbf{u}_k^{(r)}$, $\mathbf{W}_{ks}^{(r)}$, dan $\mathbf{Z}_k^{(r)}$ untuk $k = 1, 2, \dots, p$ dan $s = 1, 2, \dots, p$. Kemudian, lakukan langkah penunggalan ke- r dari algoritma Newton-Raphson sebagai berikut :
 - a. Mulai pemberian nilai awal *inner cycle* $i = 1, 2, \dots$ hingga konvergen, untuk $k = 1, 2, \dots, p$
 - i. Mulai pemberian nilai awal algoritma *backfitting* dengan $m = 1, 2, \dots$ hingga konvergen $X_k \boldsymbol{\beta}_k^{(r,i,m+1)} = H_k^{(r)} \boldsymbol{\varepsilon}_{0k}^{(r,i,m)}$ dan untuk $j = 1, 2, \dots, j_k$

$$Z_{jk} \gamma_{jk}^{(r,i,m+1)} = S_{jk}^{(r)} \varepsilon_{jk}^{(r,i,m)},$$

- ii. Selesai melakukan *backfitting*, dengan konvergen pada $\beta_k^{(r,i)}$ dan $z_{jk} \gamma_{jk}^{(r,i)}$ dan atur $\beta_k^{(r,i+1)} = \beta_k^{(r,i)}$ dan $\gamma_{jk}^{(r,i+1)} = \gamma_{jk}^{(r,i)}$ untuk $j = 1, 2, \dots, J_k$, setelah itu *update* dan lanjutkan *backfitting*, kemudian
- iii. Kalkulasi pembaruan $\eta_k^{(r,i+1)}$ dan $\theta_k^{(r,i+1)}$ kemudian *update* k ;

b. *Inner cycle* berakhir dengan didapat β_k^r yang konvergen dan prediktor aditif $Z_{jk} \gamma_{jk}^r$ dan atur $\beta_k^{r+1,i} = \beta_k^{r,i}$, $\gamma_{jk}^{(r+1,1)} = \gamma_{jk}^{(r,.)}$, $\eta_k^{(r+1)} = \eta_k^{(r,.)}$ dan $\theta_{jk}^{(r+1,1)} = \theta_k^{(r,.)}$ untuk $j = 1, 2, \dots, j_k$ dan $k = 1, 2, \dots, p$; kemudian *update* i (*inner cycle*) jika tidak konvergen.

- 4. *Outer cycle* berakhir jika didapatkan (*penalized*) likelihood yang cukup kecil. Jika tidak *update* r dan ulangi kembali ke *outer cycle*.

Matriks $\mathbf{H}_k^{(r)}$ dan $\mathbf{S}_k^{(r)}$ didefinisikan sebagai matriks proyeksi dan matriks penyusutan untuk parametrik dan komponen aditif dari setiap model pada iterasi ke- r untuk $j = 1, 2, \dots, j_k$ dan $k = 1, 2, \dots, p$. Sebagian sisa $\varepsilon_{0k}^{(r,i,m)}$ dan $\varepsilon_{jk}^{(r,i,m)}$ merupakan pekerjaan untuk *fitting* beberapa variabel parametrik dan komponen-komponen aditif (*random-effect* atau *smoothing*) dari model berturut-turut sebagai berikut :

$$\begin{aligned} \varepsilon_{0k}^{(r,i,m)} &= \mathbf{z}_k^{(r)} - \sum_{t=1}^{Jk} \mathbf{z}_{tk} \gamma_{tk}^{(r,i,c)} - \mathbf{W}_{kk}^{(r)-1} \sum_{s=1, s \neq k}^p \mathbf{W}_{ks}^{(r)} (\eta_s^{(r,c)} - \eta_s^{(r)}), \\ \varepsilon_{jk}^{(r,i,m)} &= \mathbf{z}_k^{(r)} - \mathbf{X}_k \beta_k^{(r,i,m+1)} l \\ &\quad - \sum_{t=1}^{Jk} \mathbf{z}_{tk} \gamma_{tk}^{(r,i,c)} - \mathbf{W}_{kk}^{(r)-1} \sum_{s=1, s \neq k}^p \mathbf{W}_{ks}^{(r)} (\eta_s^{(r,c)} - \eta_s^{(r)}) \end{aligned}$$

Panjang langkah dari algoritma Newton-Raphson dapat diubah dari ukuran α , dengan mengupdate prediktor linier seperti dibawah ini

$$\eta_k^{(r+1)}(\alpha) = \alpha \eta_k^{(r+1)} + (1 + \alpha) \eta_k^{(r)}$$

dengan $k = 1, 2, \dots, p$, *inner cycle* berakhir jika r *outer cycle* mengevaluasi $\mathbf{u}_k^{(r+1)}$, $\mathbf{W}_{ks}^{(r+1)}$ dan $z_k^{(r+1)}$ dengan $k = 1, 2, \dots, p$ dan $s = 1, 2, \dots, p$, dengan

menggunakan $\eta_k^{(r+1)}(\alpha)$ dengan $k = 1, 2, \dots, p$. Panjang langkah optimum untuk iterasi *rparticular* dapat diperoleh dari memaksimalkan $l_p(\alpha)$.

5. *Outer cycle* merupakan penyederhanaan dari algoritma Newton-Raphson, jika ukuran langkah optimasi dilakukan, maka *outer* akan menyimpulkan konvergen, karena standart umum dari algoritma Newton-Raphson menjamin kekonvergenan.

2.11 Pemilihan Model melalui AIC (*Akaike's Information Criterion*)

Pemilihan model dari sebuah data set yang terbaik salah satunya adalah dengan menggunakan metode AIC (*Akaike's Information Criterion*). Menurut (Tirta, 2009) besarnya AIC dihitung melalui rumus berikut :

$$AIC = -2l(\hat{\theta}) + 2q$$

dengan $l(\hat{\theta})$ adalah nilai likelihood dari model yang dihadapi dan q adalah banyaknya parameter dalam model. Pemilihan model terbaik dilihat dari nilai terkecil dari AIC.

BAB 3. METODE PENELITIAN

3.1 Data

Data dalam penelitian ini merupakan data sekunder yang diperoleh dari (Premono, 2013) yang berupa hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011 (lihat lampiran A). Pada data tersebut terdapat dua macam variabel yaitu variabel respon dan variabel eksplanatori yang tersaji pada Tabel 3.1, karakteristik respon pada data hasil Ujian Nasional ini bersifat kontinu dan non-negatif, sehingga distribusi yang akan digunakan adalah distribusi Gamma dan distribusi *Generalized* Gamma.

Tabel 3.1 Variabel prediktor(X) yang akan digunakan

Variabel	Jenis variabel
Ujian Nasional	Respon
Rapot	Eksplanatori
UjianSekolah	Eksplanatori
Try Out	Eksplanatori
IQ	Eksplanatori

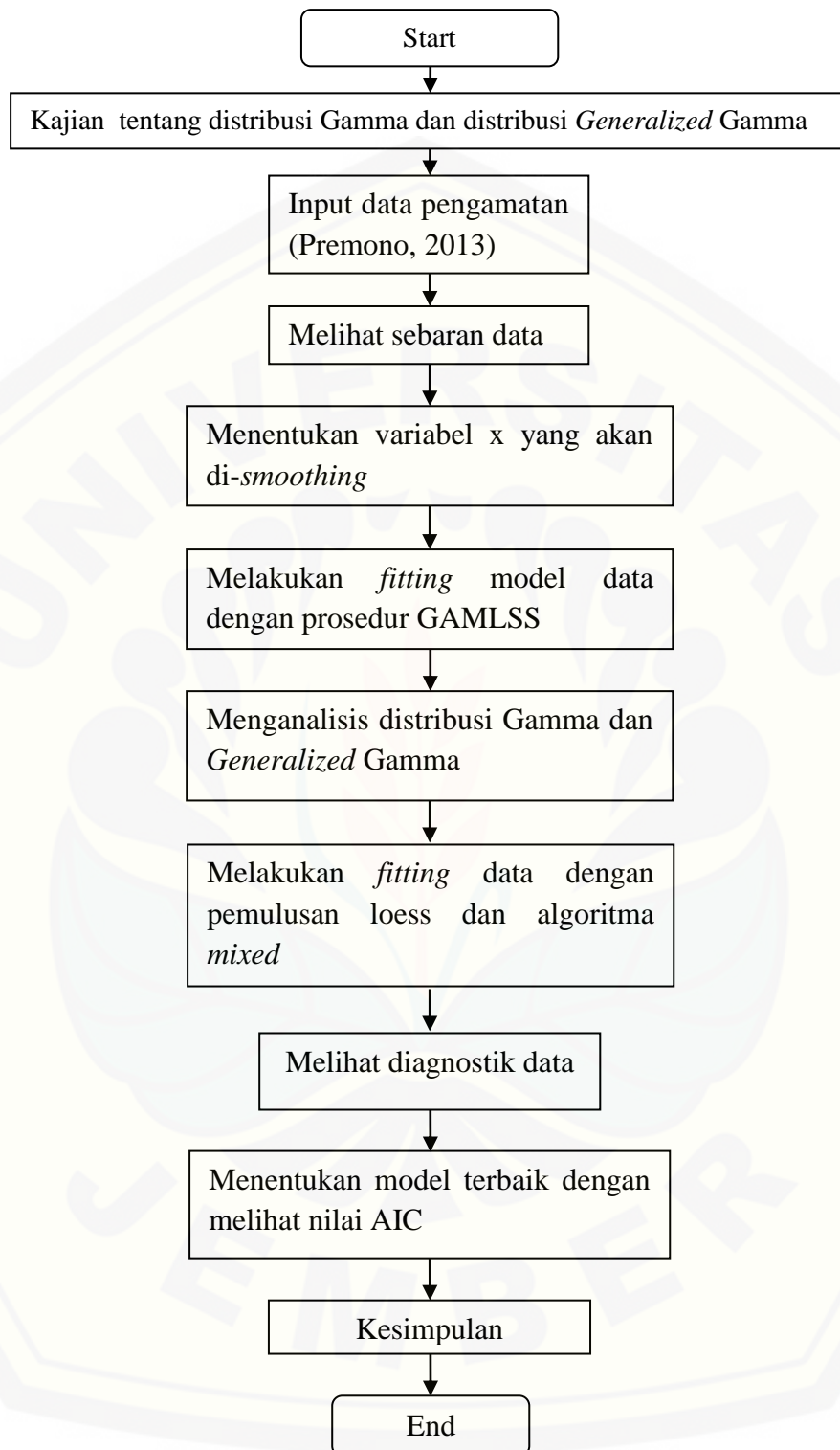
3.2 Metode Pengolahan dan Analisis Data

Dalam pengolahan data, program analisis data yang digunakan dalam menganalisis data yakni program R i386 3.1.2, dengan paket gamlss yang didapatkan dengan menginstal paket dari R-menu *packages* atau *install packages (s)*. Proses pertama yang dilakukan dalam menganalisis data yaitu diawali dengan menyusun data sampel. Data tersebut adalah data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011.

Berikut langkah-langkah yang dilakukan pada penelitian, serta analisis dengan menggunakan program R dengan paket gamlss:

1. Melakukan kajian pustaka tentang Gamma dan *Generalized* Gamma (membandingkan bentuk fungsi kepadatan peluang)
2. Menginputkan data pengamatan dalam R
3. Melihat sebaran data dengan menggunakan *scatterplot*
4. Menentukan variabel x yang akan di-*smoothing* dengan melihat hubungan kelinieran dari setiap variabel
5. Melakukan *fitting* model data menggunakan prosedur GAMLSS
 - a. Menganalisis distribusi Gamma dan *Generalized* Gamma
 - b. Melakukan *fitting* data dengan pemulusan loess dan algoritma *mixed*
6. Melihat diagnostik data
7. Menentukan model terbaik dengan melihat nilai AIC terkecil dari proses *fitting* data
8. Kesimpulan

Secara skematik, langkah-langkah yang dilakukan dalam penelitian tentang pemodelan *Generalized* gamma dengan pemulusan loess digambarkan dalam Gambar 3.2.



Gambar 3.2 Diagram Metode Penelitian

BAB 4. HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil kajian pustaka tentang Gamma dan *Generalized* Gamma, yaitu dengan membandingkan bentuk fungsi kepadatan peluang dari kedua distribusi dan model data dari hasil analisis menggunakan prosedur GAMLSS dengan pemulusan loess dan algoritma *mixed* pada data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011. Analisis data pada penelitian ini dilakukan dengan menggunakan program R 3.1.2.

Pada tahap analisis data ini dibagi beberapa tahap yaitu melihat sebaran data hasil Ujian Nasional dengan variabel hasil Ujian Nasional sebagai variabel respon, menentukan variabel eksplanatori yang akan di-*smoothing*, kemudian melakukan analisis data dengan menggunakan paket GAMLSS dengan distribusi Gamma dan *Generalized* Gamma yang dapat menggambarkan data, kemudian menganalisis data dengan pemulusan loess dengan menentukan *span* dan *degree* yang terbaik.

