



TUGAS AKHIR - SS091324

PEMODELAN JUMLAH KASUS DEMAM BERDARAH DENGUE (DBD) DENGAN PENDEKATAN REGRESI SEMIPARAMETRIK SPLINE DI KABUPATEN/KOTA MADIUN

FADLI AKBAR PAMBUDI
NRP 1309 100 095

Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si

Co.Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si

JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2014



TUGAS AKHIR - SS091324

**PEMODELAN JUMLAH KASUS DEMAM
BERDARAH DENGUE (DBD) DENGAN
PENDEKATAN REGRESI SEMIPARAMETRIK
SPLINE DI KABUPATEN/KOTA MADIUN**

FADLI AKBAR PAMBUDI
NRP 1309 100 095

Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si

Co.Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si

JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2014



FINAL PROJECT - SS091324

MODELING OF DENGUE HEMORRHAGIC FEVER (DHF) CASES IN MADIUN USING SEMIPARAMETRIC SPLINE REGRESSION

FADLI AKBAR PAMBUDI
NRP 1309 100 095

Supervisor
Dr. Dra. Ismaini Zain, M.Si

Co.Supervisor
Prof. Dr. Drs. I Nyoman Budiantara, M.Si

DEPARTMENT OF STATISTICS
Faculty of Mathematics and Natural Science
Institut Teknologi Sepuluh Nopember
Surabaya 2014

PEMODELAN JUMLAH KASUS DEMAM BERDARAH DENGUE (DBD) DENGAN PENDEKATAN REGRESI SEMIPARAMETRIK SPLINE DI KABUPATEN/KOTA MADIUN

Nama : Fadli Akbar Pambudi
NRP : 1309 100 095
Bidang/Minat : Sosial Pemerintahan
Jurusan : Statistika, FMIPA-ITS
Dosen Pembimbing : Dr. Dra. Ismaini Zain, M.Si
Co.Dosen Pembimbing : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

ABSTRAK

Demam Berdarah Dengue (DBD) merupakan salah satu penyakit menular yang sering mengakibatkan Kejadian Luar Biasa di Indonesia dikarenakan perjalanan penyakitnya cepat dan dapat menyebabkan kematian secara singkat. Data dari Kementerian Kesehatan pada tahun 2010 menunjukkan telah terjadi 156.086 kasus yang mengakibatkan kematian 1.358 orang. Selain itu juga tercatat banyaknya kejadian DBD di Kabupaten/Kota Madiun mencapai 498 kasus dan jumlah kematian sebanyak 7 orang. Hal ini cukup mengkhawatirkan karena berdasarkan beberapa indikator yang ditentukan telah melewati batas target kesehatan nasional. Untuk itu perlu dilakukan penelitian tentang pemodelan faktor-faktor yang mempengaruhi banyaknya kasus DBD Kabupaten/Kota Madiun tahun 2010. Variabel prediktor yang digunakan direduksi dengan analisis faktor untuk memilih variabel-variabel yang tidak berkorelasi satu sama lain dan menghasilkan variabel angka bebas jentik, persentase penduduk miskin, curah hujan dan kepadatan penduduk. Berdasarkan hasil penelitian didapatkan metode regresi semiparametrik spline yang digunakan untuk melakukan pemodelan tersebut. Titik knot yang digunakan adalah menggunakan kombinasi knot (322) yang menghasilkan nilai GCV sebesar 181,437 dan MSE sebesar 70,479. Model yang dihasilkan dapat menjelaskan variabilitas jumlah kasus DBD sebesar 98,351 persen dengan semua variabel berpengaruh signifikan.

Kata kunci: Demam Berdarah Dengue, Analisis Faktor, Regresi Semiparametrik Spline, Titik Knot, GCV

MODELING OF DENGUE HEMORRHAGIC FEVER (DHF) CASES IN MADIUN USING SEMIPARAMETRIC SPLINE REGRESSION

Name : Fadli Akbar Pambudi
ID Number : 1309 100 095
Laboratory : Sosial Pemerintahan
Department : Statistika, FMIPA-ITS
Supervisor : Dr. Dra. Ismaini Zain, M.Si
Co.Supervisor : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

ABSTRACT

Dengue Hemorrhagic Fever (DHF/DBD in Indonesian) is a contagious disease which causes many extraordinary events in Indonesia due to rapid transmission and high number of death. The data from ministry of health shows that in 2010, there are 156.086 cases occurred and 1.358 people died due to this disease. It also recorded the number of occurrences of dengue fever in Madiun reaches 498 cases and number of deaths are 7 people in total. It is quite serious because based on several indicators, this case have exceeded the national health target. Therefore the research was conducted to examine any factors that influence the number of cases of DBD in Madiun based on 2010 data. The predictor variables reduced by factor analysis to select variables that are not correlated with each other and produce the free rates of mosquito larvae, the percentage of poor people, rainfall and population density variable. Based on the research, the semiparametric spline regression methods were applied to perform the modeling . The knots point used are combination of knots (322) which produces GCV value of 181,437 and MSE of 70,479. The model resulted can explain the variability in the number of DBD cases at 98,351 percent with all variables significantly affects.

Keywords: Dengue Hemorrhagic Fever, Factor Analysis, Semiparametric Spline Regression, Knots, GCV.

LEMBAR PENGESAHAN

PEMODELAN JUMLAH KASUS DEMAM BERDARAH DENGUE (DBD) DENGAN PENDEKATAN REGRESI SEMIPARAMETRIK SPLINE DI KABUPATEN/KOTA MADIUN

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
Pada Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

FADLI AKBAR PAMBUDI

NRP. 1309 100 095

Disetujui oleh Pembimbing Tugas Akhir :

Dr. Dra. Ismaini Zain, M.Si

NIP. 19600525 198803 2 001

Prof. Dr. Drs. I Nyoman Budiantara, M.Si

NIP. 19650603 198903 1 003

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS

Dr. Muhammad Mashuri, M.T.

NIP. 19620408 198701 1 001

SURABAYA, FEBRUARI 2014

KATA PENGANTAR



Puji syukur dipanjatkan kehadirat **Allah SWT** atas limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan dengan baik Tugas Akhir yang berjudul

**”PEMODELAN JUMLAH KASUS
DEMAM BERDARAH DENGUE (DBD) DENGAN
PENDEKATAN REGRESI SEMIPARAMETRIK SPLINE
DI KABUPATEN/KOTA MADIUN”**

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini tidak akan berhasil dengan baik tanpa banyak pihak yang telah membantu dalam pengerjaannya. Oleh karena itu, pada kesempatan ini penulis ingin memberikan penghormatan dan terimakasih yang sebesar-besarnya kepada:

1. Bapak Dr. Muhammad Mashuri, M.T. selaku Ketua Jurusan Statistika FMIPA ITS Surabaya yang telah memberikan banyak fasilitas dalam berkarya.
2. Ibu Dra. Lucia Aridinanti, M.T. selaku Ketua Program Studi S-1 Statistika FMIPA ITS Surabaya yang memberikan kemudahan alur dan teknis pengerjaan Tugas Akhir.
3. Ibu Dr. Dra. Ismaini Zain, M.Si dan Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah bersedia memberikan waktu, kesabaran, semangat, saran dan ilmu.
4. Bapak (Alm) Ir. Arie Kismanto, M.Sc. dan Ibu Dwi Endah Kusriani, S.Si., M.Si. selaku dosen wali yang telah memberikan nasehat, motivasi dan bimbingan selama perkuliahan.
5. Seluruh dosen di Jurusan Statistika ITS atas ilmu dan pengalaman yang dibagikan.

6. Kedua orang tua, bapak (Teguh Soewarno) dan ibu (Tri Endah Listiani) yang telah memberikan doa, kasih sayang, dukungan dan kesabaran yang tiada habisnya.
7. Kakakku (Febri) yang telah memberikan semangat dan Novia yang selama ini telah menemani dalam suka dan duka.
8. Keluarga Besar Statistika ITS 2009 ($\Sigma 20$), Tim Ahli HIMASTA-ITS, Divisi SCC serta teman-teman yang tergabung dalam kepanitiaan PRS atas segala bentuk warna-warni kehidupan di kampus perjuangan Institut Teknologi Sepuluh Nopember.
9. Teman-teman spliners 107 yang telah banyak membantu dalam proses penyelesaian Tugas Akhir ini.
10. Geng Koplak (Angga, Billy, Oka, Imam, Yonkki, Mas Galon dan Mas Iwan), Penghuni Himakobus (Dima, Fikri, Mas Dhika, Mas Edwin, Arie, Dito) atas segala kebersamaannya. Semoga suatu saat dapat berkumpul kembali dalam keadaan yang lebih baik.
11. Semua pihak yang telah mendukung baik secara langsung maupun tidak langsung yang tidak dapat disebutkan satu persatu.

Penulis sangat berharap hasil Tugas Akhir ini dapat bermanfaat bagi kita semua. Penulis menyadari sepenuhnya bahwa Tugas Akhir ini masih banyak terdapat kekurangan dan ketidaksempurnaan. Oleh karena itu, penulis berterima kasih apabila ada saran dan kritik yang bersifat membangun guna perbaikan di masa mendatang demi perkembangan statistika yang lebih baik.

Surabaya, Februari 2014

Penulis

DAFTAR ISI

JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif.....	7
2.2 Uji Kecukupan Data.....	7
2.3 Uji Homogenitas Varians	8
2.4 Analisis Faktor.....	8
2.5 Analisis Regresi	9
2.6 Regresi Parametrik	10
2.7 Regresi Nonparametrik.....	11
2.8 Regresi Semiparametrik	12
2.9 Pemilihan Titik Knot Optimal	12
2.10 Kriteria Keباikan Model.....	13
2.11 Pengujian Signifikasi Parameter	14
2.11.1 Uji Serentak (simultan).....	14
2.11.2 Uji Parsial	15
2.12 Uji Asumsi Residual	16
2.12.1 Uji Identik.....	16
2.12.2 Uji Independen.....	17

2.12.3 Uji Normalitas Residual	17
2.13 Faktor – Faktor yang Mempengaruhi DBD.....	18
BAB III METODE PENELITIAN	
3.1 Sumber Data	23
3.2 Variabel Penelitian.....	23
3.3 Langkah Penelitian.....	25
3.4 Diagram Alir.....	26
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Karakteristik Kasus DBD dan Faktor – Faktor yang Mempengaruhi.....	29
4.2 Pola Hubungan antara Jumlah Kasus Demam Berdarah Dengue dengan Faktor-Faktor yang Mempengaruhi	34
4.3 Analisis Faktor	36
4.4 Pemodelan Kasus Demam Berdarah Dengue (DBD) Menggunakan Regresi Semiparametrik Spline	38
4.4.1 Pemodelan Regresi Semiparametrik Spline	38
4.4.2 Pemilihan Titik Knot Optimal	38
4.4.3 Pemodelan Regresi Semiparametrik Spline Menggunakan Titik Knot Optimal	45
4.5 Uji Signifikansi Parameter	46
4.5.1 Uji Serentak.....	46
4.5.2 Uji Parsial.....	47
4.6 Uji Asumsi Residual.....	48
4.4.1 Uji Identik	48
4.4.2 Uji Independen	49
4.4.3 Uji Distribusi Normal	50
4.7 Interpretasi Hasil Pemodelan Regresi Semiparametrik Spline.....	51
BABV KESIMPULAN DAN SARAN	
5.1 Kesimpulan	55
5.2 Saran	56
DAFTAR PUSTAKA	57
LAMPIRAN	61

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir Penelitian	27
Gambar 4.1	Diagram Batang Angka kesakitan DBD per 100.000 penduduk pada setiap kecamatan.....	30
Gambar 4.2	<i>Scatterplot</i> antara Jumlah kasus DBD dengan variabel x_1	34
Gambar 4.3	<i>Scatterplot</i> antara Jumlah Kasus DBD dengan variabel t_1, t_2, t_3, t_4 dan t_5	35
Gambar 4.4	<i>Scatterplot</i> antara Residual dengan \hat{y}	48
Gambar 4.5	<i>Scatterplot</i> antara Residual dan Order	49
Gambar 4.6	Plot ACF dari Residual	50
Gambar 4.7	Uji Normalitas Residual	51

DAFTAR TABEL

Tabel 2.1	<i>Analysis of Variance (ANOVA)</i>	15
Tabel 3.1	Variabel Penelitian	25
Tabel 3.2	Struktur Data Variabel Penelitian	25
Tabel 4.1	Karakteristik banyaknya kasus DBD dan Faktor-Faktor yang Mempengaruhi	30
Tabel 4.2	Uji Kecukupan Data dan Homogentias Varians	36
Tabel 4.3	<i>Eigenvalues</i>	37
Tabel 4.4	<i>Loading Factor</i> Penyebab DBD	37
Tabel 4.5	Nilai Knot dan Nilai GCV Untuk Satu Titik Knot...	39
Tabel 4.6	Nilai Knot dan Nilai GCV Untuk Dua Titik Knot ...	40
Tabel 4.7	Nilai Knot dan Nilai GCV Untuk Tiga Titik Knot...	41
Tabel 4.8	Nilai Knot dan Nilai GCV Untuk Kombinasi Titik Knot	42
Tabel 4.9	Hasil Kriteria Keباikan Model.....	44
Tabel 4.10	ANOVA	46
Tabel 4.11	Uji Parsial (Halaman ini sengaja dikosongkan).....	47
Tabel 4.12	ANOVA Uji Glejser.....	49

BAB I

PENDAHULUAN

1.1 Latar Belakang

Demam berdarah merupakan penyakit demam akut yang disebabkan oleh virus dengue, yang masuk ke peredaran darah manusia melalui gigitan nyamuk dari genus *Aedes*, misalnya *Aedes aegypti* atau *Aedes albopictus* (Kritina dkk, 2004). Penyakit ini endemik di Afrika, Amerika, Mediterania Timur, Asia Tenggara dan Pasifik Barat. Bahkan penyakit ini merupakan penyakit utama di Asia Tenggara dan Pasifik Barat (WHO, 1999). Menurut penjelasan Departemen Kesehatan (2005), Demam Berdarah Dengue (DBD) merupakan penyakit menular yang sering mengakibatkan Kejadian Luar Biasa dikarenakan perjalanan penyakitnya cepat dan dapat menyebabkan kematian secara singkat. Hal tersebut disebabkan nyamuk penular DBD terdapat hampir di seluruh pelosok Indonesia. Selain itu letak geografis Indonesia yang dilalui garis khatulistiwa dan beriklim tropis menyebabkan nyamuk *Aedes* dapat berkembang biak dengan pesat.

Kriteria WHO (1997) mengklasifikasikan infeksi dengue menjadi Demam Dengue, Demam Berdarah Dengue dan Sindrom Syok Dengue. Pada umumnya penderita mempunyai risiko kematian yang lebih besar ketika sudah mengalami fase Sindrom Syok Dengue (SSD). Fase tersebut merupakan keadaan dimana penderita DBD mengalami syok (renjatan) yang berkepanjangan yang disebabkan terjadinya kebocoran plasma. Sebagian besar penderita penyakit DBD di Indonesia adalah anak berusia kurang dari 15 tahun, namun dapat juga menyerang orang dewasa (Kementerian Kesehatan Republik Indonesia, 2011).

Data Profil Kesehatan Indonesia yang dikeluarkan oleh Kemenkes menunjukkan pada tahun 2010 jumlah kematian akibat kasus DBD di Indonesia sekitar 1.358 orang, hal ini menyebabkan Indonesia menduduki urutan tertinggi banyaknya kasus DBD di ASEAN. Selain itu DBD merupakan kasus kedua terbesar pada

penyakit rawat inap di rumah sakit setelah diare dan gastroenteritis. Jumlah kejadian DBD di Indonesia mencapai 156.086 kasus dimana *Incident Rate* (IR) atau angka kesakitan mencapai 65,7 kasus per 100.000 penduduk. Jumlah tersebut berada di atas target kesehatan nasional yaitu 55 kasus per 100.000 penduduk. Sementara itu *Case Fatality Rate* (CFR) atau persentase jumlah kematian dibandingkan jumlah kasus periode tahun 2006 sampai 2010 nilainya cenderung menurun setiap tahunnya dan hanya mengalami peningkatan pada tahun 2009.

