TESIS – SS 142501

MODEL REGRESI SEMIPARAMETRIK CAMPURAN SPLINE TRUNCATED DAN DERET FOURIER
(Studi Kasus : Angka Harapan Hidup Provinsi Jawa Timur)

KHAERUN NISA
NRP. 1315 201 018

DOSEN PEMBIMBING
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
Dr. Agnes Tuti Rumiati, M.Sc.

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017

THESIS - SS 142501

Semiparametric Regression Model with Combined Estimator of Truncated Spline and Fourier Series (Case Study : Life Expectancy in Province of East Java)

KHAERUN NISA
NRP. 1315 201 018

Supervisor
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
Dr. Agnes Tuti Rumiati, M.Sc.

MAGISTER PROGRAM
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY
SURABAYA
2017

MODEL REGRESI SEMIPARAMETRIK CAMPURAN SPLINE TRUNCATED DAN DERET FOURIER (Studi Kasus: Angka Harapan Hidup Provinsi Jawa Timur)

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar *Magister Sains* (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh:

KHAERUN NISA
NRP. 1315 201 018

Tanggal Ujian : 16 Januari 2017
Periode Wisuda : Maret 2017

Disetujui oleh:

1. **Prof. Dr. Drs. I Nyoman Budiantara, M.Si.** (Pembimbing I)
   NIP. 19650603 198903 1 003

2. **Dr. Agnes Tuti Rumiati, M.Sc.** (Pembimbing II)
   NIP. 19570724 198503 2 002

3. **Dr. Wahyu Wibowo, M.Si.** (Penguji)
   NIP. 19740328 199802 1 001

4. **R. Mohamad Atok, M.Si., Ph.D.** (Penguji)
   NIP. 19710915 199702 1 001

an. Direktur Program Pascasarjana
Asisten Direktur

Prof. Dr. Ir. Tri Widjaja, M.Eng.
NIP. 19611021 198603 1 001

Direktur Program Pasca Sarjana,

Prof. Ir. Djauhar Manfaat, M.Sc., Ph.D.
NIP.19601202 198701 1 001

MODEL REGRESI SEMIPARAMETRIK CAMPURAN SPLINE TRUNCATED DAN DERET FOURIER
(Studi Kasus : Angka Harapan Hidup Provinsi Jawa Timur)

Nama Mahasiswa	: Khaerun Nisa
NRP	: 1315201018
Pembimbing 1	: Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
Co-Pembimbing	: Dr.Agnes Tuti Rumiati, M.Sc.

ABSTRAK

Diberikan data berpasangan $(x_{1i},\ldots,x_{pi},t_{1i},\ldots,t_{qi},z_{1i},\ldots,z_{ri},y_i)$, hubungan antar variabel prediktor dengan variabel respon mengikuti model regresi semiparametrik campuran

$$y_i = \mu(x_{1i},\ldots,x_{pi},t_{1i},\ldots,t_{qi},z_{1i},\ldots,z_{ri}) + \varepsilon_i .$$

Kurva regresi bersifat aditif, sehingga dapat dituliskan :

$$\mu(x_{1i},\ldots,x_{pi},t_{1i},\ldots,t_{qi},z_{1i},\ldots,z_{ri}) = \sum_{j=1}^{p} f_j(x_{ji}) + \sum_{s=1}^{q} g_s(t_{si}) + \sum_{l=1}^{r} h_l(z_{li})$$

Komponen kurva regresi $f_j(x_{ji})$ didekati dengan fungsi parametrik linier, komponen kurva regresi $g_s(t_{si})$ didekati dengan fungsi Spline *Truncated* dan komponen kurva regresi $h_l(z_{li})$ didekati dengan fungsi Deret Fourier. Tujuan dari penelitian ini adalah memperoleh bentuk estimator dalam regresi semiparametrik dengan menggunakan estimator campuran Spline *Truncated* dan Deret Fourier menggunakan metode *Penalized Least Square* (PLS), serta memodelkan Angka Harapan Hidup di Provinsi Jawa Timur menggunakan model regresi semiparametrik campuran tersebut. Hasil kajian menghasilkan bahwa estimator kurva regresi parametrik linier $\hat{\hat{f}}(x) = \mathbf{A}(K,\lambda,k)\tilde{Y}$, estimator kurva regresi Spline *Truncated* $\hat{\hat{g}}(t) = \mathbf{B}(K,\lambda,k)\tilde{Y}$ dan kurva regresi Deret Fourier $\hat{\hat{h}}(z) = \mathbf{C}(K,\lambda,k)\tilde{Y}$. Selanjutnya, diperoleh estimasi model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier adalah

$$\hat{y}_i = \hat{\mu}\left(x_{1i},\ldots,x_{pi},t_{1i},\ldots,t_{qi},z_{1i},\ldots,z_{ri}\right) = \hat{\hat{f}}(x) + \hat{\hat{g}}(t) + \hat{\hat{h}}(z) = \mathbf{F}(k,\lambda,K)\tilde{Y},$$

dengan $\mathbf{F}(K,\lambda,k) = \mathbf{A}(K,\lambda,k) + \mathbf{B}(K,\lambda,k) + \mathbf{C}(K,\lambda,k).$ Model regresi semiparametrik campuran ini bergantung pada lokasi titik-titik knot *K,* parameter penghalus lambda dan osilasi *k.* Model regres semiparametrik campuran Spline *Truncated* dan Deret Fourier terbaik diperoleh dengan cara meminimumkan fungsi *Generalized Cross Validation*. Model regresi semiparametrik campuran yang diperoleh digunakan untuk memodelkan data kasus Angka Harapan Hidup (AHH) di Provinsi Jawa Timur. Model estimator campuran tersebut menghasilkan R^2 sebesar 99,62%.

Kata Kunci: Regresi Semiparametrik, Estimator Campuran, Spline *Truncated*, Deret Fourier, PLS.

Halaman ini sengaja dikosongkan

SEMIPARAMETRIC REGRESSION MODEL WITH COMBINED ESTIMATOR OF SPLINE TRUNCATED AND FOURIER SERIES

(Case Study : Life Expectancy in Province of East Java)

Name	: Khaerun Nisa
NRP	: 1315201018
Supervisor	: Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
Co-Supervisor	: Dr.Agnes Tuti Rumiati, M.Sc.

ABSTRACT

Given the data pairs $(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri}, y_i)$. The relationship between predictor variables with response variable following semiparametric regression model

$$y_i = \mu\left(x_{1i}, \ldots, x_{pi}, t_{1i}, \ldots, t_{qi}, z_{1i}, \ldots, z_{ri}\right) + \varepsilon_i.$$

The Regression curves are additive, so it can be written :

$$\mu\left(x_{1i}, \ldots, x_{pi}, t_{1i}, \ldots, t_{qi}, z_{1i}, \ldots, z_{ri}\right) = \sum_{j=1}^{p} f_j(x_{ji}) + \sum_{s=1}^{q} g_s(t_{si}) + \sum_{l=1}^{r} h_l(z_{li}).$$

Component of regression curve $f_j(x_{ji})$ approached by a linear parametric function, component regression curve $g_s(t_{si})$ approached by Spline Truncated function and component regression curve $h_l(z_{li})$ approached by Fourier Series function. The purpose of this research are to obtained the estimator of semiparametric regression model with combined estimator of Spline Truncated and Fourier Series using Penalized Least Square method (PLS), and modeling the case of life expectancy in Province of East Java using mixture semiparametric regression model. The results show that the estimator of parametric linier regression curve $\hat{\hat{f}}(x) = \mathbf{A}(K,\lambda,k)\tilde{Y}$, the estimator of Spline Truncated $\hat{\hat{g}}(t) = \mathbf{B}(K,\lambda,k)\tilde{Y}$ and the estimator of Fourier Series is $\hat{\hat{h}}(z) = \mathbf{C}(K,\lambda,k)\tilde{Y}$. Furthermore, the mix estimator of Spline Truncated and Fourier Series in semiparametric regression model is

$$\hat{y}_i = \hat{\mu}\left(x_{1i}, \ldots, x_{pi}, t_{1i}, \ldots, t_{qi}, z_{1i}, \ldots, z_{ri}\right) = \hat{\hat{f}}(x) + \hat{\hat{g}}(t) + \hat{\hat{h}}(z) = \mathbf{F}(k,\lambda,K)\tilde{Y},$$

with $\mathbf{F}(K,\lambda,k) = \mathbf{A}(K,\lambda,k) + \mathbf{B}(K,\lambda,k) + \mathbf{C}(K,\lambda,k).$ This mixture semiparametric regression model depends on the location of the dots knots , smoothing parameter lambda and oscillation. The best model of semiparametric regression model with combined estimator of Spline Truncated and Fourier series can be obtained by minimizing the function of Generalized Cross Validation. The mixture estimator model produces R^2 of 99,62%.

Key Words : Semiparametric Regression, Spline Truncated, Fourier Series, Combined Estimator, PLS.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Alhamdulillah, Puji syukur kehadirat Allah SWT karena berkat limpahan rahmat-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul :
"Model Regresi Semiparametrik Campuran *Spline Truncated* dan Deret Fourier (Studi Kasus : Angka Harapan Hidup Provinsi Jawa Timur)". Tesis ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Studi Statistika, Program Pascasarjana, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Selesainya laporan Tesis ini tak lepas dari peranan berbagai pihak. Oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih yang sedalam-dalamnya kepada :

1. Kedua orang tua yang sangat saya cintai dan hormati, Bapak Muh.Ramli dan Ibu Hasmi, terima kasih atas segala doa dan dukungan, baik moral maupun meteril yang tiada henti. Semoga selalu bisa membahagiakan Bapak dan Ibu.
2. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si dan Ibu Dr. Agnes Tuti Rumiati, M.Sc, selaku dosen pembimbing yang telah bersedia meluangkan waktu di tengah kesibukannya untuk membimbing penulis.
3. Bapak Dr.Wahyu Wibowo, S.Si., M.Si dan Bapak Dr. R. Mohamad. Atok, S.Si., M.Si selaku penguji yang telah memberikan kritik, saran serta masukan demi kesempurnaan tesis ini.
4. Bapak Dr. Suhartono, selaku ketua Jurusan Statistika ITS dan Staff, karyawan TU, RBS Jurusan Statistika ITS.
5. Bapak Dr. Heri Kuswanto, M.Si selaku Kaprodi Pascasarjana Jurusan Statistika ITS.
6. Bapak dan Ibu dosen pengajar Jurusan Statistika ITS, terima kasih ilmu yang telah diberikan.
7. Seluruh keluarga besar, terima kasih dukungan dan motivasi yang diberikan.
8. Semua rekan-rekan Pasca Sarjana Statistika ITS, terima kasih telah menjadi keluarga baru dan keluarga yang baik selama di Surabaya

9. Serta semua pihak yang tidak dapat saya sebutkan satu-persatu.

Besar harapan penulis agar Tesis ini bermanfaat dan dapat menambah wawasan keilmuan. Penulis menyadari sepenuhnya bahwa Tesis ini belum sempurna, oleh karena itu saran dan kritik yang membangun sangat penulis harapkan.

Surabaya, Januari 2017

Penulis

DAFTAR ISI

Halaman

HALAMAN 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
1.1 Latar Belakang 1
1.2 Rumusan Masalah 4
1.3 Tujuan Penelitian 5
1.4 Manfaat Penelitian 5
1.5 Batasan Masalah 5

BAB II TINJAUAN PUSTAKA 7
2.1 Analisis Regresi 7
2.2 Regresi Parametrik, Nonparametrik dan Semiparametrik 7
2.3 Regresi Semiparametrik Spline *Truncated* 8
2.4 Regresi Semiparametrik Deret Fourier 12
2.5 Koefisien Determinasi 15
2.6 *Generalized Cross Validation* (GCV) 15
2.6 *Penalized Least* Square (PLS) 16
2.7 Teorema Dasar Aljabar Matriks 16
2.8 Angka Harapan Hidup (AHH) 17

BAB III METODOLOGI PENELITIAN 21
3.1 Metode Penelitian 21
3.2 Sumber Data 25

3.3 Variabel Penelitian dan Struktur Data 26
3.4 Definisi Operasional Variabel Penelitian..26

BAB IV HASIL DAN PEMBAHASAN 29
4.1 Model Regresi Semiparametrik Campuran 29
4.2 Estimasi Kurva Regresi Semiparametrik Campuran 33
4.3 Pemodelan AHH Provinsi Jawa Timur..42
4.3.1 Eksplorasi Data 42
4.3.2 Model Umum Regresi Semiparametrik Campuran 48
4.3.3 Pemilihan Titik Knot, Parameter Penghalus
dan Osilasi *k* Optimum..49
4.3.3.1 Model dengan 1 Titik Knot dan Osilasi *k*=1 49
4.3.3.2 Model dengan 2 Titik Knot dan Osilasi *k*=1...................50
4.3.3.3 Model dengan 3 Titik Knot dan Osilasi *k*=1...................50
4.3.3.4 Model dengan 1 Titik Knot dan Osilasi *k*=2 51
4.3.3.5 Model dengan 2 Titik Knot dan Osilasi *k*=2...................52
4.3.3.6 Model dengan 3 Titik Knot dan Osilasi *k*=2...................53
4.3.3.7 Model dengan 1 Titik Knot dan Osilasi *k*=3 54
4.3.3.8 Model dengan 2 Titik Knot dan Osilasi *k*=3...................55
4.3.3.9 Model dengan 3 Titik Knot dan Osilasi *k*=3...................56
4.3.4 Estimasi Model Regresi Semiparametrik Campuran 58
4.3.5 Interpretasi Model Regresi Semiparametrik Campuran 60

BAB V KESIMPULAN DAN SARAN 67
5.1 Kesimpulan 67
5.2 Saran 69

DAFTAR PUSTAKA 71
LAMPIRAN..73
BIOGRAFI PENULIS 103

DAFTAR TABEL

		Halaman
Tabel 3.1	Variabel Penelitian	26
Tabel 3.2	Struktur Data Penelitian	26
Tabel 4.1	Deskriptif Data	42
Tabel 4.2	GCV dari model dengan 3 komponen spline *truncated* dan 1 komponen deret Fourier	46
Tabel 4.3	GCV dari model dengan 1 komponen spline *truncated* dan 3 komponen deret Fourier	47
Tabel 4.4	GCV dari model dengan dua komponen deret Fourier dan dua komponen spline *truncated*	47
Tabel 4.5	Komponen Parametrik dan Nonparametrik	48
Tabel 4.6	Nilai GCV pada Model dengan Satu Titik Knot dan osilasi $k=1$	49
Tabel 4.7	Nilai GCV pada Model dengan Satu Titik Knot dan osilasi $k=2$	50
Tabel 4.8	Nilai GCV pada Model dengan Satu Titik Knot dan osilasi $k=3$	51
Tabel 4.9	Nilai GCV pada Model dengan Dua Titik Knot dan osilasi $k=1$	52
Tabel 4.10	Nilai GCV pada Model dengan Dua Titik Knot dan osilasi $k=2$	52
Tabel 4.11	Nilai GCV pada Model dengan Dua Titik Knot dan osilasi $k=3$	53
Tabel 4.12	Nilai GCV pada Model dengan Tiga Titik Knot dan osilasi $k=1$	55
Tabel 4.13	Nilai GCV pada Model dengan Tiga Titik Knot dan osilasi $k=2$	55
Tabel 4.14	Nilai GCV pada Model dengan Tiga Titik Knot dan osilasi $k=3$	56
Tabel 4.15	Perbandingan Nilai GCV Minumum	57
Tabel 4.16	Estimasi Parameter dari Model Semiparametrik Campuran Spline *Truncated* dan Deret Fourier	58
Tabel 4.17	Perbandingan antara y dan $\hat{y}$	59

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

		Halaman
Gambar 3.1	Diagram Alir Langkah Analisis Pertama	24
Gambar 3.2	Diagram Alir Langkah Analisis Kedua	25
Gambar 3.3	Peta Administrasi Jawa Timur	25
Gambar 4.1	*Scatter Plot* antara Angka Harapan Hidup dengan Angka Kematian Bayi	43
Gambar 4.2	*Scatter plot* antara Angka Harapan Hidup dengan Persentase Bayi Berumur 0-11 Bulan yang diberi ASI 4-6 Bulan	43
Gambar 4.3	*Scatter plot* antara Angka Harapan Hidup dengan Tingat Partisipasi Angkatan Kerja	44
Gambar 4.4	*Scatter plot* antara Angka Harapan Hidup dengan Angka Melek Huruf	45
Gambar 4.5	*Scatter plot* antara Angka Harapan Hidup dengan Rata-Rata Lama Sekolah	45
Gambar 4.6	Plot data y dan $\hat{y}$	59

Halaman ini sengaja dikosongkan

DAFTAR LAMPIRAN

		Halaman
Lampiran 1	Data Angka Harapan Hidup di Jawa Timur Tahun 2013 dan Faktor-Faktor yang Diduga Berpengaruh	73
Lampiran 2	Syntax R Model Semiparametrik Campuran	74

Halaman ini sengaja dikosongkan

BAB I
PENDAHULUAN

1.1 Latar Belakang

Tujuan utama dalam analisis regresi adalah untuk mengestimasi kurva regresi. Terdapat tiga bentuk pendekatan untuk mengestimasi kurva regresi yaitu pendekatan regresi parametrik, pendekatan regresi nonparametrik, dan pendekatan regresi semiparametrik. Pendekatan regresi parametrik digunakan jika bentuk kurva regresinya diasumsikan diketahui seperti linier, kuadratik, kubik, eksponensial dan lain sebagainya (Gujarati, 2004). Namun dalam kenyataannya tidak semua data mengikuti pola-pola tertentu. Jika hubungan variabel respon dan variabel prediktor tidak diketahui bentuk polanya, maka pendekatan regresi nonparametrik sesuai untuk memodelkan hubungan variabel tersebut. Sedangkan, pendekatan regresi semiparametrik merupakan gabungan dari komponen parametrik dan komponen nonparametrik. Regresi semiparametrik muncul karena adanya kasus-kasus pemodelan dimana hubungan antar variabelnya sebagian mengikuti pola tertentu dan sebagian lainnya tidak diketahui bentuk polanya.

Model-model regresi nonparametrik maupun semiparametrik yang dikembangkan oleh para peneliti selama ini pada umumnya menggunakan jenis metode estimasi yang sama untuk sebagian bahkan pada semua variabel prediktornya. Hal ini disebabkan oleh anggapan bahwa pola data dari masing-masing prediktor dianggap sama, sehingga peneliti hanya menggunakan satu bentuk estimator model untuk semua variabel prediktor. Sementara itu, pada kenyataannya sering dijumpai kasus-kasus dengan pola data yang berbeda dari masing-masing variabel prediktor. Oleh karena itu, untuk mengatasi masalah tersebut beberapa peneliti telah mengembangkan estimator kurva regresi campuran dimana masing-masing pola data dalam model regresi dihampiri dengan estimator kurva yang sesuai dengan pola data.

Penelitian menggunakan estimator campuran pernah dilakukan oleh (Sudiarsa dkk, 2015) yaitu melibatkan estimator Deret Fourier dan Spline *Truncated* dalam regresi nonparametrik. Selain itu, (Rory dkk, 2016), (Purnomo dkk, 2016), (Rismal dkk 2016), dan (Syisliawati dkk, 2016) meneliti mengenai

estimator campuran Spline *Truncated* dan Kernel dalam regresi nonparametrik serta (Hesikumalasari dkk, 2016) yaitu estimator campuran Spline *Truncated* dan Kernel pada regresi semiparametrik. Oleh karena itu, sebagai pengembangan dari penelitian-penelitian yang telah ada sebelumnya, maka pada penelitian ini akan digunakan model estimator campuran Spline *Truncated* dan Deret Fourier pada regresi semiparametrik multivariabel menggunakan optimasi *Penalized Least Square* (PLS). Model estimator campuran Spline *Truncated* dan Deret Fourier dalam regresi semiparametrik multivariabel sangat penting peranannya dalam menyelesaikan persoalan pemodelan regresi yang memiliki hubungan antara variabel respon dan variabel-variabel prediktornya sebagian mengikuti pola tertentu, sebagian lagi memiliki pola yang berubah-ubah pada sub-sub interval tertentu, dan sebagian lagi yang lainnya mempunyai pola yang berulang. Model regresi parametrik, regresi nonparametrik dan regresi semiparametrik yang telah dikembangkan oleh penelitian-penelitian terdahulu, belum mampu menangani kasus-kasus pemodelan regresi seperti yang disebutkan di atas.

Angka Harapan Hidup merupakan salah satu indikator untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk. Angka Harapan Hidup yang tinggi di suatu daerah mengindikasikan bahwa masyarakat di daerah tersebut telah terjamin kesehatannya dan kemiskinannya sudah diatasi dengan baik, begitu pula sebaliknya. Berdasarkan data Survei Sosial Ekonomi Nasional (SUSENAS), menunjukkan Angka Harapan Hidup di Provinsi Jawa Timur dari tahun 2009 hingga 2013 mengalami peningkatan yakni dari 69,15 tahun, 69,60 tahun, 69,81 tahun, 70,09 tahun dan 70,19 tahun pada tahun 2013. Hal ini secara tidak langsung memberikan gambaran tentang adanya perbaikan kualitas hidup dan derajat kesehatan masyarakat. Secara keseluruhan Angka Harapan Hidup di Provinsi Jawa Timur mengalami peningkatan, namun masih ada beberapa daerah yang memiliki Angka Harapan Hidup dibawah 65 tahun, diantaranya adalah Kabupaten Probolinggo dengan Angka Harapan Hidup terendah yaitu hanya mencapai 61,87 tahun, kemudian diikuti Kabupaten Jember, Situbondo, Bangkalan, Bondowoso, Sampang dan Pasuruan. Perbedaan Angka Harapan Hidup pada Kabupaten/Kota di Jawa Timur tentunya tidak terlepas oleh faktor ekonomi, sosial dan budaya dari masing-masing daerah yang mempunyai

karakteristik tersendiri. Beberapa faktor yang diduga mempengaruhi Angka Harapan Hidup di Jawa Timur yaitu Angka Kematian Bayi, persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan, Tingkat Partisipasi Angkatan Kerja, Angka Melek Huruf, dan rata-rata lama sekolah.