4.1 Perbandingan Bentuk Fungsi Kepadatan dari Gamma dan *Generalized* Gamma

Distribusi Gamma (GA) tergolong dalam distribusi keluarga Eksponensial yang kontinu dan hanya menggambarkan dua parameter yaitu parameter skala (σ) dan parameter lokasi (μ). Fungsi kepadatan peluang dari Gamma dinotasikan $GA(\sigma, \mu)$ dan didefinisikan seperti pada persamaan (2.11):

$$f_y(y|\sigma, \mu) = \frac{1}{(\sigma^2\mu)\sigma^2} \frac{y\sigma^2-1 e^{-\frac{y}{\sigma^2\mu}}}{\Gamma\left(\frac{1}{\sigma^2}\right)}, \quad y > 0, \mu > 0, \sigma > 0$$

Sedangkan distribusi *Generalized* Gamma merupakan perluasan dari distribusi Gamma dan tergolong dalam distribusi keluarga non Eksponensial yang dapat menggambarkan tiga parameter yaitu σ, μ dan v yaitu parameter skala (σ), merupakan parameter lokasi (μ), dan parameter ukuran (v). Fungsi kepadatan peluang dari distribusi *Generalized* Gamma di definisikan seperti persamaan pada (2.12) sebagai berikut:

$$f_y(y|\sigma, \mu, v) = \frac{v\theta^\theta z^\theta \exp\{-\theta z\}}{\Gamma(\theta)y}, y > 0, -\infty < v < \infty, \mu > 0, \sigma > 0$$

untuk

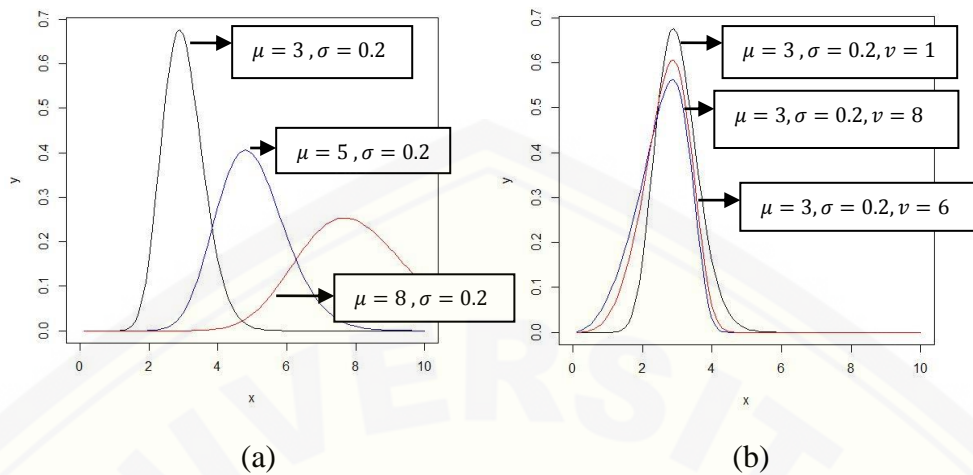
$$z = (y/\mu)^v \text{ dan}$$

$$\theta = \frac{1}{(\sigma^2 v^2)}$$

Jika parameter v pada (2.12) bernilai 1 maka didapatkan fungsi kepadatan peluang sebagai berikut:

$$\begin{aligned} f_y(y|\sigma, \mu, v) &= \frac{v\theta^\theta z^\theta \exp\{-\theta z\}}{\Gamma(\theta)y} \\ &= \frac{\left(\frac{1}{\sigma^2 v^2}\right)^{\frac{1}{\sigma^2 v^2}} \left(\frac{y}{\mu}\right)^{\frac{1}{\sigma^2 v^2}} \exp\left\{-\frac{1}{(\sigma^2 v^2)} \left(\frac{y}{\mu}\right)^1\right\}}{\Gamma\left(\frac{1}{\sigma^2 v^2}\right)y} \\ &= \frac{\left(\frac{1}{\sigma^2}\right)^{\frac{1}{\sigma^2}} \left(\frac{y}{\mu}\right)^{\frac{1}{\sigma^2}} \exp\left\{-\frac{1}{\sigma^2} \left(\frac{y}{\mu}\right)\right\}}{\Gamma\left(\frac{1}{\sigma^2}\right)y} \\ &= \frac{\left(\frac{1}{\sigma^2} \cdot \frac{y}{\mu}\right)^{\frac{1}{\sigma^2}} \exp\left\{-\frac{y}{\sigma^2 \mu}\right\}}{\Gamma\left(\frac{1}{\sigma^2}\right)y} \\ &= \frac{y^{\frac{1}{\sigma^2}} \cdot y^{-1} \exp\left\{-\frac{y}{\sigma^2 \mu}\right\}}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma\left(\frac{1}{\sigma^2}\right)} \\ &= \frac{1}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}}} \frac{y^{\frac{1}{\sigma^2}-1} e^{-\frac{y}{\sigma^2 \mu}}}{\Gamma\left(\frac{1}{\sigma^2}\right)} \end{aligned}$$

Melihat hasil uraian bentuk fungsi kepadatan peluang di atas identik dengan persamaan (2.11), sehingga dapat disimpulkan bahwa distribusi Gamma merupakan bentuk khusus dari distribusi *Generalized* Gamma dengan parameter v bernilai 1. Berikut ilustrasi Fungsi Kepadatan Peluang dari distribusi Gamma dan *Generalized* Gamma (skrip R tertera pada lampiran D1)

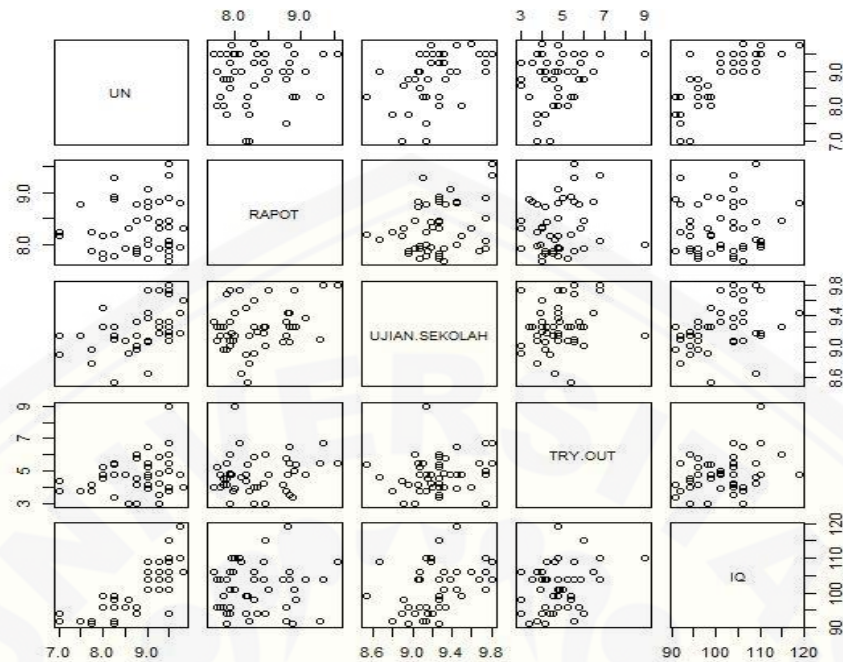


Gambar 4.1 Kurva Fungsi Kepadatan Peluang Distribusi (a) Gamma (b) *Generalized Gamma*

Ilustrasi pada Gambar (a) memperlihatkan ketika nilai μ berubah semakin besar, lokasi kurva tersebut semakin bergeser ke arah kanan. Untuk Gambar (b) saat ν bernilai 1, serta μ dan σ bernilai sama dengan nilai pada distribusi Gamma menunjukkan bahwa bentuk kurva kepadatan peluang kedua distribusi tersebut sama. Dan ketika semakin besar nilai ν , bentuk kurva menceng ke kiri yang ditandai dengan menjulurnya ekor kurva lebih ke arah kiri.

4.2 Sebaran Data

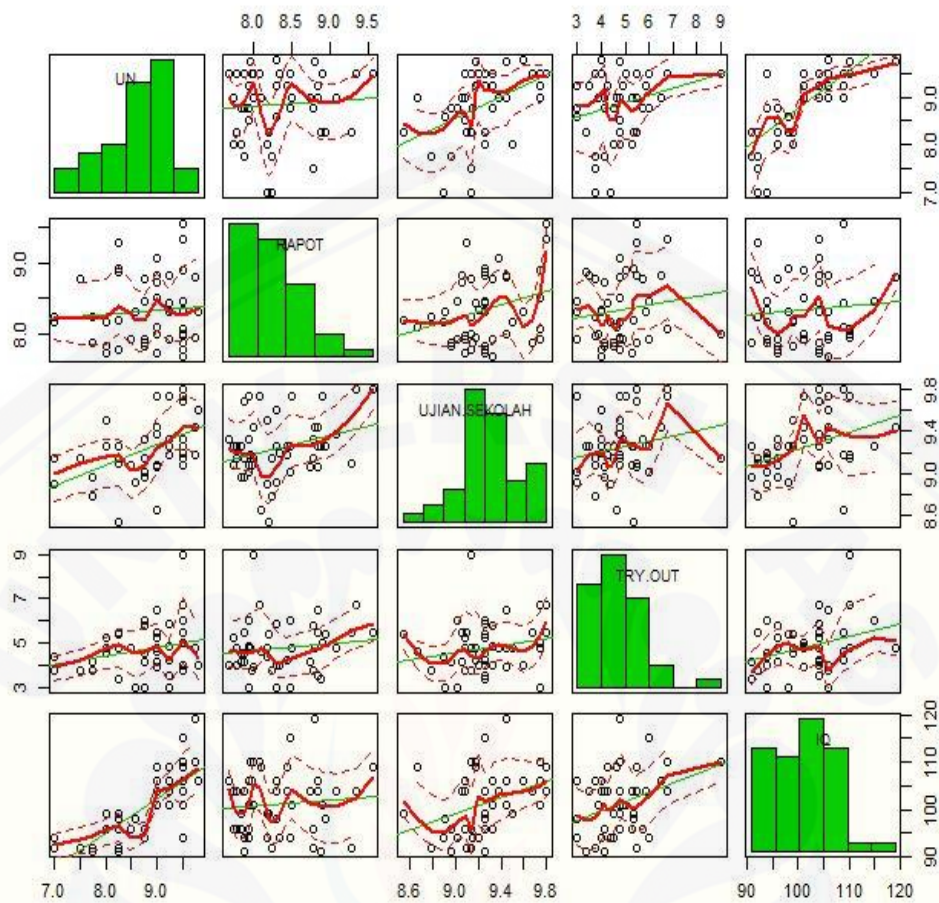
Langkah pertama setelah menginputkan data adalah melihat sebaran data dari data tersebut dengan menggunakan perintah `plot(data)`. Data yang digunakan merupakan data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011, dengan hasil Ujian Nasional sebagai variabel respon, sedangkan variabel eksplanatorinya yaitu rata-rata Raport, hasil Ujian Sekolah, rata-rata Try Out dan hasil tes IQ. Berikut grafik sebaran data hasil Ujian Nasional



Gambar 4.2 Sebaran Data Hasil Ujian Nasional

Gambar 4.2 menggambarkan sebaran data antara variabel respon (hasil Ujian Nasional) dengan masing-masing variabel eksplanatorinya. Baris pertama kolom kedua menunjukkan hubungan antara hasil Ujian Nasional dengan rata-rata Raport, baris pertama kolom ketiga dengan hasil Ujian Sekolah, baris pertama kolom keempat dengan rata-rata Try Out dan baris pertama kolom kelima dengan hasil tes IQ. Setelah melihat sebaran data, langkah selanjutnya yaitu melakukan eksplorasi data + *smoother* dengan menggunakan perintah *ScatterplotMatrix* pada R, yang mana hasil dari *ScatterplotMatrix* akan memberikan gambaran hubungan kelinieran antara variabel respon (hasil Ujian Nasional) dan masing-masing variabel eksplanatori. Sehingga dari gambaran tersebut dapat diketahui variabel mana yang dianggap tidak linier dan perlu diestimasi dengan menggunakan metode pemulusan loess. Berikut perintah *scatterplotMatrix* pada R dan visualisasi yang dihasilkan.

```
scatterplotMatrix(~UN+RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ, reg.lin
e=lm, diagonal='histogram', smooth=TRUE, data=UAN)
```

Gambar 4.3 Hubungan Kelinearan antara Variabel Respon dan Masing-masing Variabel Eksplanatori.

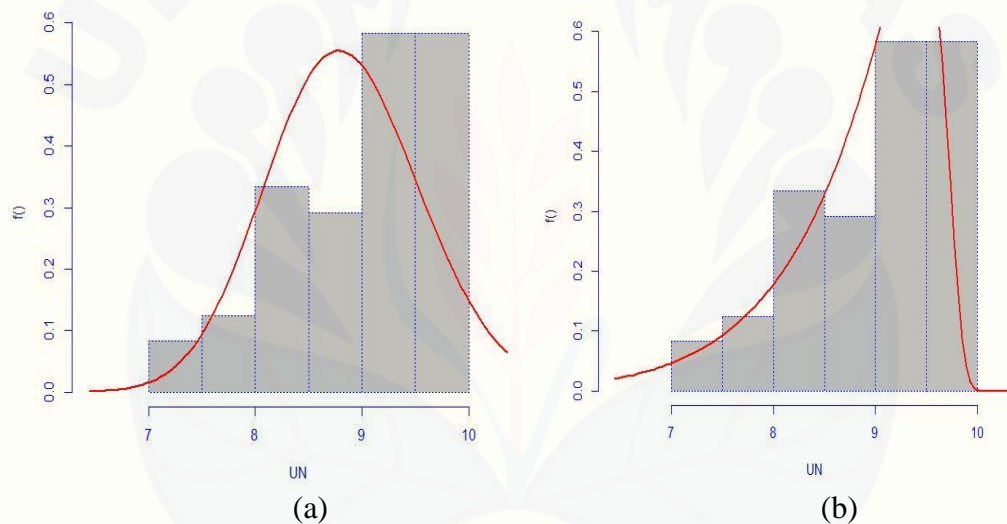
Pada Gambar 4.3 digambarkan hubungan kelinearan antara variabel respon (hasil Ujian Nasional) dengan masing-masing variabel eksplanatori (Raport,Ujian Sekolah,Try Out dan IQ). Dari gambar di atas hubungan antara hasil Ujian Nasional dengan hasil Ujian Sekolah, IQ dan rata-rata Try Out lebih cocok menggunakan pendekatan linier sehingga dimodelkan secara parametrik, sedangkan hubungan antara hasil Ujian Nasional dengan rata-rata Raport lebih cocok menggunakan pendekatan *non- linier* sehingga dapat disimpulkan bahwa rata-rata Raport dimodelkan secara *non- parametrik* dan diestimasi dengan metode pemulusan loess.

4.3 Fitting Model Data Menggunakan Prosedur GAMLSS dengan Pemulusan Loess dan Algoritma *mixed*

Untuk menentukan model terbaik akan dilihat nilai AIC yang terkecil. Langkah pertama dalam menentukan model terbaik yaitu dengan menentukan distribusi yang terbaik, kemudian memilih *span* dan *degree* polinomial yang cocok dalam melakukan pemulusan loess dan algoritma *mixed*.