Provinsi Jawa Timur menduduki peringkat pertama dalam banyaknya kasus DBD yaitu sebanyak 26.020 kasus dan menyebabkan kematian sebanyak 233 orang. Hal ini sangat mengkhawatirkan dan perlu dilakukan penanganan yang tepat mengingat kebanyakan penderita merupakan anak-anak. Kabupaten/Kota Madiun merupakan daerah di Jawa Timur yang diduga persebaran penyakitnya cukup mengkhawatirkan. Kasus DBD Kabupaten Madiun relatif besar dimana pada tahun 2010 terjadi 411 kasus dan menyebabkan kematian sebanyak 6 orang, sedangkan Kota Madiun terjadi 87 kasus dan menyebabkan kematian sebanyak 1 orang. Jika dibandingkan dengan target kesehatan nasional dimana target IR kurang dari 55 kasus per 100.000 penduduk dan CFR kurang dari 1 persen maka angka kesakitan (IR) DBD yang terjadi di Kabupaten/Kota Madiun adalah 53,42 dan 42,90 kasus per 100.000 penduduk dan berdasarkan jumlah kematian (CFR) sebesar 1,45 persen dan 1,149 persen. Hal ini mengindikasikan bahwa kasus DBD di Kabupaten/Kota Madiun cukup mengkhawatirkan dikarenakan berdasarkan jumlah kematian di Kabupaten/Kota Madiun melebihi target kesehatan nasional yaitu lebih dari 1 persen.

Penelitian tentang DBD sudah dilakukan sebelumnya seperti Ratnasari (2009) membahas faktor resiko yang mempengaruhi terjangkitnya penyakit DBD di daerah endemis. Penelitian Akhir (2013) tentang pemodelan jumlah kasus DBD di Kabupaten Bojonegoro menggunakan *Geographically Weighted Poisson Regression* (GWPR). Penelitian tersebut dalam pemodelannya

menggunakan variabel dalam menentukan faktor-faktor penyebab DBD, beberapa diantaranya adalah angka bebas jentik, persentase penduduk miskin, curah hujan, persentase rumah sehat, kepadatan penduduk dan persentase institusi yang dibina kesehatan lingkungannya dimana variabel tersebut juga digunakan dalam penelitian ini.

Selain menggunakan GWPR, pemodelan kasus DBD dapat menggunakan regresi semiparametrik. Regresi semiparametrik adalah metode statistik yang digunakan untuk mengetahui pola hubungan variabel respon dengan prediktor dimana kurva regresi diklasifikasikan menjadi 2 komponen yakni komponen parametrik (fungsi diketahui) dan komponen nonparametrik (fungsi tidak diketahui). Salah satu bentuk regresi semiparametrik yang sering digunakan adalah regresi semiparametrik spline. Spline adalah salah satu jenis *piecewise* polinomial, yaitu polinomial yang memiliki sifat tersegmen (Budiantara, 2006). Hal ini sesuai dengan realita dimana tidak semua variabel prediktor memiliki pola hubungan tertentu dengan responnya. Penelitian yang menggunakan pemodelan regresi semiparametrik spline pernah dilakukan oleh Asmin (2010) dalam bidang pendidikan, selain itu oleh Marina (2013) yang membahas tentang kriminalitas di Jawa Timur.

Dalam penelitian ini menggunakan regresi semiparametrik spline dalam melakukan pemodelan jumlah kasus DBD di Kabupaten/Kota Madiun. Variabel-variabel yang digunakan perlu direduksi terlebih dahulu menggunakan analisis faktor sehingga tidak terdapat korelasi antar variabel prediktor. Pemodelan dilakukan dengan menggunakan variabel prediktor hasil dari analisis faktor sehingga model yang dihasilkan memiliki interpretasi yang lebih sesuai. Diharapkan hasil dari penelitian ini dapat memberikan gambaran tentang DBD serta mengetahui faktor-faktor yang berpengaruh signifikan di Kabupaten/Kota Madiun. Selain itu dapat menjadi salah satu pertimbangan bagi instansi yang terkait dalam merumuskan kebijakan untuk

mengurangi terjadinya kasus DBD di Kabupaten/Kota Madiun pada khususnya dan Indonesia pada umumnya.

1.2 Perumusan Masalah

Berdasarkan latar belakang, permasalahan yang akan diselesaikan dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik kasus DBD dan faktor-faktor yang mempengaruhi di Kabupaten/Kota Madiun?
2. Apa saja faktor-faktor yang dihasilkan dari analisis faktor yang digunakan dalam pemodelan kasus DBD di Kabupaten/Kota Madiun?
3. Bagaimana bentuk pemodelan kasus DBD Kabupaten/Kota Madiun menggunakan pendekatan regresi semiparametrik spline?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik kasus DBD dan faktor-faktor yang mempengaruhi di Kabupaten/Kota Madiun.
2. Mengidentifikasi faktor-faktor yang dihasilkan dari analisis faktor yang digunakan dalam pemodelan kasus DBD di Kabupaten/Kota Madiun.
3. Mengestimasi bentuk pemodelan kasus DBD Kabupaten/Kota Madiun menggunakan pendekatan regresi semiparametrik spline.

1.4 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah sebagai berikut.

1. Dapat memberikan informasi dan pengetahuan mengenai penyakit DBD khususnya faktor-faktor yang mempunyai pengaruh signifikan dan dapat dijadikan salah satu

- pertimbangan dalam menentukan arah kebijakan di bidang kesehatan.
2. Sebagai wawasan dalam penerapan ilmu statistika dalam kehidupan nyata khususnya mengenai penggunaan metode spline dalam memodelkan kasus DBD serta dapat dijadikan salah satu referensi dalam pengembangan metode spline kedepannya.

1.5 Batasan Masalah

Penelitian ini dibatasi dengan menggunakan metode pemilihan titik knot optimal yaitu GCV (*Generalized Cross Validation*). Titik knot yang digunakan menggunakan satu, dua dan tiga serta kombinasi dari titik knot tersebut.

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistik deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna (Walpole, 1995). Statistika deskriptif memberikan informasi hanya mengenai data yang dipunyai dan sama sekali tidak menarik kesimpulan apapun tentang sekumpulan data yang lebih besar. Penyajian data yang dilakukan secara deskriptif misalnya dalam bentuk tabel, diagram, grafik, serta besaran-besaran lainnya.

2.2 Uji Kecukupan Data

Pengujian kecukupan data dapat dilakukan dengan menggunakan uji *Kaiser-Meyer-Olkin* (KMO). Uji kecukupan data dilakukan untuk mengetahui apakah data yang digunakan sudah mencukupi untuk dilakukan pempfaktoran atau tidak. Hipotesis yang digunakan adalah sebagai berikut.

H_0 : data cukup untuk dilakukan pempfaktoran

H_1 : data tidak cukup untuk dilakukan pempfaktoran

Statistik uji yang digunakan disajikan dalam persamaan (2.1).

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (2.1)$$

Dimana :

$i = 1, 2, 3, \dots, p$ dan $j = 1, 2, \dots, p$ $i \neq j$

r_{ij} = koefisien korelasi antara variabel i dan j

a_{ij} = koefisien korelasi parsial antara variabel i dan j

Menurut Hair dkk (2010), sampel data dikatakan cukup untuk dilakukan analisis faktor adalah jika nilai KMO lebih besar daripada 0,5.

2.3 Uji Homogenitas Varians

Uji homogenitas varians dilakukan untuk mengetahui adanya hubungan antarvariabel. Pengujian ini dapat dilakukan dengan menggunakan uji Bartlett. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \rho = I$$

$$H_1: \rho \neq I$$

Persamaan (2.2) menyajikan statistik uji yang digunakan, yaitu sebagai berikut.

$$\chi_{hitung}^2 = -\left(N - 1 - \frac{2p + 5}{6}\right) \sum \sum_{i < j} r_{ij}^2 \quad (2.2)$$

Dimana :

r_{ij} = koefisien korelasi antara variabel i dan j

N = banyaknya data

p = banyaknya variabel

Daerah penolakannya adalah jika $\chi_{hitung}^2 > \chi_{\alpha, p(p-1)/2}^2$ atau $p - value < \alpha$ yang berarti terdapat korelasi yang signifikan antar variabel.

2.4 Analisis Faktor

Analisis faktor dikembangkan pertama kali pada awal abad ke 20 oleh Karl Pearson, Charles Spearman dll. Analisis faktor merupakan analisis statistika yang bertujuan untuk mengelompokkan variabel-variabel yang banyak dan berbeda menjadi kumpulan variabel yang lebih kecil didasarkan pada kesamaan karakteristik variabel tersebut. Analisis faktor dapat menentukan variabel-variabel yang saling berkorelasi dengan kualitas random yang disebut faktor (Johnson dan Wichern, 2002). Menurut Wiersma (1986), analisis faktor memiliki 2 fungsi analisis, yaitu sebagai *explanatory analysis* dan *confirmatory analysis*. Persamaan (2.3) menyajikan model analisis faktor yang terbentuk, yaitu sebagai berikut.

$$\begin{aligned}
X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1j}F_j + \varepsilon_1 \\
X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2j}F_j + \varepsilon_2 \\
&\vdots && \vdots && \ddots && \vdots \\
X_i - \mu_i &= l_{i1}F_1 + l_{i2}F_2 + \dots + l_{ij}F_j + \varepsilon_i
\end{aligned} \tag{2.3}$$

Dimana :

μ_i = rata-rata variabel ke-i

F_j = *common factor* ke-j

l_{ij} = *loading faktor* ke-j pada variabel ke-i

ε_i = *specific factor* ke-i

$i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, k$

Dalam notasi matriks disajikan dalam persamaan (2.4).

$$X_{(ix1)} - \mu_{(ix1)} = l_{(ixj)}F_{(jx1)} + \varepsilon_{(ix1)} \tag{2.4}$$

2.5 Analisis Regresi

Analisis regresi merupakan suatu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor. Dengan demikian analisis regresi merupakan suatu metode inferensi statistik untuk suatu fungsi regresi atau kurva regresi (Eubank, 1999). Selain itu regresi dapat digunakan untuk melakukan peramalan atau prediksi (*forecasting*). Dalam analisis regresi akan ditentukan hubungan fungsional yang diharapkan berlaku untuk populasi berdasarkan data sampel yang diambil dari populasi yang bersangkutan. Secara umum model regresi mempunyai 3 (tiga) tujuan yaitu menjelaskan pola hubungan sebab akibat yang terjadi antara peubah respon dengan peubah bebas, mengetahui kontribusi relatif setiap peubah bebas untuk menjelaskan peubah respon, dan memprediksi nilai peubah respon untuk beberapa nilai peubah bebas tertentu (Aunuddin, 2005).

Analisis regresi dibedakan menjadi 3 pendekatan dalam mengestimasi model yaitu pendekatan regresi parametrik, regresi

nonparametrik dan regresi semiparametrik. Regresi parametrik digunakan untuk mengestimasi model ketika bentuk kurva regresi diketahui. Regresi nonparametrik digunakan ketika bentuk kurva regresi tidak diketahui. Sedangkan regresi semiparametrik digunakan ketika sebagian pola kurva regresi diketahui dan sebagian tidak diketahui.

2.6 Regresi Parametrik

Regresi parametrik merupakan metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel prediktor dengan respon dengan asumsi bentuk kurva diketahui. Pendekatan model regresi parametrik memiliki sifat yang sangat baik dari pandangan statistika inferensi, seperti parsimoni, mudah interpretasinya, estimatornya tidak bias, efisien, konsisten, sehingga metode ini sangat populer dan banyak disukai oleh berbagai kalangan (Budiantara, 2009). Secara umum, model regresi parametrik linier sejumlah k variabel prediktor disajikan dalam persamaan (2.5).

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (2.5)$$

Dimana y_i merupakan variabel responnya, $\beta_0, \beta_1, \dots, \beta_k$ merupakan parameter dari variabel prediktornya (x). Sedangkan ε_i merupakan *error* random dari model yang identik, independen dan berdistribusi normal dengan mean nol dan varian σ^2 . Bentuk regresi parametrik dapat dirumuskan dalam bentuk matrik yaitu dalam persamaan (2.6).

$$y_i = X\beta + \varepsilon \quad (2.6)$$

Dengan bentuk matriks yang diberikan yaitu.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ 1 & x_{13} & x_{23} & \dots & x_{k3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Salah satu metode yang digunakan dalam penghitungan estimasi parameter pada regresi parametrik adalah dengan menggunakan *Ordinary Least Square* (OLS) yaitu dengan meminimumkan jumlah kuadrat error.

2.7 Regresi Nonparametrik

Regresi nonparametrik merupakan salah satu model regresi dan mulai dikenal sekitar abad 19, tepatnya pada tahun 1857 (Hardle, 1990). Secara visual, model regresi ini memiliki bentuk kurva yang tidak dapat diidentifikasi. Bentuk kurva dari regresi nonparametrik hanya diasumsikan *smooth* (halus) dalam arti termuat dalam suatu ruang fungsi tertentu. Pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi faktor subyektifitas peneliti (Eubank, 1988). Model regresi nonparametrik secara umum disajikan dalam persamaan (2.7).

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.7)$$

Dimana y_i merupakan variabel responnya, kurva regresi $f(t_i)$ yang didekati dengan fungsi spline berorde p dengan titik knot k_1, k_2, \dots, k_r dan ε_i adalah *error* random yang memenuhi asumsi identik, independen dan berdistribusi normal dengan mean nol dan varian σ^2 . Bentuk persamaan regresi nonparametrik secara lengkap disajikan dalam persamaan (2.8).

$$y_i = \sum_{j=0}^p \beta_j t_i^j + \sum_{j=1}^r \beta_{p+j} (t_i - k_j)_+^p + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.8)$$

Dengan fungsi $(t_i - k_j)_+^p$ yang diberikan yaitu.

$$(t_i - k_j)_+^p = \begin{cases} (t_i - k_j)^p, & t_i \geq k_j \\ 0 & , t_i < k_j \end{cases}$$

Nilai k_j merupakan titik knot yang memperlihatkan pola perubahan dari fungsi pada sub interval yang berbeda, sedangkan nilai p merupakan derajat polinomial. Kurva polinomial derajat satu merupakan kurva linier, derajat dua adalah kurva kuadratik sedangkan derajat tiga disebut kurva kubik.

2.8 Regresi Semiparametrik

Regresi semiparametrik adalah metode analisis statistik yang merupakan gabungan dari regresi parametrik dan nonparametrik (Ruppert, 2003). Dengan kata lain regresi semiparametrik digunakan untuk mengetahui pola hubungan antara variabel respon dan prediktor dimana sebagian pola data diketahui bentuknya dan sebagian lainnya tidak. Jika terdapat data berpasangan (x_i, y_i) dan t_i yang diasumsikan telah mengikuti model regresi semiparametrik, maka persamaan yang terbentuk disajikan dalam persamaan (2.9).

$$y_i = x_i' \beta + f(t_i) + \varepsilon_i, \quad \varepsilon_i [a, b] \quad i = 1, 2, \dots, n \quad (2.9)$$

Salah satu cara dalam mengestimasi parameter dalam regresi semiparametrik dengan menggunakan metode OLS yaitu dengan memasukkan komponen nonparametrik pada matrik variabel prediktor.

2.9 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola kurva pada interval yang berlainan (Budiantara, 2006). Oleh karena itu, titik knot hanya dapat ditemukan pada data yang bentuk kurvanya tidak membentuk pola tertentu. Wahba (1990) memberikan suatu metode yang baik untuk memilih titik knot optimal yaitu metode Generalized Cross Validation (GCV). Secara teoritis, metode GCV mempunyai sifat optik optimal asimptotik (Ryan, 1997). Kebanyakan model spline terbaik dengan titik knot optimal

diperoleh dari nilai GCV terkecil. Rumus dari metode GCV didefinisikan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (2.10)$$

Berdasarkan hasil dari persamaan (2.10), nilai K_1, K_2, \dots, K_r merupakan titik knot sedangkan matriks $A(K_1, K_2, \dots, K_r)$ didapatkan dari persamaan (2.11) berikut.

$$\hat{y} = A(K_1, K_2, \dots, K_r)y \quad (2.11)$$

Persamaan (2.12) menjelaskan nilai $MSE(K_1, K_2, \dots, K_r)$ yaitu sebagai berikut.