Penelitian sebelumnya mengenai Angka Harapan Hidup di Provinsi Jawa Timur pernah dilakukan oleh (Firdial, 2010) menggunakan metode *Geographically Weighted Regression* (GWR). Metode GWR yang digunakan mempertimbangkan faktor geografis yakni lokasi daerah pengamatan dalam memodelkan angka harapan hidup. Pada penelitian yang dilakukan oleh (Firdial, 2010) faktor geografis tidak memberikan pengaruh yang signifikan sehingga pemodelan Angka Harapan Hidup dengan GWR tidak berbeda dengan regresi linier. Bila ditelusuri lebih lanjut, pola hubungan yang terbentuk antara angka harapan hidup dan faktor yang diduga mempengaruhi seperti faktor sosial, ekonomi, dan kesehatan tidak membentuk suatu pola yang diketahui fungsi kurva regresinya sehingga pemodelan dengan GWR maupun regresi linier kurang tepat untuk diterapkan.

Penelitian mengenai faktor-faktor yang mempengaruhi Angka Harapan Hidup di Provinsi Jawa Timur juga pernah dilakukan oleh (Sugiarti, 2013) menggunakan regresi semiparametrik Spline. Penelitian tersebut menggunakan enam faktor yang diduga mempengaruhi Angka Harapan Hidup di Provinsi Jawa Timur pada tahun 2010. Pola hubungan yang terbentuk antara variabel respon Angka Harapan Hidup dengan variabel prediktor Angka Kematian Bayi dan variabel prediktor Angka Buta Huruf membentuk pola hubungan linier negatif. Sedangkan bentuk pola hubungan antara variabel respon Angka Harapan Hidup dengan empat variabel prediktor lainnya yaitu persentase bayi berusia 0-11 bulan yang diberi ASI antara 4-6 bulan, laju pertumbuhan ekonomi, persentase balita usia 1-4 tahun yang mendapat imunisasi lengkap, dan Tingkat Partisipasi Angkatan Kerja tidak diketahui bentuk pola perilaku data atau fungsi kurva regresinya sehingga kondisi tersebut mengindikasikan adanya komponen nonparametrik. Dalam penelitian tersebut komponen nonparametriknya diestimasi menggunakan estimator Spline. Berdasarkan hasil penelitian tersebut adapun variabel yang memberikan pengaruh signifikan terhadap Angka Harapan Hidup di

Provinsi Jawa Timur adalah Angka Kematian Bayi, persentase bayi berusia 0-11 bulan yang diberi ASI selama 4-6 bulan dan variabel persentase balita berusia 1-4 tahun yang mendapatkan imunisasi lengkap.

Berdasarkan penelitian sebelumnya dan eksplorasi data yang dilakukan, bentuk pola hubungan antara variabel respon Angka Harapan Hidup dengan variabel prediktor Angka Kematian Bayi membentuk pola hubungan linier negatif. Variabel prediktor Angka Melek Huruf dan rata-rata lama sekolah mempunyai perilaku pola data yang diduga berubah-ubah pada sub-sub interval tertentu. Sedangkan variabel prediktor persentase bayi berusia 0-11 bulan yang diberi ASI selama 4-6 bulan dan Tingkat Partisipasi Angkatan Kerja diduga mempunyai bentuk pola perilaku data yang berulang. Sehingga, kasus Angka Harapan Hidup di Provinsi Jawa Timur ini dapat dimodelkan menggunakan analisis regresi semiparametrik dengan estimator campuran Spline *Truncated* dan Deret Fourier.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka terlihat bahwa model estimator campuran Spline *Truncated* dan Deret Fourier dalam regresi semiparametrik multivariabel sangat penting peranannya dalam menyelesaikan persoalan pemodelan regresi apabila hubungan antara variabel respon dengan variabel-variabel prediktornya sebagian mengikuti pola tertentu, sebagian lagi berubah-ubah pada sub-sub interval tertentu, dan sebagian lagi yang lainnya mempunyai pola berulang. Berdasarkan indentifikasi faktor-faktor yang mempengaruhi Angka Harapan Hidup di Provinsi Jawa Timur, maka model estimator campuran Spline *Truncated* dan Deret Fourier dalam regresi semiparametrik multivariabel ini dapat digunakan untuk memodelkan Angka Harapan Hidup di Provinsi Jawa Timur, karena diduga bentuk pola hubungan antara variabel respon dan variabel-variabel prediktor yang mempengaruhi Angka Harapan Hidup di Provinsi Jawa Timur memiliki perilaku pola data seperti yang disebutkan di atas.

1. 3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas maka tujuan yang ingin dicapai adalah

1. Memperoleh estimasi model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier
2. Memodelkan Angka Harapan Hidup (AHH) di Provinsi Jawa Timur pada tahun 2013 menggunakan regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier.

1. 4 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat, antara lain :

1. Memberikan wawasan baru mengenai pemodelan, khususnya kurva regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier.
2. Memberikan model alternatif untuk menganalisis kasus Angka Harapan Hidup (AHH) yaitu dengan menggunakan model semiparametrik campuran Spline *Truncated* dan Deret Fourier.
3. Memperoleh model Angka Harapan Hidup (AHH) di Provinsi Jawa Timur tahun 2013 menggunakan model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier yang dapat digunakan sebagai prediksi.
4. Penelitian ini dapat menjadi masukan kepada pihak pemerintah Provinsi Jawa Timur mengenai permasalahan kependudukan khususnya mengenai peningkatan kesejahteraan masyarakat.

1. 5 Batasan Masalah

Mengacu pada permasalahan diatas, ruang lingkup dalam penelitian ini dibatasi pada beberapa hal, antara lain :

1. Data yang digunakan adalah Angka Harapan Hidup di Provinsi Jawa Timur tahun 2013. Secara kesulurahan Angka Harapan Hidup di Provinsi Jawa Timur pada tahun 2013 sebesar 70,19 tahun, namun masih terdapat beberapa Kabupaten/Kota di Provindi Jawa Timur yang bahkan memiliki Angka Harapan Hidup di bawah 65 tahun, yaitu Kabupaten Probolinggo, Kabupaten Jember, Kabupaten Situbondo, Kabupaten Bangkalan, dan Kabupaten

Sampang. Data yang digunakan adalah Angka Harapan Hidup di Provinsi Jawa Timur tahun 2013.

2. Fungsi Spline yang digunakan dalam estimator campuran adalah Spline *Truncated,* karena Spline *Truncated* mempunyai interpretasi statistik dan interpretasi visual sangat khusus dan sangat baik (Eubank, 1988). Disamping itu Spline mampu menangani karakter data/fungsi yang bersifat mulus (*smooth*). Spline *Truncated* juga memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu.
3. Fungsi Deret Fourier yang digunakan dalam estimator campuran adalah Deret Fourier dengan *trend,* dikarekan bentuk perilaku data dalam penelitian ini diduga cenderung berulang dan memiliki *trend.*
4. Optimasi model menggunakan metode *Penalized Least Square* (PLS). PLS merupakan perluasan dari metode kuadrat terkecil dengan menambahkan parameter penghalus dan penalti. Parameter penghalus memiliki peranan penting, yaitu berfungsi untuk mengontrol *goodness of fit* dan penalti.
5. Pemilihan titik knot K, parameter penghalus λ, dan osilasi k optimal menggunakan metode *Generalized Cross Validation* (GCV).
6. Pemilihan titik knot optimal pada Spline *Truncated* dengan satu, dua, dan tiga titik knot.
7. Pemilihan osilasi k optimal pada Deret Fourier ditentukan yaitu $k = 1,2,3$.

BAB II
TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode Statistika yang digunakan untuk menentukan hubungan antara suatu variabel dengan variabel yang lain. Tujuan utama dalam analisis regresi adalah bagaimana mencari bentuk estimasi untuk kurva regresi. Disamping itu, analisis regresi juga dapat digunakan untuk prediksi. Misalkan terdapat sekumpulan data berpasangan (x_i, y_i) yang secara umum dapat dimodelkan dengan model regresi :

$$y_i = f(x_i) + \varepsilon_i, \; i = 1, 2, ..., n \quad (2.1)$$

dengan y_i respon ke-*i*, $f(x_i)$ kurva regresi, ε_i adalah *error* yang diasumsikan identik, independen, dan berdistribusi normal dengan mean nol dan varian σ^2 (Eubank,1999).

Berkaitan dengan model tersebut, terdapat tiga model pendekatan regresi, yaitu regresi parametrik, regresi nonparametrik dan regresi semiparametrik. Apabila dalam analisis regresi bentuk kurva regresi diketahui maka digunakan pendekatan regresi parametrik (Budiantara, 2000). Sedangkan apabila bentuk kurva regresi tidak diketahui bentuk polanya maka digunakan regresi nonparametrik. Regresi semiparametrik digunakan jika dalam model regresi terdapat komponen parametrik dan komponen nonparametrik.

2.2 Regresi Parametrik, Nonparametrik dan Semiparametrik

Regresi parametrik merupakan salah satu metode Statistika yang banyak digunakan yang dapat menggambarkan pola hubungan antara variabel respon dengan variabel prediktor. (Draper dan Smith, 1996) menyatakan bahwa bentuk regresi parametrik linier adalah :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} + \varepsilon_i, \; i = 1, 2, ..., n \quad (2.2)$$

dengan y_i sebagai variabel respon, $x_{1i}, x_{2i}, ..., x_{pi}$ sebagai variabel prediktor, $\beta_0, \beta_1, \beta_2, ..., \beta_p$ merupakan parameter yang tidak diketahui dan ε_i adalah *error*

random yang independen, berdistribusi normal dengan mean nol dan variansi σ^2. Estimator parameter model diperoleh berdasarkan berbagai metode yang telah dikenal dalam Statistika yaitu, *Least Square, Maximum Likelihood* (Wahba, 1990) dalam (Rismal, 2016).

Regresi nonparametrik merupakan regresi yang pola hubungan antara variabel respon dan variabel prediktor tidak diketahui bentuknya. Misalkan diberikan *n* pengamatan yang independen, yaitu pasangan (y_i, z_i), $i = 1, 2, ..., n$. Pola hubungan antara variabel y_i dan z_i tidak diketahui dan mengikuti model regresi :

$$y_i = g(z_i) + \varepsilon_i, \tag{2.3}$$

dengan y_i merupakan variabel respon, z_i merupakan variabel prediktor, dan ε_i adalah *error* random yang independen dan berdistribusi normal dengan mean nol dan variansi σ^2, sedangkan $g(z_i)$ merupakan fungsi regresi yang tidak diketahui bentuk polanya.

Regresi semiparametrik merupakan gabungan antara komponen parametrik dan komponen nonparametrik. Misalkan diberikan data berpasangan (x_i, z_i, y_i) dan hubungan antara x_i, z_i dan y_i diasumsikan mengikuti model regresi semiparametrik :

$$y_i = f(x_i) + g(z_i) + \varepsilon_i, i = 1, 2, ..., n, \tag{2.4}$$

dengan y_i merupakan variabel respon, x_i dan z_i merupakan variabel-variabel prediktor, dan ε_i adalah *error* random yang independen dan berdistribusi normal dengan mean nol dan variansi σ^2. $f(x_i)$ merupakan fungsi regresi yang diketahui bentuk polanya sedangkan $g(z_i)$ merupakan fungsi regresi yang tidak diketahui bentuk polanya.

2.3 Regresi Semiparametrik Spline *Truncated*

Spline merupakan salah satu teknik estimasi regresi nonparametrik yang pertama kali dikembangkan oleh Whittaker pada tahun 1923 (Hardle, 1990). Regresi nonparametrik merupakan regresi yang sangat fleksibel dalam

memodelkan pola data (Eubank, 1988). Spline juga memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu disebabkan Spline merupakan model polinomial yang tersegmen/ terpotong (Budiantara, *et al*, 2009).

Diberikan data berpasangan (x_i, t_i, y_i), $i = 1, 2, ..., n$, dimana pola hubungannya dapat dinyatakan dalam model regresi :

$$y_i = f(x_i) + g(t_i) + \varepsilon_i \tag{2.5}$$

dengan y_i merupakan variabel respon, x_i dan t_i merupakan variabel-variabel prediktor, dan ε_i adalah *error* $\varepsilon_i \sim N\left(0, \sigma^2\right)$. Kurva regresi komponen parametrik pada persamaan (2.5) $f(x_i)$ dihampiri dengan fungsi linier :

$$f(x_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} \tag{2.6}$$

Fungsi pada persamaan (2.6) dapat dituliskan dalam bentuk matriks sebagai berikut :

$$\begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} \\ \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} \\ \vdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_p x_{np} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ & & & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

atau

$$\tilde{y} = \mathbf{X}\tilde{\beta} + \tilde{\varepsilon} \tag{2.7}$$

dengan $\tilde{y}$ adalah vektor dari variabel respon berukuran $n \times 1$, $\mathbf{X}$ merupakan matriks berukuran $n \times \left(p+1\right)$ dan $\tilde{\beta}$ adalah vektor parameter yang akan diestimasi berukuran $\left(p+1\right) \times 1$.

Selanjutnya, Kurva regresi $g(t_i)$ pada persamaan (2.5) dihampiri dengan fungsi Spline *Truncated* linier dengan knot $K_1, K_2, \cdots K_m$. Secara umum, fungsi Spline *Truncated* linier dapat disajikan dalam bentuk :

$$G_i(t_i) = \sum_{j=1}^{q} \phi_{ji} t_i + \sum_{k=1}^{m} \alpha_{ki} \left(t_i - K_k\right)_+ \tag{2.8}$$

dengan,

$$\left(t_i - K_k\right)_+ = \begin{cases} \left(t_i - K_{ki}\right) & , t_i \geq K_{ki} \\ 0 & , \; t_i < K_{ki} \end{cases}$$

Dimana $\phi_1, \; \alpha_1, \ldots, \alpha_m$ merupakan parameter-parameter yang tidak diketahui, $K_1, K_2, ..., K_m$ merupakan titik knot dimana $K_1 < K_2 < \ldots < K_m$. Dari persamaan fungsi Spline *Truncated* di atas, untuk $i = 1$ sampai n, tersebut dapat digabungkan ke dalam bentuk persamaan vektor dan matriks :

$$\begin{bmatrix} g_1(t_{1i}, t_{2i}, ..., t_{qi}) \\ g_2(t_{1i}, t_{2i}, ..., t_{qi}) \\ \vdots \\ g_q(t_{1i}, t_{2i}, ..., t_{qi}) \end{bmatrix} = \begin{bmatrix} \phi_{11}t_{11} + \cdots + \phi_{1q}t_{q1} + \alpha_{11}(t_{11} - K_{11})_+ + \cdots + \alpha_{m1}(t_{11} - K_{m1})_+ + \cdots + \alpha_{1q}(t_{q1} - K_{1q})_+ + \cdots + \alpha_{mq}(t_{q1} - K_{mq})_+ \\ \phi_{11}t_{12} + \cdots + \phi_{1q}t_{q2} + \alpha_{11}(t_{12} - K_{11})_+ + \cdots + \alpha_{m1}(t_{12} - K_{m1})_+ + \cdots + \alpha_{1q}(t_{q2} - K_{1q})_+ + \cdots + \alpha_{mq}(t_{q2} - K_{mq})_+ \\ \vdots \\ \phi_{11}t_{1n} + \cdots + \phi_{1q}t_{qn} + \alpha_{11}(t_{1n} - K_{11})_+ + \cdots + \alpha_{m1}(t_{1n} - K_{m1})_+ + \cdots + \alpha_{1q}(t_{qn} - K_{1q})_+ + \cdots + \alpha_{mq}(t_{qn} - K_{mq})_+ \end{bmatrix}$$

Dalam notasi matriks dapat dituliskan menjadi :

$$\tilde{g}(t) = \mathbf{G}\tilde{\phi} \tag{2.9}$$

dengan,

$$\mathbf{G} = \begin{bmatrix} t_{11} & \cdots & t_{q1} & (t_{11} - K_{11})_+ & \cdots & (t_{11} - K_{m1})_+ & \cdots & (t_{q1} - K_{1q})_+ & \cdots & (t_{q1} - K_{mq})_+ \\ t_{12} & \cdots & t_{q2} & (t_{12} - K_{11})_+ & \cdots & (t_{12} - K_{m1})_+ & \cdots & (t_{q2} - K_{1q})_+ & \cdots & (t_{q2} - K_{mq})_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{1n} & \cdots & t_{qn} & (t_{1n} - K_{11})_+ & \cdots & (t_{1n} - K_{m1})_+ & \cdots & (t_{qn} - K_{1q})_+ & \cdots & (t_{qn} - K_{mq})_+ \end{bmatrix}$$

$$\tilde{\phi} = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1q} & \alpha_{11} & \cdots & \alpha_{m1} & \cdots & \alpha_{1q} & \cdots & \alpha_{mq} \end{bmatrix}^T$$

Vektor $\tilde{g}(t)$ berukuran $n \times 1$, matriks $\boldsymbol{G}$ berukuran $n \times ((m+1)q)$, sedangkan vektor $\tilde{\phi}$ berukuran $1 \times ((m+1)q)$.

Model regresi semiparametrik pada persamaan (2.5) dapat disajikan dalam bentuk matriks :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} t_{11} & \cdots & t_{q1} & (t_{11}-K_{11})_+ & \cdots & (t_{11}-K_{m1})_+ & \cdots & (t_{q1}-K_{1q})_+ & \cdots & (t_{q1}-K_{mq})_+ \\ t_{12} & \cdots & t_{q2} & (t_{12}-K_{11})_+ & \cdots & (t_{12}-K_{m1})_+ & \cdots & (t_{q2}-K_{1q})_+ & \cdots & (t_{q2}-K_{mq})_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{1n} & \cdots & t_{qn} & (t_{1n}-K_{11})_+ & \cdots & (t_{1n}-K_{m1})_+ & \cdots & (t_{qn}-K_{1q})_+ & \cdots & (t_{qn}-K_{mq})_+ \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \vdots \\ \phi_{1q} \\ \alpha_{11} \\ \vdots \\ \alpha_{m1} \\ \vdots \\ \alpha_{1q} \\ \vdots \\ \alpha_{mq} \end{bmatrix} +$$

$$+\begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix}^T,$$

atau dalam notasi matriks dapat dituliskan :

$$\tilde{y} = \mathbf{X}\tilde{\beta} + \mathbf{G}\tilde{\phi} + \tilde{\varepsilon} \tag{2.10}$$

dengan,

$$\tilde{y} = [y_1, y_2, \cdots, y_n]^T,$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix},$$

$$\tilde{\beta} = [\beta_0, \beta_1, \beta_2, \cdots, \beta_p]^T,$$

$$\mathbf{G} = \begin{bmatrix} t_{11} & \cdots & t_{q1} & (t_{11}-K_{11})_+ & \cdots & (t_{11}-K_{m1})_+ & \cdots & (t_{q1}-K_{1q})_+ & \cdots & (t_{q1}-K_{mq})_+ \\ t_{12} & \cdots & t_{q2} & (t_{12}-K_{11})_+ & \cdots & (t_{12}-K_{m1})_+ & \cdots & (t_{q2}-K_{1q})_+ & \cdots & (t_{q2}-K_{mq})_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{1n} & \cdots & t_{qn} & (t_{1n}-K_{11})_+ & \cdots & (t_{1n}-K_{m1})_+ & \cdots & (t_{qn}-K_{1q})_+ & \cdots & (t_{qn}-K_{mq})_+ \end{bmatrix}$$

$$\tilde{\phi} = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1q} & \alpha_{11} & \cdots & \alpha_{m1} & \cdots & \alpha_{1q} & \cdots & \alpha_{mq} \end{bmatrix}^T$$

$$\tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix}^T$$

Spline merupakan salah satu model yang mempunyai interpretasi statistik dan visual yang khusus dan baik (Eubank, 1988). Dalam fungsi Spline terdapat titik knot yang merupakan titik perpaduan yang menunjukkan perubahan perilaku kurva pada selang yang berbeda (Hardle, 1990). Titik knot merupakan bagian

yang sangat penting dalam regresi Spline. Oleh karena itu agar diperoleh Spline yang optimal perlu dipilih titik knot yang terbaik, berapa jumlahnya dan dimana letak titik-titik knot tersebut. Terdapat 2 strategi untuk menyelesaikan permasalahan ini, strategi pertama adalah memilih banyaknya knot yang relatif sedikit, sedangkan strategi kedua adalah kebalikannya yakni menggunakan knot yang relatif banyak (Wand, 2000). Dari kedua hal tersebut yang paling baik adalah lebih mengarah pada alasan kesederhanaan model (*parsimony*). Salah satu metode untuk memilih titik knot optimal adalah dengan metode *Generalized Cross Validation*.