4.3.1 Menentukan Distribusi Terbaik

Untuk menentukan distribusi yang akan digunakan untuk analisis data dengan prosedur GAMLSS, langkah awal yaitu melihat karakteristik dari setiap distribusi yang divisualisasikan dalam bentuk histogram (skrip R terdapat pada lampiran D3). Berikut visualisasi histogram dari setiap distribusi



Gambar 4.4 Histogram Distribusi (a) Gamma (b) *Generalized* Gamma

Berdasarkan histogram pada Gambar 4.4, karakteristik dalam menentukan distribusi terbaik yaitu dengan melihat sedikitnya celah yang ada pada setiap histogram. Dari histogram tersebut dapat disimpulkan bahwa *Generalized* Gamma merupakan distribusi yang cocok untuk data hasil Ujian Nasional. Namun untuk mendapatkan hasil yang lebih akurat dapat dilakukan dengan menganalisis distribusi, kemudian melihat nilai AIC terkecil dari hasil analisis masing-masing distribusi. Formula yang digunakan adalah sebagai berikut (skrip lebih lengkap dapat dilihat pada lampiran D3)

```
>gamlss1<-
gamlss (UN~RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ, data=UAN, family=GA,
method=mixed(20,20))
>gamlss2<-
gamlss (UN~RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ, data=UAN, family=GG,
method=mixed(20,20))
```

Hasil AIC dari kedua model di atas dapat dilihat pada Tabel 4.1 sebagai berikut:

Tabel 4.1 Nilai AIC Distribusi Model

Model	Distribusi	AIC
gamlss1	Gamma	73,33203
gamlss2	Generalized Gamma	71,09186

Dari grafik dan hasil AIC pada Tabel 4.1, distribusi yang lebih cocok untuk data hasil Ujian Nasional adalah distribusi *Generalized* Gamma.

4.3.2 *Fitting* data dengan pemulusan loess

Hal yang paling utama yang dilakukan proses pemulusan yaitu menentukan *degree* polinomial dan *span* yang terbaik. Sehingga dapat menghasilkan kurva yang dapat mewakili seluruh data serta nilai AIC yang terkecil. Dalam penelitian ini distribusi yang akan digunakan dalam proses *fitting* yaitu distribusi *Generalized* Gamma dan variabel eksplanatori yang akan diestimasi dengan pemulusan loess yaitu nilai rata-rata Raport.

a Memilih *degree* polinomial yang terbaik

Untuk memilih *degree* polinomial yang terbaik dilihat dari nilai AIC yang terkecil. Ada dua pilihan *degree* polinomial dalam pemulusan loess yaitu *degree*=1 dan *degree*=2. Formula yang digunakan adalah sebagai berikut :

```
>degree1<-
gamlss (UN~UJIAN.SEKOLAH+lo(~RAPOT, degree=1)+TRY.OUT+IQ, data=
UAN, family=GG, method=mixed(20,20))
refdegree1<-refit(degree1)
```

```
>degree2<-
gamlss (UN~UJIAN.SEKOLAH+lo (~RAPOT, degree=2) +TRY.OUT+IQ, data=
UAN, family=GG, method=mixed(20, 20))
```

Nilai AIC yang didapatkan dari kedua model di atas disajikan pada Tabel 4.2 (skrip dan output dari program di atas dapat dilihat pada lampiran D4) berikut ini :

Tabel 4.2 Nilai AIC Degree yang Terbaik

Model	Degree	AIC
degree1	1	71,84726
degree2	2	74,25299

Berdasarkan nilai AIC yang tersaji pada Tabel 4.2, bahwa pemilihan *degree* yang terbaik yaitu dengan melihat nilai AIC yang terkecil, dan *degree* tersebut adalah *degree* 1.

b. Memilih *span* yang terbaik

Parameter *span* digunakan untuk menentukan jumlah pemulusan yang dibutuhkan untuk proses *smoothing*. Nilai *span* di spesifikasi antara 0 sampai 1, yang mana nilai *span* tersebut mempengaruhi pemulusan terhadap kurva. Menurut (Cleveland, 1979) *span* yang terbaik ditentukan mulai *span* 0,2 sampai 0,8. Namun, karena *span* 0,2 menghasilkan *error*, sehingga *span* yang digunakan yaitu mulai 0,3 sampai 0,8. Berikut formula yang digunakan beserta nilai AIC yang tersaji pada Tabel 4.3 (skrip dan output program lebih lengkap dapat dilihat pada lampiran D5)

```
>sp3<-gamlss (UN~lo (~RAPOT, degree=1, span=.3) +UJIAN.SEKOLAH+
TRY.OUT+IQ, data=UAN, family=GG, method=mixed(20, 20))
```

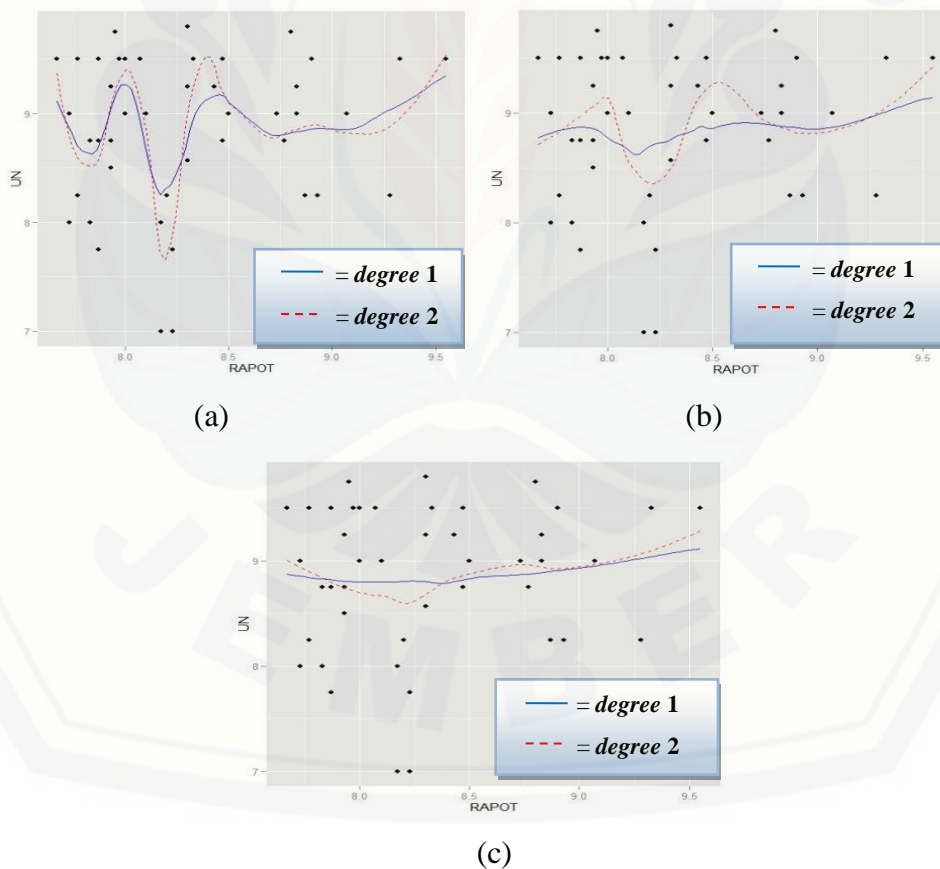
```
>sp6<-gamlss (UN~lo (~RAPOT, degree=1, span=.6) +UJIAN.SEKOLAH+
TRY.OUT+IQ, data=UAN, family=GG, method=mixed(20, 20))
```

```
>sp8<-gamlss (UN~lo (~RAPOT, degree=1, span=.8) +UJIAN.SEKOLAH+
TRY.OUT+IQ, data=UAN, family=GG, method=mixed(20, 20))
```


Tabel 4.3 Nilai AIC *Span* yang Terbaik

Model	<i>degree</i>	<i>span</i>	AIC
sp3	1	0,3	67.50023
	2	0,3	<i>error</i>
sp6	1	0,6	72.86555
	2	0,6	75.31762
sp8	1	0,8	71.90015
	2	0,8	74.11429

Dari Tabel 4.3 dapat dilihat model sp3 dengan $span=0,3$ dan $degree= 2$ menghasilkan *error* karena *global deviance* pada algoritma Cole&Green meningkat yang menyebabkan model tersebut tidak konvergen, sehingga dapat disimpulkan bahwa *span* dan *degree* yang terbaik dengan nilai AIC terkecil yaitu sp3 dengan $span = 0,3$ dan $degree = 1$. Berikut visualisasi data setelah dilakukan pemulusan beserta perbandingan dengan beberapa *span* dan *degree* terbaik:



Gambar 4.5 Hasil Visualisasi *Smoothing* dengan (a) *Span* 0,3 (b) *Span* 0,6
(c) *Span* 0,8

Berdasarkan hasil perbandingan dari ketiga gambar pada Gambar 4.5 bentuk kurva dengan nilai *span* 0,3 menunjukkan kurva tersebut *over smoothing* dan dapat mewakili data, sedangkan kurva dengan nilai *span* 0,6 dan 0,8 menunjukkan kurva tersebut *under smoothing* dan tidak dapat mewakili sebaran data. Namun untuk data ini hasil visualisasi *smoothing* terbaik yaitu kurva dengan *span* dan *degree* dengan nilai terkecil yaitu *span*=0,3 dan *degree*=1 (skrip R tertera pada lampiran D6).

4.4 Menentukan Model Terbaik dengan Menggunakan stepGAIC()

Dengan formula `stepGAIC()` akan menghasilkan beberapa *trial* untuk setiap model yang akan diuji. Berikut langkah-langkah dalam menentukan model terbaik dengan `stepGAIC()`.

- a. Memasukkan model yang akan diuji

```
>sp3<-
gamlss (UN~lo (~RAPOT, degree=1, span=.3)+UJIAN.SEKOLAH
+TRY.OUT+IQ, data=UAN, family=GG, method=mixed(20,20))
```

- b. Memasukkan fungsi `gamlss.scope()`

fungsi `gamlss.scope()` digunakan untuk mendefinisikan range yang akan diuji secara bertahap

```
>gs1<-
gamlss.scope(model.frame(UAN), smoother="lo", arg="span=.3")
```

- c. Mulai `stepGAIC()`

`stepGAIC()` merupakan fungsi yang mencakup beberapa *trial* setiap kemungkinan dari kombinasi tiap variabel yang akan diuji, dengan pemilihan model ditentukan dengan menggunakan *Generalized Akaike information criterion* (GAIC) (Stasinopoulos & Rigby, 2007). Pada fungsi ini ditambahkan dengan argumen `additive = TRUE`, untuk memungkinkan pada proses ini dapat dilakukan pemulusan

```
>step1<-stepGAIC(sp3,gs1,additive=TRUE)
```

Dari proses di atas didapatkan model terbaik adalah model berikut: $UN \sim UJIAN.SEKOLAH + IQ + lo(\sim RAPOT, degree=1, span=.3), data=UAN, family=GG, method=mixed(60, 60)$ dengan nilai AIC terkecil yang ditunjukkan pada trial ke-8. Adapun ringkasan *coefficients* dan *link function* dari beberapa parameter yang dihasilkan tersaji pada beberapa tabel dibawah ini:

Tabel 4.4 *Coefficients* Parameter μ (mu) dengan Log sebagai *Link function*

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1,182e+00	1,840e-01	6,428	1,94e-07
UJIAN.SEKOLAH	8,143e-02	1,914e-02	4,255	0,000144
IQ	4,828e-03	7,053e-07	6845,004	< 2e-16

Tabel 4.5 *Coefficients* Parameter σ (sigma) dengan Log sebagai *Link function*

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3,56466	0,09753	-36,55	< 2e-16

Tabel 4.6 *Coefficients* Parameter ν (nu) dengan Identity sebagai *Link function*

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	62,963	2,622	24,01	< 2e-16

dengan nilai AIC dan SBC sebagai berikut :

AIC : 66,95514
SBC : 89,96599

4.5 Menentukan Nilai Estimasi Parameter dari Model Terbaik

Dari subbab sebelumnya telah didapatkan model yang terbaik untuk data Hasil Ujian Nasional dengan paket GAMLSS, dengan menggunakan pemulusan loess dan algoritma *mixed* adalah:

UN~UJIAN.SEKOLAH+IQ+1o (~RAPOT, degree=1, span=.3), data=UAN, family=GG, method=mixed(60,60)) dan summary dari model diatas untuk nilai estimasi dan *p-value* dari setiap komponen disajikan pada Tabel 4.7

Tabel 4.7 Nilai Estimasi, *p-value*

	Estimasi	<i>p-value</i>
<i>Intercept</i>	1,182e+00	1,94e-07
UJIAN.SEKOLAH	8,143e-02	0,000144
IQ	4,828e-03	< 2e-16

Berdasarkan Tabel 4.7 dapat dilihat nilai *p-value* dari *intercept*, hasil Ujian Sekolah dan hasil Tes IQ menunjukkan kurang dari 0,05, sehingga dapat dikatakan bahwa kedua variabel tersebut berpengaruh secara signifikan terhadap data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011. Jadi hasil Ujian Nasional dipengaruhi secara signifikan oleh hasil Ujian Sekolah, hasil Tes IQ dan *Intercept* yang merupakan konstanta dari hasil Ujian Nasional tanpa dipengaruhi oleh variabel apapun.