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (2.12)$$

2.10 Kriteria Kebaikan Model

Penggunaan kriteria dalam menentukan model terbaik dalam regresi semiparametrik spline tidak hanya memperhatikan nilai GCV minimum. Kriteria lain yang perlu diperhatikan adalah menggunakan nilai R^2 , MSE dan jumlah parameter yang digunakan. R^2 merupakan koefisien determinasi yaitu ukuran proporsi keragaman atau variansi total di sekitar nilai tengah yang dapat dijelaskan oleh model regresi tersebut (Drapper dan Smith, 1996). Semakin besar ukuran sampel (n), maka nilai R^2 cenderung semakin kecil. Jika nilai R^2 semakin besar atau mendekati satu, maka model semakin tepat. Persamaan (2.13) menjelaskan rumus untuk menghitung nilai R^2 .

$$R^2 = \frac{\sum_i^n (\hat{y}_i - \bar{y})^2}{\sum_i^n (y_i - \bar{y})^2} \quad (2.13)$$

Selain menggunakan nilai R^2 , kriteria lain yang digunakan dalam menentukan model terbaik dengan memperhatikan nilai MSE. Semakin kecil MSE semakin baik, hal ini dikarenakan

penyimpangan yang terjadi semakin kecil. Perhitungan MSE disajikan dalam persamaan (2.14)

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.14)$$

Jumlah parameter merupakan banyaknya parameter yang digunakan dalam pemodelan. Jumlah parameter yang dihasilkan bergantung pada banyak variabel dan kombinasi titik knot yang digunakan. Semakin sedikit jumlah parameter yang dihasilkan maka semakin baik model yang digunakan dengan asumsi nilai R^2 yang dihasilkan cukup besar dan MSE yang kecil.

2.11 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter model dilakukan untuk menentukan variabel prediktor yang memberikan pengaruh yang signifikan dalam model. Adapun pengujian yang dilakukan adalah sebagai berikut

2.11.1. Uji Serentak (simultan)

Uji serentak merupakan pengujian signifikansi parameter β secara bersama-sama terhadap variabel respon dengan menggunakan ANOVA dan uji F. Hipotesis yang digunakan adalah sebagai berikut.

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

$$H_1: \text{Minimal ada satu } \beta_j \neq 0 ; j = 1, 2, \dots, k$$

Persamaan (2.15) menyajikan statistik uji yang digunakan, yaitu sebagai berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \quad (2.15)$$

Metode perhitungan tabel *Analysis of Variance* (ANOVA) disajikan dalam Tabel 2.1.

Tabel 2.1 Analysis of Variance (ANOVA)

Sumber Variasi	Derajat Bebas (Df)	Jumlah Kuadrat	Rata-rata Kuadrat	F hitung
Regresi	k	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SSR}{k}$	$\frac{MSR}{MSE}$
Error	$n - k - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SSE}{n - k - 1}$	
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Daerah penolakannya adalah tolak H_0 jika F_{hitung} lebih besar daripada $F_{\alpha;(k,n-k-1)}$ atau p - value kurang dari α dan mengindikasikan terdapat minimal satu parameter yang tidak sama dengan nol atau dapat dikatakan signifikan.

2.11.2. Uji Parsial

Uji parsial merupakan pengujian untuk mengetahui parameter yang berpengaruh signifikan secara individu terhadap model. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0 ; j = 1, 2, \dots, k$$

Statistik uji yang digunakan disajikan dalam persamaan (2.16).

$$t_{hitung} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (2.16)$$

Daerah penolakan: tolak H_0 jika $|t_{hitung}|$ lebih besar daripada $t_{tabel} (t_{\frac{\alpha}{2}, n-k})$, dimana n adalah jumlah pengamatan dan k adalah banyaknya parameter.

2.12 Uji Asumsi Residual

Pemeriksaan asumsi residual harus dipenuhi setelah didapatkan model regresi terbaik. Asumsi residual yang digunakan sama dengan asumsi residual pada regresi parametrik. Berikut ini asumsi residual yang harus dipenuhi.

2.12.1 Uji Identik

Uji asumsi identik digunakan untuk melihat homogenitas dari variansi residual. Jika terjadi pelanggaran dalam asumsi residual identik disebut heteroskedastisitas yang mengakibatkan kerugian bagi efisiensi estimator (Eubank, 1988). Salah satu cara mendeteksi adanya kasus heteroskedastisitas residual dengan membuat *scatterplot* antara residual dengan estimasi variabel respon. Jika hasil sebaran plot menunjukkan suatu pola yang membentuk corong atau garis melengkung, maka ada indikasi adanya heteroskedastisitas.

Selain itu cara lain untuk mendeteksi adanya heteroskedastisitas adalah menggunakan Uji *Glejser*. Pengujian ini dilakukan dengan cara meregresikan absolut dari residual dengan variabel prediktornya (Gujarati, 2006). Hipotesis yang digunakan adalah sebagai berikut.

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

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

Statistik uji yang digunakan dijelaskan dalam persamaan (2.17).

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} = \frac{\frac{\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2}{k-1}}{\frac{\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2}{n-k}} \quad (2.17)$$

Daerah penolakannya adalah tolak H_0 jika nilai $F_{hitung} > F_{tabel} (F_{\alpha; (k-1, n-k)})$ atau $p - value < \alpha$.

2.12.2 Uji Independen

Pengujian independensi residual bertujuan untuk mengetahui korelasi antar residual apakah sama dengan nol atau tidak. Salah satu cara mendeteksi adanya kasus autokorelasi pada residual dengan membuat *scatterplot* antara residual dengan order. Jika hasil sebaran plot tidak membentuk pola tertentu atau tersebar secara acak maka dapat diasumsikan residual telah memenuhi asumsi independen. Selain itu dapat dideteksi dengan membuat plot *Autocorrelation Function (ACF)*. Jika dalam plot ACF tidak terdapat lag yang keluar dari batas signifikansi, maka terdapat indikasi tidak adanya kasus autokorelasi atau memenuhi asumsi independen. Persamaan (2.18) menjelaskan rumus yang digunakan dalam menghitung ACF.

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{i=1}^{n-k} (e_i - \bar{e})(e_{i+k} - \bar{e})}{\sum_{i=1}^n (e_i - \bar{e})^2}, k = 1, 2, 3 \dots \quad (2.18)$$

Dimana $\hat{\rho}_k$ merupakan korelasi diantara e_i dengan e_{i+k} dan k merupakan lag yang digunakan. Nilai tersebut harus berada diantara batas atas dan bawah signifikansi ACF agar tidak diindikasikan kasus autokorelasi. Nilai batas atas dan bawah dari signifikansi ACF disajikan dalam persamaan (2.19).

$$\text{Batas atas} = t_{(1-\frac{\alpha}{2}, n-1)} SE(\hat{\rho}_k) \quad (2.19)$$

$$\text{Batas bawah} = t_{(\frac{\alpha}{2}, n-1)} SE(\hat{\rho}_k)$$

2.12.3 Uji Normalitas Residual

Pengujian normalitas residual dilakukan untuk melihat apakah residual mengikuti distribusi normal atau tidak. Pengujian ini dapat dilakukan melihat hasil dari *normal probability plot residual*. Dikatakan memenuhi asumsi normal apabila plot tersebar mengikuti garis distribusi normal atau membentuk sudut 45°. Selain itu dapat juga menggunakan uji Kolmogorov-

Smirnov. Hipotesis yang digunakan untuk pengujian normalitas residual adalah sebagai berikut.

H_0 : Residual mengikuti distribusi normal

H_1 : Residual tidak mengikuti distribusi normal

Persamaan (2.20) menjelaskan statistik uji yang digunakan yaitu menggunakan uji *Kolmogorov-Smirnov*.

$$z_{hitung} = \sup_x |F_n(x) - F_0(x)| \quad (2.20)$$

Daerah penolakannya adalah tolak H_0 jika $z_{hitung} > z_\alpha$ atau $p - value < \alpha$.

2.13 Faktor - Faktor yang Mempengaruhi DBD

Permasalahan DBD menjadi semakin kompleks apabila tidak dapat didefinisikan faktor-faktor yang mempengaruhi persebaran penyakit tersebut. Berbagai penelitian sudah dilakukan untuk mengetahui faktor-faktor yang diduga mempengaruhi persebaran DBD. Salah satunya adalah Akhir (2013) yang menggunakan beberapa variabel untuk dijadikan faktor penyebab DBD yang kemudian akan digunakan dalam penelitian ini untuk memodelkan banyaknya kasus DBD di Kabupaten/Kota Madiun. Adapun faktor-faktornya adalah sebagai berikut.

1. Persentase rumah atau bangunan bebas jentik nyamuk *Aedes aegypti* atau Angka Bebas Jentik

Angka bebas jentik didapatkan melalui kegiatan pemeriksaan jentik secara berkala (PJB) pada rumah dan kontainer (tempat penampungan air) di dalam dan diluar rumah yang positif atau negatif terhadap jentik (Yuswulandary, 2008). ABJ dapat dijadikan tolok ukur upaya pemberantasan vektor nyamuk dan tingkat partisipasi masyarakat dalam pencegahan penyakit DBD (Kemenkes RI, 2011). Salah satu upaya dalam meminimumkan ABJ yaitu dengan melakukan metode Pemberantasan Sarang Nyamuk (PSN) melalui 3M plus

(Menguras, Menutup dan Mengubur) plus menabur larvasida, penyebaran ikan pada tempat penampungan air serta kegiatan lain yang dapat mencegah perkembangbiakan nyamuk *Aedes*. Target Nasional untuk ABJ pada tahun 2010 adalah lebih dari 60%.

2. Persentase penduduk miskin

Kemiskinan adalah suatu kondisi dimana tidak memiliki akses ke prasarana dan sarana dasar lingkungan yang memadai, kualitas perumahan dan pemukiman yang jauh di bawah standar kelayakan serta mata pencaharian yang tidak menentu yang mencakup seluruh multidimensi, yaitu dimensi politik, dimensi sosial, dimensi lingkungan, dimensi ekonomi dan dimensi aset. Kemiskinan menjadi hambatan besar dalam pemenuhan kebutuhan terhadap makanan yang sehat sehingga dapat melemahkan daya tahan tubuh yang dapat berdampak pada kerentanan untuk terserang penyakit tertentu (Kemenkes RI, 2011). Menurut BPS, definisi penduduk miskin adalah mereka yang hanya berpenghasilan Rp 233 ribu per bulan, atau Rp 7.000 per hari. Bank Dunia menetapkan bahwa penduduk miskin adalah orang yang memiliki penghasilan di bawah \$2 per hari.

3. Curah hujan.

Curah hujan adalah jumlah air yang jatuh di permukaan tanah datar selama periode tertentu yang diukur dengan satuan tinggi (mm^3) di atas permukaan horizontal bila tidak menguap, meresap dan mengalir. Jumlah curah hujan dihitung secara aktif oleh radar cuaca dan secara pasif oleh satelit cuaca. Curah hujan merupakan faktor penentu tersedianya tempat perindukan bagi nyamuk vektor. Berdasarkan penelitian yang dilakukan oleh Iriani (2012) di Palembang tentang pengaruh curah hujan dan peningkatan kasus DBD menghasilkan kesimpulan, salah satunya adalah tingkat curah hujan berkorelasi dengan jumlah kasus DBD dimana korelasi paling kuat terjadi pada bulan puncak curah hujan. Selain itu hujan dengan intensitas yang cukup akan menimbulkan genangan air di tempat penampungan air sekitar rumah maupun di cekungan dan menjadi sarang berkembangbiak

bagi nyamuk *Aedes*. Curah hujan yang besar menyebabkan genangan air ini melimpah sehingga larva atau pupa nyamuk tersebar ke tempat-tempat lain yang sesuai atau tidak sesuai untuk menyelesaikan siklus kejadian timbulnya atau menularnya penyakit (Hidayati, 2008).

4. Persentase rumah atau bangunan sehat.

Rumah adalah bangunan gedung yang berfungsi sebagai tempat tinggal yang layak huni, sarana pembinaan keluarga, cerminan harkat dan martabat penghuninya, serta aset bagi pemiliknya. Menurut UU RI No.36 Tahun 2009 Tentang Kesehatan, sehat adalah keadaan sehat, baik secara fisik, mental, spritual maupun sosial yang memungkinkan setiap orang untuk hidup produktif secara sosial dan ekonomis. Menurut Mubarak dan Chayatin (2009) rumah sehat adalah adalah sebuah rumah yang dekat dengan air bersih, berjarak lebih dari seratus meter dari tempat pembuangan sampah, dekat dengan sarana pembersihan, serta berada di tempat dimana air hujan dan air kotor tidak menggenang.

5. Kepadatan penduduk

Kepadatan penduduk diukur dari jumlah penduduk di suatu wilayah dibandingkan dengan luas wilayah tersebut per km². Wilayah yang kepadatan penduduknya tinggi seringkali terjadi masalah sosial diantaranya kurang peduli terhadap kebersihan lingkungan sekitar. Hal itu menjadi awal dari perkembangbiakan jentik nyamuk *Aedes*. Selain itu semakin padat penduduk, semakin mudah nyamuk *Aedes* menularkan virusnya dari satu orang ke orang lainnya. Pertumbuhan penduduk yang tidak memiliki pola tertentu dan urbanisasi yang tidak terencana serta tidak terkontrol merupakan salah satu faktor yang berperan dalam munculnya kembali kejadian luar biasa penyakit DBD (WHO, 2000).

6. Persentase institusi yang dibina kesehatan lingkungannya.

Penyebaran penyakit seperti DBD tidak hanya terjadi di lingkungan perumahan tetapi juga di tempat umum. Tempat umum yang sering dikunjungi tetapi tidak terjaga kebersihan lingkungannya bisa menjadi sarang perkembangbiakan nyamuk *Aedes* sekaligus penyebaran penyakit DBD. Oleh karena itu, Kementerian Kesehatan melalui Dinas Kesehatan di masing-masing wilayah perlu melakukan pembinaan kesehatan lingkungan terhadap institusi dan tempat umum untuk meminimalisir dampak penyebaran tersebut. Institusi yang dibina kesehatan lingkungannya meliputi sarana kesehatan, sarana pendidikan, sarana ibadah, perkantoran, instalasi pengolahan air minum, hotel, pasar, tempat wisata, terminal dan stasiun (Dinas Kesehatan Provinsi Jawa Timur, 2009).

(Halaman ini sengaja dikosongkan)

BAB III

METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapat dari Dinas Kesehatan dan Badan Pusat Statistik Kabupaten/Kota Madiun tahun 2010. Rinciannya adalah data banyaknya kasus DBD, persentase rumah atau bangunan bebas jentik nyamuk atau Angka Bebas Jentik, persentase penduduk miskin, persentase rumah atau bangunan sehat, kepadatan penduduk dan persentase institusi dibina kesehatan lingkungannya didapatkan dari Dinas Kesehatan sedangkan data curah hujan berasal dari Badan Pusat Statistik. Unit pengamatan yang digunakan adalah kecamatan yang terdiri dari 18 kecamatan dengan rincian Kabupaten Madiun terdapat 15 kecamatan dan Kota Madiun terdapat 3 kecamatan.

3.2 Variabel Penelitian

Pada penelitian ini variabel yang diteliti dibedakan menjadi dua yaitu variabel respon dan variabel prediktor. Pada variabel respon adalah banyaknya kasus DBD di Kabupaten/Kota Madiun. Variabel ini diwakilkan dengan *Incident Rate (IR)* yaitu angka kesakitan DBD yang terjadi per 100.000 penduduk dan juga merupakan salah satu penghitungan dalam menetapkan target kesehatan nasional. Terdapat enam variabel prediktor yang diduga mempengaruhi banyaknya kasus DBD tersebut. Berdasarkan hasil sebaran plot pada Bab 4.2, variabel tersebut dibedakan menjadi dua yaitu variabel x yang mengandung unsur parametrik dan variabel t yang mengandung unsur nonparametrik. Variabel prediktor yang diduga berpengaruh dan digunakan pada penelitian ini adalah sebagai berikut (per kecamatan).

- a. Variabel x_1 yaitu persentase rumah atau bangunan bebas jentik nyamuk *Aedes aegypti* atau angka bebas jentik.

Angka bebas jentik merupakan persentase dari jumlah rumah yang tidak ditemukan jentik dibanding dengan jumlah rumah yang diperiksa. Apabila angka bebas jentik suatu daerah tertentu rendah, maka kemungkinan penduduk daerah tersebut untuk terkena DBD adalah lebih besar dibanding daerah lain yang angka bebas jentiknya lebih besar.