2.4 Regresi Semiparametrik Deret Fourier

Deret Fourier umumnya digunakan apabila data yang diselidiki polanya ada kecenderungan pola berulang. Deret Fourier merupakan polinomial trigonometri yang mempunyai fleksibilitas, sehingga dapat menyesuaikan diri secara efektif terhadap sifat lokal data.

Misalkan diberikan data berpasangan $(x_1, x_2, ..., x_{p,} z_1, z_2, ..., z_q, y),$ dengan x dan z merupakan variabel prediktor dan y adalah variabel respon. Hubungan antara x dan y diketahui bentuk polanya, sementara hubungan antara variabel z dan y tidak diketahui bentuk polanya. Oleh karena itu hubungan antara x_i, z_i dan y_i diasumsikan mengikuti model regresi semiparametrik. Dalam penelitian ini model regresi semiparametrik diasumsikan terdapat p variabel prediktor $x_1, x_2, ..., x_{p,}$ merupakan komponen parametrik linier dan n variabel komponen nonparametrik $z_1, z_2, ..., z_n$. Hubungan antara variabel respon dan variabel prediktor mengikuti model regresi semiparametrik :

$$y_i = \tilde{x}_i \tilde{\beta} + h(z_1, z_2, ..., z_n) + \varepsilon_i, i = 1, 2, ..., n. \tag{2.11}$$

Bentuk kurva regresi $h(z)$ tidak diketahui dan termuat dalam ruang fungsi kontinu $C(0,\pi)$. *Error* random ε_i diasumsikan berdistribusi normal independen dengan *mean* nol dan variansi σ^2. Karena $h(z)$ kontinu pada interval $(0,\pi)$ maka dapat dihampiri oleh fungsi Deret Fourier $h(z),$ dengan :

$$h(z)=\sum_{i=1}^{n} bz_i+\frac{1}{2}a_0+\sum_{k=1}^{K} a_k\cos kz_i \qquad (2.12)$$

Dimana $b,\ a_0,\ a_k,\ k=1,2,...,K$ merupakan parameter-parameter model. Dari persamaan fungsi Deret Fourier tersebut, dapat digabungkan ke dalam bentuk persamaan vektor dan matriks :

$$\begin{bmatrix} h_1(z_{1i},z_{2i},\ldots,z_{ri}) \\ h_2(z_{1i},z_{2i},\ldots,z_{ri}) \\ \vdots \\ h_r(z_{1i},z_{2i},\ldots,z_{ri}) \end{bmatrix} = \begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \cdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn} \end{bmatrix} \begin{bmatrix} b_1 \\ a_0 \\ a_{11} \\ \vdots \\ a_{K1} \\ \vdots \\ b_r \\ a_0 \\ a_{r1} \\ \vdots \\ a_{Kr} \end{bmatrix}$$

Dalam notasi matriks dapat dituliskan menjadi :

$$\tilde{h}(z)=\mathbf{D}\tilde{a} \qquad (2.13)$$

dengan,

$$\mathbf{D}=\begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \cdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn} \end{bmatrix},$$

dan $\begin{bmatrix} b_1 & a_0 & a_{11} & \cdots & a_{K1} & \cdots & b_r & a_0 & a_{1r} & \cdots & a_{Kr} \end{bmatrix}^T$.

Model regresi semiparametrik pada persamaan (2.11) dapat disajikan dalam bentuk matriks:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \cdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn} \end{bmatrix}$$

$$\begin{bmatrix} b_1 & a_{01} & a_{11} & \cdots & a_{K1} & \cdots & b_{q1} & a_{0q} & a_{1q} & \cdots & a_{Kq} \end{bmatrix}^T + \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix}^T$$

atau :

$$\tilde{y} = \mathbf{X}\tilde{\beta} + \mathbf{D}\tilde{a} + \tilde{\varepsilon} \tag{2.14}$$

dengan

$\tilde{y} = [y_1, y_2, \cdots, y_n]^T$,

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{q1} & \frac{1}{2} & \cos z_{q1} & \cdots & \cos Kz_{q1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{q2} & \frac{1}{2} & \cos z_{q2} & \cdots & \cos Kz_{q2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{1n} & \cdots & z_{qn} & \frac{1}{2} & \cos z_{qn} & \cdots & \cos Kz_{qn} \end{bmatrix}$$

$\tilde{\beta} = [\beta_0, \beta_1, \beta_2, \cdots, \beta_p]^T$,

$\tilde{a} = \begin{bmatrix} b_1 & a_{01} & a_{11} & \cdots & a_{K1} & \cdots & b_{q1} & a_{0q} & a_{1q} & \cdots & a_{Kq} \end{bmatrix}^T$ dan

$\tilde{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]^T$.

Menurut (Tripena, 2007), dalam analisis regresi untuk mengestimasi kurva g dapat digunakan metode *least square* (LS), yaitu meminimumkan jumlah kuadrat *error*. Dengan kata lain, penduga untuk g dapat diperoleh dari optimasi :

$$\min_{g \in C(0,\pi)} \left\{ \sum_{i=1}^{n} \varepsilon_i^2 \right\} = \min_{g \in C(0,\pi)} \left\{ \sum_{i=1}^{n} \left(y_i - f(x_i) - h(z_i) \right)^2 \right\}, \tag{2.15}$$

2.5 Koefisien Determinasi (R^2)

Salah satu tujuan analisis regresi adalah mendapatkan model terbaik yang mampu menjelaskan hubungan antara variabel prediktor dan variabel respon berdasarkan kriteria tertentu. Salah satu kriteria yang digunakan dalam pemilihan model terbaik adalah dengan menggunakan koefisien determinasi R^2. Koefisien determinasi (R^2) adalah besaran yang menggambarkan besarnya persentase variasi dalam variabel respon yang dijelaskan oleh variabel prediktor.

Menurut (Eubank, 1988), koefisien determinasi didefinisikan sebagai berikut:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{2.16}$$

dengan y_i merupakan variabel respon ke-*i*, $\hat{y}_i$ merupakan penduga variabel respon ke-*i*, dan $\bar{y}$ merupakan rata-rata dari variabel respon.

2.6 *Generalized Cross Validation (GCV)*

Spline terbaik bergantung pada pemilihan titik–titik knot *k*. Titik knot *k* merupakan perpaduan bersama antara perubahan fungsi pada interval yang berlainan. Sementara itu Deret Fourier bergantung pada parameter penghalus λ dan osilasi *k*. Parameter penghalus λ berfungsi untuk mengontrol kemulusan dari kurva yang diestimasi, dan osilasi *k* merupakan banyaknya osilasi dari gelombang cosinus pada model. Titik knot dan parameter penghalus λ yang terlalu kecil akan menghasilkan kurva yang *under smoothing* yaitu sangat kasar dan sangat fluktuatif, dan sebaliknya titik knot maupun parameter penghalus λ yang terlalu besar/lebar akan menghasilkan kurva yang *over smoothing* yaitu sangat mulus, tetapi tidak sesuai dengan pola data. Sedangkan osilasi *k*, semakin besar nilai *k* akan menyebabkan model semakin kompleks dan osilasi dari kurva semakin rapat serta mengikuti pola data aktual, sehingga bias semakin kecil dan varian semakin besar. Oleh karena itu perlu dipilih titik knot, parameter penghalus λ maupun osilasi *k* yang optimal. Nilai K, λ, k optimal diperoleh dari nilai $GCV\ (K, \lambda, k)$ yang terkecil.

Salah satu metode yang sering digunakan dalam memilih titik knot, parameter penghalus λ maupun osilasi k yang optimal adalah *Generalized Cross Validation* (GCV). Menurut (Purnomo, 2016) jika dibandingkan dengan metode lain, misalnya *Cross Validation* (CV) dan metode *Unbiased Risk* (UBR) ataupun *Generalized Maximum Likelihood* (GML), GCV secara teoritis memiliki sifat *optimal asymtotik*. Metode GCV juga memiliki kelebihan tidak memerlukan pengetahuan terhadap variansi populasi σ^2 serta metode GCV invarians terhadap transformasi. Fungsi GCV untuk pemilihan titik knot optimal dapat ditunjukkan dalam persamaan berikut:

$$\text{GCV}(K,\lambda,k)=\frac{\text{MSE}(K,\lambda,k)}{\left(n^{-1}tr(\text{I}-\text{S}(K,\lambda,k))\right)^2}, \tag{2.17}$$

dengan

$$MSE(K,\lambda,k)=n^{-1}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2.$$

2.7 *Penalized Least Square* (PLS)

Penalized Least Square (PLS) atau metode kuadrat terkecil terpenalti merupakan perluasan metode kuadrat terkecil yaitu dengan menambahkan parameter penghalus dan penalti pada fungsi yang akan dipergunakan. Bentuknya diberikan sebagai berikut :

$$\underset{f,g,h}{Min}\left\{n^{-1}\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{p}f_j(x_{ji})-\sum_{s=1}^{q}g_s(t_{si})-\sum_{l=1}^{r}h_l(\text{z}_{li})\right)^2+\sum_{l=1}^{r}\lambda_l\int_0^{\pi}\frac{2}{\pi}\left(h_l''(\text{z}_l)\right)^2dz_l\right\}, \tag{2.18}$$

dalam hal ini λ merupakan parameter penghalus, sedangkan penalti diberikan $J_l(h_l)=\sum_{l=1}^{r}\lambda_l\int_0^{\pi}\frac{2}{\pi}\left(h_l''(\text{z}_l)\right)^2dz_l$.

2.8 Teorema Dasar Terkait dengan Aljabar Matriks

Beberapa konsep dasar yang digunakan dalam proses mendapatkan estimator model regresi semiparametrik Spline *Truncated* dan Deret Fourier multivariabel berikut ini.

Teorema 2.5.1 (Rencher dan Schaalje 2008: 9)

Diberikan matriks **A** dan **B**, maka berlaku sifat – sifat sebagai berikut :

a. Jika matriks **A** simetris, maka $\mathbf{A}^T = \mathbf{A}$.

b. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

c. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

Teorema selanjutnya berkaitan dengan penurunan matriks dan vektor.

Teorema 2.5.2 (Rencher dan Schaalje 2008: 56)

Diberikan vektor $\mathbf{a}$ dan $\mathbf{x}$, dimana $\mathbf{a}^T\mathbf{x} = \mathbf{x}^T\mathbf{a}$, $\mathbf{a}^T = (a_1, a_2, ..., a_p)$ konstanta, maka :

$$\frac{\partial(\mathbf{a}^T\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}^T\mathbf{a})}{\partial\mathbf{x}} = a.$$

Teorema 2.5.3 (Rencher dan Schaalje 2008: 56)

Apabila vektor $\mathbf{x}$ dan A merupakan suatu matriks simetri, maka :

$$\frac{\partial(\mathbf{x}^T\mathbf{A}\mathbf{x})}{\partial\mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$

2.9 Angka Harapan Hidup

Angka Harapan Hidup adalah rata-rata perkiraan atau ekspektasi dari usia bayi yang baru lahir hingga mencapai kematiannya. Angka Harapan Hidup merupakan salah satu komponen pembentuk indeks pembangunan manusia sekaligus sebagai indikator dampak peningkatan derajat kesejahteraan masyarakat, khususnya di bidang kesehatan (BPS, 2013b). Keberhasilan program kesehatan dan program sosial ekonomi pada umumnya dapat dilihat dari peningkatan angka harapan hidup penduduk di suatu wilayah. Semakin tinggi Angka Harapan Hidup di suatu wilayah mengindikasikan pembangunan sosial ekonomi di wilayah tersebut semakin maju.

Sejak tahun 2009 angka harapan hidup penduduk Jawa Timur mengalami peningkatan yaitu dari 69,15 tahun menjadi 70,19 tahun pada tahun 2013. Hal ini secara tidak langsung memberikan gambaran tentang adanya perbaikan kualitas hidup dan derajat kesehatan masyarakat. Angka Harapan Hidup diatas

memberikan gambaran bahwa bayi-bayi yang lahir pada tahun 2013 mempunyai usia harapan hidup lebih panjang yakni 70,19 tahun, dibandingkan dengan bayi-bayi yang lahir di sekitar tahun 2009 yang mempunyai usia harapan hidup hanya hingga 69,15 tahun. Angka harapan hidup dari setiap daerah biasanya berbeda-beda. Perbedaan tersebut disebabkan oleh faktor ekonomi, sosial, dan budaya dari masing–masing daerah yang mempunyai karakteristik tersendiri, sehingga berpengaruh terhadap kebijakan peningkatan kualitas kesehatan dan ekonomi untuk menaikkan angka harapan hidup penduduknya. Beberapa faktor yang diduga mempengaruhi Angka Harapan Hidup di Jawa Timur 2013 yaitu sebagai berikut :

Angka Kematian Bayi menggambarkan tingkat kesehatan masyarakat pada umumnya disamping untuk menilai keberhasilan pelayanan kesehatan serta program pembangunan kesehatan pada wilayah tersebut. Perkembangan selama enam tahun terakhir menunjukkan bahwa tren Angka Kematian Bayi di Jawa Timur cenderung menurun yaitu 31,44 per 1000 kelahiran hidup di tahun 2009 menurun hingga 26,66 per seribu kelahiran hidup pada tahun 2013. Sehingga bisa dikatakan bahwa sejak lima tahun yang lalu, jumlah bayi meninggal pada setiap 1000 kelahiran hidup sekitar 31 sampai 27 bayi. Terdapat sebanyak 6 daerah di Jawa Timur dengan AKB diatas 50, antara lain Kabupaten Probolinggo (61,66), Kabupaten Jember (54,99), Kabupaten Situbondo (53,37), Kabupaten Bangkalan (53,21), Kabupaten Bondowoso (51,75), Kabupaten Sampang (50,74). Hal ini sejalan dengan penelitian sebelumnya yang pernah dilakukan oleh (Riskiyanti, 2010) yang menyatakan bahwa Angka Kematian Bayi berpengaruh terhadap Angka Harapan Hidup.

ASI (Air Susu Ibu) merupakan makanan alamiah yang mudah diserap oleh bayi dengan komposisi nutrisi yang sesuai untuk perkembangan bayi. Nutrisi yang terkandung pada ASI kaya akan antibodi (zat kekebalan tubuh) yang membantu tubuh bayi untuk melawan infeksi dan penyakit lainnya. Pemberian ASI terbaik adalah pemberian ASI eksklusif yaitu hanya memberi ASI tanpa makanan/minuman tambahan, bahkan tanpa air putih sekalipun sampai bayi berusia 6 bulan. Memberikan ASI eksklusif berarti menjamin ketersedian sumber daya manusia yang berkualitas di masa depan. Bayi-bayi yang diberi ASI

diharapkan mampu meningkatkan angka harapan hidup, dengan argumentasi bahwa dengan semakin terpenuhinya bayi-bayi yang diberi ASI, maka kelangsungan hidup anak dapat ditingkatkan. Hal ini sejalan dengan penelitian sebelumnya yang menemukan hubungan bayi-bayi yang diberi ASI berpengaruh terhadap angka harapan hidup (Sugiarti, 2013).

Menurut penelitian sebelumnya (Sigh, 2006) menyatakan bahwa salah satu faktor yang mempengaruhi Angka Harapan Hidup adalah pendidikan. Pendidikan memberikan sumbangan yang besar terhadap perkembangan kehidupan sosial ekonomi masyarakat. Dengan pendidikan, diharapkan mampu meningkatkan Angka Melek Huruf. Bertambahnya kapasitas seseorang akibat dari mengenyam pendidikan, diharapkan mampu mendapatkan pekerjaan dan tingkat pendapatan yang lebih baik. Dengan Tingkat pendapatan yang baik, seseorang dapat memperoleh pelayanan kesehatan yang baik pula, dengan demikian akan meningkatkan Angka Harapan Hidup. Hal ini sejalan dengan penelitian sebelumnya yang menemukan bahwa Angka Melek Huruf berpengaruh terhadap Angka Harapan Hidup (Riskiyanti, 2010).

Rata-rata lama sekolah adalah rata-rata jumlah tahun yang dihabiskan oleh penduduk di seluruh jenjang pendidikan formal yang pernah dijalani. Secara umum rata-rata lama sekolah di Provinsi Jawa Timur pada daerah kota lebih tinggi, jika dibandingkan dengan rata-rata lama sekolah pada daerah-daerah Kabupaten wilayah Madura dan sekitar kawasan tapal kuda (Kabupaten Bondowoso, Kabupaten Probolinggo, Kabupaten Situbondo, Kabupaten Jember, dll) yang berkisar hanya 5–6 tahun saja. Rata-rata lama sekolah dapat menunjukkan kualitas penduduk dalam hal mengenyam pendidikan formal, semakin tinggi angka rata-rata lama sekolah maka semakin tinggi tingkat/jenjang pendidikan yang ditamatkannya, begitu pula sebaliknya. Rata-rata lama sekolah tertinggi di Provinsi Jawa Timur adalah Kota Madiun yaitu sebesar 10,94 tahun sedangkan yang terendah adalah Kabupaten Sampang yaitu sebesar 4,39 tahun.

Tingkat Partisipasi Angkatan Kerja (TPAK) adalah suatu indikator ketenagakerjaan yang memberikan gambaran tentang penduduk yang aktif secara ekonomi dalam kegiatan sehari-hari merujuk pada suatu waktu dalam periode survei. Pada tahun 2013, TPAK di Jawa Timur tercatat 69,78 persen atau terdapat

sebanyak 67 sampai 68 orang angkatan kerja untuk setiap 100 penduduk usia kerja. Angka ini meningkat sebesar 0,21 persen dari TPAK tahun 2012. TPAK tertinggi di Jawa Timur tahun 2013 adalah Kabupaten Pacitan yaitu sebesar 79,44 persen sedangkan TPAK terendah adalah Kota Probolinggo sebesar 63,7 persen. Semakin tinggi TPAK menunjukkan bahwa semakin tinggi pula pasokan tenaga kerja *(labour supply)* yang tersedia untuk memproduksi barang dan jasa dalam suatu perekonomian. Penelitian sebelumnya (Sugiarti, 2013) menyatakan bahwa terdapat pengaruh TPAK terhadap Angka Harapan Hidup.

BAB III
METODOLOGI PENELITIAN

3.1 Metode Penelitian

Untuk menyelesaikan tujuan penelitian, dibuat langkah-langkah yang disusun mengikuti tahapan-tahapan berikut.

1. Estimasi kurva regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier
   a. Mendefinisikan model regresi semiparametrik multivariabel prediktor mengikuti model aditif. Diberikan respon y_i dengan variabel prediktor komponen parametrik x_i. Sedangkan variabel prediktor t_i dan z_i merupakan komponen nonparametrik :

   $$y_i = \mu(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri}) + \varepsilon_i,\ i = 1, 2, ..., n$$

   $$= \sum_{j=1}^{p} f_j(x_{ji}) + \sum_{s=1}^{p} g_s(t_{si}) + \sum_{l=1}^{r} h_l(z_{li}) + \varepsilon_i,\ i = 1, 2, ..., n$$

   b. Menyajikan kurva-kurva regresi dalam bentuk matriks :

   $$\tilde{f}(x) = \mathbf{X}\tilde{\beta},\ \tilde{g}(t) = \mathbf{G}\tilde{\phi},\ \text{dan}\ \tilde{h}(z) = \mathbf{D}\tilde{a}$$

   c. Membuat pendekatan untuk kurva-kurva tersebut dengan fungsi-fungsi sebagai berikut :
      1) Menghampiri kurva regresi $f_j(x_{ji})$ menggunakan fungsi linier pada persamaan (2.6)
      2) Menghampiri fungsi $g_s(t_{si})$ menggunakan Spline *Truncated* linier pada persamaan (2.8)
      3) Menghampiri fungsi $h_l(z_{li})$ menggunakan Deret Fourier dengan *trend* pada persamaan (2.12)

   d. Model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier dapat dituliskan:

   $$\tilde{y} = \mathbf{X}\tilde{\beta} + \mathbf{G}\tilde{\phi} + \mathbf{D}\tilde{a} + \tilde{\varepsilon}$$

dengan

$$\mathbf{X}=\begin{bmatrix}1 & x_{11} & x_{12} & \cdots & x_{1p}\\ 1 & x_{21} & x_{22} & \cdots & x_{2p}\\ 1 & x_{31} & x_{23} & \cdots & x_{3p}\\ \vdots & & & \ddots & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np}\end{bmatrix},\ \tilde{\beta}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \vdots\\ \beta_p\end{bmatrix}$$

$$\mathbf{G}=\begin{bmatrix}t_{11} & \cdots & t_{q1} & (t_{11}-K_{11})_+ & \cdots & (t_{11}-K_{m1})_+ & \cdots & (t_{q1}-K_{1q})_+ & \cdots & (t_{q1}-K_{mq})_+\\ t_{12} & \cdots & t_{q2} & (t_{12}-K_{11})_+ & \cdots & (t_{12}-K_{m1})_+ & \cdots & (t_{q2}-K_{1q})_+ & \cdots & (t_{q2}-K_{mq})_+\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ t_{1n} & \cdots & t_{qn} & (t_{1n}-K_{11})_+ & \cdots & (t_{1n}-K_{m1})_+ & \cdots & (t_{qn}-K_{1q})_+ & \cdots & (t_{qn}-K_{mq})_+\end{bmatrix},$$

$$\tilde{\phi}=\begin{bmatrix}\phi_{11} & \cdots & \phi_{1q} & \alpha_{11} & \cdots & \alpha_{m1} & \cdots & \alpha_{1q} & \cdots & \alpha_{mq}\end{bmatrix}^T \text{ dan}$$

$$\mathbf{D}=\begin{bmatrix}z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \vdots & \cdots & \vdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1}\\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \vdots & \cdots & \vdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2}\\ \vdots & & \vdots & \ddots & \vdots & & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots\\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \vdots & \cdots & \vdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn}\end{bmatrix},$$

$$\tilde{a}=\begin{bmatrix}b_1 & a_0 & a_{11} & \cdots & a_{K1} & \vdots & \cdots & \vdots & b_r & a_0 & a_{1r} & \cdots & a_{Kr}\end{bmatrix}^T$$

e. Membangun komponen *goodness of fit*

$$\varepsilon^T\varepsilon=(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^T(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})$$

f. Membangun komponen *penalty*

$$J_l(h_l)=\sum_{l=1}^{r}\lambda_l\int_0^{\pi}\frac{2}{\pi}\left(h''(\mathrm{z}_i)\right)^2 dz_i=\tilde{a}^T\mathbf{U}\tilde{a}$$

g. Menyelesaikan optimasi PLS dengan menggabungkan komponen *goodness of fit* dan komponen *penalty* $J_l(h_l)$ pada persamaan (2.19), dapat diuraikan sebagai berikut :

$$\underset{f,g,h}{Min}\left\{n^{-1}\sum_{i=1}^{n}\left(y_i-f(x_i)-g(t_i)-h(\mathrm{z}_i)\right)^2+\sum_{l=1}^{r}\lambda_l\int_0^{\pi}\frac{2}{\pi}\left(h''(\mathrm{z}_i)\right)^2 dz_i\right\}$$

$$=\underset{\tilde{\beta},\tilde{\phi},\tilde{a}}{Min}\left\{n^{-1}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^T(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})+\tilde{a}^T\mathbf{U}\tilde{a}\right\}$$

$$=\underset{f,g,h}{Min}\left\{Q\left(\tilde{\beta},\tilde{\phi},\tilde{a}\right)\right\}$$

h. Mengestimasi parameter $\tilde{\beta}$, $\tilde{\phi}$ dan $\tilde{a}$ menggunakan derivatif parsial, kemudian menyamakan derivatif parsial tersebut dengan nol

$$\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial \tilde{\beta}} = 0, \quad \frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial \tilde{\phi}} = 0, \quad \frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial \tilde{a}} = 0$$

i. Berdasarkan (f) diharapkan diperoleh estimator komponen parametrik linier $\hat{\tilde{f}}(x)$, komponen Spline *Truncated* $\hat{\tilde{g}}(t)$, komponen Deret Fourier $\hat{\tilde{h}}(z)$ dan estimator semiparametrik campuran Spline *Truncated* dan Deret Fourier $\hat{\mu}(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri}) = \hat{\tilde{f}}(x) + \hat{\tilde{g}}(t) + \hat{\tilde{h}}(z)$.