Estimasi dari hasil *summary* tersebut didapatkan hasil sebagai berikut :

$$\log(\hat{\mu}_{\text{Ujian Nasional}}) = \hat{\beta}_0 + \hat{\beta}_1 \text{Ujian Sekolah} + \hat{\beta}_2 \text{IQ}$$

dengan

$$\begin{aligned}\hat{\beta}_0 &= 1,182 \\ \hat{\beta}_1 &= 8,143 \times 10^{-2} \\ \hat{\beta}_2 &= 4,828 \times 10^{-3}\end{aligned}$$

Sehingga untuk parameter lokasi (μ) mempunyai nilai estimasi :

$$\hat{\mu}_{\text{Ujian Nasional}} = \exp(1,182 + 8,143 \times 10^{-2} \text{Ujian Sekolah} + 4,828 \times 10^{-3} \text{IQ})$$

Untuk parameter *scale* (σ) diperoleh :

$$\log(\hat{\sigma}_{\text{Ujian Nasional}}) = -3,56466$$

$$\hat{\sigma}_{\text{Ujian Nasional}} = \exp(-3,56466)$$

Dan untuk parameter *shape* (v) diperoleh :

$$\hat{v}_{\text{Ujian Nasional}} = 62,963$$

BAB. 5 PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah dan hasil pembahasan pada Bab 4, aplikasi GAMLSS dalam pemodelan *Generalized Gamma* dengan algoritma *mixed* pada pemulusan loess dapat diambil kesimpulan sebagai berikut:

1. telah terverifikasi bahwa distribusi Gamma merupakan bentuk khusus dari distribusi *Generalized Gamma* dengan parameter ν bernilai 1
2. distribusi yang lebih baik dari dua distribusi untuk analisis data Hasil Ujian Nasional SMA Negeri 1 Situbondo adalah *Generalized Gamma*
3. model terbaik dari analisis *Generalized Gamma* dengan algoritma *mixed* pada pemulusan loess dengan *degree* 1 dan *span* 0,3 untuk data hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun Ajaran 2010/2011 adalah :

a. $\hat{\mu}_{\text{Ujian Nasional}} = \exp(1,182 + 8,143 \times 10^{-2} \text{ Ujian Sekolah} + 4,828 \times 10^{-3} \text{ IQ})$

- b. Untuk parameter *scale* (σ) dan *shape* (ν) diperoleh :

$$\hat{\sigma}_{\text{Ujian Nasional}} = \exp(-3,56466)$$

$$\hat{\nu}_{\text{Ujian Nasional}} = 62,963$$

5.2 Saran

Penelitian ini masih terbuka peluang untuk mengaplikasikan metode GAMLSS dengan data lain menggunakan distribusi-distribusi kontinu lainnya yang memiliki 4 parameter seperti *Box-Cox Power Exponential* dan *Box-Cox-t*.

DAFTAR SUMBER RUJUKAN

- Brown, S.2011.Measures of Shape: Skewness and Kurtosis
<http://www.tc3.edu/instruct/sbrown/stat/shape.htm> [22 November 2011].
- Cleveland, W.S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of The American Statistical Association*.**74**:829-836.
- Jacoby, W.G. 2000. Loess : a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*.**19**:577-613.
- Premono, I.S.K. 2013. “Model Aditif Terampat dengan Pendekatan Spline Studi Kasus Menduga Nilai Ujian Nasional.” Tidak Diterbitkan.Tesis. Jember: Program Pascasarjana Universitas Jember.
- Rigby,R.A& Stasinopoulos D.M.2005 Generalized Additive Models for Location, and Shape. *Applied Statistic*. **54**.507-554.
- Stasinopoulos, D. M. & Rigby, R.A. 2007.Generalized Additive Models for Location, Scale and Shape (GAMLSS) in R. *J.R.Statisc.* [**23**]:7
- Stasinopoulos, D. M. & Rigby, R.A. 2010.*A flexible regression approach using GAMLSS in R*. University of Athena.
- Stasinopoulos, D. M. & Rigby, R.A. 2014.Package ‘gamlss.dist’.<http://cran.r-project.org/web/packages/gamlss.dist/gamlss.dist.pdf>
- Tirta, I.M.2009.Analisis Regresi dengan R (ANRER).Jember:UPT Penerbitan Universitas Jember.
- Wandi, W. F.2012.”Aplikasi GAMLSS dengan Pemulusan Cubic-Spline dan Algoritma Rigby & Stasinopoulus pada Data Agroklimatologi. “Tidak Diterbitkan.Skripsi.Jember:Jurusan Matematika FMIPA Universitas Jember.

**Lampiran A. Data Hasil Ujian Nasional SMA Negeri 1 Situbondo Tahun
Ajaran 2010/2011.**

UN	RAPOT	UJIAN SEKOLAH	TRY OUT	IQ
9,5	9,33	9,8	6,75	104
9	7,73	9,08	5,25	104
9,8	8,3	9,6	4	106
8	7,83	9,26	4,5	99
9,5	7,87	9,68	5,5	104
9,25	8,43	9,18	4,25	109
9,25	7,93	9,74	3	106
9,25	8,83	9,26	3,5	104
8,57	8,3	8,91	3	98
9	8,	9,44	3,875	101
9,5	8,9	9,74	5	101
9,75	8,8	9,44	4,75	119
9,25	8,3	9,26	5,875	104
8,25	8,93	9,26	5,375	98
9,5	8,07	9,74	6,75	110
8	8,17	9,5	4,75	99
9	8,5	9,74	4,75	101
7,75	7,87	8,96	4,125	91
7,5	8,77	9,14	3,75	92
9,5	7,67	9,32	4	106
7	8,23	9,14	3,75	92
9	8,73	9,06	4,125	104
9	9,07	9,38	4,75	104
8,75	8,47	9,02	3	94
9,5	8,33	9,2	4	94
9,5	8,47	9,26	6	115
9,75	7,95	9,18	4,75	110
7,75	8,23	8,79	3,75	92
9,5	8	9,14	9	110
8,75	7,93	9,02	6	94
8,5	7,93	9,08	4,75	96
8,25	7,77	9,14	4,75	96
8,75	7,87	9,15	4,5	94
8	7,73	9,26	5,25	96
8,75	8,77	9,33	5,75	96
9,5	7,77	9,26	4	104
9,5	7,97	9,08	3,75	106
9,25	7,93	9,32	4,875	101
9,25	8,43	9,26	4,25	104

9,5	9,55	9,8	5,5	109
8,25	8,2	8,54	5,375	99
8,25	8,87	9,26	3,375	91
8,25	9,28	9,1	5,5	92
9	8,83	9,06	5,5	104
8,75	7,83	8,96	4,125	96
9	8,83	9,44	6,5	106
7	8,17	8,9	4,375	94
9	8,1	8,66	4,625	109



Lampiran B. Pencocokan *Generalized Additive Models for Location Scale & Shape* (GAMLSS) Secara Umum

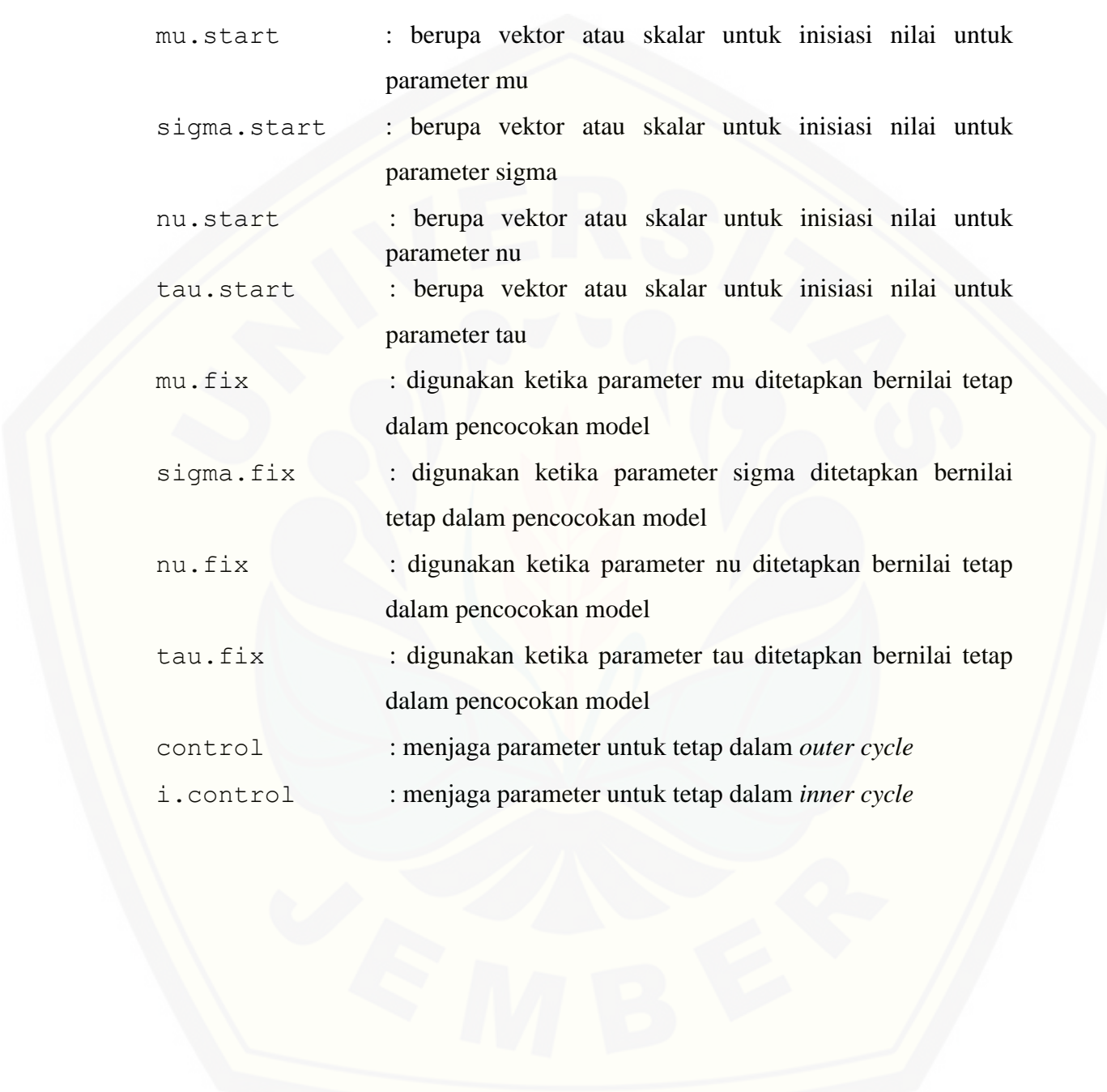
Formula berikut digunakan untuk pencocokan model objek menggunakan prosedur GAMLSS

Formula

```
gamlss(formula = formula(data), sigma.formula = ~1,
       nu.formula = ~1, tau.formula = ~1, family = NO(),
       data = sys.parent(), weights = NULL,
       contrasts = NULL, method = RS(), start.from = NULL,
       mu.start = NULL, sigma.start = NULL,
       nu.start = NULL, tau.start = NULL,
       mu.fix = FALSE, sigma.fix = FALSE, nu.fix = FALSE,
       tau.fix = FALSE, control = gamlss.control(...),
       i.control = glim.control(...), ...)
```

Keterangan

Formula	: rumus dari model. Contoh penulisan $y \sim x$
Mu.formula	: rumus model untuk parameter mu (lokasi)
sigma.formula	: rumus model untuk parameter sigma (skala)
nu.formula	: rumus model untuk parameter nu (ukuran) /skewness
tau.formula	: rumus model untuk parameter tau (ukuran) /kurtosis
family	:distribusi yang masuk ke dalam gamlss family
method	: algoritma yang digunakan antara lain algoritma Rigby & Stasinopoulus ditulis RS() algoritma Cole & Green ditulis CG() atau algoritma mixed ditulis mixed ()
data	: frame data yang berisi variabel yang terjadi dalam formula
weights	: suatu vektor dari bobot (weights) yang digunakan dalam proses pencocokan
contrast	: daftar kontras yang digunakan untuk memunculkan variabel yang hilang dalam formula



<code>start.from</code>	: pencocokan model GAMLSS dengan nilai pencocokan akan digunakan untuk memulai nilai awal pada pencocokan model tersebut
<code>mu.start</code>	: berupa vektor atau skalar untuk inisiasi nilai untuk parameter mu
<code>sigma.start</code>	: berupa vektor atau skalar untuk inisiasi nilai untuk parameter sigma
<code>nu.start</code>	: berupa vektor atau skalar untuk inisiasi nilai untuk parameter nu
<code>tau.start</code>	: berupa vektor atau skalar untuk inisiasi nilai untuk parameter tau
<code>mu.fix</code>	: digunakan ketika parameter mu ditetapkan bernilai tetap dalam pencocokan model
<code>sigma.fix</code>	: digunakan ketika parameter sigma ditetapkan bernilai tetap dalam pencocokan model
<code>nu.fix</code>	: digunakan ketika parameter nu ditetapkan bernilai tetap dalam pencocokan model
<code>tau.fix</code>	: digunakan ketika parameter tau ditetapkan bernilai tetap dalam pencocokan model
<code>control</code>	: menjaga parameter untuk tetap dalam <i>outer cycle</i>
<code>i.control</code>	: menjaga parameter untuk tetap dalam <i>inner cycle</i>

Lampiran C. Tabel Distribusi Kontinu yang Dapat Diimplementasikan dalam Paket GAMLSS

Berikut merupakan daftar distribusi kontinu yang dapat diimplementasikan ke dalam paket gamlss. R Name merupakan penamaan distribusi tersebut dalam program R.