- b. Variabel t_1 yaitu persentase penduduk miskin.
Penduduk miskin adalah penduduk yang berada di bawah garis kemiskinan.
- c. Variabel t_2 yaitu curah hujan.
Curah hujan adalah jumlah air yang jatuh di permukaan tanah datar selama periode tertentu yang diukur dengan satuan tinggi (mm^3). Pada penelitian ini curah hujan yang didapatkan merupakan jumlah curah hujan selama satu tahun dibagi 12 yaitu banyaknya bulan.
- d. Variabel t_3 yaitu persentase rumah atau bangunan sehat.
Rumah sehat adalah bangunan rumah tinggal yang memenuhi syarat kesehatan diantaranya memiliki jamban sehat, tempat pembuangan sampah, sarana air bersih, sarana pembuangan air limbah, ventilasi baik, kepadatan hunian rumah sesuai dan lantai rumah tidak tanah.
- e. Variabel t_4 yaitu kepadatan penduduk.
Kepadatan penduduk merupakan banyaknya penduduk yang mendiami suatu wilayah per km^2 .
- f. Variabel t_5 yaitu persentase institusi yang dibina kesehatan lingkungannya.
Persentase ini didapatkan dengan membandingkan banyaknya institusi yang dibina kesehatan lingkungannya dibandingkan dengan jumlah institusi yang dimiliki wilayah tersebut Adapun pembinaan yang dilakukan antara lain melakukan penyuluhan kesehatan, pemberian obat-obatan tertentu, pengawasan dll.

Variabel–variabel tersebut disajikan pada Tabel 3.1 berikut.

Tabel 3.1 Variabel penelitian

Variabel	Keterangan	Tipe Variabel
y	Banyaknya kasus DBD	Kontinu
x_1	Persentase rumah atau bangunan bebas jentik nyamuk <i>Aedes aegypti</i> atau angka bebas jentik	Kontinu
t_1	Persentase penduduk miskin	Kontinu
t_2	Curah hujan	Kontinu
t_3	Persentase rumah atau bangunan sehat	Kontinu
t_4	Kepadatan penduduk	Kontinu
t_5	Persentase institusi yang dibina kesehatan lingkungannya	Kontinu

Struktur data dari variabel yang digunakan dalam penelitian disajikan pada Tabel 3.2.

Tabel 3.2 Struktur Data Variabel Penelitian

Kecamatan	y	x_1	t_1	t_2	...	t_5
1	y_1	$x_{1.1}$	$t_{1.1}$	$t_{2.1}$...	$t_{5.1}$
2	y_2	$x_{1.2}$	$t_{1.2}$	$t_{2.2}$...	$t_{5.2}$
3	y_3	$x_{1.3}$	$t_{1.3}$	$t_{2.3}$...	$t_{5.3}$
4	y_4	$x_{1.4}$	$t_{1.4}$	$t_{2.4}$...	$t_{5.4}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
18	y_{18}	$x_{1.18}$	$t_{1.18}$	$t_{2.18}$...	$t_{5.18}$

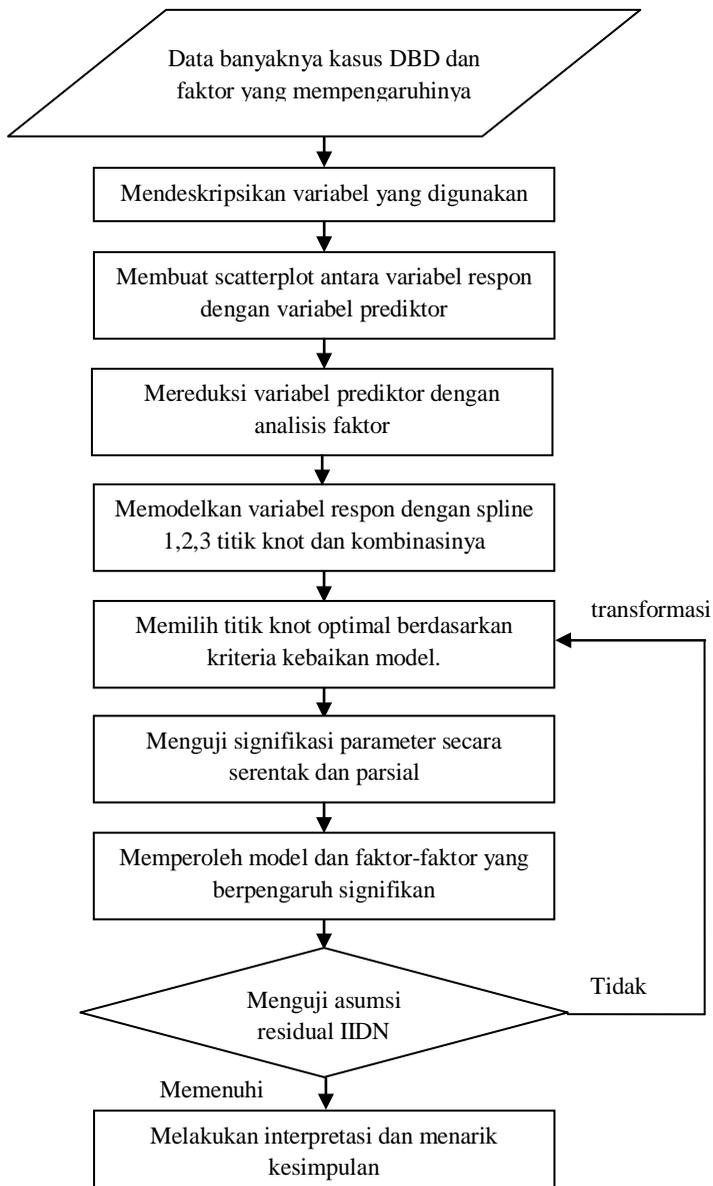
3.3 Langkah Penelitian

Data tersebut harus mengikuti langkah penelitian yang sistematis agar menghasilkan output yang valid. Adapun langkah-langkah dalam penelitian ini adalah sebagai berikut.

- a. Mendeskripsikan variabel yang digunakan dalam penelitian yaitu variabel respon yang terdiri dari angka penderita DBD dan variabel prediktornya yang terdiri dari angka bebas jentik, persentase penduduk miskin, curah hujan, persentase rumah atau bangunan sehat, kepadatan penduduk dan persentase institusi yang dibina kesehatan lingkungannya meliputi nilai rata-rata, variansi, minimum dan maksimum.
- b. Membuat *scatterplot* antara variabel respon dengan masing-masing variabel prediktor. Tujuannya untuk mengetahui bentuk pola data dan menentukan pendekatan regresi yang digunakan yaitu regresi parametrik, nonparametrik atau semiparametrik spline.
- c. Mereduksi variabel prediktor dengan analisis faktor untuk memilih variabel-variabel yang tidak berkorelasi satu sama lain dan sebelumnya dilakukan uji kecukupan data yaitu dengan uji *Kaiser-Meyer-Olkin* (KMO) dan uji homogenitas varians menggunakan uji Bartlett.
- d. Melakukan pemodelan menggunakan spline satu knot, dua knot dan tiga knot serta kombinasinya.
- e. Memilih titik knot optimal berdasarkan kriteria kebaikan model yaitu GCV, R^2 , MSE dan jumlah parameter.
- f. Melakukan pengujian signifikansi parameter regresi semiparametrik spline menggunakan uji serentak dan parsial.
- g. Memperoleh hasil model jumlah kasus DBD dan faktor-faktor yang berpengaruh signifikan.
- h. Melakukan pengujian asumsi residual IIDN pada model spline.
- i. Melakukan interpretasi terhadap model dan menarik kesimpulan.

3.4 Diagram Alir

Untuk mempermudah dalam memahami langkah - langkah penelitian yang digunakan, maka akan disajikan diagram alir pada Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram Alir

(Halaman ini sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

Dalam bab ini dibahas dan dianalisa mengenai banyaknya kasus DBD di Kabupaten/Kota Madiun beserta faktor-faktor yang mempengaruhinya. Analisis yang digunakan dalam penelitian ini menggunakan statistika deskriptif untuk mengetahui karakteristik dari variabel – variabel yang digunakan, analisis faktor untuk mereduksi variabel prediktor sehingga menghasilkan variabel yang tidak berkorelasi satu sama lain, pendekatan regresi semiparametrik spline dengan menggunakan fungsi spline linear satu, dua, dan tiga knot, serta kombinasinya untuk mengetahui pemodelan kasusnya dan melakukan pengujian signifikansi parameter hingga melakukan pengujian asumsi residual IIDN agar model yang dihasilkan memenuhi asumsi yang ditetapkan sehingga model dapat diinterpretasikan dan dipertanggung-jawabkan hasilnya.

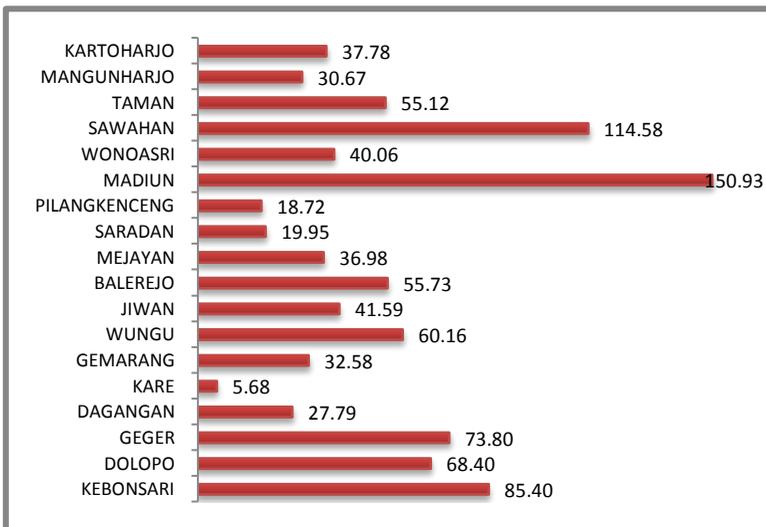
4.1 Karakteristik Kasus DBD dan Faktor - Faktor yang Mempengaruhi

Provinsi Jawa Timur pada tahun 2010 menduduki peringkat pertama dalam banyaknya kasus DBD yaitu sebanyak 26.020 kasus dengan kematian sebanyak 233 orang. Jika berdasarkan banyaknya kasus, Kabupaten/Kota Madiun berada di peringkat 28 dan 34 dari 38 daerah di Jawa Timur. Meskipun demikian jika dilihat berdasarkan target *Case Fatality Rate* (CFR) nasional, Kabupaten/Kota Madiun sudah melebihi target kesehatan nasional. Hal ini menjadi semacam peringatan awal bahwa kasus DBD dapat berkembang secara cepat dan mengkhawatirkan sehingga diperlukan tindakan pencegahan secara dini oleh masyarakat. Dalam bagian ini disajikan karakteristik banyaknya kasus DBD beserta faktor-faktor yang mempengaruhinya yang terdiri dari enam variabel dan disajikan pada Tabel 4.1.

Tabel 4.1 Karakteristik banyaknya kasus DBD dan Faktor-Faktor yang Mempengaruhi

Variabel	Rata-rata	Varians	Minimum	Maksimum
y	53,11	1294,66	5,68	150,93
x_1	70,29	254,03	44,21	98,93
t_1	26,62	117,41	9,67	44,66
t_2	210,1	4192,5	30,3	289,6
t_3	65,10	363,50	27,86	93,21
t_4	1856	4022216	184	6844
t_5	90,88	143,97	60,58	100,00

Tabel 4.1 menunjukkan bahwa rata-rata banyaknya kasus DBD pada tahun 2010 sebesar 53,11 yang menunjukkan bahwa rata-rata terdapat 53,11 kasus DBD per 100.000 penduduk. Untuk lebih jelasnya dalam variabel ini, disajikan dalam Gambar 4.1



Gambar 4.1 Diagram Batang Angka kesakitan DBD per 100.000 penduduk pada setiap kecamatan

Gambar 4.1 menunjukkan angka kesakitan DBD per 100.000 penduduk pada setiap kecamatan. Kasus DBD tertinggi terjadi di

Kecamatan Madiun dimana dalam satu tahun terjadi 61 kasus atau 150,93 kasus per 100.000 penduduk. Kecamatan Sawahan menempati urutan kedua dengan 30 kasus atau 114,58 kasus per 100.000 penduduk. Kasus terendah terjadi di Kecamatan Kare dimana hanya terjadi 2 kasus atau 5,68 kasus per 100.000 penduduk.

Pada variabel x_1 yaitu angka bebas jentik (ABJ) menunjukkan bahwa rata-rata ABJ di Kabupaten/Kota Madiun sebesar 70,29 persen. Hal itu menunjukkan bahwa ABJ cukup baik karena sudah diatas target nasional yaitu sebesar 60 persen. Nilai ABJ diperoleh melalui kegiatan pemeriksaan jentik yang dilakukan secara berkala pada rumah penduduk dan tempat penampungan air yang terdapat disekitar rumah tersebut. Berdasarkan hasil analisis, ABJ terbesar terdapat di Kecamatan Mangunharjo sebesar 98,93 persen dan terkecil terdapat di Kecamatan Madiun sebesar 44,21 persen. Selain itu terdapat 5 kecamatan yang nilai ABJ masih dibawah target nasional atau kurang dari 60 persen yaitu di Kecamatan Kebonsari, Dolopo, Geger, Mejayan dan Madiun. ABJ mempunyai pengaruh yang sangat kuat terhadap persebaran penyakit DBD karena suatu wilayah yang kurang terawat lingkungannya akan menjadi sarang perkembangbiakan nyamuk atau dapat dikatakan ABJ di wilayah tersebut rendah. Semakin pesatnya perkembangbiakan nyamuk terutama jenis *Aedes*, maka peluang terjadinya banyak kasus DBD semakin besar.

Salah satu faktor yang berpengaruh pada banyaknya kasus DBD adalah variabel t_1 yaitu persentase penduduk miskin. Penduduk miskin yang kemakmuran hidupnya dibawah rata-rata akan mendapat perlindungan dari negara yang salah satunya yaitu mendapat Jaminan Kesehatan Masyarakat. Upaya tersebut dilakukan agar masyarakat tersebut masih dapat menikmati haknya yaitu hak untuk hidup dan mendapatkan fasilitas kesehatan. Selain itu pemerintah dapat menekan jumlah kematian berbagai macam penyakit yang salah satunya disebabkan oleh penyakit DBD. Nilai rata-rata persentase penduduk miskin

mencapai 26,616 persen atau mencapai 192.555 jiwa. Jumlah tersebut relatif besar dan perlu dilakukan usaha di bidang ekonomi agar kemakmuran penduduk meningkat. Persentase penduduk miskin terbesar terdapat di Kecamatan Kare sebesar 44,657 persen dan terkecil di Kecamatan Jiwan sebesar 9,67 persen.

Selain itu terdapat variabel t_2 yaitu curah hujan yang mempunyai pengaruh cukup besar dalam persebaran penyakit DBD. Seperti diketahui bahwa nyamuk memilih tempat dimana terdapat air yang menggenang untuk berkembangbiak. Air tersebut berasal dari air hujan yang mengalir dan tertampung di wadah-wadah yang tidak dibersihkan oleh masyarakat. Berdasarkan nilai rata-rata curah hujan di Kabupaten/Kota Madiun mencapai 210,143 mm. Dengan curah hujan yang cukup tinggi tersebut dikhawatirkan terdapat banyak lokasi strategis bagi nyamuk untuk berkembang biak. Untuk curah hujan tertinggi terletak di Kecamatan Pilangkenceng sebesar 289,583 mm dan terkecil terletak di Kecamatan Kartoharjo dengan curah hujan 30,250 mm.

Rumah atau bangunan sehat merupakan sebuah kondisi tempat tinggal yang memenuhi syarat kesehatan yaitu dekat dengan air bersih, berjarak lebih dari seratus meter dari tempat pembuangan sampah, dekat dengan sarana pembersihan, serta berada di tempat dimana air hujan dan air kotor tidak menggenang. Rumah yang sehat membuat penghuninya menjadi nyaman, sehat dan dapat mengurangi permasalahan sosial yang terjadi di masyarakat. Variabel t_3 menjelaskan banyaknya persentase rumah atau bangunan sehat. Hasil analisis menunjukkan bahwa rata-rata rumah atau bangunan sehat di Kabupaten/Kota Madiun sebesar 65,09 persen. Daerah yang memiliki persentase rumah sehat tertinggi adalah Kecamatan Kare dengan 93,20 persen dan daerah yang memiliki persentase terkecil adalah Kecamatan Sawahan yaitu sebesar 27,86 persen. Adapun nilai keragaman yang dimiliki sebesar 363,5.