2. Memodelkan Angka Harapan Hidup di Jawa Timur tahun 2013 menggunakan estimasi regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan langkah-langkah sebagai berikut :

a. Melakukan eksplorasi data untuk mengetahui gambaran umum tentang data Angka Harapan Hidup di Jawa Timur.

b. Membuat *scatter plot* secara parsial antar variabel respon dengan masing-masing variabel prediktor.

c. Menentukan variabel prediktor yang dihampiri dengan fungsi linier, fungsi Spline *Truncated* dan yang dihampiri dengan fungsi Deret Fourier.

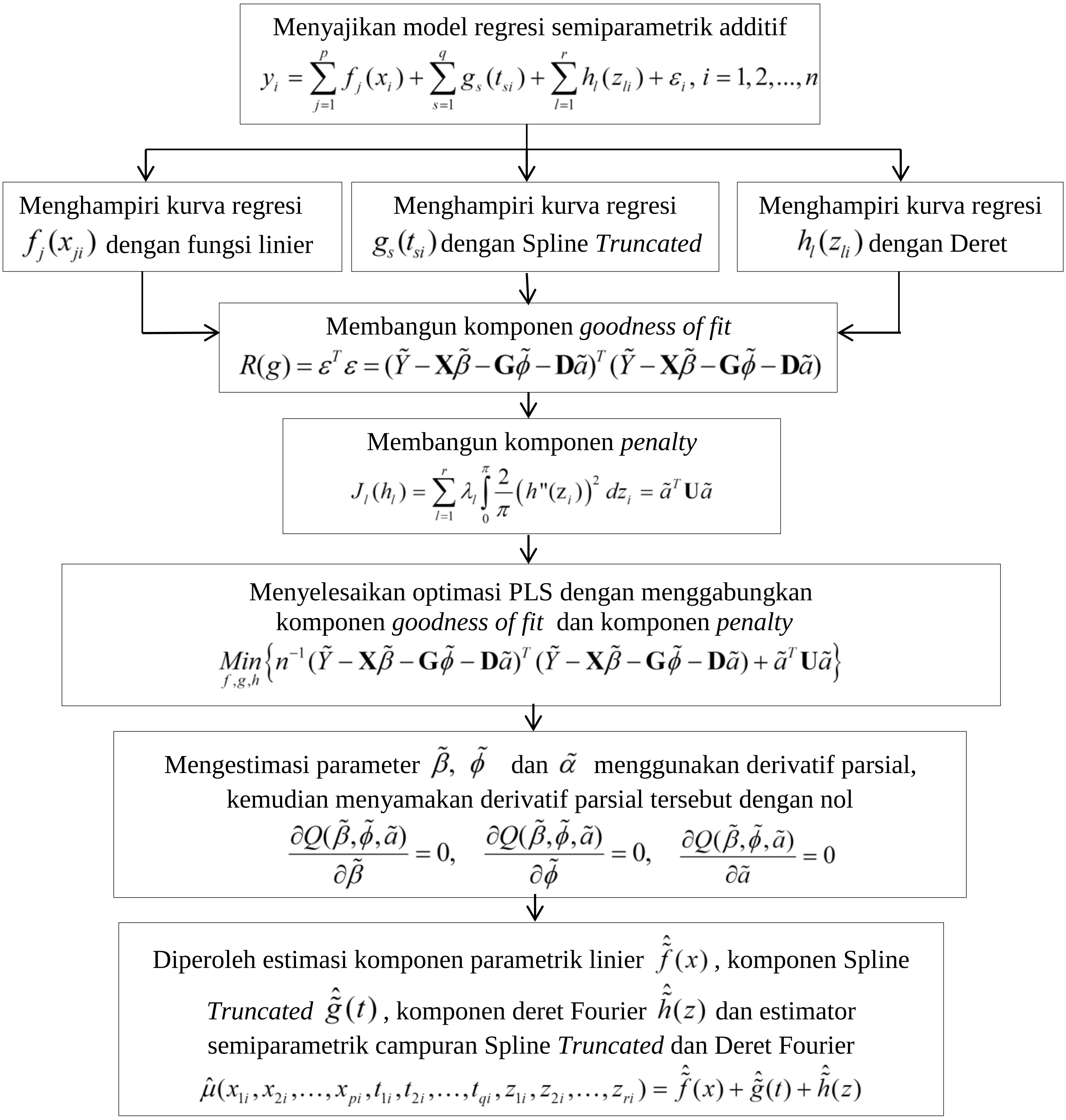
d. Memodelkan data Angka Harapan Hidup di Jawa Timur dengan menggunakan pendekatan semiparametrik campuran Spline *Truncated* dan Deret Fourier.

e. Memilih titik knot *K*, parameter penghalus λ dan parameter osilasi *k* optimum menggunakan metode GCV dan selanjutnya menetapkan model regresi campuran semiparametrik terbaik.

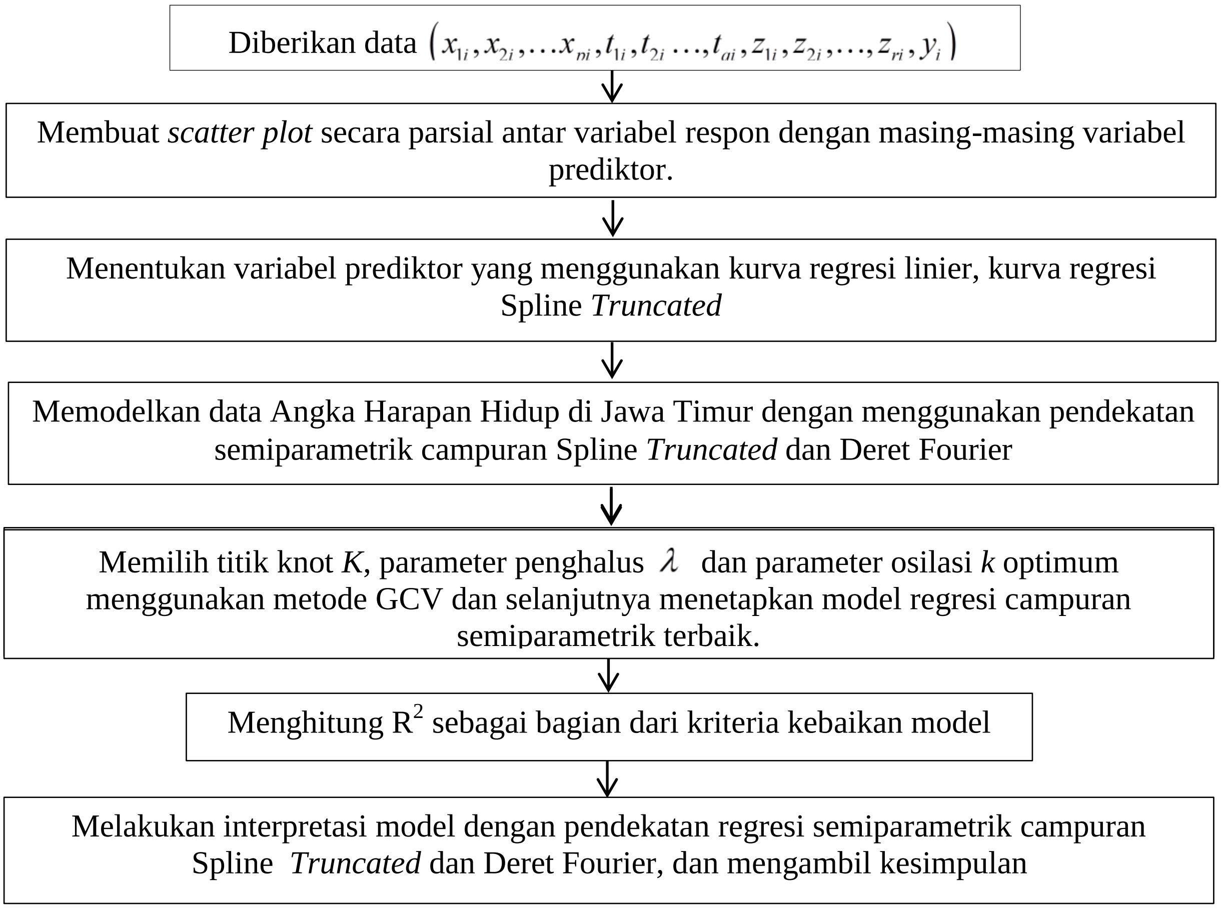
f. Menghitung R^2 sebagai bagian dari kriteria kebaikan model.

g. Melakukan interpretasi model dengan pendekatan regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dan mengambil kesimpulan.

Langkah-langkah analisis dalam penelitian ini dapat digambarkan pada diagram alir seperti Gambar 3.1 dan Gambar 3.2



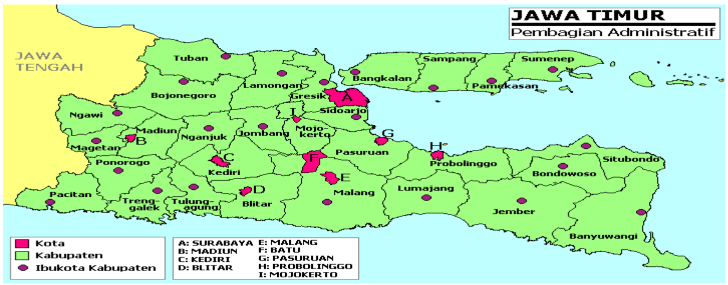
Gambar 3.1 Diagram Alir Langkah Analisis untuk Tujuan Pertama



Gambar 3.2 Diagram Alir Langkah Analisis Data untuk Tujuan Kedua

3.2 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Publikasi Badan Pusat Statistik tahun 2013, yakni Laporan Eksekutif Kesehatan Jawa Timur 2013, Laporan Eksekutif Pendidikan Jawa Timur 2013 dan Survei Ekonomi Nasional Jawa Timur (SUSENAS) 2013. Unit observasi yang digunakan dalam penelitian ini adalah 38 Kabupaten/Kota di Provinsi Jawa Timur yang ditunjukkan pada Gambar 3.3.



Gambar 3.3 Peta Administrasi Jawa Timur

3.3 Variabel Penelitian dan Struktur Data

Variabel-variabel yang akan digunakan dalam penelitian ini yaitu pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Variabel	Simbol Variabel	Nama Variabel
Respon	y	Angka Harapan Hidup (AHH)
Prediktor	x_1	Angka Kematian Bayi (AKB)
	x_2	Persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan
	x_3	Tingkat Partisipasi Angkatan Kerja (TPAK)
	x_4	Angka Melek Huruf (AMH)
	x_5	Rata-rata lama sekolah

Adapun struktur data pada penelitian ini adalah sebagai berikut :

Tabel 3.2 Struktur Data Penelitian

Kabupaten/Kota di Jawa Timur	x_1	x_2	x_3	x_4	x_5	y
1	$x_{1(1)}$	$x_{2(1)}$	$x_{3(1)}$	$x_{4(1)}$	$x_{5(1)}$	$y_{1(1)}$
2	$x_{1(2)}$	$x_{2(2)}$	$x_{3(2)}$	$x_{4(2)}$	$x_{5(2)}$	$y_{1(2)}$
3	$x_{1(3)}$	$x_{2(3)}$	$x_{3(3)}$	$x_{4(3)}$	$x_{5(3)}$	$y_{1(3)}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
38	$x_{1(38)}$	$x_{2(38)}$	$x_{3(38)}$	$x_{4(38)}$	$x_{5(38)}$	$y_{1(38)}$

3.3 Definisi Operasional

a. Angka Harapan Hidup (AHH) adalah perkiraan rata-rata lamanya hidup sejak lahir yang mungkin akan dicapai oleh sekelompok penduduk (BPS, 2012a)
b. Angka Kematian Bayi (AKB) adalah besarnya kemungkinan bayi meninggal sebelum mencapai usia satu tahun, dinyatakan dalam per seribu kelahiran hidup (BPS, 2012b)
c. Persentase bayi berumur 0-11 bulan yang diberi ASI selama 4-6 bulan adalah hasil dari banyaknya bayi berumur 0-11 bulan yang diberi ASI selama 4-6 bulan dengan banyaknya bayi yang berumur 0-11 bulan, dinyatakan dalam persentase (BPS, 2012a)

d. Tingkat Partisipasi Angkatan Kerja (TPAK) adalah rasio antara angkatan kerja dengan jumlah penduduk usia kerja (BPS, 2013a)
e. Angka Melek Huruf (AMH) adalah persentase penduduk usia 15 tahun ke atas yang dapat membaca dan menulis huruf latin dan atau huruf lainnya (BPS, 2012a)
f. Rata-rata lama sekolah adalah rata-rata jumlah tahun yang dihabiskan oleh penduduk pada seluruh jenjang pendidikan formal yang pernah dijalani oleh penduduk 15 tahun ke atas di Jawa timur (BPS, 2013b)

Halaman ini sengaja dikosongkan

BAB IV
HASIL DAN PEMBAHASAN

Berdasarkan tujuan penelitian, pada bagian ini dibahas mengenai estimasi regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier. Selanjutnya, Memodelkan Angka Harapan Hidup (AHH) di Provinsi Jawa Timur pada tahun 2013 menggunakan regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier.

4.1 Model Regresi Semiparametrik Campuran Spline *Truncated* dan Deret Fourier

Berikut ini dibahas mengenai model regresi semiparametrik campuran Spline *Truncated* linear dan Deret Fourier. Diberikan data berpasangan $\left(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri}, y_i\right), i = 1, 2, ..., n,$ yang diasumsikan mengikuti model regresi semiparametrik :

$$y_i = \mu(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri}) + \varepsilon_i \tag{4.1}$$

kurva regresi μ pada persamaan (4.1) diasumsikan bersifat additif sehingga dapat ditulis menjadi :

$$\begin{aligned} y_i &= f(x_{1i}, x_{2i}, \ldots, x_{pi}) + g(t_{1i}, t_{2i}, \ldots, t_{qi}) + h(z_{1i}, z_{2i}, \ldots, z_{ri}) + \varepsilon_i \\ &= f_1(x_{1i}) + f_2(x_{2i}) + \cdots + f_p(x_{pi}) + g_1(t_{1i}) + g_2(t_{2i}) + \cdots + g_q(t_{qi}) \\ &\quad + h_1(z_{1i}) + h_2(z_{2i}) + \cdots + h_r(z_{ri}) + \varepsilon_i . \end{aligned} \tag{4.2}$$

Persamaan 4.2 dapat dituliskan sebagai berikut :

$$y_i = \sum_{j=1}^{p} f_j(x_{ji}) + \sum_{s=1}^{q} g_s(t_{si}) + \sum_{l=1}^{r} h_l(z_{li}) + \varepsilon_i,\ i = 1, 2, ..., n \tag{4.3}$$

Bentuk pola hubungan antara variabel respon y_i dengan variabel prediktor $\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right)$ diasumsikan mengikuti pola linear. Sementara itu, pola hubungan antara variabel respon y_i dengan variabel prediktor $\left(t_{1i}, t_{2i}, \ldots, t_{qi}\right)$, dan variabel prediktor $\left(z_{1i}, z_{2i}, \ldots, z_{ri}\right)$ diasumsikan tidak diketahui, dimana bentuk pola hubungan variabel respon y_i dengan variabel prediktor $\left(t_{1i}, t_{2i}, \ldots, t_{qi}\right)$

diasumsikan berubah-ubah pada sub-sub interval tertentu. Sementara itu, bentuk pola hubungan variabel respon y_i dengan variabel prediktor $(z_{1i}, z_{2i}, \ldots, z_{ri})$ diasumsikan memiliki pola berulang. Sehingga secara teoritis kurva regresi $f_j(x_{ji})$ dapat didekati dengan fungsi linier, kurva regresi $g_s(t_{si})$ didekati dengan fungsi Spline *Truncated* linear. Sementara itu, kurva regresi $h_l(z_{li})$ dapat dihampiri dengan fungsi Deret Fourier. Dengan demikian, kurva campuran $\mu(x_{1i}, x_{2i}, \ldots, x_{pi}, t_{1i}, t_{2i}, \ldots, t_{qi}, z_{1i}, z_{2i}, \ldots, z_{ri})$ secara keseluruhan dapat didekati dengan kurva regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier.