Distribusi	R Name	μ	σ	ν	τ	Rentang
beta	BE ()	Logit	Logit	-	-	$y = (0,1),$ $\mu > 0, \sigma > 0$
Beta inflated(at 0 and 1)	BEINF ()	Logit	Logit	Log	Log	$y = (0,1),$ $\mu > 0, \sigma > 0,$ $v = \frac{p0}{p2}, \tau = \frac{p1}{p2}.$ $p2$ $= 1 - p0 - p1$
Box-Cox Cole and Green	BCCG()	identity	Log	Identity	-	$y > 0, \mu > 0,$ $\sigma > 0,$ $v = (-\infty, +\infty)$
Box-Cox Power exponential	BCPE()	identity	Log	Identity	Log	$y > 0, \mu > 0,$ $\sigma > 0,$ $v = (-\infty, +\infty),$ $\tau > 0$
Box-Cox-t	BCT()	identity	Log	Identity	Log	$y > 0, \mu > 0,$ $\sigma > 0,$ $v = (-\infty, +\infty),$ $\tau > 0$
Exponensial	EXP()	Log	-	-	-	$y > 0, \mu > 0$
Exponensial Gaussian	exGAUS()	identity	Log	Log	-	$-\infty < y < \infty,$ $-\infty < \mu < \infty,$ $\sigma > 0, v > 0$
Exponensial gen. Beta type 2	EGB2()	identity	Identity	Log	Log	$-\infty < y < \infty,$ $-\infty < \mu < \infty,$ $-\infty < \sigma < \infty,$ $v > 0, \tau > 0$
Gamma	GA()	Log	Log	-	-	$y > 0, \mu > 0,$ $\sigma > 0$
Generalized beta type1	GB1()	Logit	Logit	Log	Log	$0 < y < 1,$ $\mu > 0, \sigma > 0$
Generalized beta type 2	GB2()	Log	Identity	Log	Log	$y > 0, \mu > 0,$ $-\infty < \sigma < \infty,$ $v > 0, \tau > 0$
Generalized gamma	GG()	Log	Log	Identity	-	$y > 0, \mu > 0,$ $\sigma > 0,$ $-\infty < v < +\infty,$
Generalized inverse gaussian	GIG()	Log	Log	Identity	-	$y > 0, \mu > 0,$ $\sigma > 0,$ $-\infty < v < +\infty,$

Generalized t	GT()	Identity	Log	Log	Log	$-\infty < y < \infty,$ $\mu = (-\infty, +\infty),$ $\sigma > 0, v > 0,$ $\tau > 0$
Gumbel	GU()	Identity	Log	-	-	$y = (-\infty, \infty),$ $\mu = (-\infty, +\infty),$ $\sigma > 0$
Inverse gaussian	IG()	Log	log	-	-	$y > 0, \mu > 0,$ $\sigma > 0$
Johnson's original SU	JSUo()	Identity	log	Identity	Log	$-\infty < y < \infty,$ $\mu = (-\infty, +\infty),$ $\sigma > 0,$ $v = (-\infty, +\infty),$ $\tau > 0$
Logistic	LO()	Identity	Log	-	-	$y = (-\infty, \infty),$ $\mu = (-\infty, \infty),$ $\sigma > 0$
Log normal	LOGNO()	Log	Log	-	-	$y > 0,$ $\mu = (-\infty, +\infty),$ $\sigma > 0$
Log normal(Box-Cox)	LNO()	Log	Log	Fixed	-	$y > 0,$ $\mu > 0, \sigma > 0,$ $v = (-\infty, +\infty)$

Lampiran D. PROGRAM

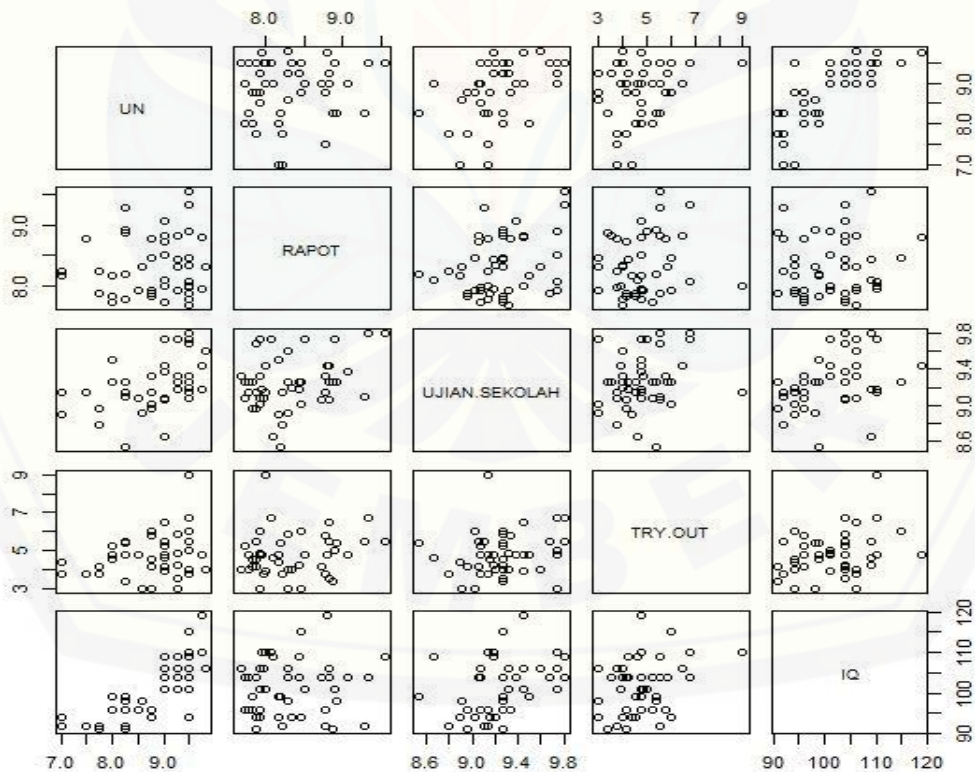
Berikut program analisis data Ujian Nasional dengan menggunakan formula GAMLSS dengan metode pemulusan loess dan algoritma *mixed*

D1. ilustrasi Fungsi Kepadatan Peluang dari distribusi Gamma dan Generalized Gamma

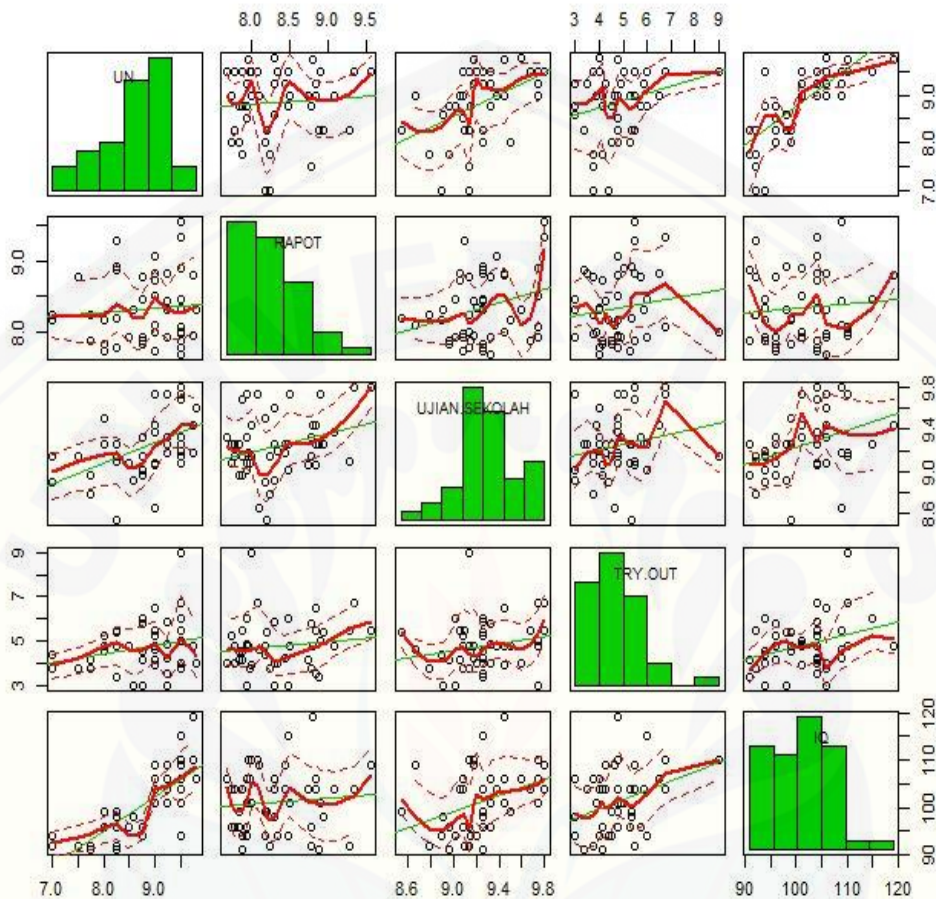
```
>x<-seq(0,10,0.1)
>y<-dGG(x,3,.2,1)
>plot(x,y,type='l')
>y<-dGG(x,3,.2,8)
>lines(x,y,col="blue")
>y<-dGG(x,3,.2,6)
>lines(x,y,col="red")
>y<-dGA(x,3,.2)
>plot(x,y,type='l')
>y<-dGA(x,5,.2)
>lines(x,y,col="blue")
>y<-dGA(x,8,.2)
>lines(x,y,col="red")
```

D2. Sebaran Data

```
>data(UAN)
>plot(UAN)
```



```
>library(car)
>scatterplotMatrix(~UN+RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ,reg.line=lm,diagonal='histogram',smooth=TRUE,data=UAN)
```



D3. Analisis Distribusi Terbaik

```
#melihat distribusi terbaik dengan histogram
```

```
>histDist(UN,family=GA,data=UAN)
```

```
>histDist(UN,family=GG,data=UAN)
```

```
#menganalisis dengan GAMLSS
```

```
>gamlss1<-
```

```
gamlss(UN~RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ,data=UAN,family=GA,
method=mixed(20,20))
```

```
GAMLSS-RS iteration 1: Global Deviance = 61.332
```

```
GAMLSS-RS iteration 2: Global Deviance = 61.332
```

```
GAMLSS-CG iteration 1: Global Deviance = 61.332
```

```
>gamlss2<-
```

```
gamlss(UN~RAPOT+UJIAN.SEKOLAH+TRY.OUT+IQ,data=UAN,family=GG,
method=mixed(20,20))
```

```
GAMLSS-RS iteration 1: Global Deviance = 60.262
```

```
GAMLSS-RS iteration 2: Global Deviance = 58.8692
```

```
GAMLSS-RS iteration 3: Global Deviance = 58.1089
```

```
GAMLSS-RS iteration 4: Global Deviance = 57.6872
```

```

GAMLSS-RS iteration 5: Global Deviance = 57.4466
GAMLSS-RS iteration 6: Global Deviance = 57.3023
GAMLSS-RS iteration 7: Global Deviance = 57.2179
GAMLSS-RS iteration 8: Global Deviance = 57.1679
GAMLSS-RS iteration 9: Global Deviance = 57.138
GAMLSS-RS iteration 10: Global Deviance = 57.1199
GAMLSS-RS iteration 11: Global Deviance = 57.1101
GAMLSS-RS iteration 12: Global Deviance = 57.1037
GAMLSS-RS iteration 13: Global Deviance = 57.0995
GAMLSS-RS iteration 14: Global Deviance = 57.0967
GAMLSS-RS iteration 15: Global Deviance = 57.095
GAMLSS-RS iteration 16: Global Deviance = 57.0933
GAMLSS-RS iteration 17: Global Deviance = 57.0924
GAMLSS-CG iteration 1: Global Deviance = 57.0919

> GAIC(gamlss1,gamlss2)
      df      AIC
gamlss2  7 71.09186
gamlss1  6 73.33203

> summary(gamlss2)
*****
*****
Family:  c("GG", "generalised Gamma Lopatatsidis-Green")

Call:  gamlss(formula = UN ~ RAPOT + UJIAN.SEKOLAH + TRY.OUT
+ IQ, family = GG,
      data = UAN, method = mixed(20, 20))

Fitting method: mixed(20, 20)

-----
Mu link function:  log
Mu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.006061  0.233040  4.317 9.76e-05 ***
RAPOT        -0.014181  0.015584 -0.910 0.368144
UJIAN.SEKOLAH 0.076363  0.026138  2.922 0.005641 **
TRY.OUT      -0.004178  0.005839 -0.716 0.478352
IQ           0.006142  0.001533  4.006 0.000254 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
1

-----
Sigma link function:  log
Sigma Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3.104      0.127 -24.43 <2e-16 ***
---

```


Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 Nu link function: identity

Nu Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	17.780	9.225	1.927	0.0609 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 No. of observations in the fit: 48
 Degrees of Freedom for the fit: 7
 Residual Deg. of Freedom: 41
 at cycle: 1