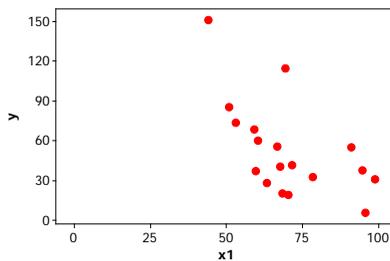
Variabel t_4 menunjukkan rata-rata kepadatan penduduk yang mencapai 1.856 jiwa per km^2 . Angka tersebut menunjukkan bahwa penduduk di Kabupaten/Kota Madiun cukup padat dan memudahkan nyamuk *Aedes* dalam menularkan virus penyebab penyakit DBD. Nilai varians yang cukup besar yaitu 4022216 menunjukkan terjadi ketimpangan dalam persebaran penduduk. Hal itu dibuktikan dengan kepadatan penduduk terkecil yang hanya 184 jiwa per km^2 di Kecamatan Kare bila dibandingkan dengan kepadatan penduduk terbesar di Kecamatan Taman yang mencapai 6844 jiwa per km^2 . Selain Kecamatan Taman terdapat beberapa daerah yang kepadatan penduduknya sangat tinggi yaitu Kecamatan Mangunharjo dan Kartoharjo dimana ketiga kecamatan yang mempunyai kepadatan penduduk tertinggi merupakan kecamatan di Kota Madiun.

Variabel terakhir yang mempengaruhi adalah persentase institusi yang dibina kesehatan lingkungannya (t_5). Institusi dan tempat umum merupakan sasaran bagi Dinas Kesehatan dalam melakukan kerjasama menjaga kebersihan dan kesehatan lingkungannya untuk mencegah persebaran berbagai jenis penyakit. Hal ini disebabkan tidak sedikit masyarakat yang menjadi korban seperti sebagian besar penderita DBD yang umumnya anak-anak tertular virus atau mendapat gigitan nyamuk *Aedes* ketika melakukan aktivitas belajar mengajar di sekolah. Oleh karena itu semakin banyak institusi yang dibina kesehatan lingkungannya diharapkan dapat menurunkan banyaknya kasus DBD. Berdasarkan hasil analisis didapatkan rata-rata institusi yang dibina mencapai 90,83 persen. Nilai tersebut sangat tinggi dan menunjukkan bahwa kepedulian terhadap masalah kesehatan lingkungan cukup tinggi. Tetapi hal tersebut belum menjadi jaminan bahwa persebaran penyakit DBD menjadi rendah karena perlu dilihat keefektifan dan pengawasannya. Terdapat beberapa wilayah yang semua institusinya berhasil dibina kesehatan lingkungannya yaitu Kecamatan Dagangan, Kare, Gemarang, Jiwan, Pilangkenceng dan Sawahan. Untuk wilayah yang institusi

yang dibina paling kecil adalah Kecamatan Kebonsari yaitu 60,58 persen.

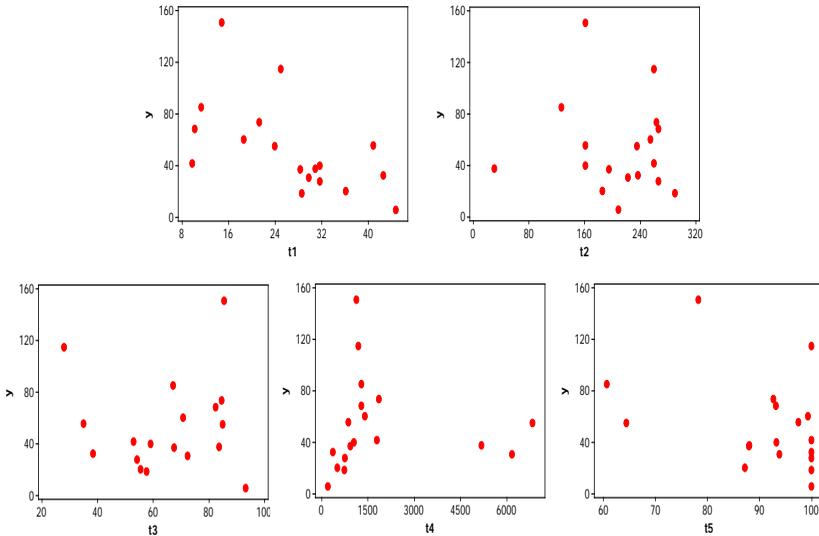
4.2 Pola Hubungan antara Jumlah Kasus Demam Berdarah Dengue dengan Faktor-Faktor yang Mempengaruhi

Salah satu cara untuk mengetahui pola hubungan variabel-variabel tersebut adalah menggunakan *scatterplot* atau sebaran plot. Dengan melihat sebaran plot yang dibentuk maka dapat ditentukan metode regresi yang digunakan. Gambar 4.2 menunjukkan pola hubungan antara jumlah kasus DBD dan Angka Bebas Jentik (ABJ) mempunyai kecenderungan adanya hubungan linier negatif. Hal tersebut dapat dilihat pada saat ABJ rendah maka jumlah kasus DBD sangat tinggi. Semakin besar nilai ABJ, terdapat kecenderungan jumlah kasusnya semakin kecil dan begitu juga sebaliknya. Berdasarkan hasil pola hubungan antara kedua variabel tersebut maka dapat digunakan pendekatan regresi parametrik untuk menganalisis data tersebut.



Gambar 4.2 Scatterplot antara Jumlah Kasus DBD dengan variabel x_1

Setelah mengetahui pola hubungan antara variabel y dengan x_1 yaitu Angka Bebas Jentik yang berbentuk linier negatif dimana semakin rendah nilai ABJ maka semakin tinggi kasus DBD yang terjadi. Selanjutnya dilakukan hal serupa untuk mengetahui pola hubungan antara jumlah kasus DBD dengan variabel t_1 , t_2 , t_3 , t_4 dan t_5 yang disajikan dalam Gambar 4.3.



Gambar 4.3 Scatterplot antara Jumlah Kasus DBD dengan variabel t_1 , t_2 , t_3 , t_4 dan t_5

Gambar 4.3 menunjukkan pola hubungan variabel respon yaitu jumlah kasus DBD dengan masing-masing lima variabel prediktor lainnya yaitu persentase penduduk miskin (t_1), curah hujan (t_2), persentase rumah atau bangunan sehat (t_3), kepadatan penduduk (t_4), dan persentase institusi yang dibina kesehatan lingkungannya (t_5) yang tidak membentuk pola tertentu. Sebaran data yang acak tersebut mengindikasikan pendekatan regresi nonparametrik sebagai metode yang tepat untuk menganalisis data tersebut.

Berdasarkan Gambar 4.2 dan Gambar 4.3 dimana terdapat variabel yang pola hubungannya membentuk pola tertentu yaitu linier dan sebagian lainnya tidak membentuk pola tertentu atau sebaran datanya acak, maka dapat digunakan pendekatan metode regresi semiparametrik spline untuk menganalisis kasus tersebut.

4.3 Analisis Faktor

Analisis faktor merupakan analisis statistika yang bertujuan untuk mengelompokkan variabel-variabel yang banyak dan berbeda menjadi kumpulan variabel yang lebih kecil didasarkan pada kesamaan karakteristik variabel tersebut. Dengan kata lain, tujuan dilakukan analisis faktor dalam pembahasan ini adalah untuk mereduksi variabel-variabel yang digunakan sehingga dapat dipilih variabel-variabel yang tidak terdapat korelasi antar variabelnya. Analisis faktor dilakukan dengan melakukan uji kecukupan data menggunakan uji KMO dan uji homogenitas varians menggunakan uji Bartlett terlebih dahulu. Hasil dari kedua uji tersebut disajikan dalam Tabel 4.2

Tabel 4.2 Uji Kecukupan Data dan Homogenitas Varians

Pengujian	Nilai	Pengujian	P-value
<i>Kaiser Mayer Olkin</i>	0,533	<i>Bartlett</i>	0,002

Tabel 4.2 menunjukkan bahwa hasil pengujian KMO sebesar 0,533. Nilai tersebut lebih besar dari 0,5 sehingga data yang digunakan telah mencukupi untuk dilakukan pemfaktoran. Pada pengujian *Bartlett* menghasilkan nilai 0,002 dimana nilai tersebut lebih kecil dari alpha yaitu 0,05 yang berarti terdapat korelasi yang signifikan antar variabel. Berdasarkan kedua uji tersebut, data telah memenuhi untuk dilakukan analisis faktor.

Tahapan selanjutnya dalam analisis faktor adalah menentukan banyaknya faktor yang terbentuk. Ekstraksi variabel tersebut dilakukan dengan metode *principal component* dengan analisis matriks korelasi dan dilakukan rotasi *varimax*. Banyaknya faktor dapat ditentukan dari *eigenvalues* yang lebih dari 1. Tabel 4.3 menunjukkan bahwa dari 6 faktor awal yang digunakan terdapat tiga faktor yang *eigenvalues* lebih dari 1, sehingga faktor-faktor penyebab DBD direduksi menjadi tiga faktor. Ketiga faktor tersebut mampu menjelaskan variabilitas data sebesar 82,398 persen.

Tabel 4.3 Eigenvalues

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.166	36.104	36.104
2	1.767	29.453	65.557
3	1.011	16.842	82.398
4	.613	10.211	92.609
5	.332	5.535	98.144
6	.111	1.856	100.000

Dalam melakukan pengelompokkan faktor, dapat dilihat dari nilai *loading factor* yang terbentuk dan disajikan pada Tabel 4.4.

Tabel 4.4 Loading Factor Penyebab DBD

Variabel	Component		
	1	2	3
x_1	0.591	0.772	0.060
t_1	-0.211	0.900	-0.097
t_2	-0.077	-0.176	0.935
t_3	0.668	-0.296	-0.192
t_4	0.906	0.117	-0.124
t_5	-0.393	0.474	0.648

Tabel 4.4 menunjukkan dari 3 faktor yang terbentuk masing-masing memiliki dua variabel pembentuk. Hal tersebut dapat dilihat dari nilai *loading factor* terbesar dari setiap variabel. Nilai tersebut menunjukkan bahwa variabel tersebut masuk ke dalam faktor baru yang terbentuk. Selain itu variabel pembentuk dalam satu faktor menunjukkan bahwa variabel-variabel tersebut berkorelasi. Sehingga perlu dipilih masing-masing satu variabel dalam setiap faktor untuk digunakan sebagai variabel prediktor dalam pemodelan kasus DBD. Pemilihan tersebut dapat dilihat dari nilai *loading factor* tertinggi yang terdapat dalam setiap faktor. Hasil reduksi tersebut menghasilkan variabel prediktor yaitu kepadatan penduduk (t_4) yang mewakili faktor pertama, ABJ (x_1) dan persentase penduduk miskin (t_1) yang mewakili

faktor kedua dan curah hujan (t_2) mewakili faktor ketiga. Pemilihan dua variabel dalam faktor kedua dikarenakan variabel ABJ mengandung unsur parametrik sedangkan persentase penduduk miskin mengandung unsur nonparametrik.

4.4 Pemodelan Kasus Demam Berdarah Dengue (DBD) Menggunakan Regresi Semiparametrik Spline

Setelah mengetahui bentuk sebaran plot dan hasil reduksi variabel prediktor menggunakan analisis faktor maka dilakukan pemodelan dengan menggunakan pendekatan dalam analisis regresi. Tahapan awal yang dilakukan adalah membentuk persamaan model regresi semiparametrik spline.

4.4.1 Pemodelan Regresi Semiparametrik Spline

Aplikasi pemodelan regresi semiparametrik spline yang digunakan terdiri dari 4 model yaitu satu titik knot, dua titik knot, tiga titik knot dan kombinasi dari ketiga titik knot tersebut. Adapun bentuk persamaan model regresi semiparametrik spline yang terdiri dari 1 komponen parametrik dan 3 komponen nonparametrik dijabarkan seperti berikut.

$$\begin{aligned}
 y_i &= \beta_0 + \beta_1 x_{i1} + f(t_{i1}) + f(t_{i2}) + f(t_{i3}) + \varepsilon_i \\
 &= \beta_0 + \beta_1 x_{i1} + \sum_{j=0}^p \beta_j t_{i1}^j + \sum_{l=1}^r \beta_{p+l} (t_{i1} - k_l)_+^p + \sum_{j=0}^p \beta_j t_{i2}^j + \\
 &\quad \sum_{l=1}^r \beta_{p+l} (t_{i2} - k_l)_+^p + \sum_{j=0}^p \beta_j t_{i3}^j + \sum_{l=1}^r \beta_{p+l} (t_{i3} - k_l)_+^p + \varepsilon_i
 \end{aligned}$$

4.4.2 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan pola perilaku fungsi atau kurva. Titik knot yang digunakan dalam analisis terdiri dari satu, dua dan tiga titik

knot serta kombinasinya. Titik knot tersebut berasal dari variabel, persentase penduduk miskin (t_1), curah hujan (t_2) dan kepadatan penduduk (t_4). Setiap titik menghasilkan GCV dan dibandingkan untuk memilih titik knot optimal. Berikut disajikan nilai GCV menggunakan satu titik knot, dua titik knot, tiga titik knot dan kombinasi dari ketiga titik knot serta kriteria pemilihan titik knot optimal secara berurutan pada Tabel 4.5, Tabel 4.6, Tabel 4.7, Tabel 4.8 dan Tabel 4.9.

Tabel 4.5 Nilai Knot dan GCV Untuk Satu Titik Knot

No.	Knot			GCV
	t_1	t_2	t_4	
1	10,387	35,543	319,918	1421,397
2	11,101	40,835	455,837	1340,252
3	11,815	46,128	591,755	1293,189
4	12,529	51,420	727,673	1532,091
5	13,243	56,713	863,592	1536,448
6	13,957	62,005	999,510	1518,008
7	14,671	67,298	1135,429	1454,192
8	15,385	72,590	1271,347	1556,073
9	16,099	77,883	1407,265	1650,741
10	16,813	83,175	1543,184	1703,895

Berdasarkan Tabel 4.5 maka dapat dibandingkan nilai GCV yang dihasilkan pada satu titik knot. Adapun nilai GCV minimum yang dihasilkan adalah 1293,189 dengan titik knot pada variabel t_1 sebesar 11,815; variabel t_2 sebesar 46,128; dan variabel t_4 sebesar 591,755. Selanjutnya untuk nilai GCV yang dihasilkan menggunakan dua titik knot disajikan dalam Tabel 4.6.

Tabel 4.6 Nilai Knot dan GCV Untuk Dua Titik Knot

No.	Knot			GCV
	t_1	t_2	t_4	
1	26,808	157,270	3446,041	571,185
	28,236	167,855	3717,878	
2	26,094	151,977	3310,122	581,277
	28,9505	173,147	3853,796	
3	26,809	157,270	3446,041	484,917
	28,951	173,148	3853,796	
4	27,522	162,562	3581,959	532,527
	28,236	167,855	3717,878	
5	26,094	151,977	3310,122	747,267
	29,664	178,440	3989,714	
6	25,380	146,685	3174,204	813,716
	29,664	178,440	3989,714	
7	25,380	146,685	3174,204	783,787
	28,950	173,147	3853,796	
8	26,094	151,977	3310,122	761,633
	28,236	167,855	3717,878	
9	26,808	157,270	3446,041	826,801
	27,522	162,562	3581,959	
10	27,522	162,562	3581,959	672,340
	28,950	173,147	3853,796	

Tabel 4.6 menjelaskan nilai GCV yang dihasilkan dengan menggunakan dua titik knot. Nilai GCV minimum yang didapatkan sebesar 484,917 dengan titik knot sebesar 26,809 dan 28,951 pada variabel t_1 ; titik knot 157,27 dan 173,148 pada variabel t_2 dan pada variabel t_4 sebesar 3446,041 dan 3853,796.

Nilai GCV yang dihasilkan menggunakan tiga titik knot disajikan dalam Tabel 4.7.