Komponen parametrik linier $f_j(x_{ji})$ pada persamaan (4.3) dapat didekati dengan fungsi linier pada persamaan (2.6). Regresi linier untuk masing-masing variabel prediktor dapat dituliskan dalam bentuk matriks :

$$\begin{bmatrix} f_1\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right) \\ f_2\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right) \\ \vdots \\ f_p\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right) \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \ldots + \beta_p x_{p1} \\ \beta_0 + \beta_1 x_{12} + \ldots + \beta_p x_{p2} \\ \vdots \\ \beta_0 + \beta_1 x_{1n} + \ldots + \beta_p x_{pn} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Jika dituliskan dalam notasi matriks dapat disajikan sebagai berikut:

$$\tilde{f}(x)_{n\times 1} = \mathbf{X}_{n\times(p+1)} \tilde{\beta}_{(p+1)\times 1}, \tag{4.4}$$

dengan :

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix} \text{ dan } \tilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Sedangkan komponen nonparametrik $g_s(t_{si})$ pada persamaan (4.3) dapat didekati dengan fungsi Spline *Truncated* pada persamaan (2.9). Dari persamaan fungsi

Spline *Truncated* pada persamaan (2.9), fungsi Spline *Truncated* linier jika dituliskan dapat dalam bentuk matriks diperoleh sebagai berikut :

$$\begin{bmatrix} g_1(t_{1i},t_{2i},\dots,t_{qi}) \\ g_2(t_{1i},t_{2i},\dots,t_{qi}) \\ \vdots \\ g_q(t_{1i},t_{2i},\dots,t_{qi}) \end{bmatrix} = \begin{bmatrix} \phi_{11}t_{11}+\dots+\phi_{1q}t_{q1}+\alpha_{11}(t_{11}-K_{11})_+ +\dots+\alpha_{m1}(t_{11}-K_{m1})_+ +\dots+\alpha_{1q}(t_{q1}-K_{1q})_+ +\dots+\alpha_{mq}(t_{q1}-K_{mq})_+ \\ \phi_{11}t_{12}+\dots+\phi_{1q}t_{q2}+\alpha_{11}(t_{12}-K_{11})_+ +\dots+\alpha_{m1}(t_{12}-K_{m1})_+ +\dots+\alpha_{1q}(t_{q2}-K_{1q})_+ +\dots+\alpha_{mq}(t_{q2}-K_{mq})_+ \\ \vdots \qquad\qquad \vdots \\ \phi_{11}t_{1n}+\dots+\phi_{1q}t_{qn}+\alpha_{11}(t_{1n}-K_{11})_+ +\dots+\alpha_{m1}(t_{1n}-K_{m1})_+ +\dots+\alpha_{1q}(t_{qn}-K_{1q})_+ +\dots+\alpha_{mq}(t_{qn}-K_{mq})_+ \end{bmatrix}$$

$$= \begin{bmatrix} t_{11} & \cdots & t_{q1} & (t_{11}-K_{11})_+ & \cdots & (t_{11}-K_{m1})_+ & \cdots & (t_{q1}-K_{1q})_+ & \cdots & (t_{q1}-K_{mq})_+ \\ t_{12} & \cdots & t_{q2} & (t_{12}-K_{11})_+ & \cdots & (t_{12}-K_{m1})_+ & \cdots & (t_{q2}-K_{1q})_+ & \cdots & (t_{q2}-K_{mq})_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{1n} & \cdots & t_{qn} & (t_{1n}-K_{11})_+ & \cdots & (t_{1n}-K_{m1})_+ & \cdots & (t_{qn}-K_{1q})_+ & \cdots & (t_{qn}-K_{mq})_+ \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \vdots \\ \phi_{1q} \\ \alpha_{11} \\ \vdots \\ \alpha_{m1} \\ \vdots \\ \alpha_{1q} \\ \vdots \\ \alpha_{mq} \end{bmatrix},$$

jika dituliskan dalam notasi matriks dapat sajikan sebagai berikut :

$$\tilde{g}(t)_{n\times 1} = \mathbf{G}_{n\times((m+1)q)}\tilde{\phi}_{((m+1)q)\times 1} \tag{4.5}$$

dengan,

$$\mathbf{G} = \begin{bmatrix} t_{11} & \cdots & t_{q1} & (t_{11}-K_{11})_+ & \cdots & (t_{11}-K_{m1})_+ & \cdots & (t_{q1}-K_{1q})_+ & \cdots & (t_{q1}-K_{mq})_+ \\ t_{12} & \cdots & t_{q2} & (t_{12}-K_{11})_+ & \cdots & (t_{12}-K_{m1})_+ & \cdots & (t_{q2}-K_{1q})_+ & \cdots & (t_{q2}-K_{mq})_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{1n} & \cdots & t_{qn} & (t_{1n}-K_{11})_+ & \cdots & (t_{1n}-K_{m1})_+ & \cdots & (t_{qn}-K_{1q})_+ & \cdots & (t_{qn}-K_{mq})_+ \end{bmatrix},$$

$$\tilde{\phi} = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1q} & \alpha_{11} & \cdots & \alpha_{m1} & \cdots & \alpha_{1q} & \cdots & \alpha_{mq} \end{bmatrix}^T.$$

Berdasarkan persamaan (4.3) komponen $h_l(z_{li})$ merupakan komponen nonparametrik yang didekati dengan Deret Fourier dapat didefinisikan seperti persamaan (2.13), dapat ditulis dalam bentuk matriks sebagai berikut :

$$\begin{bmatrix} h_1(z_{1i}, z_{2i}, \ldots, z_{ri}) \\ h_2(z_{1i}, z_{2i}, \ldots, z_{ri}) \\ \vdots \\ h_r(z_{1i}, z_{2i}, \ldots, z_{ri}) \end{bmatrix} = \begin{bmatrix} b_1 z_{11} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{11} & \cdots & b_r z_{r1} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{r1} \\ b_1 z_{12} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{12} & \cdots & b_r z_{r2} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{r2} \\ \vdots & \vdots & \vdots \\ b_1 z_{1n} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{1n} & \cdots & b_r z_{rn} + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kz_{rn} \end{bmatrix}$$

$$= \begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \cdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn} \end{bmatrix} \begin{bmatrix} b_1 \\ a_0 \\ a_{11} \\ \vdots \\ a_{K1} \\ \vdots \\ b_r \\ a_0 \\ a_{r1} \\ \vdots \\ a_{Kr} \end{bmatrix}.$$

Jika dituliskan dalam notasi matriks dapat sajikan sebagai berikut :

$$\tilde{h}(z)_{nx1} = \mathbf{D}_{n\times((k+2)r)}\tilde{a}_{(k+2)r)\times 1} \tag{4.6}$$

dengan,

$$\mathbf{D} = \begin{bmatrix} z_{11} & \frac{1}{2} & \cos z_{11} & \cdots & \cos Kz_{11} & \cdots & z_{r1} & \frac{1}{2} & \cos z_{r1} & \cdots & \cos Kz_{r1} \\ z_{12} & \frac{1}{2} & \cos z_{12} & \cdots & \cos Kz_{12} & \cdots & z_{r2} & \frac{1}{2} & \cos z_{r2} & \cdots & \cos Kz_{r2} \\ \vdots & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{1n} & \frac{1}{2} & \cos z_{1n} & \cdots & \cos Kz_{1n} & \cdots & z_{rn} & \frac{1}{2} & \cos z_{rn} & \cdots & \cos Kz_{rn} \end{bmatrix},$$

$$\tilde{a} = \begin{bmatrix} b_1 & a_0 & a_{11} & \cdots & a_{K1} & \cdots & b_r & a_0 & a_{1r} & \cdots & a_{Kr} \end{bmatrix}^T .$$

Berdasarkan uraian untuk komponen parametrik dan komponen nonparametrik di atas, maka model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier dan pada persamaan (4.3) dapat dinyatakan menjadi :

$$\tilde{y}_{n\times 1} = \mathbf{X}_{n\times(p+1)}\tilde{\beta}_{(p+1)\times 1} + \mathbf{G}_{n\times((m+1)q)}\tilde{\phi}_{((m+1)q)\times 1} + \mathbf{D}_{n\times((k+2)r)}\tilde{a}_{(k+2)r)\times 1} + \tilde{\varepsilon}_{n\times 1} \tag{4.7}$$

4.2 Estimasi Kurva Regresi Campuran Semiparametrik Spline *Truncated* dan Deret Fourier

Teorema 1

Diberikan model regresi semiparametrik campuran pada persamaan (4.3), dimana kurva regresi $f_j(x_{ji})$ didekati oleh persamaan (2.6), kurva regresi $g_s(t_{si})$ didekati oleh persamaan (2.9), dan kurva regresi $h_l(\mathrm{z}_{li})$ didekati oleh persamaan (2.13), maka estimator model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier dan dapat diperoleh dengan menggunakan metode *Penalized Least Square*

$$\underset{f,g,h}{Min}\left\{ n^{-1}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} f_j(x_{ji}) - \sum_{s=1}^{q} g_s(t_{si}) - \sum_{l=1}^{r} h_l(\mathrm{z}_{li}) \right)^2 + \sum_{l=1}^{r}\lambda_l \int_0^{\pi} \frac{2}{\pi}\left(h_l''(\mathrm{z}_l)\right)^2 dz_l \right\}. \quad (4.8)$$

Dalam menyelesaikan optimasi PLS ini, dibutuhkan beberapa lemma.

Lemma 1

Apabila fungsi $f_j(x_{ji})$ didekati oleh persamaan (2.6), kurva regresi $g_s(t_{si})$ didekati oleh persamaan (2.9), dan kurva regresi $h_l(\mathrm{z}_{li})$ didekati oleh persamaan (2.13), maka *goodness of fit* $R(f,g,h)$ diberikan oleh :

$$R(f,g,h) = n^{-1}(\tilde{Y} - \mathbf{X}\tilde{\beta} - \mathbf{G}\tilde{\phi} - \mathbf{D}\tilde{a})^T(\tilde{Y} - \mathbf{X}\tilde{\beta} - \mathbf{G}\tilde{\phi} - \mathbf{D}\tilde{a})$$

dengan

$\tilde{Y} = \left[y_1, y_2, \ldots, y_n\right]^T$, $\tilde{\beta} = \left[\beta_0, \beta_1, \ldots, \beta_p\right]^T$,

$\tilde{\phi} = \left[\phi_{11} \ \cdots \ \phi_{1q} \ \ \alpha_{11} \ \cdots \ \alpha_{m1} \ \cdots \ \alpha_{1q} \ \cdots \ \alpha_{mq}\right]^T$, dan

$\tilde{a} = \left[b_1 \ \ a_0 \ \ a_{11} \ \cdots \ a_{K1} \ \vdots \ \cdots \ \vdots \ b_r \ \ a_0 \ \ a_{1r} \ \cdots \ a_{Kr}\right]^T$

Bukti :

Goodness of fit didefinisikan sebagai berikut :

$$R(f,g,h) = n^{-1}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} f_j(x_{ji}) - \sum_{s=1}^{q} g_s(t_{si}) - \sum_{l=1}^{r} h_l(\mathrm{z}_{li}) \right)^2. \quad (4.9)$$

Selanjutnya, didapat :

$$R(f,g,h)=n^{-1}\sum_{i=1}^{n}\left(y_i-\beta_0+\beta_1x_{i1}+\beta_2x_{i2}+\cdots+\beta_px_{ip}-\sum_{j=1}^{q}\phi_{ji}t_i^j+\sum_{k=1}^{m}\alpha_{ki}\left(t_i-K_k\right)_+ + \right.$$

$$\left.-b_lz_{li}+\frac{1}{2}a_{0l}+\sum_{k=1}^{K}a_{kl}\cos kz_{li}\right)^2.$$

Akibatnya, *goodness of fit* dari $R(f,g,h)$, adalah :

$$R(f,g,h)=n^{-1}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^T(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a}). \qquad (4.10)$$

Lemma 2

Apabila fungsi $h_l(z_l)$ diberikan persamaan (2.13) maka komponen *penalty* dari $J_l(h_l)$ pada optimasi (4.8) diberikan sebagai berikut :

$$J_l(h_l)=\sum_{k=1}^{K}k^4a_{kl}^2$$

Bukti :

Fungsi $h_l\left(z_1,z_2,\ldots,z_r\right)=b_lz_l+\frac{1}{2}a_{0l}+\sum_{k=1}^{K}a_{kl}\cos kz_l$, maka derivatif kedua dari fungsi tersebut adalah

$$h_l''(z_l)=\frac{d}{dz_l}\left[\frac{d}{dz_l}\left(bz_l+\frac{1}{2}a_{0l}+\sum_{k=1}^{K}a_{kl}\cos kz_l\right)\right]$$

$$=\frac{d}{dz_l}\left(b-\sum_{k=1}^{K}ka_{kl}\sin kz_l\right)$$

$$=-\sum_{k=1}^{K}k^2a_{kl}\cos kz_l.$$

Diperoleh *penalty* dari $J_l(h_l)$ sebagai berikut :

$$J_l(h_l)=\int_0^{\pi}\frac{2}{\pi}\left(\sum_{k=1}^{K}k^2a_{kl}\cos kz_l\right)^2dz_l$$

$$=\frac{2}{\pi}\int_0^{\pi}\left\{\sum_{k=1}^{K}\left(k^2a_{kl}\cos kz_l\right)+2\sum_{k<j}^{K}\left(k^2a_{kl}\cos kz_l\right)\left(j^2a_{jl}\cos jz_l\right)\right\}dz_l,$$

selanjutnya, diasumsikan bahwa

$$A=\frac{2}{\pi}\sum_{k=1}^{K}\int_0^{\pi}\left(k^2a_{kl}\cos kz_l\right)^2dz_l \text{ dan } B=\frac{2}{\pi}2\sum_{k<j}^{K}\int_0^{\pi}\left(k^2a_{kl}\cos kz_l\right)^2\left(j^2a_{jl}\cos jz_l\right)dz_l,$$

kemudian, A dan B dapat dihitung dan dijabarkan seperti berikut :

$$A = \frac{2}{\pi}\sum_{k=1}^{K}\int_{0}^{\pi}\left(k^2 a_{kl}\cos kz_l\right)^2 dz_l$$

$$= \frac{1}{\pi}\sum_{k=1}^{K}k^4 a_{kl}^2\left[\frac{1}{2}z_l + \frac{1}{4k}\sin kz_l\right]_0^{\pi}$$

$$= \sum_{k=1}^{K}k^4 a_{kl}^2 \tag{4.11}$$

$$B = \frac{4}{\pi}\sum_{k<j}^{K}\int_{0}^{\pi}\left(k^2 a_{kl}\cos kz_l\right)\left(j^2 a_{jl}\cos jz_l\right)dz_l$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_{kl}a_{jl}\int_{0}^{\pi}\cos kz_l\cos jz_l dz_l$$

$$= 0. \tag{4.12}$$

Berdasarkan persamaan (4.11) dan (4.12), maka *penalty* $J_l(h_l)$ diperoleh sebagai berikut :

$$J_l(h_l) = \sum_{k=1}^{K}k^4 a_{kl}^2 \qquad \blacksquare \tag{4.13}$$

Berdasarkan persamaan (4.13), *penalty* dapat dituliskan :

$$\sum_{l=1}^{r}\lambda_l J_l(h_l) = \sum_{l=1}^{r}\lambda_l\left(\sum_{k=1}^{K}k^4 a_{kl}^2\right)$$

$$= \lambda_1\left(\sum_{k=1}^{K}k^4 a_{k1}^2\right) + \lambda_2\left(\sum_{k=1}^{K}k^4 a_{k2}^2\right) + \cdots + \lambda_r\left(\sum_{k=1}^{K}k^4 a_{kr}^2\right)$$

$$= \lambda_1\left(1^4 a_{11}^2 + 2^4 a_{21}^2 + 3^4 a_{31}^2 + \cdots + K^4 a_{K1}^2\right) + \lambda_2\left(1^4 a_{12}^2 + 2^4 a_{22}^2 + 3^4 a_{32}^2 + \cdots + K^4 a_{K2}^2\right) + \cdots +$$
$$\lambda_r\left(1^4 a_{1r}^2 + 2^4 a_{2r}^2 + 3^4 a_{3r}^2 + \cdots + K^4 a_{Kr}^2\right),$$

diberikan :

$$\mathbf{U}(\lambda_1,\lambda_2,\ldots,\lambda_r) = diag\left(\mathbf{U}_1(\lambda_1) \quad \mathbf{U}_2(\lambda_2) \quad \cdots \quad \mathbf{U}_r(\lambda_r)\right),$$

$$\mathbf{U}_l(\lambda_l) = \left(0 \quad 0 \quad \lambda_l 1^4 \quad \lambda_l 2^4 \quad \cdots \quad \lambda_l K^4\right), l = 1,2,\ldots,r$$

Berdasarkan uraian tersebut, diperoleh $\sum_{l=1}^{r}\lambda_l J_l(h_l) = \tilde{a}^T\mathbf{U}\tilde{a}.$ (4.14)

Dengan menggabungkan *goodness of fit* (4.10) dan *penalty* (4.13), maka optimisasi dari (4.8) diberikan sebagai berikut :

$$\underset{\tilde{\beta},\tilde{\phi},\tilde{a}}{Min}\left\{n^{-1}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^T(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})+\tilde{a}^T\mathbf{U}\tilde{a}\right\}=\underset{\tilde{\beta},\tilde{\phi},\tilde{a}}{Min}\left\{Q(\tilde{\beta},\tilde{\phi},\tilde{a})\right\}$$

$$\begin{aligned}Q(\tilde{\beta},\tilde{\phi},\tilde{a}) &= n^{-1}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^T(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})+\tilde{a}^T\mathbf{U}\tilde{a}\\
&= n^{-1}(\tilde{Y}^T-\tilde{\beta}^T\mathbf{X}^T-\tilde{\phi}^T\mathbf{G}^T-\tilde{a}^T\mathbf{D}^T)(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})+\tilde{a}^T\mathbf{U}\tilde{a}\\
&= n^{-1}(\tilde{Y}^T\tilde{Y}-\tilde{Y}^T\mathbf{X}\tilde{\beta}-\tilde{Y}^T\mathbf{G}\tilde{\phi}-\tilde{Y}^T\mathbf{D}\tilde{a}-\tilde{\beta}^T\mathbf{X}^T\tilde{Y}+\tilde{\beta}^T\mathbf{X}^T\mathbf{X}\tilde{\beta}+\\
&\quad \tilde{\beta}^T\mathbf{X}^T\mathbf{G}\tilde{\phi}+\tilde{\beta}^T\mathbf{X}^T\mathbf{D}\tilde{a}-\tilde{\phi}^T\mathbf{G}^T\tilde{Y}+\tilde{\phi}^T\mathbf{G}^T\mathbf{X}\tilde{\beta}+\tilde{\phi}^T\mathbf{G}^T\mathbf{G}\tilde{\phi}+\\
&\quad \tilde{\phi}^T\mathbf{G}^T\mathbf{D}\tilde{a}-\tilde{a}^T\mathbf{D}^T\tilde{Y}+\tilde{a}^T\mathbf{D}^T\mathbf{X}\tilde{\beta}+\tilde{a}^T\mathbf{D}^T\mathbf{G}\tilde{\phi}+\tilde{a}^T\mathbf{D}^T\mathbf{D}\tilde{a})+\tilde{a}^T\mathbf{U}\tilde{a}\\
&= n^{-1}\tilde{Y}^T\tilde{Y}-2n^{-1}\tilde{\beta}^T\mathbf{X}^T\tilde{Y}-2n^{-1}\tilde{\phi}^T\mathbf{G}^T\tilde{Y}-2n^{-1}\tilde{a}^T\mathbf{D}^T\tilde{Y}+n^{-1}\tilde{\beta}^T\mathbf{X}^T\mathbf{X}\tilde{\beta}+\\
&\quad 2n^{-1}\tilde{\phi}^T\mathbf{G}^T\mathbf{X}\tilde{\beta}+2n^{-1}\tilde{a}^T\mathbf{D}^T\mathbf{X}\tilde{\beta}+n^{-1}\tilde{\phi}^T\mathbf{G}^T\mathbf{G}\tilde{\phi}+2n^{-1}\tilde{a}^T\mathbf{D}^T\mathbf{G}\tilde{\phi}+\\
&\quad n^{-1}\tilde{a}^T\mathbf{D}^T\mathbf{D}\tilde{a}+\tilde{a}^T\mathbf{U}\tilde{a}\end{aligned}$$

Derivatif parsial dari $Q(\tilde{\beta},\tilde{\phi},\tilde{a})$ terhadap $\tilde{\beta}$, diperoleh :

$$\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{\beta}}=-2n^{-1}\mathbf{X}^T\tilde{Y}+2n^{-1}\mathbf{X}^T\mathbf{X}\hat{\tilde{\beta}}+2n^{-1}\mathbf{X}^T\mathbf{G}\hat{\tilde{\phi}}+2n^{-1}\mathbf{X}^T\mathbf{D}\hat{\tilde{a}}. \qquad (4.15)$$

Derivatif parsial dari $Q(\tilde{\beta},\tilde{\phi},\tilde{a})$ terhadap $\tilde{\phi}$, diperoleh :

$$\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{\phi}}=-2n^{-1}\mathbf{G}^T\tilde{Y}+2n^{-1}\mathbf{G}^T\mathbf{X}\hat{\tilde{\beta}}+2n^{-1}\mathbf{G}^T\mathbf{G}\hat{\tilde{\phi}}+2n^{-1}\mathbf{G}^T\mathbf{D}\hat{\tilde{a}}. \qquad (4.16)$$

Derivatif parsial dari $Q(\tilde{\beta},\tilde{\phi},\tilde{a})$ terhadap $\tilde{a}$, diperoleh :

$$\begin{aligned}\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{a}} &= -2n^{-1}\mathbf{D}^T\tilde{Y}+2n^{-1}\mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}}+2n^{-1}\mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}}+2n^{-1}\mathbf{D}^T\mathbf{D}\hat{\tilde{a}}+\mathbf{U}\hat{\tilde{a}}\\
&= -2n^{-1}\mathbf{D}^T\tilde{Y}+2n^{-1}\mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}}+2n^{-1}\mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}}+\left(2n^{-1}\mathbf{D}^T\mathbf{D}+\mathbf{U}\right)\hat{\tilde{a}}. \qquad (4.17)\end{aligned}$$

Hasil derivatif-derivatif yang diperoleh pada persamaan (4.15), (4.16) dan (4.7) disamadengankan dengan nol, sebagai berikut :

$$\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{\beta}}=0,$$

$$-2n^{-1}\mathbf{X}^T\tilde{Y}+2n^{-1}\mathbf{X}^T\mathbf{X}\hat{\tilde{\beta}}+2n^{-1}\mathbf{X}^T\mathbf{G}\hat{\tilde{\phi}}+2n^{-1}\mathbf{X}^T\mathbf{D}\hat{\tilde{a}}=0$$

$$\mathbf{X}^T\mathbf{X}\hat{\tilde{\beta}}=\mathbf{X}^T\tilde{Y}-\mathbf{X}^T\mathbf{G}\hat{\tilde{\phi}}-\mathbf{X}^T\mathbf{D}\hat{\tilde{a}}$$

$$\hat{\tilde{\beta}} = \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\left(\mathbf{X}^T\tilde{Y} - \mathbf{X}^T\mathbf{G}\hat{\tilde{\phi}} - \mathbf{X}^T\mathbf{D}\hat{\tilde{a}}\right)$$

$$\hat{\tilde{\beta}} = \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\mathbf{X}^T\tilde{Y} - \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{G}\hat{\tilde{\phi}} - \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{D}\hat{\tilde{a}}. \quad (4.18)$$