Global Deviance: 57.09186
 AIC: 71.09186
 SBC: 84.19026

D4. Analisis Derajat Polinomial (*degree*) Terbaik

```
>degree1<-
```

```
gamlss(UN~UJIAN.SEKOLAH+lo(~RAPOT, degree=1)+TRY.OUT+IQ, data=
UAN, family=GG, method=mixed(20,20))
```

```
GAMLSS-RS iteration 1: Global Deviance = 60.629
GAMLSS-RS iteration 2: Global Deviance = 59.3467
GAMLSS-RS iteration 3: Global Deviance = 58.3718
GAMLSS-RS iteration 4: Global Deviance = 57.6374
GAMLSS-RS iteration 5: Global Deviance = 57.0859
GAMLSS-RS iteration 6: Global Deviance = 56.6688
GAMLSS-RS iteration 7: Global Deviance = 56.3382
GAMLSS-RS iteration 8: Global Deviance = 56.0902
GAMLSS-RS iteration 9: Global Deviance = 55.8986
GAMLSS-RS iteration 10: Global Deviance = 55.751
GAMLSS-RS iteration 11: Global Deviance = 55.6359
GAMLSS-RS iteration 12: Global Deviance = 55.5454
GAMLSS-RS iteration 13: Global Deviance = 55.4735
GAMLSS-RS iteration 14: Global Deviance = 55.4201
GAMLSS-RS iteration 15: Global Deviance = 55.3758
GAMLSS-RS iteration 16: Global Deviance = 55.3393
GAMLSS-RS iteration 17: Global Deviance = 55.309
GAMLSS-RS iteration 18: Global Deviance = 55.2837
GAMLSS-RS iteration 19: Global Deviance = 55.2624
GAMLSS-RS iteration 20: Global Deviance = 55.2445
GAMLSS-CG iteration 1: Global Deviance = 55.206
GAMLSS-CG iteration 2: Global Deviance = 55.1935
```



```
GAMLSS-CG iteration 3: Global Deviance = 55.1825
GAMLSS-CG iteration 4: Global Deviance = 55.1713
GAMLSS-CG iteration 5: Global Deviance = 55.1648
GAMLSS-CG iteration 6: Global Deviance = 55.1574
GAMLSS-CG iteration 7: Global Deviance = 55.1544
GAMLSS-CG iteration 8: Global Deviance = 55.1456
GAMLSS-CG iteration 9: Global Deviance = 55.1456
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> redegree1<-refit(degree1)
GAMLSS-RS iteration 10: Global Deviance = 55.1368
GAMLSS-RS iteration 11: Global Deviance = 55.1398
GAMLSS-RS iteration 12: Global Deviance = 55.1395
GAMLSS-CG iteration 10: Global Deviance = 55.1366
GAMLSS-CG iteration 11: Global Deviance = 55.1352
GAMLSS-CG iteration 12: Global Deviance = 55.1353

>degree2<-
gamlss(UN~UJIAN.SEKOLAH+lo(~RAPOT, degree=2)+TRY.OUT+IQ, data=
UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 58.2488
GAMLSS-RS iteration 2: Global Deviance = 57.6493
GAMLSS-RS iteration 3: Global Deviance = 57.1398
GAMLSS-RS iteration 4: Global Deviance = 56.6994
GAMLSS-RS iteration 5: Global Deviance = 56.3174
GAMLSS-RS iteration 6: Global Deviance = 55.9857
GAMLSS-RS iteration 7: Global Deviance = 55.6984
GAMLSS-RS iteration 8: Global Deviance = 55.4508
GAMLSS-RS iteration 9: Global Deviance = 55.2383
GAMLSS-RS iteration 10: Global Deviance = 55.0569
GAMLSS-RS iteration 11: Global Deviance = 54.9028
GAMLSS-RS iteration 12: Global Deviance = 54.7723
GAMLSS-RS iteration 13: Global Deviance = 54.6621
GAMLSS-RS iteration 14: Global Deviance = 54.5694
GAMLSS-RS iteration 15: Global Deviance = 54.4916
GAMLSS-RS iteration 16: Global Deviance = 54.4264
GAMLSS-RS iteration 17: Global Deviance = 54.3719
GAMLSS-RS iteration 18: Global Deviance = 54.3265
GAMLSS-RS iteration 19: Global Deviance = 54.2899
GAMLSS-RS iteration 20: Global Deviance = 54.2592
GAMLSS-CG iteration 1: Global Deviance = 54.2998
GAMLSS-CG iteration 2: Global Deviance = 54.2124
GAMLSS-CG iteration 3: Global Deviance = 54.1816
GAMLSS-CG iteration 4: Global Deviance = 54.1813
Warning messages:
```

```

1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> redegree2<-refit(degree2)
GAMLSS-RS iteration 5: Global Deviance = 54.1661
GAMLSS-RS iteration 6: Global Deviance = 54.1564
GAMLSS-RS iteration 7: Global Deviance = 54.1497
GAMLSS-RS iteration 8: Global Deviance = 54.1476
GAMLSS-RS iteration 9: Global Deviance = 54.1466
GAMLSS-RS iteration 10: Global Deviance = 54.1472
GAMLSS-CG iteration 5: Global Deviance = 54.1484
GAMLSS-CG iteration 6: Global Deviance = 54.1481

> GAIC(redegree1, redegree2)
      df      AIC
redegree1  8.355979 71.84726
redegree2 10.052445 74.25299

```

D5. Analisis Span Terbaik

```

>sp3<-
gamlss(UN~lo(~RAPOT, degree=1, span=.3)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 42.9564
GAMLSS-RS iteration 2: Global Deviance = 42.2644
GAMLSS-RS iteration 3: Global Deviance = 41.822
GAMLSS-RS iteration 4: Global Deviance = 41.5351
GAMLSS-RS iteration 5: Global Deviance = 41.3539
GAMLSS-RS iteration 6: Global Deviance = 41.2438
GAMLSS-RS iteration 7: Global Deviance = 41.1812
GAMLSS-RS iteration 8: Global Deviance = 41.1504
GAMLSS-RS iteration 9: Global Deviance = 41.1406
GAMLSS-RS iteration 10: Global Deviance = 41.1447
GAMLSS-RS iteration 11: Global Deviance = 41.1579
GAMLSS-RS iteration 12: Global Deviance = 41.1771
GAMLSS-RS iteration 13: Global Deviance = 41.2013
GAMLSS-RS iteration 14: Global Deviance = 41.2264
GAMLSS-RS iteration 15: Global Deviance = 41.2526
GAMLSS-RS iteration 16: Global Deviance = 41.2794
GAMLSS-RS iteration 17: Global Deviance = 41.3065
GAMLSS-RS iteration 18: Global Deviance = 41.3336
GAMLSS-RS iteration 19: Global Deviance = 41.3604
GAMLSS-RS iteration 20: Global Deviance = 41.387
GAMLSS-CG iteration 1: Global Deviance = 40.8069
GAMLSS-CG iteration 2: Global Deviance = 40.8739
GAMLSS-CG iteration 3: Global Deviance = 40.8724
GAMLSS-CG iteration 4: Global Deviance = 40.8776
GAMLSS-CG iteration 5: Global Deviance = 40.8821