Tabel 4.7 Nilai Knot dan GCV Untuk Tiga Titik Knot

No.	Knot			GCV
	t_1	t_2	t_4	
1	19,669	104,345	2086,857	246,010
	24,667	141,393	3038,286	
	28,951	173,148	3853,796	
2	18,955	99,053	1950,939	245,599
	25,381	146,685	3174,204	
	28,951	173,148	3853,796	
3	19,669	104,345	2086,857	244,135
	25,381	146,685	3174,204	
	28,951	173,148	3853,796	
4	20,383	109,638	2222,776	245,828
	23,953	136,100	2902,367	
	28,951	173,148	3853,796	
5	20,383	109,638	2222,776	259,626
	25,381	146,685	3174,204	
	28,951	173,148	3853,796	
6	20,383	109,638	2222,776	255,802
	24,667	141,393	3038,286	
	28,951	173,148	3853,796	
7	18,955	99,053	1950,939	256,978
	24,667	141,393	3038,286	
	28,951	173,148	3853,796	
8	20,383	109,638	2222,776	275,058
	23,239	130,808	2766,449	
	28,951	173,148	3853,796	

Tabel 4.7 Nilai Knot dan GCV Untuk Tiga Titik Knot (Lanjutan)

No.	Knot			GCV
	t_1	t_2	t_4	
9	19,668	104,345	2086,857	277,204
	26,094	151,977	3310,122	
	28,950	173,147	3853,796	
10	19,668	104,345	2086,857	278,359
	23,952	136,100	2902,367	
	28,950	173,147	3853,796	

Berdasarkan Tabel 4.7 didapatkan nilai GCV minimum dengan menggunakan tiga titik knot sebesar 244,135. Dengan titik knot optimal yang dihasilkan dari GCV tersebut adalah 19,669; 25,381 dan 28,951 pada variabel persentase penduduk miskin (t_1); titik 104,345; 146,685 dan 173,148 pada variabel curah hujan (t_2); serta titik 2086,857; 3174,204 dan 3853,796 pada variabel kepadatan penduduk (t_4).

Selain menggunakan satu titik knot, dua titik knot dan tiga titik knot, dalam mencari model terbaik perlu dibandingkan nilai GCV minimum yang terbentuk dengan menggunakan kombinasi knot dari ketiga titik knot tersebut. Kombinasi tersebut dapat dihasilkan setelah diketahui titik knot optimum dari satu titik knot, dua titik knot dan tiga titik knot. Adapun nilai GCV yang dihasilkan dari kombinasi knot disajikan dalam Tabel 4.8.

Tabel 4.8 Nilai Knot dan GCV Untuk Kombinasi Titik Knot

No.	Knot			GCV
	t_1	t_2	t_4	
1	26,809	104,345	2086,857	647,757
	28,951	146,685	3174,204	
		173,148	3853,796	
2	19,669	104,345	2086,857	244,135
	25,381	146,685	3174,204	
	28,951	173,148	3853,796	

Tabel 4.8 Nilai Knot dan GCV Untuk Kombinasi Titik Knot (Lanjutan)

No.	Knot			GCV
	t_1	t_2	t_4	
3	19,669	157,270	3446,041	181,437
	25,381	173,148	3853,796	
	28,951			
4	26,809	157,270	3446,041	484,917
	28,951	173,148	3853,796	
5	19,669	46,128	3446,041	1027,291
	25,381		3853,796	
	28,951			
6	26,809	46,128	3446,041	1141,444
	28,951		3853,796	
7	11,815	46,128	591,755	1293,189
8	26,809	46,128	591,755	1338,070
	28,951			
9	19,669	46,128	591,755	1527,643
	25,381			
	28,951			
10	26,809	157,270	591,755	1439,236
	28,951	173,148		

Tabel 4.8 menjelaskan nilai GCV yang dihasilkan dengan menggunakan kombinasi titik knot. Nilai GCV minimum yang dihasilkan adalah 181,437 yaitu hasil kombinasi knot (322). Titik

knot yang terbentuk dari nilai GCV minimum tersebut adalah titik 19,669; 25,381 dan 28,951 pada variabel t_1 ; titik 157,27 dan 173,148 pada variabel t_2 dan titik 3446,041 dan 3853,796 pada variabel t_4 . Penentuan pemilihan model terbaik tidak hanya membandingkan nilai GCV minimum, tetapi juga nilai R^2 , MSE dan jumlah parameter yang dihasilkan. Berdasarkan kriteria kebaikan model tersebut dapat ditentukan model terbaik yang dapat digunakan. Adapun hasil kriteria untuk tiap model disajikan dalam Tabel 4.9.

Tabel 4.9 Hasil Kriteria Kebaikan Model

Kriteria kebaikan Banyak model Knot optimal	GCV	R^2	MSE	Jumlah parameter
kombinasi (233)	647,757	94,114	43,984	12
3 titik knot	244,135	98,459	25,742	13
kombinasi (322)	181,437	98,351	70,479	11
2 titik knot	484,917	94,002	90,735	10
kombinasi (312)	1027,291	87,294	123,578	10
kombinasi (212)	1141,444	81,560	142,595	9
1 titik knot	1293,189	60,438	198,702	7
kombinasi (211)	1338,070	72,642	168,395	8
kombinasi (311)	1527,643	75,321	157,981	9
kombinasi (221)	1439,236	76,749	1734,713	9

Tabel 4.9 menjelaskan nilai yang dihasilkan tiap kombinasi knot yang dianalisis sebelumnya berdasarkan GCV, R^2 , MSE dan jumlah parameter. Indikatornya adalah untuk nilai GCV, MSE dan jumlah parameter, semakin kecil nilainya maka semakin baik. Untuk nilai R^2 semakin besar nilainya (mendekati atau bernilai 100) maka semakin baik.

Berdasarkan Tabel 4.9 terlihat kombinasi knot (322) merupakan kombinasi yang berdasarkan kriteria kebaikan

modelnya paling baik. Hal tersebut didasarkan nilai GCV yang dihasilkan paling kecil yaitu 181,437 dan R^2 yang besar yaitu 98,351. Nilai MSE yang dihasilkan juga relatif kecil yaitu 70,479 dengan jumlah parameter sebanyak 11. Meskipun jumlah parameter yang dihasilkan 1 titik knot jauh lebih sedikit yaitu sebanyak 7 parameter, tetapi jika dibandingkan secara keseluruhan kombinasi knot (322) merupakan model yang lebih baik. Oleh karena itu kombinasi knot (322) digunakan sebagai model dalam menganalisis kasus DBD dan faktor-faktor yang mempengaruhinya. Titik knot yang terbentuk dari model tersebut adalah titik 19,669; 25,381 dan 28,951 pada variabel persentase penduduk miskin; titik 157,27 dan 173,148 pada variabel curah hujan dan titik 3446,041 dan 3853,796 pada variabel kepadatan penduduk.

4.4.3 Pemodelan Regresi Semiparametrik Spline dengan Menggunakan Titik Knot Optimal

Persamaan yang terbentuk dari regresi semiparametrik spline dengan menggunakan titik knot optimal yang menghasilkan kombinasi (322) adalah sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 t_{i1} + \hat{\beta}_3 (t_{i1} - k_1)_+ + \hat{\beta}_4 (t_{i1} - k_2)_+ + \hat{\beta}_5 (t_{i1} - k_3)_+ + \hat{\beta}_6 t_{i2} + \hat{\beta}_7 (t_{i2} - k_1)_+ + \hat{\beta}_8 (t_{i2} - k_2)_+ + \hat{\beta}_9 t_{i4} + \hat{\beta}_{10} (t_{i4} - k_1)_+ + \hat{\beta}_{11} (t_{i4} - k_2)_+$$

Jika memasukkan nilai estimasi parameter, bentuk model regresi semiparametrik spline yang dihasilkan adalah sebagai berikut.

$$\hat{y} = 22,1045 - 1,6967x_1 - 0,6612t_1 + 11,8618(t_1 - 19,669)_+ - 48,3093(t_1 - 25,381)_+ + 38,5898(t_1 - 28,951)_+ + 1,691t_2 + 5,2401(t_2 - 157,27)_+ + 3,6136(t_2 - 173,148)_+ - 0,0419t_4 + 1,4117(t_4 - 3446,041)_+ - 1,5288(t_4 - 3853,796)_+$$

Model tersebut merupakan model regresi semiparametrik spline terbaik untuk kasus DBD dan faktor-faktor yang mempengaruhinya. Adapun nilai R^2 yang dihasilkan adalah 98,351 persen yang berarti variabel prediktor yang digunakan mampu menjelaskan variabilitas banyaknya kasus DBD sebesar 98,351 persen.

4.5 Uji Signifikasi Parameter

Uji signifikasi parameter pada model dilakukan untuk mengetahui variabel prediktor yang memiliki pengaruh yang signifikan terhadap banyaknya kasus DBD yang terjadi. Pengujian parameter pada model terdapat dua macam yaitu pengujian secara serentak dan secara parsial.

4.5.1 Uji Serentak

Hipotesis yang digunakan dalam pengujian signifikasi parameter secara serentak adalah sebagai berikut.

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_{11} = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0, k = 0, 1, \dots, 11.$$

Dengan menggunakan nilai alpha sebesar 0,05 maka didapatkan tabel ANOVA yang disajikan dalam Tabel 4.10

Tabel 4.10 ANOVA

Sumber Variasi	Df	SS	MS	F_{hitung}	$P\text{-value}$
Regresi	11	21646,26	1967,842		
<i>Error</i>	6	362,874	60,479	32,5376	0,0001
Total	17	22009,13			

Berdasarkan Tabel 4.10 yaitu hasil tabel ANOVA dari uji signifikasi parameter secara serentak didapatkan nilai F_{hitung} sebesar 32,5376. Berdasarkan $p\text{-value}$ menghasilkan nilai sebesar 0,0001. Jika dibandingkan dengan nilai alpha yang ditentukan yaitu 0,05 maka dapat diputuskan tolak H_0 . Oleh karena itu dapat

dsimpulkan berdasarkan uji serentak, minimal terdapat satu parameter yang berpengaruh signifikan terhadap model.

4.5.2 Uji Parsial

Berdasarkan uji serentak yang menghasilkan kesimpulan yaitu terdapat minimal satu parameter yang berpengaruh maka perlu dilakukan uji parsial untuk mengetahui parameter mana saja yang berpengaruh signifikan. Adapun hipotesis yang digunakan dalam pengujian signifikansi parameter secara parsial adalah sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 0,1, \dots, 11$$

Dengan menggunakan nilai alpha sebesar 0,05 maka didapatkan tabel pengujian parameter secara parsial yang disajikan dalam Tabel 4.11

Tabel 4.11 Uji Parsial

Variabel	Parameter	Koefisien	t_{hit}	<i>P-value</i>
-	β_0	22,1045	0,6079	0,5655
x_1	β_1	-1,6967	-4,7289	0,0032*
	β_2	-0,6612	-0,6536	0,5376
t_1	β_3	11,8618	3,7843	0,0091*
	β_4	-48,3093	-9,8485	0,0000*
	β_5	38,5898	11,5687	0,0000*
t_2	β_6	1,6910	7,8715	0,0002*
	β_7	5,2401	-6,6473	0,0005*
	β_8	3,6136	4,9524	0,0026*
t_4	β_9	-0,0419	-4,5164	0,0040*
	β_{10}	1,4117	8,393	0,0001*
	β_{11}	-1,5288	-8,4613	0,0001*

*variabel memberikan pengaruh yang signifikan

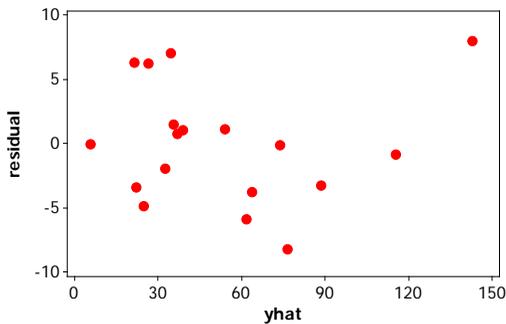
Berdasarkan Tabel 4.11 dihasilkan nilai koefisien, t_{hit} dan P -value dari setiap parameter. P -value yang dihasilkan setiap parameter dibandingkan dengan nilai alpha yaitu 0,05. Dikatakan signifikan apabila p -value yang dihasilkan setiap parameter kurang dari nilai alpha (0,05). Adapun hampir semua parameter signifikan kecuali β_0 dan β_2 , sedangkan semua variabel prediktor berpengaruh signifikan terhadap banyaknya kasus DBD di Kabupaten/Kota Madiun.

4.6 Uji Asumsi Residual

Setelah melakukan uji signifikansi parameter, maka dilakukan uji asumsi residual. Pengujian asumsi residual dilakukan untuk mengetahui kebaikan model regresi. Adapun uji asumsi residual yang dilakukan yaitu residual harus memenuhi asumsi identik, independen dan berdistribusi normal.

4.6.1 Uji Identik

Pengujian asumsi identik bertujuan untuk mengetahui apakah varians dari residual yang dihasilkan terdapat heteroskedastisitas atau tidak. Adapun cara untuk melakukan asumsi identik adalah dengan melihat sebaran plot antara residual dengan \hat{y} yang dihasilkan serta melakukan uji Glejser.



Gambar 4.4 Scatterplot antara residual dan \hat{y}

Berdasarkan hasil sebaran plot pada Gambar 4.4 terlihat bahwa uji identik telah terpenuhi. Hal tersebut disebabkan plot tersebar secara acak atau tidak membentuk suatu pola tertentu seperti bentuk pola corong, garis lurus ataupun melengkung. Selanjutnya dilakukan uji Glejser menggunakan tabel ANOVA yang disajikan dalam Tabel 4.12.

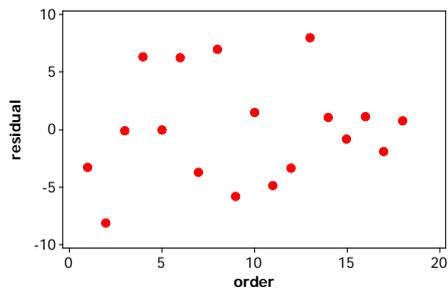
Tabel 4.12 ANOVA Uji Glejser

Sumber Variasi	Df	SS	MS	F	<i>P-value</i>
Regresi	11	104,3506	9,4864		
<i>Error</i>	6	29,3344	4,889	1,9403	0,2146
Total	17	133,685	-		

Berdasarkan Tabel 4.12 didapatkan nilai F_{hitung} sebesar 1,9403. Berdasarkan *p-value* menghasilkan nilai sebesar 0,2146. Jika dibandingkan dengan nilai alpha yang ditentukan yaitu 0,05 maka dapat disimpulkan bahwa tidak terdapat heteroskedastisitas atau asumsi identik pada residual telah terpenuhi.

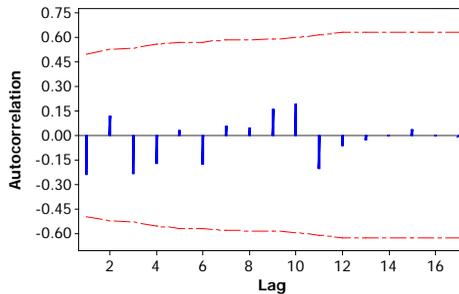
4.6.2 Uji Independen

Uji asumsi independen bertujuan untuk mengetahui ada tidaknya korelasi antar residual. Adapun cara untuk melakukan pengujian asumsi independen adalah dengan melihat sebaran plot antara residual dan order serta plot ACF (Autocorrelation Function) dari residual.



Gambar 4.5 Scatterplot antara residual dan order

Gambar 4.5 menunjukkan bahwa sebaran plot yang dihasilkan antara residual dengan order menyebar secara acak dan tidak membentuk suatu pola tertentu. Berdasarkan sebaran plotnya, maka asumsi residual independen telah terpenuhi. Plot ACF dari residual yang terbentuk disajikan dalam Gambar 4.6



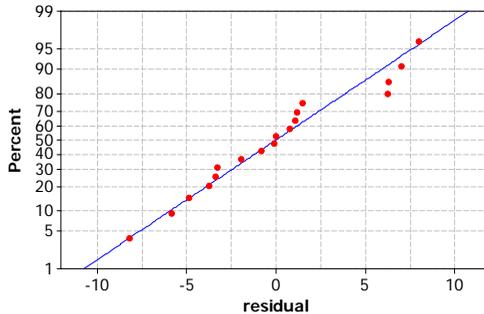
Gambar 4.6 Plot ACF dari Residual

Berdasarkan plot ACF yang terbentuk dari residual yang digunakan, maka uji asumsi independen telah terpenuhi. Hal tersebut dikarenakan tidak adanya lag yang keluar dari garis batas toleransi.

4.6.3 Uji Distribusi Normal

Uji asumsi distribusi normal bertujuan untuk mengetahui apakah residual yang dihasilkan mengikuti distribusi normal atau tidak. Pengujian ini dilakukan dengan cara menggunakan uji Kolmogorov-Smirnov.