Selanjutnya untuk $\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{\phi}} = 0,$ dapat dijabarkan sebagai berikut :

$$-2n^{-1}\mathbf{G}^T\tilde{Y} + 2n^{-1}\mathbf{G}^T\mathbf{X}\hat{\tilde{\beta}} + 2n^{-1}\mathbf{G}^T\mathbf{G}\hat{\tilde{\phi}} + 2n^{-1}\mathbf{G}^T\mathbf{D}\hat{\tilde{a}} = 0$$

$$\mathbf{G}^T\mathbf{G}\hat{\tilde{\phi}} = \mathbf{G}^T\tilde{Y} - \mathbf{G}^T\mathbf{X}\hat{\tilde{\beta}} - \mathbf{G}^T\mathbf{D}\hat{\tilde{a}}$$

$$\hat{\tilde{\phi}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T(\tilde{Y} - \mathbf{X}\hat{\tilde{\beta}} - \mathbf{D}\hat{\tilde{a}})$$

$$\hat{\tilde{\phi}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\tilde{Y} - \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{X}\hat{\tilde{\beta}} - \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{D}\hat{\tilde{a}}. \quad (4.19)$$

Sementara untuk $\frac{\partial Q(\tilde{\beta},\tilde{\phi},\tilde{a})}{\partial\tilde{a}} = 0$, dapat diuraikan seperti berikut :

$$-2n^{-1}\mathbf{D}^T\tilde{Y} + 2n^{-1}\mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}} + 2n^{-1}\mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}} + \left(2n^{-1}\mathbf{D}^T\mathbf{D} + \mathbf{U}\right)\hat{\tilde{a}} = 0$$

$$(n^{-1}\mathbf{D}^T\mathbf{D} + \mathbf{U})\hat{\tilde{a}} = (\mathbf{D}^T\tilde{Y} - \mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}} - \mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}})$$

$$\hat{\tilde{a}} = (\mathbf{D}^T\mathbf{D} + n\mathbf{U})^{-1}(\mathbf{D}^T\tilde{Y} - \mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}} - \mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}})$$

$$\hat{\tilde{a}} = (\mathbf{D}^T\mathbf{D} + n\mathbf{U})^{-1}\mathbf{D}^T\tilde{Y} - (\mathbf{D}^T\mathbf{D} + n\mathbf{U})^{-1}\mathbf{D}^T\mathbf{X}\hat{\tilde{\beta}} - (\mathbf{D}^T\mathbf{D} + n\mathbf{U})^{-1}\mathbf{D}^T\mathbf{G}\hat{\tilde{\phi}}. \quad (4.20)$$

Berdasarkan hasil yang diperoleh pada persamaan (4.18), (4.19) dan (4.20) terlihat bahwa $\hat{\tilde{\beta}}$, $\hat{\tilde{\phi}}$ dan $\hat{\tilde{a}}$ masih mengandung parameter, maka untuk menyelesaikan permasalahan tersebut digunakan metode eleminasi substitusi. Untuk memudahkan proses perhitungan maka persamaan (4.18), (4.19) dan (4.20) dapat dituliskan sebagai berikut :

$$\hat{\tilde{\beta}} = P\tilde{Y} - PG\hat{\tilde{\phi}} - PD\hat{\tilde{a}} \quad (4.17)$$

dengan

$$P = \left(X^TX\right)^{-1}X^T,$$

$$\hat{\tilde{\phi}} = Q\tilde{Y} - QX\hat{\tilde{\beta}} - QD\hat{\tilde{a}} \quad (4.18)$$

dengan

$$Q = \left(G^TG\right)^{-1}G^T$$

$$\hat{\tilde{a}} = R\tilde{Y} - RX\hat{\tilde{\beta}} - RG\hat{\tilde{\phi}}, \tag{4.19}$$

dengan

$$R = \left(D^T D + nU\right)^{-1} D^T .$$

Langkah pertama adalah menggunakan metode eliminasi untuk menemukan persamaan yang memuat dua parameter antara persamaan (4.17) dan (4.19), dengan cara mengalikan kedua kedua ruas persamaan (4.19) dengan *PD*. Kemudian, persamaan (4.17) diselisihkan dengan hasil perkalian persamaan (4.19) dengan *PD,* diperoleh :

$$\hat{\tilde{\beta}} - PD\hat{\tilde{a}} + PG\hat{\tilde{\phi}} + PDRX\hat{\tilde{\beta}} + PD\hat{\tilde{a}} - PDRG\hat{\tilde{\phi}} = P\tilde{Y} - PDR\tilde{Y},$$

persamaan tersebut dapat dituliskan seperti berikut :

$$\left[I - PDRX\right]\hat{\tilde{\beta}} + \left[PG - PDRG\right]\hat{\tilde{\phi}} = \left[P - PDR\right]\tilde{Y}. \tag{4.20}$$

Melakukan metode eliminasi pada persamaan (4.18) dan (4.19) untuk mendapatkan sebuah persamaan yang memuat dua parameter, dengan cara mengalikan persamaan (4.19) dengan *QD*. Kemudian persamaan (4.18) diselisihkan dengan hasil perkalian persamaan (4.19) dengan *QD,* diperoleh :

$$\hat{\tilde{\phi}} - QD\hat{\tilde{a}} + QX\hat{\tilde{\beta}} - QDRX\hat{\tilde{\beta}} + QD\hat{\tilde{a}} - QDRG\hat{\tilde{\phi}} = Q\tilde{Y} - PDR\tilde{Y},$$

persamaan tersebut dapat dituliskan seperti berikut :

$$\left[I - QDRG\right]\hat{\tilde{\phi}} + \left[QX - QDRX\right]\hat{\tilde{\beta}} = \left[Q - PDR\right]\tilde{Y}. \tag{4.21}$$

Berdasarkan hasil dari dua metode eliminasi yang telah dilakukan, diperoleh persamaan (4.20) dan (4.21), dimana kedua persamaan tersebut hanya mengandung dua parameter. Selanjutnya, menggunakan metode eliminasi pada kedua persamaan tersebut, dengan mengalikan persamaan (4.21) dengan $\left[PG - PDRG\right]\left[I - QDRG\right]^{-1}$. Kemudian persamaan (4.20) diselisihkan dengan hasil perkalian persamaan (4.12) dengan $\left[PG - PDRG\right]\left[I - QDRG\right]^{-1}$ diperoleh :

$$\left[I - PDRX\right]\hat{\tilde{\beta}} - \left[PG - PDRG\right]\left[I - QDRG\right]^{-1}\left[I - QDRG\right]\hat{\tilde{\phi}} + \left[PG - PDRG\right]\hat{\tilde{\phi}} -$$
$$\left[PG - PDRG\right]\left[I - QDRG\right]^{-1}\left[QX - QDRX\right]\hat{\tilde{\beta}} = \left[P - PDR\right]\tilde{\tilde{Y}} - \left[PG - PDRG\right]$$
$$\left[I - QDRG\right]^{-1}\left[Q - QRD\right]\tilde{Y}$$

Persamaan tersebut dapat dituliskan seperti berikut :

$$\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)\hat{\tilde{\beta}}=\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right)\tilde{Y}.$$

Kedua ruas persamaan tersebut dikalikan dengan $\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}$, sehingga diperoleh :

$$\hat{\tilde{\beta}}=\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right)\tilde{Y} \tag{4.22}$$

$$\hat{\tilde{\beta}}=\mathbf{M}\left(K,\lambda,k\right)\tilde{Y},$$

dengan

$$\mathbf{M}\left(K,\lambda,k\right)=\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right).$$

Mensubstitusi persamaan (4.22) pada persamaan (4.20), seperti pada uraian berikut :

$$\left[I-PDRX\right]\left[\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right)\tilde{Y}\right]+\left[PG-PDRG\right]\hat{\tilde{\phi}}=\left[P-PDR\right]\tilde{Y},$$

$$\hat{\tilde{\phi}}=\left(\left[I-PDRX\right]\left[\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right)\tilde{Y}\right]+\left[PG-PDRG\right]\right)^{-1}\left[P-PDR\right]\tilde{Y} \tag{4.23}$$

$$\hat{\tilde{\phi}}=\mathbf{N}\left(K,\lambda,k\right)\tilde{Y},$$

dengan

$$\mathbf{N}\left(K,\lambda,k\right)=\left(\left[I-PDRX\right]\left[\left(\left[I-PDRX\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[QX-QDRX\right]\right)^{-1}\left(\left[P-PDR\right]-\left[PG-PDRG\right]\left[I-QDRG\right]^{-1}\left[Q-QRD\right]\right)\tilde{Y}\right]+\left[PG-PDRG\right]\right)^{-1}\left[P-PDR\right].$$

Mensubstitusi persamaan (4.22) dan persamaan (4.23) pada persamaan (4.19)

$$
\begin{aligned}
\hat{\hat{a}} = {} & R\tilde{Y} - RX\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}\Big([P-PDR]-[PG-PDRG] \\
& [I-QDRG]^{-1}[Q-QRD]\tilde{Y}\Big] - RG\Big[\Big([I-PDRX]\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1} \\
& \Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\tilde{Y}\Big]+[PG-PDRG]\Big)^{-1}[P-PDR]\tilde{Y}\Big]
\end{aligned}
$$

$$
\begin{aligned}
\hat{\hat{a}} = {} & \Big(R - RX\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}\Big([P-PDR]-[PG-PDRG] \\
& [I-QDRG]^{-1}[Q-QRD]\Big)\Big] - RG\Big[\Big([I-PDRX]\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-} \\
& \Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\tilde{Y}\Big]+[PG-PDRG]\Big)^{-1}[P-PDR]\Big]\Big)\tilde{Y}
\end{aligned}
\qquad (4.24)
$$

$$\hat{\hat{a}} = \mathbf{O}(K,\lambda,k)\tilde{Y},$$

dengan

$$
\begin{aligned}
\mathbf{O}(K,\lambda,k) = {} & \Big(R - RX\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}\Big([P-PDR]-[PG-PDRG] \\
& [I-QDRG]^{-1}[Q-QRD]\Big)\Big] - RG\Big[\Big([I-PDRX]\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1} \\
& \Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\tilde{Y}\Big]+[PG-PDRG]\Big)^{-1}[P-PDR]\Big]\Big).
\end{aligned}
$$

Berdasarkan persamaan (4.22), maka diperoleh estimator kurva regresi parametrik linier :

$$
\begin{aligned}
\hat{\hat{f}}(x) &= \mathbf{X}\hat{\hat{\beta}} \\
&= \mathbf{A}(K,\lambda,k)\tilde{Y},
\end{aligned}
\qquad (4.25)
$$

dengan

$$
\begin{aligned}
\mathbf{A}(K,\lambda,k) = \mathbf{X}\Big[& \Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1} \\
& \Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\Big].
\end{aligned}
$$

Berdasarkan persamaan (4.23), diperoleh estimator kurva regresi Spline *Truncated* :

$$
\begin{aligned}
\hat{\hat{g}}(t) &= \mathbf{G}\hat{\hat{\phi}} \\
&= \mathbf{B}(K,\lambda,k)\tilde{Y},
\end{aligned}
\qquad (4.26)
$$

dengan

$$\mathbf{B}(K,\lambda,k)=\mathbf{G}\Bigg[\Big([I-PDRX]\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}$$

$$\Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\tilde{Y}\Big]+[PG-PDRG]\Big)^{-1}$$

$$[P-PDR]\Bigg].$$

Berdasarkan $\hat{\tilde{a}}$, diperoleh estimator kurva regresi Deret Fourier :

$$\begin{aligned}\hat{\tilde{h}}(z)&=\mathbf{D}\hat{\tilde{a}}\\&=\mathbf{C}(K,\lambda,k)\tilde{Y},\end{aligned}\qquad(4.27)$$

dengan

$$\mathbf{C}(K,\lambda,k)=\mathbf{D}\Bigg(R-RX\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}$$

$$\Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\Big]$$

$$-RG\Big[\Big([I-PDRX]\Big[\Big([I-PDRX]-[PG-PDRG][I-QDRG]^{-1}[QX-QDRX]\Big)^{-1}$$

$$\Big([P-PDR]-[PG-PDRG][I-QDRG]^{-1}[Q-QRD]\Big)\tilde{Y}\Big]+[PG-PDRG]\Big)^{-1}[P-PDR]\Big]\Bigg).$$

Diperoleh estimasi model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, sebagai berikut:

$$\begin{aligned}\hat{y}_i&=\hat{\mu}(x_{1i},x_{2i},\ldots,x_{pi},t_{1i},t_{2i},\ldots,t_{qi},z_{1i},z_{2i},\ldots,z_{ri})\\&=\hat{\tilde{f}}(x)+\hat{\tilde{g}}(t)+\hat{\tilde{h}}(z)\\&=\mathbf{A}(k,\lambda,K)\tilde{Y}+\mathbf{B}(k,\lambda,K)\tilde{Y}+\mathbf{C}(k,\lambda,K)\tilde{Y}\\&=\big[\mathbf{A}(k,\lambda,K)+\mathbf{B}(k,\lambda,K)+\mathbf{C}(k,\lambda,K)\big]\tilde{Y}\\&=\mathbf{F}(K,\lambda,k)\tilde{Y}\end{aligned}\qquad(4.28)$$

dengan

$$\mathbf{F}(K,\lambda,k)=\mathbf{A}(K,\lambda,k)+\mathbf{B}(K,\lambda,k)+\mathbf{C}(K,\lambda,k).$$

4.3 Pemodelan Angka Harapan Hidup di Provinsi Jawa Timur Menggunakan Regresi Semiparametrik dengan estimator Campuran Spline Truncated dan Deret Fourier

Berikut ini akan dibahas mengenai pemodelan Angka Harapan Hidup di Jawa Timur pada tahun 2013 menggunakan regresi semiparametrik dengan estimator campuran Spline *Truncated* dan Deret Fourier.

4.3.1 Eksplorasi Data

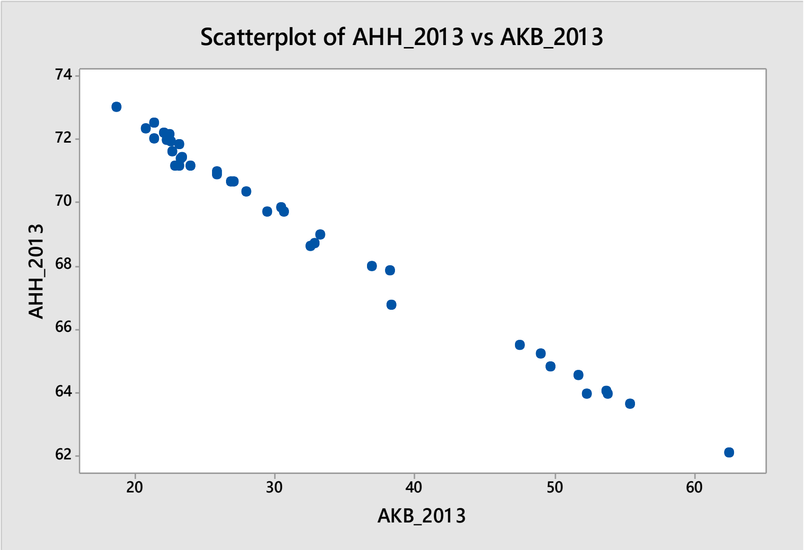
Unit Observasi yang digunakan dalam penelitian ini adalah 38 Kabupaten/Kota di Provinsi Jawa Timur. Banyak variabel yang digunakan adalah enam variabel, yaitu satu variabel respon dan lima variabel prediktor. Variabel respon yang digunakan adalah Angka Harapan Hidup, sedangkan variabel-variabel prediktornya adalah Angka Kematian Bayi, persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan, Tingkat Partisipasi Angkatan Kerja, Angka Melek huruf dan rata-rata lama sekolah. Berikut statistik deskriptif variabel respon dan variabel prediktor yang digunakan :

Tabel 4.1 Deskriptif Data

Variabel	Jumlah data	Min	Maks	Range	Rata-Rata	Variansi
AHH	2,626,660	62,1	73,000	10,9	69,123	9,865
AKB	1246,01	18,71	62,45	43,74	32,79	155,25
ASI	127,060	1,14	4,810	3,670	3,344	0,624
TPAK	2,658,230	63,7	79,440	15,74	69,953	14,082
AMH	3425,37	69,47	98,40	28,93	90,14	45,34
RLS	295,430	4,39	10,890	6,5	7,774	2,368

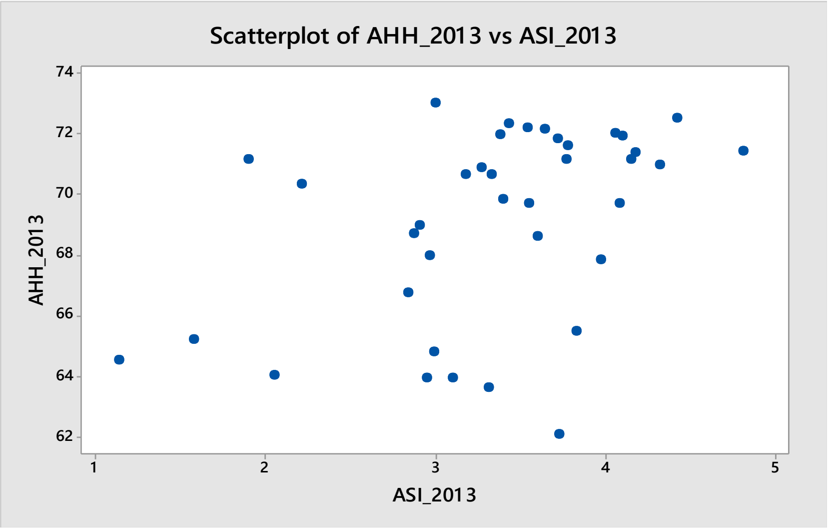
Tahap awal sebelum melakukan pemodelan regresi dalam hal ini memodelkan kasus Angka Harapan Hidup di Jawa Timur adalah mengetahui bentuk pola hubungan antara variabel respon dengan setiap variabel prediktor. Informasi mengenai bentuk pola hubungan tersebut digunakan untuk menentukan jenis kurva regresi yang sesuai dalam menghampiri pola data. Informasi mengenai bentuk pola hubungan antara variabel respon dan variabel prediktor dapat diperoleh dari sebaran data yang disajikan dalam bentuk *scatter plot*.

Hasil diagram pencar untuk variabel respon terhadap masing-masing variabel prediktor adalah sebagai berikut :



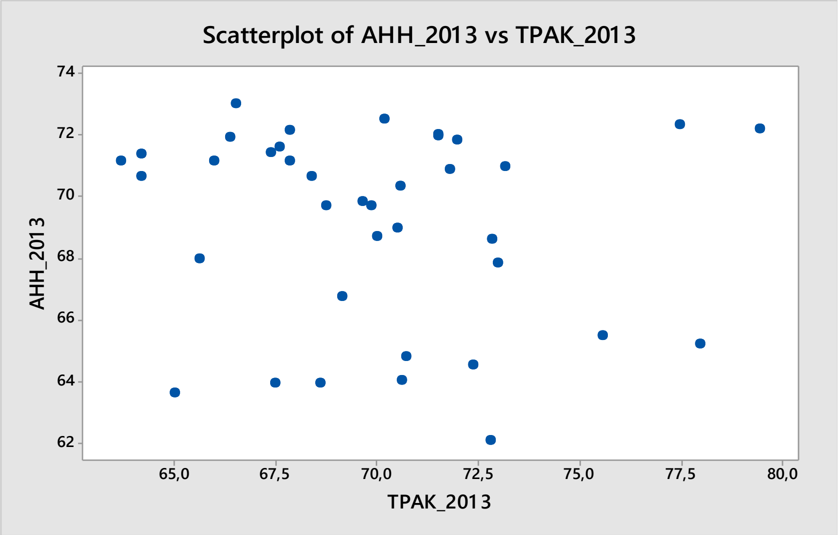
Gambar 4.1 *Scatter Plot* antara Angka Harapan Hidup dengan Angka Kematian Bayi

Gambar 4.1 menggambarkan pola hubungan antara Angka Harapan Hidup dengan Angka Kematian Bayi. Pola hubungan yang terbentuk antara variabel tersebut yakni membentuk hubungan linier negatif. Hal ini terlihat dari pergerakan plot secara umum yakni semakin tinggi Angka Kematian Bayi maka Angka Harapan Hidup cenderung semakin menurun. Begitu pula sebaliknya, semakin rendah Angka Kematian Bayi maka Angka Harapan Hidup cenderung meningkat.



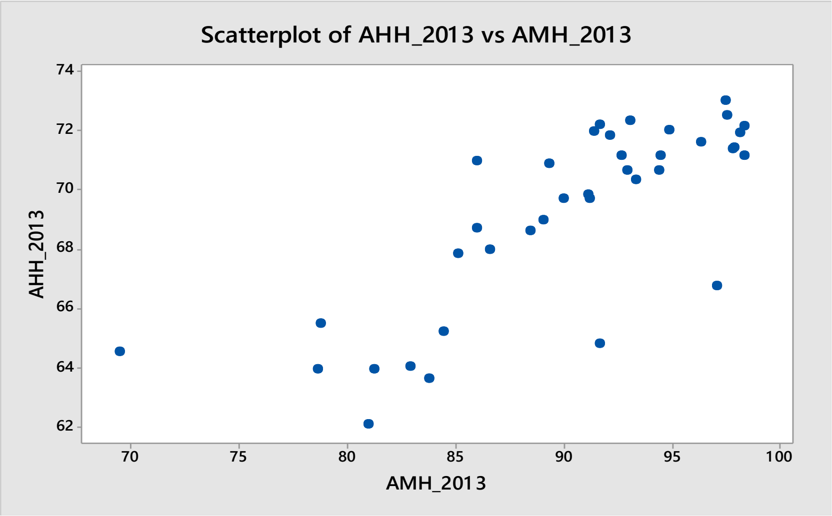
Gambar 4.2 *Scatter plot* antara Angka Harapan Hidup dengan Persentase Bayi Berumur 0-11 Bulan yang diberi ASI 4-6 Bulan

Gambar 4.2 menunjukkan bahwa bentuk pola hubungan antara variabel respon Angka Harapan Hidup (y) dan variabel prediktor persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan (x_2) cenderung tidak mengikuti bentuk pola, sehingga dapat dimodelkan secara nonparametrik. Bentuk pola hubungan antara Angka Harapan Hidup (y) dengan persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan (x_2) cenderung mengalami perubahan perilaku pola data yang berulang, sehingga dimodelkan dengan fungsi Deret Fourier.