```

```
GAMLSS-CG iteration 6: Global Deviance = 40.8857
GAMLSS-CG iteration 7: Global Deviance = 40.889
GAMLSS-CG iteration 8: Global Deviance = 40.8917
GAMLSS-CG iteration 9: Global Deviance = 40.8939
GAMLSS-CG iteration 10: Global Deviance = 40.896
GAMLSS-CG iteration 11: Global Deviance = 40.8979
GAMLSS-CG iteration 12: Global Deviance = 40.8995
GAMLSS-CG iteration 13: Global Deviance = 40.901
GAMLSS-CG iteration 14: Global Deviance = 40.9024
GAMLSS-CG iteration 15: Global Deviance = 40.9035
GAMLSS-CG iteration 16: Global Deviance = 40.9046
GAMLSS-CG iteration 17: Global Deviance = 40.9055

>sp3<-
gamlss (UN~lo(~RAPOT, degree=2, span=.3)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 33.5437
GAMLSS-RS iteration 2: Global Deviance = 33.1269
GAMLSS-RS iteration 3: Global Deviance = 32.7854
GAMLSS-RS iteration 4: Global Deviance = 32.4984
GAMLSS-RS iteration 5: Global Deviance = 32.2569
GAMLSS-RS iteration 6: Global Deviance = 32.0543
GAMLSS-RS iteration 7: Global Deviance = 31.8893
GAMLSS-RS iteration 8: Global Deviance = 31.7458
GAMLSS-RS iteration 9: Global Deviance = 31.6206
GAMLSS-RS iteration 10: Global Deviance = 31.5118
GAMLSS-RS iteration 11: Global Deviance = 31.4166
GAMLSS-RS iteration 12: Global Deviance = 31.3337
GAMLSS-RS iteration 13: Global Deviance = 31.2585
GAMLSS-RS iteration 14: Global Deviance = 31.1906
GAMLSS-RS iteration 15: Global Deviance = 31.1338
GAMLSS-RS iteration 16: Global Deviance = 31.0902
GAMLSS-RS iteration 17: Global Deviance = 31.0629
GAMLSS-RS iteration 18: Global Deviance = 31.056
GAMLSS-RS iteration 19: Global Deviance = 31.075
GAMLSS-RS iteration 20: Global Deviance = 31.1266
GAMLSS-CG iteration 1: Global Deviance = 30.1517
GAMLSS-CG iteration 2: Global Deviance = 32.1066
GAMLSS-CG iteration 3: Global Deviance = 31.3218
GAMLSS-CG iteration 4: Global Deviance = 30.4069
GAMLSS-CG iteration 5: Global Deviance = 30.3525
GAMLSS-CG iteration 6: Global Deviance = 30.3414
GAMLSS-CG iteration 7: Global Deviance = 30.3451
GAMLSS-CG iteration 8: Global Deviance = 30.3486
GAMLSS-CG iteration 9: Global Deviance = 30.3448
GAMLSS-CG iteration 10: Global Deviance = 30.3444
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
```

```
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
> resp3<-refit(sp3)
GAMLSS-RS iteration 11: Global Deviance = 32.216
GAMLSS-RS iteration 12: Global Deviance = 33.3721
GAMLSS-RS iteration 13: Global Deviance = 34.2554
GAMLSS-RS iteration 14: Global Deviance = 34.9238
GAMLSS-RS iteration 15: Global Deviance = 35.4265
GAMLSS-RS iteration 16: Global Deviance = 35.8085
GAMLSS-RS iteration 17: Global Deviance = 36.1049
GAMLSS-RS iteration 18: Global Deviance = 36.3396
GAMLSS-RS iteration 19: Global Deviance = 36.5286
GAMLSS-RS iteration 20: Global Deviance = 36.6828
Error in CG(n.cyc = n2) :
  The global deviance is increasing in the inner CG loop
  Try different steps for the parameters or the model maybe
  inappropriate

>sp4<-
gamlss(UN~lo(~RAPOT,degree=1,span=.4)+UJIAN.SEKOLAH+TRY.OUT+
IQ,data=UAN,family=GG,method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 51.4954
GAMLSS-RS iteration 2: Global Deviance = 51.2487
GAMLSS-RS iteration 3: Global Deviance = 51.0965
GAMLSS-RS iteration 4: Global Deviance = 50.986
GAMLSS-RS iteration 5: Global Deviance = 50.9032
GAMLSS-RS iteration 6: Global Deviance = 50.8427
GAMLSS-RS iteration 7: Global Deviance = 50.7973
GAMLSS-RS iteration 8: Global Deviance = 50.7629
GAMLSS-RS iteration 9: Global Deviance = 50.7364
GAMLSS-RS iteration 10: Global Deviance = 50.7176
GAMLSS-RS iteration 11: Global Deviance = 50.7029
GAMLSS-RS iteration 12: Global Deviance = 50.6922
GAMLSS-RS iteration 13: Global Deviance = 50.6845
GAMLSS-RS iteration 14: Global Deviance = 50.6792
GAMLSS-RS iteration 15: Global Deviance = 50.6761
GAMLSS-RS iteration 16: Global Deviance = 50.6749
GAMLSS-RS iteration 17: Global Deviance = 50.6736
GAMLSS-RS iteration 18: Global Deviance = 50.6731
GAMLSS-CG iteration 1: Global Deviance = 50.6225
GAMLSS-CG iteration 2: Global Deviance = 50.645
GAMLSS-CG iteration 3: Global Deviance = 50.6419
GAMLSS-CG iteration 4: Global Deviance = 50.6421
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
```



```
> respa4<-refit(spa4)
GAMLSS-RS iteration 5: Global Deviance = 50.6906
GAMLSS-RS iteration 6: Global Deviance = 50.6814
GAMLSS-RS iteration 7: Global Deviance = 50.683
GAMLSS-RS iteration 8: Global Deviance = 50.6851
GAMLSS-RS iteration 9: Global Deviance = 50.689
GAMLSS-RS iteration 10: Global Deviance = 50.6926
GAMLSS-RS iteration 11: Global Deviance = 50.6963
GAMLSS-RS iteration 12: Global Deviance = 50.6999
GAMLSS-RS iteration 13: Global Deviance = 50.7035
GAMLSS-RS iteration 14: Global Deviance = 50.7068
GAMLSS-RS iteration 15: Global Deviance = 50.7101
GAMLSS-RS iteration 16: Global Deviance = 50.7133
GAMLSS-RS iteration 17: Global Deviance = 50.7112
GAMLSS-RS iteration 18: Global Deviance = 50.7149
GAMLSS-RS iteration 19: Global Deviance = 50.7187
GAMLSS-RS iteration 20: Global Deviance = 50.7225
GAMLSS-CG iteration 5: Global Deviance = 50.7184
GAMLSS-CG iteration 6: Global Deviance = 50.7187

>sp4<-
gamlss (UN~lo (~RAPOT, degree=2, span=.4)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(80,80) )
GAMLSS-RS iteration 1: Global Deviance = 42.9642
GAMLSS-RS iteration 2: Global Deviance = 42.6155
GAMLSS-RS iteration 3: Global Deviance = 42.3528
GAMLSS-RS iteration 4: Global Deviance = 42.1355
GAMLSS-RS iteration 5: Global Deviance = 41.9554
GAMLSS-RS iteration 6: Global Deviance = 41.8053
GAMLSS-RS iteration 7: Global Deviance = 41.684
GAMLSS-RS iteration 8: Global Deviance = 41.5815
GAMLSS-RS iteration 9: Global Deviance = 41.494
GAMLSS-RS iteration 10: Global Deviance = 41.4196
GAMLSS-RS iteration 11: Global Deviance = 41.3561
GAMLSS-RS iteration 12: Global Deviance = 41.3015
GAMLSS-RS iteration 13: Global Deviance = 41.2534
GAMLSS-RS iteration 14: Global Deviance = 41.2125
GAMLSS-RS iteration 15: Global Deviance = 41.176
GAMLSS-RS iteration 16: Global Deviance = 41.1449
GAMLSS-RS iteration 17: Global Deviance = 41.1166
GAMLSS-RS iteration 18: Global Deviance = 41.0926
GAMLSS-RS iteration 19: Global Deviance = 41.0704
GAMLSS-RS iteration 20: Global Deviance = 41.0506
GAMLSS-RS iteration 21: Global Deviance = 41.0339
GAMLSS-RS iteration 22: Global Deviance = 41.018
GAMLSS-RS iteration 23: Global Deviance = 41.0047
GAMLSS-RS iteration 24: Global Deviance = 40.9917
GAMLSS-RS iteration 25: Global Deviance = 40.9801
GAMLSS-RS iteration 26: Global Deviance = 40.9705
GAMLSS-RS iteration 27: Global Deviance = 40.9618
```



```
GAMLSS-RS iteration 28: Global Deviance = 40.9548
GAMLSS-RS iteration 29: Global Deviance = 40.9485
GAMLSS-RS iteration 30: Global Deviance = 40.9428
GAMLSS-RS iteration 31: Global Deviance = 40.9325
GAMLSS-RS iteration 32: Global Deviance = 40.9276
GAMLSS-RS iteration 33: Global Deviance = 40.9228
GAMLSS-RS iteration 34: Global Deviance = 40.9187
GAMLSS-RS iteration 35: Global Deviance = 40.915
GAMLSS-RS iteration 36: Global Deviance = 40.9116
GAMLSS-RS iteration 37: Global Deviance = 40.9085
GAMLSS-RS iteration 38: Global Deviance = 40.9058
GAMLSS-RS iteration 39: Global Deviance = 40.9032
GAMLSS-RS iteration 40: Global Deviance = 40.9009
GAMLSS-RS iteration 41: Global Deviance = 40.8988
GAMLSS-RS iteration 42: Global Deviance = 40.8969
GAMLSS-RS iteration 43: Global Deviance = 40.8951
GAMLSS-RS iteration 44: Global Deviance = 40.8935
GAMLSS-RS iteration 45: Global Deviance = 40.892
GAMLSS-RS iteration 46: Global Deviance = 40.8906
GAMLSS-RS iteration 47: Global Deviance = 40.8894
GAMLSS-RS iteration 48: Global Deviance = 40.8882
GAMLSS-RS iteration 49: Global Deviance = 40.8871
GAMLSS-RS iteration 50: Global Deviance = 40.8861
GAMLSS-CG iteration 1: Global Deviance = 40.8839
GAMLSS-CG iteration 2: Global Deviance = 40.8844

> resp4<-refit(sp4)
GAMLSS-RS iteration 3: Global Deviance = 40.8881
GAMLSS-RS iteration 4: Global Deviance = 40.8917
GAMLSS-RS iteration 5: Global Deviance = 40.8877
GAMLSS-RS iteration 6: Global Deviance = 40.8837
GAMLSS-RS iteration 7: Global Deviance = 40.8826
GAMLSS-RS iteration 8: Global Deviance = 40.8814
GAMLSS-RS iteration 9: Global Deviance = 40.8806
GAMLSS-CG iteration 3: Global Deviance = 40.8791
GAMLSS-CG iteration 4: Global Deviance = 40.8792

>sp5<-
gamlss(UN~lo(~RAPOT,degree=1,span=.5)+UJIAN.SEKOLAH+TRY.OUT+
IQ,data=UAN,family=GG,method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 54.0798
GAMLSS-RS iteration 2: Global Deviance = 53.7396
GAMLSS-RS iteration 3: Global Deviance = 53.5172
GAMLSS-RS iteration 4: Global Deviance = 53.351
GAMLSS-RS iteration 5: Global Deviance = 53.2256
GAMLSS-RS iteration 6: Global Deviance = 53.1295
GAMLSS-RS iteration 7: Global Deviance = 53.0546
GAMLSS-RS iteration 8: Global Deviance = 52.9954
GAMLSS-RS iteration 9: Global Deviance = 52.9489
GAMLSS-RS iteration 10: Global Deviance = 52.9122
GAMLSS-RS iteration 11: Global Deviance = 52.8822
```

```
GAMLSS-RS iteration 12: Global Deviance = 52.8592
GAMLSS-RS iteration 13: Global Deviance = 52.84
GAMLSS-RS iteration 14: Global Deviance = 52.8247
GAMLSS-RS iteration 15: Global Deviance = 52.8124
GAMLSS-RS iteration 16: Global Deviance = 52.8039
GAMLSS-RS iteration 17: Global Deviance = 52.7972
GAMLSS-RS iteration 18: Global Deviance = 52.7909
GAMLSS-RS iteration 19: Global Deviance = 52.7867
GAMLSS-RS iteration 20: Global Deviance = 52.7783
GAMLSS-CG iteration 1: Global Deviance = 52.7309
GAMLSS-CG iteration 2: Global Deviance = 52.7666
GAMLSS-CG iteration 3: Global Deviance = 52.7508
GAMLSS-CG iteration 4: Global Deviance = 52.7508
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> resp5<-refit(sp5)
GAMLSS-RS iteration 5: Global Deviance = 52.7701
GAMLSS-RS iteration 6: Global Deviance = 52.7836
GAMLSS-RS iteration 7: Global Deviance = 52.7906
GAMLSS-RS iteration 8: Global Deviance = 52.7889
GAMLSS-RS iteration 9: Global Deviance = 52.7911
GAMLSS-RS iteration 10: Global Deviance = 52.7928
GAMLSS-RS iteration 11: Global Deviance = 52.7945
GAMLSS-RS iteration 12: Global Deviance = 52.7922
GAMLSS-RS iteration 13: Global Deviance = 52.7941
GAMLSS-RS iteration 14: Global Deviance = 52.7962
GAMLSS-RS iteration 15: Global Deviance = 52.7982
GAMLSS-RS iteration 16: Global Deviance = 52.8001
GAMLSS-RS iteration 17: Global Deviance = 52.8018
GAMLSS-RS iteration 18: Global Deviance = 52.8035
GAMLSS-RS iteration 19: Global Deviance = 52.805
GAMLSS-RS iteration 20: Global Deviance = 52.8064
GAMLSS-CG iteration 5: Global Deviance = 52.8029
GAMLSS-CG iteration 6: Global Deviance = 52.803

>sp5<-
gamlss (UN~lo (~RAPOT, degree=2, span=.5)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 48.736
GAMLSS-RS iteration 2: Global Deviance = 48.466
GAMLSS-RS iteration 3: Global Deviance = 48.252
GAMLSS-RS iteration 4: Global Deviance = 48.0871
GAMLSS-RS iteration 5: Global Deviance = 47.9549
GAMLSS-RS iteration 6: Global Deviance = 47.8461
GAMLSS-RS iteration 7: Global Deviance = 47.7595
```

```
GAMLSS-RS iteration 8: Global Deviance = 47.6876
GAMLSS-RS iteration 9: Global Deviance = 47.627
GAMLSS-RS iteration 10: Global Deviance = 47.5747
GAMLSS-RS iteration 11: Global Deviance = 47.53
GAMLSS-RS iteration 12: Global Deviance = 47.4905
GAMLSS-RS iteration 13: Global Deviance = 47.4574
GAMLSS-RS iteration 14: Global Deviance = 47.4286
GAMLSS-RS iteration 15: Global Deviance = 47.4044
GAMLSS-RS iteration 16: Global Deviance = 47.3831
GAMLSS-RS iteration 17: Global Deviance = 47.3645
GAMLSS-RS iteration 18: Global Deviance = 47.348
GAMLSS-RS iteration 19: Global Deviance = 47.3334
GAMLSS-RS iteration 20: Global Deviance = 47.3205
GAMLSS-CG iteration 1: Global Deviance = 47.305
GAMLSS-CG iteration 2: Global Deviance = 47.2896
GAMLSS-CG iteration 3: Global Deviance = 47.2898

> resp5<-refit(sp5)
GAMLSS-RS iteration 4: Global Deviance = 47.2984
GAMLSS-RS iteration 5: Global Deviance = 47.2918
GAMLSS-RS iteration 6: Global Deviance = 47.2694
GAMLSS-RS iteration 7: Global Deviance = 47.2643
GAMLSS-RS iteration 8: Global Deviance = 47.2577
GAMLSS-RS iteration 9: Global Deviance = 47.2477
GAMLSS-RS iteration 10: Global Deviance = 47.2425
GAMLSS-RS iteration 11: Global Deviance = 47.2377
GAMLSS-RS iteration 12: Global Deviance = 47.2335
GAMLSS-RS iteration 13: Global Deviance = 47.2296
GAMLSS-RS iteration 14: Global Deviance = 47.2261
GAMLSS-RS iteration 15: Global Deviance = 47.223
GAMLSS-RS iteration 16: Global Deviance = 47.2202
GAMLSS-RS iteration 17: Global Deviance = 47.2177
GAMLSS-RS iteration 18: Global Deviance = 47.2154
GAMLSS-RS iteration 19: Global Deviance = 47.2133
GAMLSS-RS iteration 20: Global Deviance = 47.2115
GAMLSS-CG iteration 4: Global Deviance = 47.2089
GAMLSS-CG iteration 5: Global Deviance = 47.209

>sp6<-
gamlss(UN~lo(~RAPOT,degree=1,span=.6)+UJIAN.SEKOLAH+TRY.OUT+
IQ,data=UAN,family=GG,method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 57.9084
GAMLSS-RS iteration 2: Global Deviance = 57.1143
GAMLSS-RS iteration 3: Global Deviance = 56.5244
GAMLSS-RS iteration 4: Global Deviance = 56.0619
GAMLSS-RS iteration 5: Global Deviance = 55.69
GAMLSS-RS iteration 6: Global Deviance = 55.3978
GAMLSS-RS iteration 7: Global Deviance = 55.1645
GAMLSS-RS iteration 8: Global Deviance = 54.9788
GAMLSS-RS iteration 9: Global Deviance = 54.829
GAMLSS-RS iteration 10: Global Deviance = 54.7091
```

```
GAMLSS-RS iteration 11: Global Deviance = 54.611
GAMLSS-RS iteration 12: Global Deviance = 54.5318
GAMLSS-RS iteration 13: Global Deviance = 54.4659
GAMLSS-RS iteration 14: Global Deviance = 54.4115
GAMLSS-RS iteration 15: Global Deviance = 54.3746
GAMLSS-RS iteration 16: Global Deviance = 54.3416
GAMLSS-RS iteration 17: Global Deviance = 54.314
GAMLSS-RS iteration 18: Global Deviance = 54.2905
GAMLSS-RS iteration 19: Global Deviance = 54.2705
GAMLSS-RS iteration 20: Global Deviance = 54.2534
GAMLSS-CG iteration 1: Global Deviance = 54.203
GAMLSS-CG iteration 2: Global Deviance = 54.1933
GAMLSS-CG iteration 3: Global Deviance = 54.1863
GAMLSS-CG iteration 4: Global Deviance = 54.1653
GAMLSS-CG iteration 5: Global Deviance = 54.1645

>sp6<-
gamlss (UN~lo (~RAPOT, degree=2, span=.6)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 53.6331
GAMLSS-RS iteration 2: Global Deviance = 53.4563
GAMLSS-RS iteration 3: Global Deviance = 53.3074
GAMLSS-RS iteration 4: Global Deviance = 53.1855
GAMLSS-RS iteration 5: Global Deviance = 53.0838
GAMLSS-RS iteration 6: Global Deviance = 52.998
GAMLSS-RS iteration 7: Global Deviance = 52.9247
GAMLSS-RS iteration 8: Global Deviance = 52.8617
GAMLSS-RS iteration 9: Global Deviance = 52.8112
GAMLSS-RS iteration 10: Global Deviance = 52.7658
GAMLSS-RS iteration 11: Global Deviance = 52.7253
GAMLSS-RS iteration 12: Global Deviance = 52.6891
GAMLSS-RS iteration 13: Global Deviance = 52.6565
GAMLSS-RS iteration 14: Global Deviance = 52.627
GAMLSS-RS iteration 15: Global Deviance = 52.6002
GAMLSS-RS iteration 16: Global Deviance = 52.5758
GAMLSS-RS iteration 17: Global Deviance = 52.5534
GAMLSS-RS iteration 18: Global Deviance = 52.5329
GAMLSS-RS iteration 19: Global Deviance = 52.514
GAMLSS-RS iteration 20: Global Deviance = 52.4966
GAMLSS-CG iteration 1: Global Deviance = 52.4762
GAMLSS-CG iteration 2: Global Deviance = 52.4587
GAMLSS-CG iteration 3: Global Deviance = 52.4484
GAMLSS-CG iteration 4: Global Deviance = 52.4252
GAMLSS-CG iteration 5: Global Deviance = 52.421
GAMLSS-CG iteration 6: Global Deviance = 52.396
GAMLSS-CG iteration 7: Global Deviance = 52.396

>sp7<-
gamlss (UN~lo (~RAPOT, degree=1, span=.7)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 60.0802
```



```
GAMLSS-RS iteration 2: Global Deviance = 58.9343
GAMLSS-RS iteration 3: Global Deviance = 58.0589
GAMLSS-RS iteration 4: Global Deviance = 57.3771
GAMLSS-RS iteration 5: Global Deviance = 56.8529
GAMLSS-RS iteration 6: Global Deviance = 56.4417
GAMLSS-RS iteration 7: Global Deviance = 56.1269
GAMLSS-RS iteration 8: Global Deviance = 55.8852
GAMLSS-RS iteration 9: Global Deviance = 55.6986
GAMLSS-RS iteration 10: Global Deviance = 55.5538
GAMLSS-RS iteration 11: Global Deviance = 55.4406
GAMLSS-RS iteration 12: Global Deviance = 55.3513
GAMLSS-RS iteration 13: Global Deviance = 55.2803
GAMLSS-RS iteration 14: Global Deviance = 55.2235
GAMLSS-RS iteration 15: Global Deviance = 55.1808
GAMLSS-RS iteration 16: Global Deviance = 55.1449
GAMLSS-RS iteration 17: Global Deviance = 55.1154
GAMLSS-RS iteration 18: Global Deviance = 55.0906
GAMLSS-RS iteration 19: Global Deviance = 55.0699
GAMLSS-RS iteration 20: Global Deviance = 55.0525
GAMLSS-CG iteration 1: Global Deviance = 55.0199
GAMLSS-CG iteration 2: Global Deviance = 54.9961
GAMLSS-CG iteration 3: Global Deviance = 54.9937
GAMLSS-CG iteration 4: Global Deviance = 54.9917
GAMLSS-CG iteration 5: Global Deviance = 54.9799
GAMLSS-CG iteration 6: Global Deviance = 54.9774
GAMLSS-CG iteration 7: Global Deviance = 54.9769
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> resp7<-refit(sp7)
GAMLSS-RS iteration 8: Global Deviance = 54.9576
GAMLSS-RS iteration 9: Global Deviance = 54.9629
GAMLSS-RS iteration 10: Global Deviance = 54.9587
GAMLSS-RS iteration 11: Global Deviance = 54.9574
GAMLSS-RS iteration 12: Global Deviance = 54.9551
GAMLSS-RS iteration 13: Global Deviance = 54.9535
GAMLSS-RS iteration 14: Global Deviance = 54.952
GAMLSS-RS iteration 15: Global Deviance = 54.9508
GAMLSS-RS iteration 16: Global Deviance = 54.9497
GAMLSS-RS iteration 17: Global Deviance = 54.9488
GAMLSS-CG iteration 8: Global Deviance = 54.9464
GAMLSS-CG iteration 9: Global Deviance = 54.9438
GAMLSS-CG iteration 10: Global Deviance = 54.9439
```



```
>sp7<-
gamlss (UN~lo (~RAPOT, degree=2, span=.7)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 55.1834
GAMLSS-RS iteration 2: Global Deviance = 54.8914
GAMLSS-RS iteration 3: Global Deviance = 54.6773
GAMLSS-RS iteration 4: Global Deviance = 54.5085
GAMLSS-RS iteration 5: Global Deviance = 54.3724
GAMLSS-RS iteration 6: Global Deviance = 54.2607
GAMLSS-RS iteration 7: Global Deviance = 54.1675
GAMLSS-RS iteration 8: Global Deviance = 54.0889
GAMLSS-RS iteration 9: Global Deviance = 54.0219
GAMLSS-RS iteration 10: Global Deviance = 53.9643
GAMLSS-RS iteration 11: Global Deviance = 53.9146
GAMLSS-RS iteration 12: Global Deviance = 53.8714
GAMLSS-RS iteration 13: Global Deviance = 53.8354
GAMLSS-RS iteration 14: Global Deviance = 53.8037
GAMLSS-RS iteration 15: Global Deviance = 53.7755
GAMLSS-RS iteration 16: Global Deviance = 53.7506
GAMLSS-RS iteration 17: Global Deviance = 53.7284
GAMLSS-RS iteration 18: Global Deviance = 53.7087
GAMLSS-RS iteration 19: Global Deviance = 53.6912
GAMLSS-RS iteration 20: Global Deviance = 53.6755
GAMLSS-CG iteration 1: Global Deviance = 53.6277
GAMLSS-CG iteration 2: Global Deviance = 53.6551
GAMLSS-CG iteration 3: Global Deviance = 53.5937
GAMLSS-CG iteration 4: Global Deviance = 53.5939
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> resp7<-refit(sp7)
GAMLSS-RS iteration 5: Global Deviance = 53.593
GAMLSS-CG iteration 5: Global Deviance = 53.5871
GAMLSS-CG iteration 6: Global Deviance = 53.5791
GAMLSS-CG iteration 7: Global Deviance = 53.5651
GAMLSS-CG iteration 8: Global Deviance = 53.5653

>sp8<-
gamlss (UN~lo (~RAPOT, degree=1, span=.8)+UJIAN.SEKOLAH+TRY.OUT+
IQ, data=UAN, family=GG, method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 60.7059
GAMLSS-RS iteration 2: Global Deviance = 59.422
GAMLSS-RS iteration 3: Global Deviance = 58.4519
GAMLSS-RS iteration 4: Global Deviance = 57.7273
GAMLSS-RS iteration 5: Global Deviance = 57.187
GAMLSS-RS iteration 6: Global Deviance = 56.7757
```

```
GAMLSS-RS iteration 7: Global Deviance = 56.4706
GAMLSS-RS iteration 8: Global Deviance = 56.2271
GAMLSS-RS iteration 9: Global Deviance = 56.0424
GAMLSS-RS iteration 10: Global Deviance = 55.899
GAMLSS-RS iteration 11: Global Deviance = 55.7878
GAMLSS-RS iteration 12: Global Deviance = 55.7005
GAMLSS-RS iteration 13: Global Deviance = 55.6313
GAMLSS-RS iteration 14: Global Deviance = 55.5802
GAMLSS-RS iteration 15: Global Deviance = 55.538
GAMLSS-RS iteration 16: Global Deviance = 55.5032
GAMLSS-RS iteration 17: Global Deviance = 55.4743
GAMLSS-RS iteration 18: Global Deviance = 55.4502
GAMLSS-RS iteration 19: Global Deviance = 55.43
GAMLSS-RS iteration 20: Global Deviance = 55.4131
GAMLSS-CG iteration 1: Global Deviance = 55.3748
GAMLSS-CG iteration 2: Global Deviance = 55.3676
GAMLSS-CG iteration 3: Global Deviance = 55.3603
GAMLSS-CG iteration 4: Global Deviance = 55.3594
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> resp8<-refit(sp8)
GAMLSS-RS iteration 5: Global Deviance = 55.3683
GAMLSS-RS iteration 6: Global Deviance = 55.3637
GAMLSS-RS iteration 7: Global Deviance = 55.3531
GAMLSS-RS iteration 8: Global Deviance = 55.3461
GAMLSS-RS iteration 9: Global Deviance = 55.34
GAMLSS-RS iteration 10: Global Deviance = 55.3352
GAMLSS-RS iteration 11: Global Deviance = 55.3312
GAMLSS-RS iteration 12: Global Deviance = 55.328
GAMLSS-RS iteration 13: Global Deviance = 55.3254
GAMLSS-RS iteration 14: Global Deviance = 55.3232
GAMLSS-RS iteration 15: Global Deviance = 55.3214
GAMLSS-RS iteration 16: Global Deviance = 55.3199
GAMLSS-RS iteration 17: Global Deviance = 55.3187
GAMLSS-RS iteration 18: Global Deviance = 55.3177
GAMLSS-RS iteration 19: Global Deviance = 55.3169
GAMLSS-CG iteration 5: Global Deviance = 55.3134
GAMLSS-CG iteration 6: Global Deviance = 55.3123
GAMLSS-CG iteration 7: Global Deviance = 55.3123

>sp8<-
gamlss(UN~lo(~RAPOT,degree=2,span=.8)+UJIAN.SEKOLAH+TRY.OUT+
IQ,data=UAN,family=GG,method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 58.9409
GAMLSS-RS iteration 2: Global Deviance = 58.2773
```

```

GAMLSS-RS iteration 3: Global Deviance = 57.7009
GAMLSS-RS iteration 4: Global Deviance = 57.1967
GAMLSS-RS iteration 5: Global Deviance = 56.7555
GAMLSS-RS iteration 6: Global Deviance = 56.3704
GAMLSS-RS iteration 7: Global Deviance = 56.0361
GAMLSS-RS iteration 8: Global Deviance = 55.7479
GAMLSS-RS iteration 9: Global Deviance = 55.501
GAMLSS-RS iteration 10: Global Deviance = 55.2908
GAMLSS-RS iteration 11: Global Deviance = 55.1127
GAMLSS-RS iteration 12: Global Deviance = 54.9625
GAMLSS-RS iteration 13: Global Deviance = 54.8361
GAMLSS-RS iteration 14: Global Deviance = 54.7301
GAMLSS-RS iteration 15: Global Deviance = 54.6412
GAMLSS-RS iteration 16: Global Deviance = 54.5668
GAMLSS-RS iteration 17: Global Deviance = 54.5046
GAMLSS-RS iteration 18: Global Deviance = 54.4527
GAMLSS-RS iteration 19: Global Deviance = 54.4111
GAMLSS-RS iteration 20: Global Deviance = 54.3758
GAMLSS-CG iteration 1: Global Deviance = 54.3398
GAMLSS-CG iteration 2: Global Deviance = 54.2791
GAMLSS-CG iteration 3: Global Deviance = 54.2719
GAMLSS-CG iteration 4: Global Deviance = 54.2476
GAMLSS-CG iteration 5: Global Deviance = 54.2478
Warning messages:
1: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations
2: In additive.fit(x = X, y = wv, w = wt * w, s = s, who =
who, smooth.frame, :
  additive.fit convergence not obtained in 30 iterations

> resp8<-refit(sp8)
GAMLSS-RS iteration 6: Global Deviance = 54.2547
GAMLSS-RS iteration 7: Global Deviance = 54.2428
GAMLSS-RS iteration 8: Global Deviance = 54.2365
GAMLSS-RS iteration 9: Global Deviance = 54.2339
GAMLSS-RS iteration 10: Global Deviance = 54.2323
GAMLSS-RS iteration 11: Global Deviance = 54.2323
GAMLSS-CG iteration 6: Global Deviance = 54.2331

>GAIC(sp3, respa4, resp4, resp5, resp5, sp6, sp6, resp7, resp7, resp8
, resp8)
      df      AIC
sp3    13.297367 67.50023
resp4   14.001524 68.88221
resp5   12.316587 71.84214
resp8    8.293904 71.90015
resp7    8.745578 72.43502
respa4  10.924549 72.56781
resp5    9.939043 72.68112
sp6     9.350540 72.86555