Berdasarkan Gambar 4.7 terlihat bahwa sebaran plot tersebar dan mengikuti garis distribusi normal. Selain itu berdasarkan hasil uji Kolmogorov Smirnov yang menghasilkan nilai sebesar 0,15 dimana nilai yang dihasilkan lebih kecil daripada nilai D_{tabel} yaitu sebesar 0,2785. P -value yang dihasilkan lebih dari 0,15 sehingga nilai tersebut lebih besar daripada nilai alpha (0,05). Berdasarkan hasil dari pengujian diatas maka residual dari model telah memenuhi asumsi distribusi normal.



Gambar 4.7 Uji Normalitas Residual

4.7 Interpretasi Hasil Model Regresi Semiparametrik Spline

Setelah melakukan pengujian asumsi residual, dimana sudah memenuhi asumsi IIDN ($0, \sigma^2$) maka perlu dilakukan interpretasi pada model. Nilai koefisien determinasi yang dihasilkan oleh model regresi sebesar 98,351 persen yang berarti variabel angka bebas jentik (x_1), persentase penduduk miskin (t_1), curah hujan (t_2), dan kepadatan penduduk (t_4) dapat menjelaskan variabilitas dari jumlah kasus DBD di Kabupaten/Kota Madiun sebesar 98,351 persen. Nilai tersebut menunjukkan bahwa model yang terbentuk sangat layak dalam memodelkan pola data. Selanjutnya dilakukan interpretasi terhadap variabel yang hanya memberikan pengaruh signifikan terhadap model. Adapun hasil model regresi semiparametrik spline yang terbentuk dari faktor-faktor yang mempengaruhi banyaknya kasus DBD di Kabupaten/Kota Madiun pada tahun 2010 adalah sebagai berikut.

$$\hat{y} = 22,1045 - 1,6967x_1 - 0,6612t_1 + 11,8618(t_1 - 19,669)_+ - 48,3093(t_1 - 25,381)_+ + 38,5898(t_1 - 28,951)_+ + 1,691t_2 + 5,2401(t_2 - 157,27)_+ + 3,6136(t_2 - 173,148)_+ - 0,0419t_4 + 1,4117(t_4 - 3446,041)_+ - 1,5288(t_4 - 3853,796)_+$$

Interpretasi model terhadap variabel x_1 yaitu Angka Bebas jentik (ABJ) adalah apabila ABJ naik sebesar satu persen, maka jumlah kasus DBD cenderung menurun sebesar 1,6967 kasus per 100.000 penduduk dengan asumsi variabel lain konstan. Interpretasi model terhadap variabel t_1 yaitu persentase penduduk miskin adalah sebagai berikut.

$$\hat{y} = \begin{cases} -233,3097 + 11,2006t_1 & ; 19,669 < t_1 < 25,381 \\ 992,8286 - 37,1087t_1 & ; 25,381 < t_1 < 28,951 \\ -124,3847 + 1,4811t_1 & ; t_1 \geq 28,951 \end{cases}$$

Pada saat persentase penduduk miskin berada diantara 19,669 sampai 25,381 persen, apabila persentase penduduk miskin naik sebesar satu persen maka jumlah kasus DBD naik sebesar 11,2006 kasus per 100.000 penduduk. Kecamatan yang persentase penduduk miskinnya berada diantara 19,669 sampai 25,381 persen adalah Geger, Sawahan dan Taman. Ketika persentase penduduk miskin diantara 25,381 sampai 28,951 persen, maka apabila persentase penduduk miskin naik sebesar satu persen maka jumlah kasus DBD turun sebesar 37,1087 kasus per 100.000 penduduk. Hal tersebut sesuai pada Kecamatan Mejayan dan Pilangkenceng. Saat persentase penduduk miskin lebih dari 28,951 persen, apabila persentase penduduk miskin naik sebesar satu persen maka jumlah kasus DBD naik sebesar 1,4811 kasus per 100.000 penduduk. Kecamatan yang termasuk dalam kategori tersebut adalah Dagangan, Kare, Gemarang, Balerejo, Saradan, Wonoasri, Mangunharjo dan Kartoharjo. Hal ini berlaku apabila variabel lain yang mempengaruhinya bernilai konstan.

Interpretasi model terhadap variabel t_2 yaitu curah hujan adalah sebagai berikut.

$$\hat{y} = \begin{cases} 1,691 t_2 & ; t_2 < 157,27 \\ -824,1105 + 6,9311t_2 & ; 157,27 < t_2 < 173,148 \\ -1449,7981 + 10,5447t_2 & ; t_2 \geq 173,148 \end{cases}$$

Saat curah hujan kurang dari $157,27 \text{ mm}^3$, apabila curah hujan naik sebesar satu mm^3 maka jumlah kasus DBD naik sebesar 1,691 kasus per 100.000 penduduk. Kecamatan yang curah hujan kurang dari $157,27 \text{ mm}^3$ adalah Kebonsari dan Kartoharjo. Ketika curah hujan diantara $157,27$ sampai $173,148 \text{ mm}^3$, maka apabila curah hujan naik sebesar satu mm^3 maka jumlah kasus DBD naik sebesar 6,9311 kasus per 100.000 penduduk. Hal tersebut sesuai pada Kecamatan Balerejo, Madiun dan Wonoasri. Sedangkan pada saat curah hujan lebih dari $173,148 \text{ mm}^3$, apabila curah hujan naik sebesar satu mm^3 maka jumlah kasus DBD naik sebesar 10,5447 kasus per 100.000 penduduk. Hal ini berlaku apabila variabel lain yang mempengaruhinya bernilai konstan.

Interpretasi model terhadap variabel t_4 yaitu kepadatan penduduk adalah sebagai berikut.

$$\hat{y} = \begin{cases} -0,0419 t_4 ; t_4 < 3446,041 \\ -4864,7760 + 1,3698t_4 ; 3446,041 < t_4 < 3853,796 \\ 1026,9073 - 0,1590t_4 ; t_4 \geq 3853,796 \end{cases}$$

Pada saat kepadatan penduduk kurang dari $3446,041$ jiwa/ km^2 , apabila kepadatan penduduk naik sebesar satu satuan maka jumlah kasus DBD cenderung menurun sebesar 0,0419 kasus per 100.000 penduduk. Interpretasi tersebut sesuai pada kecamatan yang berada pada lingkup Kabupaten Madiun. Kepadatan penduduk berada diantara $3446,041$ sampai $3853,796$ jiwa/ km^2 , apabila kepadatan penduduk naik sebesar satu satuan mengakibatkan jumlah kasus DBD cenderung naik sebesar 1,3698 kasus per 100.000 penduduk. Pada saat kepadatan penduduk lebih dari $3853,796$ jiwa/ km^2 , apabila kepadatan penduduk naik sebesar satu satuan mengakibatkan jumlah kasus DBD cenderung turun sebesar 0,159 kasus per 100.000 penduduk. Daerah yang kepadatan penduduknya lebih dari $3853,796$ jiwa/ km^2 adalah Kecamatan Taman, Kartoharjo dan Mangunharjo. Hal ini berlaku apabila variabel lain yang mempengaruhinya bernilai konstan.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dari analisis dan pembahasan yang dilakukan sebelumnya, maka dapat diambil beberapa kesimpulan dari pemodelan angka penderita Demam Berdarah Dengue (DBD) di Kabupaten/Kota Madiun pada tahun 2010 dengan pendekatan regresi semiparametrik spine yaitu sebagai berikut.

- 1 Berdasarkan hasil dari deskripsi permasalahan yaitu banyaknya jumlah kasus DBD yang terjadi tahun 2010 sebesar 498 kasus dan mengakibatkan 7 orang meninggal. Berdasarkan wilayahnya, di Kabupaten Madiun terdapat 411 kasus yang menyebabkan 6 orang meninggal dan Kota Madiun terdapat 87 kasus dan menyebabkan 1 orang meninggal dunia. Rata-rata kasus DBD yang terjadi sebesar 53,11 kasus per 100.000 penduduk dengan jumlah kasus tertinggi terjadi di Kecamatan Madiun dimana terjadi 61 kasus atau 150,93 kasus per 100.000 penduduk. Sedangkan kasus terendah terjadi di Kecamatan Kare dimana hanya terjadi 2 kasus atau 5,68 kasus per 100.000 penduduk. Faktor yang diduga mempengaruhi kasus DBD antara lain variabel angka bebas jentik (x_1) yang mempunyai rata-rata 70,29 persen, persentase penduduk miskin (t_1) yang rata-ratanya 26,62 persen, curah hujan (t_2) dengan rata-rata 210,1 mm³, persentase rumah atau bangunan sehat (t_3) yang mempunyai rata-rata sebesar 65,1 persen, kepadatan penduduk (t_4) dengan rata-rata 1856 jiwa/km² dan institusi yang dibina kesehatan lingkungannya (t_5) yang rata-ratanya 90,88 persen.
- 2 Berdasarkan hasil pengujian kecukupan data dan homogenitas varians menunjukkan bahwa data telah memenuhi untuk dilakukan analisis faktor. Analisis faktor membentuk 3 faktor baru dan dipilih masing-masing satu

variabel dalam setiap faktor untuk digunakan sebagai variabel prediktor dalam pemodelan kasus DBD. Hal tersebut dilakukan untuk menjamin bahwa variabel prediktor hasil reduksi tidak berkorelasi satu sama lain. Hasil reduksi tersebut menghasilkan variabel prediktor yang digunakan dalam pemodelan banyaknya kasus DBD yaitu kepadatan penduduk (t_4) yang mewakili faktor pertama, ABJ (x_1) dan persentase penduduk miskin (t_1) yang mewakili faktor kedua dan curah hujan (t_2) mewakili faktor ketiga.

- 3 Model regresi yang digunakan dalam melakukan analisis adalah dengan menggunakan regresi semiparametrik spline dengan titik knot optimal yang diperoleh pada kombinasi knot (322) dengan nilai GCV 181,437 dan MSE sebesar 70,479 adalah sebagai berikut.

$$\hat{y} = 22,1045 - 1,6967x_1 - 0,6612t_1 + 11,8618(t_1 - 19,669)_+ - 48,3093(t_1 - 25,381)_+ + 38,5898(t_1 - 28,951)_+ + 1,691t_2 + 5,2401(t_2 - 157,27)_+ + 3,6136(t_2 - 173,148)_+ - 0,0419t_4 + 1,4117(t_4 - 3446,041)_+ - 1,5288(t_4 - 3853,796)_+$$

Model yang dihasilkan telah memenuhi asumsi IIDN ($0, \sigma^2$). Selain itu model tersebut dapat menjelaskan variabilitas jumlah kasus DBD sebesar 98,351 persen dengan semua variabel prediktor yang digunakan berpengaruh signifikan terhadap model.

5.2 Saran

Untuk penelitian selanjutnya disarankan melakukan penambahan orde pada model regresi semiparametrik spline baik kuadratik maupun kubik dengan berbagai kombinasi titik knot. Tidak menutup kemungkinan penggunaan analisis statistik lain selain analisis faktor dalam pemilihan variabel prediktor dapat membantu menghasilkan model dengan interpretasi yang lebih sesuai.

DAFTAR PUSTAKA

- Akhir, Y.S. (2013). *Pemodelan Jumlah Kasus Demam Berdarah Dengue (DBD) di Kabupaten Bojonegoro Jawa Timur Tahun 2011 dengan Geographically Weighted Poisson Regression*. Tugas Akhir Jurusan Statistika ITS, Surabaya.
- Asmin, S. (2010). *Pemodelan Nilai Unas IPA dengan Pendekatan Regresi Semiparametrik Spline di SMAN 1 Grati Pasuruan*. Tugas Akhir Jurusan Statistika ITS, Surabaya.
- Aunuddin. (2005). *Statistika: Rancangan dan Analisis Data*. IPB Press, Bogor.
- Budiantara, I.N. (2006). *Model Spline dengan Knots Optimal*. Jurnal Ilmu Dasar FMIPA Universitas Jember, Vol.7, Hal.77-85 Jember.
- Budiantara, I.N. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Institut Teknologi Sepuluh Nopember, Surabaya.
- Departemen Kesehatan RI. (2005). *Pencegahan dan Pemberantasan Demam Berdarah Dengue di Indonesia*. Dirjen Pengendalian Penyakit dan Penyehatan Lingkungan, Jakarta.
- Dinas Kesehatan Provinsi Jawa Timur. (2009). *Profil Kesehatan Provinsi Jawa Timur*. Pemerintah Provinsi Jawa Timur, Jawa Timur.
- Draper, N. & Smith, H. (1996). *Analisis Regresi Terapan Edisi Kedua*. PT. Gramedia Pustaka Utama, Jakarta.
- Eubank, R.L. (1988), *Spline Smoothing and Nonparametric Regression*. Marcel Dekker, New York.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing 2nd Edition*. Marcel Dekker, New York.
- Gujarati, D.N. (2006). *Dasar-Dasar Ekonometrika Edisi Pertama*. Jakarta, Erlangga.

- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2010). *Multivariate Data Analysis (7th Edition ed.)*. Pearson Prentice Hall, New Jersey.
- Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, New York.
- Hidayati, R. (2008). *Model Peringatan Dini Penyakit Demam Berdarah dengan Informasi Unsur Iklim*. Disertasi Sekolah Pascasarjana IPB, Bogor.
- Iriani, Y. (2012). *Hubungan antara Curah Hujan dan Peningkatan Kasus Demam Berdarah Dengue Anak di Kota Palembang*. Sari Pediatri Vol. 13, No 6, Palembang.
- Johnson, R., & Wichern, D. (2002). *Applied Multivariate Statistical Analysis, 5th edition*. Prentice Hall International Inc, New Jersey.
- Kementerian Kesehatan Republik Indonesia (2011). *Profil Kesehatan Indonesia 2010*. Kemenkes RI, Jakarta.
- Kristina, Isminah & Wulandari, L. (2004). *Kajian Masalah Kesehatan: Demam Berdarah Dengue*. Badan Penelitian dan Pengembangan Kesehatan Departemen Kesehatan, Jakarta.
- Marina, S.M.T. (2013). *Pemodelan Faktor-faktor yang Mempengaruhi Persentase Kriminalitas di Jawa Timur Dengan Pendekatan Regresi Semiparameterik Spline*. Tugas Akhir Jurusan Statistika ITS, Surabaya.
- Mubarak, W.I & Chayatin, N. (2009). *Ilmu Kesehatan Masyarakat : Teori dan Aplikasi*. Salemba Medika, Jakarta.
- Ratnasari, A. (2009). *Faktor Resiko yang Mempengaruhi Terjangkitnya Penyakit Demam Berdarah Dengue di Daerah Endemis (DKI Jakarta)*. Tugas Akhir. Jurusan Statistika ITS, Surabaya.
- Ruppert, M. & Carroll, R. (2003). *Semiparametric Regression*. Cambridge University Press, New York.
- Ryan, T.P. (1997). *Modern Regression Method*. John Wiley & Sons. Inc, Canada.

- Wahba, G. (1990). *Spline Models for Observation Data*. SIAM, Pennsylvania.
- Walpole, R.E. (1995). *Pengantar Statistika*. PT Gramedia Pustaka Utama, Jakarta.
- WHO. (1999). *Strengthening Implementation of the Global Strategy for Dengue Fever and Dengue Hemorrhagic Fever, prevention and control*. WHO HQ, Geneva.
- WHO. (1997). *Dengue Hemorrhagic Fever: diagnosis, treatment, prevention and control, Second edition*. WHO HQ, Geneva.
- WHO. (2000). *Pencegahan dan Penanggulangan Penyakit Demam Berdarah Dengue*, Terjemahan dari WHO Regional Publication SEARO No.29: Prevention Control of Dengue and Dengue Haemorrhagic Fever. Depkes RI, Jakarta.
- Wiersma, W. (1986). *Research methods in education : an Introduction 4th ed*. Allyn and Bacon Inc, Boston.
- Yuswulandary, V. (2008). *Karakteristik Penderita Demam Berdarah Dengue di Wilayah Kerja Dinas Kesehatan Kota Lhoksumawe dan Kegiatan Pemberantasannya Tahun 2003 – 2007*. Program Sarjana Fakultas Kesehatan Masyarakat Universitas Sumatera Utara, Sumatera Utara.