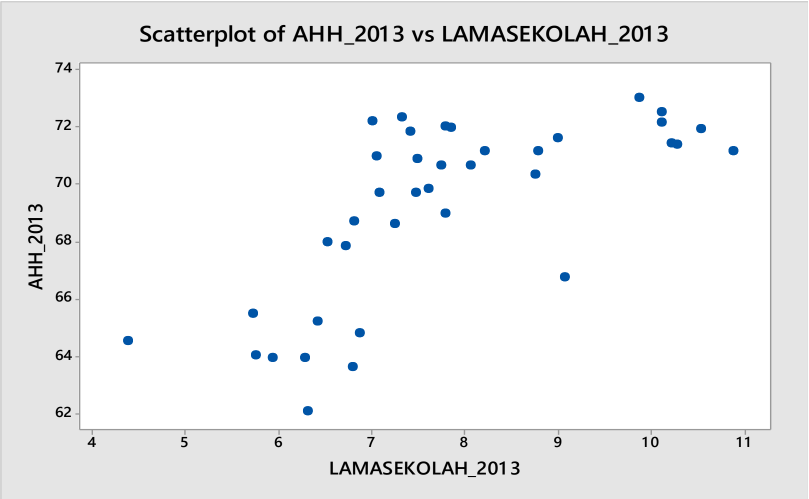
Gambar 4.3 *Scatter plot* antara Angka Harapan Hidup dengan Tingat Partisipasi Angkatan Kerja

Gambar 4.2 menunjukkan bahwa bentuk pola hubungan antara variabel respon Angka Harapan Hidup (y) dan variabel prediktor Tingkat Partisipasi Angkatan Kerja (x_3) tidak diketahui bentuk pola hubungannya, sehingga dapat dimodelkan secara nonparametrik. Bentuk pola hubungan antara Angka Harapan Hidup (y) dengan Tingkat Partisipasi Angkatan Kerja (x_3) cenderung mengalami perubahan perilaku pola data yang berulang, sehingga dapat dimodelkan dengan fungsi Deret Fourier.



Gambar 4.4 *Scatter plot* antara Angka Harapan Hidup dengan Angka Melek Huruf

Gambar 4.4 menunjukkan bahwa bentuk pola hubungan antara variabel respon Angka Harapan Hidup (y) dan variabel prediktor Angka Melek Huruf (x_4) tidak diketahui bentuk pola hubungannya, sehingga dapat dimodelkan secara nonparametrik. Bentuk pola hubungan antara Angka Harapan Hidup (y) dan variabel prediktor Angka Melek Huruf (x_4) cenderung mengalami perubahan perilaku pada titik sekitar 85. Terlihat bahwa pola data pada interval sebelum 85 cenderung turun sementara pada interval setelah 85 cenderung naik, sehingga dapat dimodelkan dengan fungsi Spline *Truncated*.



Gambar 4.5 *Scatter plot* antara Angka Harapan Hidup dengan Rata-Rata Lama Sekolah

Gambar 4.5 menunjukkan bahwa bentuk pola hubungan antara variabel respon Angka Harapan Hidup (y) dan variabel prediktor rata-rata lama sekolah (x_5) tidak diketahui bentuk pola hubungannya, sehingga dapat dimodelkan secara nonparametrik. Bentuk pola hubungan antara Angka Harapan Hidup (y) dengan rata-rata lama sekolah (x_5) cenderung mengalami perubahan perilaku pada titik sekitar 8, 9 dan 10. Terlihat bahwa pola data pada interval sebelum 8 cenderung naik sementara pada interval setelah 8 cenderung turun, kemudian pada interval setelah 9 cenderung naik dan pada interval setelah 10 cenderung turun, sehingga dapat dimodelkan dengan fungsi Spline *Truncated*.

Dalam menentuan variabel-variabel mana yang didekati dengan Spline *Truncated* dan Deret Fourier selain menggunakan *scatter plot* juga dapat dilakukan dengan memeriksa nilai GCV dari masing-masing variabel prediktor menggunakan persamaan (2.18). Berikut ini adalah hasil perhitungan nilai GCV untuk semua kemungkinan model campuran Spline *Truncated* dan Deret Fourier.

Tabel 4.2 GCV dari Model dengan Satu Komponen Spline *Truncated* dan Tiga Komponen Deret Fourier

No	Variabel		GCV
	Spline *Truncated*	Deret Fourier	
1	x_2	x_3, x_4, x_5	0,069865*
2	x_3	x_2, x_4, x_5	0,074269
3	x_4	x_2, x_3, x_5	0,089038
4	x_5	x_2, x_3, x_4	0,083343

Berdasarkan Tabel 4.2, model regresi semiparametrik campuran dengan satu komponen Spline *Truncated* dan tiga komponen Deret Fourier diperoleh nilai GCV minimum 0,069865 dengan komponen Spline *Truncated* adalah variabel dan rata-rata lama sekolah dan komponen Deret Fourier adalah variabel Angka Melek Huruf, persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan dan Tingkat Partisipasi Angkatan Kerja.

Tabel 4.3 GCV dari Model dengan Dua Komponen Spline *Truncated* dan Dua Deret Fourier Komponen

No	Variabel		GCV
	Spline *Truncated*	Deret Fourier	
1	x_4, x_5	x_2, x_3	0,0516573*
2	x_3, x_5	x_2, x_4	0,0527419
3	x_3, x_4	x_2, x_5	0,0763684
4	x_2, x_5	x_3, x_4	0,0532863
5	x_2, x_4	x_3, x_5	0,0684562
6	x_2, x_3	x_4, x_5	0,0517909

Berdasarkan Tabel 4.3, model regresi semiparametrik campuran dengan dua komponen Spline *Truncated* dan dua komponen Deret Fourier diperoleh nilai GCV minimum 0,0516573, dengan komponen Spline *Truncated* adalah variabel Angka Melek Huruf dan variabel Tingkat Partisipasi Angkatan Kerja dan komponen Deret Fourier adalah variabel persentase banyaknya bayi yang diberi ASI dan rata-rata lama sekolah.

Tabel 4.4 GCV dari Model dengan Tiga Komponen Spline *Truncated* dan Satu Komponen Deret Fourier

No	Variabel		GCV
	Spline *Truncated*	Deret Fourier	
1	x_3, x_4, x_5	x_2	0,0538918
2	x_2, x_4, x_5	x_3	0,053148*
3	x_2, x_3, x_5	x_4	0,0535301
4	x_2, x_3, x_4	x_5	0,0640583

Berdasarkan Tabel 4.3, model regresi semipametrik campuran dengan tiga komponen Spline *Truncated* dan satu komponen Deret Fourier diperoleh nilai GCV minimum 0,053148, dengan komponen Spline *Truncated* adalah variabel persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan, Angka Melek Huruf dan rata-rata lama sekolah dan Deret Fourier adalah variabel Tingkat Partisipasi Angkatan Kerja.

Berdasarkan Tabel 4.2, 4.3 dan 4.4, diperoleh nilai GCV minimum 0,0516573. Dengan demikian data Angka Harapan Hidup di Jawa Timur tahun 2013 dapat didekati dengan regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dimana terdapat satu variabel yang didekati dengan parametrik linier yaitu Angka Kematian Bayi, dua variabel prediktor yang didekati dengan Spline *Truncated* yaitu variabel Angka Melek Huruf dan variabel Tingkat Partisipasi Angkatan Kerja, serta dua variabel prediktor yang didekati dengan Deret Fourier yaitu variabel persentase bayi yang diberi ASI dan variabel rata-rata lama sekolah.

Variabel yang didekati dengan parametrik linear disimbolkan dengan x, variabel yang didekati dengan fungsi Spline *Truncated* disimbolkan dengan t, dan variabel yang didekati dengan fungsi Deret Fourier disimbolkan dengan z. Daftar keterangan lengkap dari masing-masing variabel prediktor dapat dilihat pada Tabel 4.5.

Tabel 4.5 Komponen Parametrik dan Nonparametrik

No	Variabel Prediktor	Kurva Regresi	Pendekatan yang digunakan	Simbol Variabel
1	Angka Kematian Bayi	Parametrik	Parametrik Linear	x
2	Angka Melek Huruf	Nonparametrik	Fungsi Spline *Truncated*	t_1
3	Rata-rata lama sekolah		Fungsi Spline *Truncated*	t_2
4	Persentase Bayi yang diberi ASI		Fungsi Deret Fourier	z_1
5	Tingkat Partisipasi Angkatan Kerja		Fungsi Deret Fourier	z_2

4.3.2 Model Umum Regresi Campuran Semiparametrik Spline *Truncated* dan Deret Forier

Berdasarkan Tabel 4.5 menunjukkan bahwa terdapat satu variabel yang didekati dengan parametrik linear, dua variabel yang didekati dengan fungsi Spline *Truncated* dan dua variabel yang didekati dengan fungsi Deret Fourier, sehingga model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier pada data Angka Harapan Hidup dengan *m* titik knot dan *K* osilasi dapat ditulis sebagai berikut :

$$y_i = \mu(x_i, t_{1i}, t_{2i}, z_{1i}, z_{2i}) + \varepsilon_i$$
$$= f(x_i) + g_1(t_{1i}) + g_2(t_{2i}) + h_1(z_{1i}) + h_2(z_{2i}) + \varepsilon_i$$

Kemudian $f(x_i)$, $g_1(t_{1i})$, $g_2(t_{2i})$, $h_1(z_{1i})$ dan $h_2(z_{2i})$ dihampiri dengan fungsi sebagai berikut :

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \ldots + \alpha_{1m}(t_{1i} - K_{1m}) + \\ & \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \alpha_{22}(t_{2i} - K_{22})_+ + \ldots + \alpha_{2m}(t_{2i} - K_{2m})_+ + \\ & b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + \cdots + a_{1K}\cos K z_{1i} + \\ & b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + \cdots + a_{2K}\cos K z_{2i} + \varepsilon_i . \end{aligned}$$

4.3.3 Pemilihan Titik Knot, Parameter Penghalus dan Osilasi *k* Optimum

4.3.3.1 Model dengan Komponen Spline *Truncated* 1 Titik Knot dan Deret Fourier Osilasi $k = 1$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing satu titik knot pada komponen Spline *Truncated* dan osilasi $k = 1$ pada masing-masing komponen Deret Fourier.

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \\ & b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + \varepsilon_i \end{aligned}$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan satu titik knot dan osilasi $k = 1$ dapat dilihat pada Tabel 4.6

Tabel 4.6 Nilai GCV pada Model dengan Satu Titik Knot dan Osilasi $k = 1$

No	Titik Knot		Lambda		GCV
	K_{11}	K_{21}	(λ_1)	(λ_2)	
1	97,2873	10,64	0,9	1	0,0789276*
2	96,1746	10,39	1	0,07	0,0853312
3	95,0619	10,14	0,8	0,04	0,0874518
4	93,9492	9,89	0,7	0,9	0,0913926
5	88,3858	8,64	0,6	0,8	0,0915393

Berdasarkan Tabel 4.6 diperoleh GCV minimum adalah 0,0765056. GCV minimum tersebut diperoleh ketika lokasi titik knot (K_{11}, K_{21}) berada pada 97,2873, 10,64 dengan lambda (λ_1, λ_2) yaitu 0,9 dan 1.

4.3.3.2 Model dengan Komponen Spline *Truncated* 2 Titik Knot dan Deret Fourier Osilasi $k=1$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing dua titik knot pada komponen Spline *Truncated* dan osilasi $k=1$ pada masing-masing komponen Deret Fourier.

$$y_i = \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ +$$
$$\alpha_{22}(t_{2i} - K_{22})_+ + b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + \varepsilon_i .$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k=1$ dapat dilihat pada Tabel 4.7.

Tabel 4.7 Nilai GCV pada Model dengan Dua Titik Knot dan Osilasi $k=1$

No	Titik Knot		Lambda (λ)	GCV
1	$K_{11}=90{,}1343$	$K_{21}=9{,}03286$	$\lambda_1=0{,}0004$	0,0603866*
	$K_{12}=94{,}2671$	$K_{22}=9{,}96143$	$\lambda_2=1$	
2	$K_{11}=88{,}7567$	$K_{21}=8{,}72333$	$\lambda_1=0{,}00007$	0,0606185
	$K_{12}=94{,}2671$	$K_{22}=9{,}96143$	$\lambda_2=0{,}9$	
3	$K_{11}=91{,}5119$	$K_{21}=9{,}34238$	$\lambda_1=0{,}0007$	0,0618078
	$K_{12}=94{,}2671$	$K_{22}=9{,}96143$	$\lambda_2=0{,}8$	
4	$K_{11}=92{,}8895$	$K_{21}=9{,}65190$	$\lambda_1=0{,}0001$	0,0637706
	$K_{12}=94{,}2671$	$K_{22}=9{,}96143$	$\lambda_2=0{,}7$	
5	$K_{11}=90{,}1343$	$K_{21}=9{,}03286$	$\lambda_1=1$	0,0638043
	$K_{12}=92{,}8895$	$K_{22}=9{,}65190$	$\lambda_2=0{,}6$	

Berdasarkan Tabel 4.7 diperoleh GCV minimum adalah 0,0556817. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 90,1343, 94,2671, 9,03286, 9,96143 dengan lambda (λ_1, λ_2) yaitu 0,0004 dan 1.

4.3.3.3 Model dengan Komponen Spline *Truncated* 3 Titik Knot dan Deret Fourier Osilasi $k=1$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing tiga titik knot pada komponen Spline *Truncated* dan osilasi $k=1$ pada masing-masing komponen Deret Fourier.

$$y_i = \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \alpha_{13}(t_{1i} - K_{13}) +$$
$$\phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \alpha_{22}(t_{2i} - K_{22})_+ + \alpha_{23}(t_{2i} - K_{23})_+ +$$
$$b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + \varepsilon_i .$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 1$ dapat dilihat pada Tabel 4.8.

Tabel 4.8 Nilai GCV pada Model dengan Tiga Titik Knot dan Osilasi $k = 1$

No	Titik Knot		Lambda (λ)	GCV
1	$K_{11} = 85,7431$ $K_{12} = 87,5513$ $K_{13} = 94,9756$	$K_{21} = 8,04625$ $K_{22} = 8,4525$ $K_{23} = 10,0775$	$\lambda_1 = 0,00007$ $\lambda_2 = 1$	0,0546567*
2	$K_{11} = 83,9350$ $K_{12} = 87,5513$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,4525$ $K_{23} = 10,0775$	$\lambda_1 = 1$ $\lambda_2 = 0,9$	0,0555913
3	$K_{11} = 89,3594$ $K_{12} = 91,1675$ $K_{13} = 92,9756$	$K_{21} = 8,85875$ $K_{22} = 9,265$ $K_{23} = 9,67125$	$\lambda_1 = 0,9$ $\lambda_2 = 0,8$	0,0569536
4	$K_{11} = 83,9350$ $K_{12} = 89,5513$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,85875$ $K_{23} = 10,07750$	$\lambda_1 = 0,0001$ $\lambda_2 = 0,7$	0,0598037
5	$K_{11} = 85,7431$ $K_{12} = 87,5513$ $K_{13} = 92,9756$	$K_{21} = 8,04625$ $K_{22} = 8,45250$ $K_{23} = 9,67125$	$\lambda_1 = 0,8$ $\lambda_2 = 0,6$	0,0601007

Berdasarkan Tabel 4.8 diperoleh GCV minimum adalah 0,0546567. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 85,7431, 87,5513, 94,7837, 8,04625, 8,4525, 10,0775 dan lambda (λ_1, λ_2) yaitu 0,00007 dan 1.

4.3.3.4 Model dengan Komponen Spline *Truncated* 1 Titik Knot dan Deret Fourier Osilasi $k = 2$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing satu titik knot pada komponen Spline *Truncated* dan osilasi $k = 2$ pada masing-masing komponen Deret Fourier.

$$y_i = \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ +$$
$$b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} +$$
$$b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + \varepsilon_i .$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan satu titik knot dan osilasi $k = 2$ dapat dilihat pada Tabel 4.9.

Tabel 4.9 Nilai GCV pada Model dengan Satu Titik Knot dan Osilasi $k = 2$

No	Titik Knot		Lambda		GCV
	K_{11}	K_{21}	λ_1	λ_2	
1	97,2873	10,64	0,9	1	0,0789381*
2	96,1746	10,39	1	0,2	0,0854198
3	95,0619	10,14	0,8	0,04	0,0876832
4	88,3858	8,64	0,7	0,9	0,0914063
5	87,2731	8,89	0,6	0,8	0,0915547

Berdasarkan Tabel 4.9 diperoleh GCV minimum adalah 0,0789381. GCV minimum tersebut diperoleh ketika lokasi titik knot (K_{11}, K_{21}) berada pada 97,2873, 10,64 dengan lambda (λ_1, λ_2) yaitu 0,9 dan 1.

4.3.3.5 Model dengan Komponen Spline *Truncated* 2 Titik Knot dan Deret Fourier Osilasi $k = 2$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing dua titik knot pada komponen Spline *Truncated* dan osilasi $k = 2$ pada masing-masing komponen Deret Fourier.

$$y_i = \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ +$$
$$\alpha_{22}(t_{2i} - K_{22})_+ + b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + b_2 z_{2i} + a_{02} +$$
$$a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + \varepsilon_i .$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 2$ dapat dilihat pada Tabel 4.10.

Tabel 4.10 Nilai GCV pada Model dengan Dua Titik Knot dan Osilasi $k = 2$

No	Titik Knot		Lambda	GCV
1	$K_{11} = 90{,}1343$	$K_{21} = 9{,}03286$	$\lambda_1 = 7 \times 10^{-8}$	0,0576090*
	$K_{12} = 94{,}2671$	$K_{22} = 9.96143$	$\lambda_2 = 1$	

Tabel 4.10 Nilai GCV pada Model dengan Dua Titik Knot dan Osilasi $k = 2$ (Lanjutan)

No	Titik Knot		Lambda	GCV
2	$K_{11} = 91,5119$ $K_{12} = 94,2671$	$K_{21} = 9,34238$ $K_{22} = 9.96143$	$\lambda_1 = 4\times10^{-8}$ $\lambda_2 = 0,9$	0,0579542
3	$K_{11} = 88,7567$ $K_{21} = 94,2671$	$K_{21} = 8,72333$ $K_{22} = 9.96143$	$\lambda_1 = 1\times10^{-7}$ $\lambda_2 = 0,8$	0,0584851
4	$K_{11} = 92,8895$ $K_{21} = 94,2671$	$K_{21} = 9,65190$ $K_{22} = 9,96143$	$\lambda_1 = 1\times10^{-8}$ $\lambda_2 = 0,7$	0,0597267
5	$K_{11} = 91,5119$ $K_{12} = 92,8895$	$K_{21} = 9,34238$ $K_{22} = 9,65190$	$\lambda_1 = 4\times10^{-7}$ $\lambda_2 = 0,6$	0,0634929

Berdasarkan Tabel 4.10 diperoleh GCV minimum adalah 0,0561610. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 90,1343, 94,2671, 9,03286, 9,96143 dengan lambda (λ_1, λ_2) yaitu 7×10^{-8} dan 1.

4.3.3.6 Model dengan Komponen Spline *Truncated* 3 Titik Knot dan Deret Fourier Osilasi $k = 2$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing tiga titik knot pada komponen Spline *Truncated* dan osilasi $k = 2$ pada masing-masing komponen Deret Fourier.

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \alpha_{13}(t_{1i} - K_{13}) + \\ & \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \alpha_{22}(t_{2i} - K_{22})_+ + \alpha_{23}(t_{2i} - K_{23})_+ + b_1 z_{1i} + a_{01} + \\ & a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + \varepsilon_i . \end{aligned}$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 2$ dapat dilihat pada Tabel 4.11.

Tabel 4.11 Nilai GCV pada Model dengan Tiga Titik Knot dan Osilasi $k = 2$

No	Titik Knot		Lambda (λ)	GCV
1	$K_{11} = 85,7431$ $K_{12} = 87,5513$ $K_{13} = 94,7837$	$K_{21} = 8,04625$ $K_{22} = 8,45250$ $K_{23} = 10,0775$	$\lambda_1 = 7\times10^{-8}$ $\lambda_2 = 1$	0,0516263*

Tabel 4.11 Nilai GCV pada Model dengan Tiga Titik Knot dan Osilasi $k = 2$ (Lanjutan)

No	Titik Knot		Lambda (λ)	GCV
2	$K_{11} = 83,935$ $K_{12} = 87,5513$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,4525$ $K_{23} = 10,0775$	$\lambda_1 = 1\times10^{-7}$ $\lambda_2 = 0,9$	0,0538531
3	$K_{11} = 89,3594$ $K_{12} = 91,1675$ $K_{13} = 92,9756$	$K_{21} = 8,85875$ $K_{22} = 9,265$ $K_{23} = 9,67125$	$\lambda_1 = 1$ $\lambda_2 = 0,9$	0,0569596
4	$K_{11} = 83,9350$ $K_{12} = 89,3594$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,85875$ $K_{23} = 10,0775$	$\lambda_1 = 4\times10^{-8}$ $\lambda_2 = 0,7$	0,0573984
5	$K_{11} = 71,2781$ $K_{12} = 89,3594$ $K_{13} = 94,7837$	$K_{21} = 4,79625$ $K_{22} = 8,85875$ $K_{23} = 10,0775$	$\lambda_1 = 1\times10^{-8}$ $\lambda_2 = 0,6$	0,0587330

Berdasarkan Tabel 4.11 diperoleh GCV minimum adalah 0,516263. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 85,7431, 87,5513, 94,7837, 8,04625, 8,4525, 10,0775 dengan lambda (λ_1, λ_2) yaitu 7×10^{-8} dan 1.