```

```
resp8 9.940584 74.11429
resp7 10.669725 74.90471
sp6 11.460823 75.31762
```

D6. Visualisasi Model Sesudah dilakukan *Smoothing* dengan Beberapa Pilihan Span

```
library(ggplot2)
>p<-
ggplot(data=UAN,aes_string(x="RAPOT",y="UN"))+geom_point(,col="black")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.3,degree=1,lty=4,se=FALSE,col="blue")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.3,degree=2,lty=2,se=FALSE,col="red")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.6,degree=1,lty=4,se=FALSE,col="blue")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.6,degree=2,lty=2,se=FALSE,col="red")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.8,degree=1,lty=4,se=FALSE,col="blue")
>print(p)

>p<-
p+geom_smooth(,method="loess",span=0.8,degree=2,lty=2,se=FALSE,col="green")
>print(p)
```

D7. Analisis Model Terbaik dengan Menggunakan *stepGAIC()*

```
>sp3<-
gamlss(UN~lo(~RAPOT,degree=1,span=.3)+UJIAN.SEKOLAH+TRY.OUT+IQ,data=UAN,family=GG,method=mixed(20,20))
GAMLSS-RS iteration 1: Global Deviance = 42.9564
GAMLSS-RS iteration 2: Global Deviance = 42.2644
GAMLSS-RS iteration 3: Global Deviance = 41.822
GAMLSS-RS iteration 4: Global Deviance = 41.5351
GAMLSS-RS iteration 5: Global Deviance = 41.3539
```



```
GAMLSS-RS iteration 6: Global Deviance = 41.2438
GAMLSS-RS iteration 7: Global Deviance = 41.1812
GAMLSS-RS iteration 8: Global Deviance = 41.1504
GAMLSS-RS iteration 9: Global Deviance = 41.1406
GAMLSS-RS iteration 10: Global Deviance = 41.1447
GAMLSS-RS iteration 11: Global Deviance = 41.1579
GAMLSS-RS iteration 12: Global Deviance = 41.1771
GAMLSS-RS iteration 13: Global Deviance = 41.2013
GAMLSS-RS iteration 14: Global Deviance = 41.2264
GAMLSS-RS iteration 15: Global Deviance = 41.2526
GAMLSS-RS iteration 16: Global Deviance = 41.2794
GAMLSS-RS iteration 17: Global Deviance = 41.3065
GAMLSS-RS iteration 18: Global Deviance = 41.3336
GAMLSS-RS iteration 19: Global Deviance = 41.3604
GAMLSS-RS iteration 20: Global Deviance = 41.387
GAMLSS-CG iteration 1: Global Deviance = 40.8069
GAMLSS-CG iteration 2: Global Deviance = 40.8739
GAMLSS-CG iteration 3: Global Deviance = 40.8724
GAMLSS-CG iteration 4: Global Deviance = 40.8776
GAMLSS-CG iteration 5: Global Deviance = 40.8821
GAMLSS-CG iteration 6: Global Deviance = 40.8857
GAMLSS-CG iteration 7: Global Deviance = 40.889
GAMLSS-CG iteration 8: Global Deviance = 40.8917
GAMLSS-CG iteration 9: Global Deviance = 40.8939
GAMLSS-CG iteration 10: Global Deviance = 40.896
GAMLSS-CG iteration 11: Global Deviance = 40.8979
GAMLSS-CG iteration 12: Global Deviance = 40.8995
GAMLSS-CG iteration 13: Global Deviance = 40.901
GAMLSS-CG iteration 14: Global Deviance = 40.9024
GAMLSS-CG iteration 15: Global Deviance = 40.9035
GAMLSS-CG iteration 16: Global Deviance = 40.9046
GAMLSS-CG iteration 17: Global Deviance = 40.9055
```

```
gs1<-
gamlss.scope(model.frame(UAN), smoother="lo", arg="span=.3")
step1<-stepGAIC(resp3,go,additive=TRUE)
Distribution parameter:mu
Start:UN~lo(~RAPOT,degree=1,span=.3)+UJIAN.SEKOLAH+TRY.OUT+IQ,
data=UAN,family=GG,method=mixed(20,20)); AIC=67,50023
Trial1:UN~TRY.OUT+lo(~RAPOT,degree=1,span=.3),data=UAN,family=GG,
method=mixed(60,60);AIC=103.42723
Trial2:UN~lo(~RAPOT,degree=1,span=.3),data=UAN,family=GG,method=
mixed(80,80); AIC: 102.088
Trial3:UN~UJIAN.SEKOLAH+TRY.OUT+lo(~RAPOT,degree=1,span=.3),
data=UAN,family=GG,method=mixed(80,80); AIC=91.84
Trial4:UN~UJIAN.SEKOLAH+lo(~RAPOT,degree=1,span=.3),data=UAN,
family=GG,method=mixed(80,80); AIC=90,0038
Trial5:UN~lo(~RAPOT,degree=1,span=.3)+UJIAN.SEKOLAH,data=UAN,
family=GG,method=mixed(60,60);AIC= 90.0038
Trial6:UN~IQ+TRY.OUT+lo(~RAPOT,degree=1,span=.3),data=UAN,family=GG,
method=mixed(20,20); AIC=69.94553
```



```

Trial7:UN~IQ+lo(~RAPOT,degree=1,span=.3),data=UAN,family=GG,
method=mixed(80,80) :AIC=67.93908
Trial8:UN~UJIAN.SEKOLAH+IQ+lo(~RAPOT,degree=1,span=.3),data=
UAN,family=GG,method=mixed(60,60); AIC= 66.95514
>step1
Family: c("GG", "generalised Gamma Lopatatsidis-Green")
Fitting method: mixed(20, 20)

Call:  gamlss(formula = UN ~ lo(~RAPOT, degree = 1, span =
0.3) + UJIAN.SEKOLAH + IQ, family = GG, data = UAN, method
= mixed(20, 20), trace = FALSE)

Mu Coefficients:
              (Intercept)
UJIAN.SEKOLAH              1.182419
0.081429                    0.004828
lo(~RAPOT, degree = 1, span = 0.3)
                                     NA
Sigma Coefficients:
(Intercept)
      -3.565
Nu Coefficients:
(Intercept)
      62.96

Degrees of Freedom for the fit: 12.29737 Residual Deg. of
Freedom 35.70263
Global Deviance: 42.3604
                AIC: 66.9551
                SBC: 89.966

```