(Halaman ini sengaja dikosongkan)

BIOGRAFI PENULIS



Penulis dengan nama lengkap Fadli Akbar Pambudi lahir di Madiun, pada tanggal 10 Mei 1991. Penulis merupakan anak kedua dari dua bersaudara dari pasangan Drs. Teguh Soewarno dan Tri Endah Listiani, ST. Penulis telah menempuh pendidikan formal di SDN 01 Klegen Madiun, SMP Negeri 1 Madiun dan SMA Negeri 2 Madiun. Penulis diterima di Jurusan Statistika FMIPA ITS pada tahun 2009 melalui jalur Seleksi

Nasional Masuk Perguruan Tinggi Negeri (SNMPTN). Selama menjadi mahasiswa, penulis pernah tergabung dalam organisasi HIMASTA-ITS sebagai Tim Ahli dan Divisi SCC selama dua periode yaitu 2010/2011 dan 2011/2012. Penulis pernah mengemban amanah sebagai Ketua PRS 2011 dan menjadi SC (*Steering Comitee*) PRS periode 2012 dan 2013. Selain itu penulis merupakan delegasi ITS dalam organisasi IHMSI (Ikatan Himpunan Mahasiswa Statistika Indonesia) dan menjadi Ketua Bidang di Wilayah IV periode 2010-2012. Ketika senggang, penulis mengisi waktu dengan menyalurkan hobinya yaitu bermain badminton dan membaca buku. Bagi pembaca yang memiliki saran, kritik atau ingin berdiskusi lebih lanjut terkait dengan metode pada Tugas Akhir ini dengan penulis dapat disampaikan melalui email fadli.akbar.p@gmail.com

DAFTAR LAMPIRAN

Lampiran 1	Data Kasus Demam Berdarah Dengue (DBD) dan Faktor-Faktor yang Diduga Mempengaruhi Di Madiun Tahun 2010	61
Lampiran 2	Output Analisis Faktor	63
Lampiran 3	Program Regresi Semiparametrik Spline Menggunakan GCV Satu, Dua, Tiga dan Kombinasi Knot Menggunakan <i>Software R</i>	64
Lampiran 4	Program Uji Signifikansi Parameter Menggunakan <i>Software R</i>	77
Lampiran 5	Program Uji Glejser Menggunakan <i>Software R</i>	80
Lampiran 6	<i>Output</i> Uji Signifikansi Parameter	82
Lampiran 7	<i>Output</i> Uji Glejser	85
Lampiran 8	<i>Output</i> Uji Normalitas Residual Menggunakan Kolmogorov-Smirnov	86

LAMPIRAN 1. Data Kasus Demam Berdarah Dengue (DBD) dan Faktor-Faktor yang Diduga Mempengaruhi Di Madiun Tahun 2010.

Kecamatan	y	x_1	t_1	t_2
KEBONSARI	85.404	50.880	11.222	126.583
DOLOPO	68.403	59.246	10.142	266.500
GEGER	73.800	53.159	21.197	263.583
DAGANGAN	27.786	63.333	31.618	266.500
KARE	5.683	95.675	44.657	208.416
GEMARANG	32.580	78.571	42.505	236.333
WUNGU	60.157	60.544	18.541	255.000
JIWAN	41.588	71.732	9.673	259.500
BALEREJO	55.730	66.711	40.834	161.000
MEJAYAN	36.980	59.813	28.295	195.167
SARADAN	19.953	68.466	36.116	185.500
PILANGKENCENG	18.718	70.543	28.565	289.583
MADIUN	150.930	44.213	14.739	161.000
WONOASRI	40.058	67.855	31.626	161.000
SAWAHAN	114.582	69.538	24.902	259.500
TAMAN	55.117	91.224	23.891	234.917
MANGUNHARJO	30.670	98.934	29.749	222.250
KARTOHARJO	37.7752	94.716	30.811	30.250

LAMPIRAN 1. Data Kasus Demam Berdarah Dengue (DBD) dan Faktor-Faktor yang Diduga Mempengaruhi Di Madiun Tahun 2010 (Lanjutan).

Kecamatan	t_3	t_4	t_5
KEBONSARI	67.02	1283	60.58
DOLOPO	82.34	1287	93.15
GEGER	84.49	1851	92.69
DAGANGAN	54.16	746	100
KARE	93.20	184	100
GEMARANG	38.39	351	100
WUNGU	70.70	1387	99.3
JIWAN	52.82	1781	100
BALEREJO	34.89	863	97.42
MEJAYAN	67.50	930	87.93
SARADAN	55.36	492	87.02
PILANGKENCENG	57.57	722	100
MADIUN	85.42	1125	78.26
WONOASRI	58.99	1030	93.2
SAWAHAN	27.86	1182	100
TAMAN	84.98	6844	64.34
MANGUNHARJO	72.38	6170	93.83
KARTOHARJO	83.60	5181	88

Keterangan :

- y : Persentase jumlah kasus DBD di Madiun
- x_1 : Persentase rumah atau bangunan bebas jentik nyamuk *Aedes aegypti* atau angka bebas jentik
- t_1 : Persentase penduduk miskin
- t_2 : Curah hujan
- t_3 : Persentase rumah atau bangunan sehat
- t_4 : Kepadatan penduduk
- t_5 : Institusi yang dibina kesehatan lingkungannya

LAMPIRAN 2. Output Analisis Faktor.**KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.533
Bartlett's Test of Sphericity	Approx. Chi-Square	44.536
	Df	21
	Sig.	.002

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.166	36.104	36.104	2.166	36.104	36.104
2	1.767	29.453	65.557	1.767	29.453	65.557
3	1.011	16.842	82.398	1.011	16.842	82.398
4	.613	10.211	92.609			
5	.332	5.535	98.144			
6	.111	1.856	100.000			

Extraction Method : Principal Component Analysis

Rotated Component Matrix^a

	Component		
	1	2	3
x1	.591	.772	.060
t1	-.211	.900	-.097
t2	-.077	-.176	.935
t3	.668	-.296	-.192
t4	.906	.117	-.124
t5	-.393	.474	.648

LAMPIRAN 3. Program Regresi Semiparametrik Spline Menggunakan GCV Satu, Dua, Tiga dan Kombinasi Knot Menggunakan *Software R*.

Program GCV Untuk 1 Titik Knot

```
GCV1=function(para)
{
  data=read.table("d:/spline.txt")
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot1[j,i]=a[j]
    }
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
```

```

    for (k in 1:p)
    {
        if (data[k,(j+para+1)]<knot1[i,j])    data1[k,j]=0    else
data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
    }
}
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (Rsqr)

```

```

cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d:/output GCV1fix.csv")
write.csv(Rsq, file="d:/output Rsq1fix.csv")
write.csv(knot1, file="d:/output knot1fix.csv")
}

```

Program GCV Untuk 2 Titik Knot

```

GCV2=function(para)
{
data=read.table("d:/spline.txt")
data=as.matrix(data)
p=length(data[,1])
q=length(data[,1])
m=ncol(data)-para-1
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
}
}

```

```

a1=nrow(knot)
knot=knot[2:(a1-1),]
a2=nrow(knot)
z=(a2*(a2-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
      for (j in 1:(a2-1))
      {
          for (k in (j+1):a2)
          {
              xx=cbind(knot[j,i],knot[k,i])
              knot1=rbind(knot1,xx)
          }
      }
knot2=cbind(knot2,knot1)
}
knot2=knot2[2:(z+1),2:(2*m+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[,:(para+2):q]
data3=data[,2:q]
a3=length(knot2[,1])
GCV=rep(NA,a3)
Rsq=rep(NA,a3)
for (i in 1:a3)
{
  for (j in 1:(2*m))
  {
    if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
    for (k in 1:p)
    {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else
data2[k,j]=data1[k,b]-knot2[i,j]
    }
  }
}

```

```

    }
    mx=cbind(aa,data3,data2)
    mx=as.matrix(mx)
    C=pinv(t(mx)%*%mx)
    B=C%*%(t(mx)%*%data[,1])
    yhat=mx%*%B
    SSE=0
    SSR=0
    for (r in (1:p))
    {
        sum=(data[r,1]-yhat[r,])^2
        sum1=(yhat[r,]-mean(data[,1]))^2
        SSE=SSE+sum
        SSR=SSR+sum1
    }
    Rsq[i]=(SSR/(SSE+SSR))*100
    MSE=SSE/p
    A=mx%*%C%*%t(mx)
    A1=(F-A)
    A2=(sum(diag(A1))/p)^2
    GCV[i]=MSE/A2
    }
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print(knot1)
print (knot2)
    cat("=====", "\n")
cat("Rsq dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (GCV)

```

```

s1=min(GCV)
print(max(Rsq))
cat("=====","\n")
cat("HASIL GCV terkecil dengan Spline linear 2

knot","\n")
cat("=====","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV2.csv")
write.csv(Rsq,file="d:/output Rsq2.csv")
write.csv(knot2,file="d:/output knot2.csv")
}

```

Program GCV Untuk 3 Titik Knot

```

GCV3=function(para)
{
data=read.table("d:/spline.txt")
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[(para+2):q]
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
}
knot=knot[2:(nk-1),]
a2=nrow(knot)

```

```

z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
      for (j in 1:(a2-2))
      {
        for (k in (j+1):(a2-1))
        {
          for (g in (k+1):a2)
          {
            xx=cbind(knot[j,i],knot[k,i],knot[g,i])
            knot2=rbind(knot2,xx)
          }
        }
      }
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsq=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else
data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
}

```

```

mx=cbind(aa,data[,2:q],data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====\n")
cat("Nilai Knot dengan Spline linear 3 knot","\n")
cat("=====\n")
print (knot1)
cat("=====\n")
cat("Rsq dengan Spline linear 3 knot","\n")
cat("=====\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=====\n")
cat("HASIL GCV dengan Spline linear 3 knot","\n")
cat("=====\n")
print (GCV)

```

```

s1=min(GCV)
cat("=====","\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot","\n")
cat("=====","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV3fx.csv")
write.csv(Rsq,file="d:/output Rsq3fx.csv")
write.csv(knot1,file="d:/output knot3fx.csv")
}

```

#PROGRAM GCV UNTUK KNOT KOMBINASI

```

GCVkom=function(para)
{
data=read.table("d:spline.txt")
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("d:/GCVvar1.txt")
x2=read.table("d:/GCVvar2.txt")
x3=read.table("d:/GCVvar3.txt")
n2=nrow(x1)
a=matrix(nrow=3,ncol=3^(v))
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
{
m=m+1
a[,m]=c(i,j,k)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^(v))
for (i in 1:3^(v))

```

```

{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
}
}

```

```

if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[, (v+1)])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )

```

```

gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[, 2:3])
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[, 4:6])
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa, bb, cc))
mx=cbind(rep(1, nrow(data)), data[, 2:q1], na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[, 1])
yhat=mx%*%B

```

```

SSE=0
SSR=0
for (r in 1:nrow(data))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}
Rsq=(SSR/(SSE+SSR))*100
MSE=SSE/p1
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}
if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
kkk=cbind(sp,spl,splin)
cat("=====", "\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="d:/GCVkom.csv")
write.csv(Rsq,file="d:/Rsqkom.csv")
}

```

LAMPIRAN 4. Program Uji Signifikasi Parameter Menggunakan *Software R*.

#UJI SIGNIFIKASI PARAMETER

```

uji=function(alpha,para)
{
data=read.table("d:/spline.txt")
knot=read.table("d:/knot 322.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],
data[,m+2],data[,m+2])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
      for(j in 1:p)
      {
            if      (dataA[j,i]<knot[1,i])      data.knot[j,i]=0      else
data.knot[j,i]=dataA[j,i]-knot[1,i]
      }
}
mx=cbind(satu,
data[,2:m],data.knot[,1:3],data[,m+1],data.knot[,4:5],data[,m+2],
data.knot[,6:7])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")

```

```

print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/(SSR+SSE))*100
#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan", "\n")
cat("", "\n")
}
#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*% mx))))
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")

```

```

cat("-----","\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]
pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan dengan
pvalue",pval[i],"\\n") else cat("Gagal tolak Ho yakni prediktor tidak
signifikan dengan pvalue",pval[i],"\\n")
}
thit=as.matrix(thit)
cat("=====","\\n")
cat("nilai t hitung","\\n")
cat("=====","\\n")
print (thit)
cat("Analysis of Variance","\\n")
cat("=====","\\n")
      cat("Sumber      df    SS    MS    Fhit","\\n")
      cat("Regresi      ",(n1-1)," ",SSR," ",MSR," ",Fhit,"\\n")
      cat("Error        ",p-n1," ",SSE," ",MSE,"\\n")
      cat("Total        ",p-1," ",SST,"\\n")
cat("=====","\\n")
      cat("s=",sqrt(MSE),"      Rsq=",Rsq,"\\n")
      cat("pvalue(F)=",pvalue,"\\n")
write.csv(res,file="d:/residual.csv")
write.csv(pval,file="d:/pvalue.csv")
write.csv(mx,file="d:/mx.csv")
write.csv(yhat,file="d:/yhat.csv")
}

```

LAMPIRAN 5. Program Uji Glejser Menggunakan *Software R***#UJI GLEJSER**

```

uji=function(alpha)
{
matX=read.csv("d:/mx.csv",header=FALSE)
res=read.csv("d:/residual.csv",header=FALSE)
matX=as.matrix(matX)
res=abs(res)
res=as.matrix(res)
resbar=mean(res)
p=nrow(res)
B=(pinv(t(matX)%*%matX))%*%t(matX)%*%res
n1=nrow(B)
yhat=matX%*%B
res1=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-resbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan","\n")
cat("","\n")
}
else
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")

```

```

cat("-----","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan","\n")
cat("","\n")
}
cat("Analysis of Variance","\n")
cat("=====","\n")
cat("Sumber    df    SS    MS    Fhit","\n")
cat("Regresi   ",n1-1," ",SSR," ",MSR,"",Fhit,"\n")
cat("Error     ",p-n1," ",SSE,"",MSE,"\n")
cat("Total     ",p-1," ",SST,"\n")
cat("=====","\n")
cat("Fhit","\n")
print(Fhit)
cat("pvalue","\n")
print(pvalue)
}

```

LAMPIRAN 6. *Output Uji Signifikansi Parameter*

=====
 Estimasi Parameter
 =====

[,1]

[1,] 22.10453598

[2,] -1.69674018

[3,] -0.66123968

[4,] 11.86178552

[5,] -48.30927785

[6,] 38.58983998

[7,] 1.69097752

[8,] -5.24005794

[9,] 3.61363033

[10,] -0.04188424

[11,] 1.41168867

[12,] -1.52880283

 Kesimpulan hasil uji serentak

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

Kesimpulan hasil uji individu

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue
 0.5654949

Tolak Ho yakni prediktor signifikan dengan pvalue
 0.00322803

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue
 0.5375629

Tolak Ho yakni prediktor signifikan dengan pvalue
0.009133

Tolak Ho yakni prediktor signifikan dengan pvalue
6.318209e-05

Tolak Ho yakni prediktor signifikan dengan pvalue
2.509369e-05

Tolak Ho yakni prediktor signifikan dengan pvalue
0.00022255

Tolak Ho yakni prediktor signifikan dengan pvalue
0.0005597

Tolak Ho yakni prediktor signifikan dengan pvalue
0.0025715

Tolak Ho yakni prediktor signifikan dengan pvalue
0.0040325

Tolak Ho yakni prediktor signifikan dengan pvalue
0.0001557

Tolak Ho yakni prediktor signifikan dengan pvalue
0.0001488

=====
nilai t hitung
=====

[,1]

[1,] 0.6079689

[2,] -4.7289475

[3,] -0.6536567

[4,] 3.7843738

[5,] -9.8485435

[6,] 11.5687374

[7,] 7.8715971

[8,] -6.6473559

[9,] 4.9524875

[10,] -4.5164931

[11,] 8.3930299

[12,] -8.4613007

Analysis of Variance

```
=====
Sumber      df      SS      MS      Fhit
Regresi    11    21646.26  1967.842  32.5376
Error       6      362.87   60.479
Total      17    22009.13
=====
```

s= 7.776824 Rsq= 98.35126

pvalue(F)= 0.0001893929

LAMPIRAN 7. *Output Uji Glejser*

 Kesimpulan hasil uji serentak

 Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan

Analysis of Variance

```
=====
Sumber      df      SS      MS      Fhit
Regresi    11  104.3506  9.486415  1.940328
Error       6   29.3347  4.889078
Total      17  133.685
=====
```

Fhit

[1] 1.940328

pvalue

[1] 0.2146641

LAMPIRAN 8. *Output Uji Normalitas Residual Menggunakan Kolmogorov-Smirnov*