4.3.3.7 Model dengan Komponen Spline *Truncated* 1 Titik Knot dan Deret Fourier Osilasi $k = 3$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing satu titik knot pada komponen Spline *Truncated* dan osilasi $k = 3$ pada masing-masing komponen Deret Fourier.

$$
\begin{aligned}
y_i = {} & \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \\
& b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + a_{13}\cos 3z_{1i} + \\
& b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + a_{23}\cos 3z_{2i} + \varepsilon_i
\end{aligned}
$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 3$ dapat dilihat pada Tabel 4.12.

Tabel 4.12 Nilai GCV pada Model dengan Satu Titik Knot dan osilasi $k = 3$

No	Titik Knot		Lambda		GCV
	K_{11}	K_{21}	λ_1	λ_2	
1	97,2873	10,64	0,9	1	0,0789485
2	96,1746	10,39	1	0,9	0,0854457
3	95,0619	10,14	0,8	0,7	0,0878278
4	88,3858	8,64	0,7	0,8	0,0914284
5	8,64	8,89	0,6	0,7	0,0915825

Berdasarkan Tabel 4.12 diperoleh GCV minimum adalah 0,0789485. GCV minimum tersebut diperoleh ketika lokasi titik knot (K_{11}, K_{21}) berada pada 97,2873, 10,64 dengan lambda (λ_1, λ_2) yaitu 0,9 dan 1.

4.3.3.8 Model dengan Komponen Spline *Truncated* 2 Titik Knot dan Deret Fourier Osilasi $k = 3$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing dua titik knot pada komponen Spline *Truncated* dan osilasi $k = 3$ pada masing-masing komponen Deret Fourier.

$$y_i = \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ +$$
$$\alpha_{22}(t_{2i} - K_{22})_+ + b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + a_{13}\cos 3z_{1i} +$$
$$b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + a_{23}\cos 3z_{2i} + \varepsilon_i .$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 3$ dapat dilihat pada Tabel 4.13.

Tabel 4.13 Nilai GCV pada Model dengan Dua Titik Knot dan Osilasi $k = 3$

No	Titik Knot		Lambda	GCV
1	$K_{11} = 90{,}1343$ $K_{12} = 94{,}2671$	$K_{21} = 9{,}03286$ $K_{22} = 9{,}96143$	$\lambda_1 = 4 \times 10^{-7}$ $\lambda_1 = 1$	0,0577269*
2	$K_{11} = 88{,}7567$ $K_{12} = 94{,}2671$	$K_{21} = 9{,}34238$ $K_{22} = 9{,}96143$	$\lambda_1 = 7 \times 10^{-8}$ $\lambda_1 = 0{,}9$	0,0580481
3	$K_{11} = 92{,}8895$ $K_{12} = 94{,}2671$	$K_{21} = 8{,}72333$ $K_{22} = 9{,}96143$	$\lambda_1 = 7 \times 10^{-7}$ $\lambda_1 = 0{,}8$	0,0586129
4	$K_{11} = 92{,}8895$ $K_{12} = 94{,}2671$	$K_{21} = 9{,}65190$ $K_{22} = 9{,}96143$	$\lambda_1 = 4 \times 10^{-8}$ $\lambda_1 = 0{,}7$	0,0597626
5	$K_{11} = 91{,}5119$ $K_{12} = 92{,}8895$	$K_{21} = 9{,}03286$ $K_{22} = 9{,}65190$	$\lambda_1 = 1 \times 10^{-8}$ $\lambda_1 = 0{,}6$	0,0627881

Berdasarkan Tabel 4.13 diperoleh GCV minimum adalah 0,0577269. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 90,1343, 94,2671, 9,03286, 9,96143 dengan lambda (λ_1, λ_2) yaitu 4×10^{-7} dan 1.

4.3.3.9 Model dengan Komponen Spline *Truncated* 3 Titik Knot dan Deret Fourier Osilasi $k = 3$

Berikut ini model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, dengan masing-masing tiga titik knot pada komponen Spline *Truncated* dan osilasi $k = 3$ pada masing-masing komponen Deret Fourier.

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_i + \phi_{11} t_{1i} + \alpha_{11}(t_{1i} - K_{11})_+ + \alpha_{12}(t_{1i} - K_{12})_+ + \alpha_{13}(t_{1i} - K_{13}) + \\ & \phi_{21} t_{2i} + \alpha_{21}(t_{2i} - K_{21})_+ + \alpha_{22}(t_{2i} - K_{22})_+ + \alpha_{23}(t_{2i} - K_{23})_+ + \\ & b_1 z_{1i} + a_{01} + a_{11}\cos z_{1i} + a_{12}\cos 2z_{1i} + a_{13}\cos 3z_{1i} + \\ & b_2 z_{2i} + a_{02} + a_{21}\cos z_{2i} + a_{22}\cos 2z_{2i} + a_{23}\cos 3z_{2i} + \varepsilon_i . \end{aligned}$$

Lokasi titik-titik knot, lambda dan nilai GCV yang diperoleh dari model dengan dua titik knot dan osilasi $k = 3$ dapat dilihat pada Tabel 4.14

Tabel 4.14 Nilai GCV pada Model dengan Tiga Titik Knot dan Osilasi $k = 3$

No	Titik Knot		Lambda	GCV
1	$K_{11} = 85,7431$ $K_{12} = 87,5513$ $K_{13} = 94,7837$	$K_{21} = 8,04625$ $K_{22} = 8,4525$ $K_{23} = 10,0775$	$\lambda_1 = 4 \times 10^{-7}$ $\lambda_1 = 1$	0,0516573*
2	$K_{11} = 89,3594$ $K_{12} = 87,5513$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,4525$ $K_{23} = 10,0775$	$\lambda_1 = 7 \times 10^{-7}$ $\lambda_1 = 0,9$	0,0537743
3	$K_{11} = 89,3594$ $K_{12} = 91,1675$ $K_{13} = 92,9756$	$K_{21} = 8,85875$ $K_{22} = 9,265$ $K_{23} = 9,67125$	$\lambda_1 = 1$ $\lambda_1 = 0,8$	0,0569700
4	$K_{11} = 83,935$ $K_{12} = 89,3594$ $K_{13} = 94,7837$	$K_{21} = 7,64$ $K_{22} = 8,85875$ $K_{23} = 10,0775$	$\lambda_1 = 1 \times 10^{-6}$ $\lambda_1 = 0,7$	0,0574898
5	$K_{11} = 73,0863$ $K_{12} = 89,3594$ $K_{13} = 94,7837$	$K_{21} = 5,2025$ $K_{22} = 8,85875$ $K_{23} = 10,0775$	$\lambda_1 = 1 \times 10^{-7}$ $\lambda_1 = 0,6$	0,0586574

Berdasarkan Tabel 4.14 diperoleh GCV minimum adalah 0,0516573. GCV minimum tersebut diperoleh ketika lokasi titik knot $(K_{11}, K_{12}, K_{21}, K_{22})$ berada pada 85,7431, 87,5513, 94,7837, 8,04625, 8,4525, 10,0075 dengan lambda (λ_1, λ_2) yaitu 4×10^{-7} dan 1.

Perbandingan nilai GCV minimum antara model dengan satu titik knot dan osilasi $k=1$, model dengan satu titik knot dan osilasi $k=2$, model dengan satu titik knot dan osilasi $k=3$, model dengan dua titik knot dan osilasi $k=1$, model dengan dua titik knot dan osilasi $k=2$, model dengan dua titik knot dan osilasi $k=3$, model dengan tiga titik knot dan osilasi $k=1$, model dengan tiga titik knot dan osilasi $k=2$, dan model dengan tiga titik knot dan osilasi $k=3$ ditunjukkan oleh tabel 4.15.

Tabel 4.15 Perbandingan Nilai GCV Minumum

No	Model	GCV
1	1 titik knot dan osilasi $k=1$	0,0789276
2	1 titik knot dan osilasi $k=2$	0,0789381
3	1 titik knot dan osilasi $k=3$	0,0789485
4	2 titik knot dan osilasi $k=1$	0,0603866
5	2 titik knot dan osilasi $k=2$	0,057609
6	2 titik knot dan osilasi $k=3$	0,0577269
7	3 titik knot dan osilasi $k=1$	0,0546567
8	3 titik knot dan osilasi $k=2$	0,0516263
9	3 titik knot dan osilasi $k=3$	0,0516573

Berdasarkan Tabel 4.15 terlihat bahwa model regresi yang memiliki GCV minimum adalah model dengan tiga titik knot dibandingkan dengan satu dan dua titik knot. Meskipun berdasarkan tabel tersebut diperoleh nilai GCV minimum adalah model dengan tiga titik knot dan $k=2$ yaitu sebesar 0,0516263. Namun, karena perbedaan nilai GCV pada tiga titik knot dengan osilasi $k=1, k=2$ dan $k=3$ tidak signifikan berbeda dan selain itu karena pertimbangan parsimoni model maka dipilih model terbaik dengan tiga titik knot dan osilasi $k=1$.

4.3.4 Estimasi Model Regresi Semiparametrik Campuran Spline *Truncated* dan Deret Fourier

Model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier terbaik diperoleh ketika model menggunakan banyaknya titik knot, lokasi titik-titik knot dan lambda yang optimum. Berdasarkan pemilihan titik-titik knot dan lambda, model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier terbaik diperoleh dengan tiga titik knot dan osilasi $k = 1$.

Estimasi parameter dari model semiparametrik campuran Spline *Truncated* dan Deret fourier dengan tiga titik knot dan osilasi $k = 1$, ditunjukkan pada Tabel 4.16.

Tabel 4.16 Estimasi Parameter dari Model Semiparametrik Campuran Spline *Truncated* dan Deret Fourier dengan 3 Titik Knot dan Osilasi $k = 1$.

Variabel	Parameter	Estimasi
x	β_0	-0,00000000000568583
	β_1	-0,250025
t_1	ϕ_{11}	-0,0315882
	α_{11}	-0,182079
	α_{12}	0,277273
	α_{13}	-0,508491
t_2	ϕ_{21}	0,176008
	α_{21}	-2,05685
	α_{22}	3,30714
	α_{23}	-2,09454
z_1	b_1	0,162176
	a_{01}	74,4552
	a_{11}	2,23833
z_2	b_1	0,0278582
	a_{02}	74,4552
	a_{21}	-0,00201112

Berdasarkan hasil estimasi pada Tabel 4.16 dan Tabel 4.12 maka model regresi semiparametrik campuran Spline *Truncated* dan Deret fourier dengan 3 titik knot dan osilasi $k = 1$ dapat ditulis menjadi :

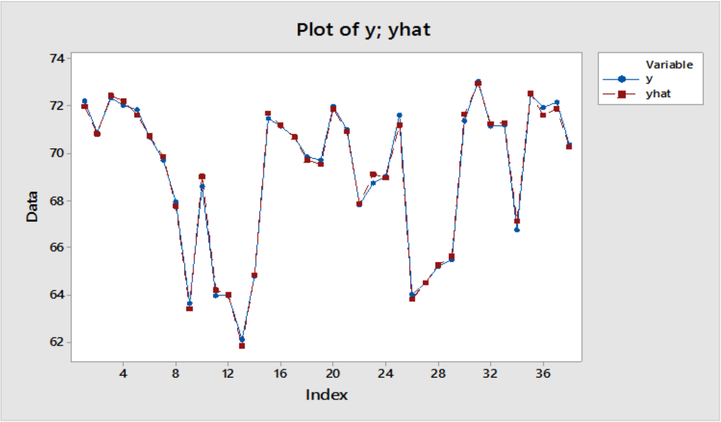
$$\hat{y}_i = 148.9104 - 0,250025 x_i - 0,0315882 t_{1i} - 0,182079(t_{1i} - 85,7431)_+ +$$
$$0,277273(t_{1i} - 87,5513)_+ - 0,508491(t_{1i} - 94,7837)_+ + 0,176008 t_{2i} +$$
$$-2,05685(t_{2i} - 8,04625)_+ + 3,30714(t_{2i} - 8,4525)_+ - 2,09454(t_{2i} - 10,0775)_+ +$$
$$0,162176 z_{1i} + 2,23833 \cos z_{1i} + 0,0278582 z_{2i} - 0,00201112 \cos z_{2i}$$

Perbandingan nilai Angka Harapan Hidup (y) dan $\hat{y}$ ditunjukkan oleh Tabel 4.17.

Tabel 4.17 Perbandingan antara y dan $\hat{y}$

No	y	$\hat{y}$	No	y	$\hat{y}$
1	72.18	71.97682	20	71,96	71,85316
2	70.85	70.78455	21	70,97	70,90683
3	72,33	72,40115	22	67,81	67,83652
4	72,02	72,18239	23	68,71	69,0821
5	71,8	71,60996	24	68,98	68,93494
6	70,65	70,71313	25	71,57	71,19129
7	69,7	69,8278	26	64,02	63,83304
8	67,95	67,75799	27	64,52	64,50017
9	63,64	63,40865	28	65,19	65,25421
10	68,58	69,01682	29	65,49	65,63421
11	63,95	64,19676	30	71,36	71,64537
12	63,95	63,98739	31	73	72,92531
13	62,1	61,85219	32	71,14	71,20413
14	64,81	64,82368	33	71,16	71,24814
15	71,43	71,68051	34	66,75	67,09958
16	71,13	71,18682	35	72,48	72,50979
17	70,64	70,66676	36	71,89	71,5859
18	69,82	69,69965	37	72,13	71,85784
19	69,68	69,52753	38	70,32	70,25694

Plot antara y dan $\hat{y}$ untuk 38 observasi dapat dilihat pada Gambar 4.6.



Gambar 4.6 Plot data y dan $\hat{y}$

Berdasarkan Tabel 4.17 dan Gambar 4.6 terlihat bahwa grafik $\hat{y}$ mendekati grafik data sebenarnya. Dari perhitungan, model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier dengan 3 titik knot dan osilasi $k=1$ mempunyai nilai R^2 sebesar 99,62%, yang berarti variabel Angka Kematian Bayi, persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan, Tingkat Partisipasi Angkatan Kerja, Angka Melek Huruf, dan rata-rata lama sekolah mampu menjelaskan variabel respon Angka Harapan Hidup sebesar 99,62%.

4.3.5 Interpretasi Model Regresi Semiparametrik Campuran Spline *Truncated* dan Deret Fourier

Berikut ini adalah model masing-masing kelompok data dan interpretasinya :

1) Model untuk variabel Angka Kematian Bayi

Mengasumsikan bahwa variabel independen lainnya selain Angka Kematian Bayi konstan, maka diperoleh model umumnya :

$$\hat{y}_i = -0,250025 x_i + c_1$$

dengan,

$$\begin{aligned} c_1 = {} & 148.9104_i - 0,0315882 t_{1i} - 0,182079(t_{1i} - 85,7431)_+ + \\ & 0,277273(t_{1i} - 87,5513)_+ - 0,508491(t_{1i} - 94,7837)_+ + 0,176008 t_{2i} + \\ & -2,05685(t_{2i} - 8,04625)_+ + 3,30714(t_{2i} - 8,4525)_+ - 2,09454(t_{2i} - 10,0775)_+ + \\ & 0,162176 z_{1i} + 2,23833 \cos z_{1i} + 0,0278582 z_{2i} - 0,00201112 \cos z_{2i} \end{aligned}$$

Koefisien regresi variabel Angka Kematian Bayi sebesar $-0,250025$, artinya jika variabel independen lainnya konstan dan variabel Angka Kematian Bayi mengalami kenaikan sebesar 1%, maka Angka Harapan Hidup akan mengalami penurunan sebesar $0,250025$ tahun. Koefisien bernilai negatif artinya terjadi hubungan negatif antara variabel Angka Kematian Bayi dan Angka Harapan Hidup, semakin tinggi Angka Kematian Bayi maka Angka Harapan Hidup semakin menurun.

2) Model untuk variabel Angka Melek Huruf

Mengasumsikan bahwa variabel independen lainnya selain Angka Melek Huruf konstan, maka diperoleh model umumnya :

$$\hat{y}_i =_i -0,0315882t_{1i} - 0,182079(t_{1i} - 85,7431)_+ + 0,277273(t_{1i} - 87,5513)_+ + \\ -0,508491(t_{1i} - 94,7837)_+ + c_2$$

dengan,

$$c_2 = 148.9104 - 0,250025x_i + 0,176008t_{2i} - 2,05685(t_{2i} - 8,04625)_+ + \\ 3,30714(t_{2i} - 8,4525)_+ - 2,09454(t_{2i} - 10,0775)_+ + 0,162176z_{1i} + \\ 2,23833\cos z_{1i} + 0,0278582z_{2i} - 0,00201112\cos z_{2i}$$

Model tersebut dapat diinterpretasi dengan menggunakan fungsi berikut :

$$\hat{y}_i = \begin{cases} -0,0315882t_{1i} + c_2, & t_{1i} \le 85,7431 \\ 15,612017 - 0,2136672t_{1i} + c_2, & 85,7431 < t_{1i} \le 87,5513 \\ -8,663583 + 0,0636058t_{1i} + c_2, & 87,5513 < t_{1i} \le 94,7837 \\ 39,53317 - 0,4448852t_{1i} + c_2, & t_{1i} > 94,7837, \end{cases}$$

a. Angka Melek Huruf Provinsi Jawa Timur yang kurang dari 85,7431% meliputi Kabupaten Jember, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Bojonogoro, Kabupaten Tuban, Kabupaten Bangkalan, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Sumenep, ketika Angka Melek Huruf dinaikkan sebesar 1%, maka Angka Harapan Hidup di wilayah-wilayah tersebut akan mengalami penurunan sebesar 0,0315882 tahun.
b. Angka Melek Huruf Provinsi Jawa Timur pada rentang nilai 85,7431% sampai 87,5513%, meliputi Kabupaten Lumajang dan Kabupaten Ngawi. Jika Angka Melek Huruf mengalami kenaikan sebesar 1%, maka Angka Harapan Hidup di wilayah-wilayah tersebut akan mengalami penurunan sebesar 0,2136672 tahun.
c. Angka Melek Huruf Provinsi Jawa Timur pada rentang nilai 87,5513% sampai 94,7837% meliputi Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Banyuwangi, Kabupaten Pasuruan, Kabupaten Mojekerto, Kabupaten Jombang, Kabupaten Nganjuk, Kabupaten Madiun,

Kabupaten Magetan, Kabupaten Lamongan, dan Kota Probolinggo, serta Kota Batu. Jika Angka Melek Huruf mengalami kenaikan sebesar 1%, maka Angka Harapan Hidup di wilayah-wilayah tersebut akan mengalami kenaikan sebesar 0,0636 tahun.

d. Angka Melek Huruf Provinsi Jawa Timur yang lebih besar dari 94,7837% meliputi Kabupaten Tulungagung, Kabupaten Sidoarjo, Kabupaten Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Kota Mojokerto, dan Kota Surabaya. Jika Angka Melek Huruf mengalami kenaikan sebesar 1%, maka Angka Harapan Hidup di wilayah-wilayah tersebut akan mengalami penurunan sebesar 0,4448 tahun.

3) Model untuk variabel rata-rata lama sekolah

Mengasumsikan bahwa variabel independen lainnya selain rata-rata lama sekolah adalah konstan, maka diperoleh model umumnya :

$$\hat{y}_i = 0{,}176008 t_{2i} - 2{,}05685(t_{2i} - 8{,}04625)_+ + 3{,}30714(t_{2i} - 8{,}4525)_+ + \\ -2{,}09454(t_{2i} - 10{,}0775)_+ + c_3$$

dengan,

$$c_3 = 148.9104 - 0{,}250025 x_i - 0{,}0315882 t_{1i} - 0{,}182079(t_{1i} - 85{,}7431)_+ + \\ 0{,}277273(t_{1i} - 87{,}5513)_+ - 0{,}508491(t_{1i} - 94{,}7837)_+ + \\ 0{,}162176 z_{1i} + 2{,}23833 \cos z_{1i} + 0{,}0278582 z_{2i} - 0{,}00201112 \cos z_{2i}$$

Model tersebut dapat diinterpretasi dengan menggunakan fungsi berikut :

$$\hat{y}_i = \begin{cases} 0{,}176008 t_{2i} + c_3, & t_{1i} \le 8{,}04625 \\ 16{,}5499 - 1{,}880842 t_{2i} + c_3, & 8{,}04625 < t_{1i} \le 8{,}4525 \\ -11{,}4037 + 1{,}42629 t_{2i} + c_3, & 8{,}4525 < t_{1i} \le 10{,}0775 \\ 9{,}704 - 0{,}668242 t_{2i} + c_3, & t_{1i} > 10{,}0775 \end{cases}$$

a. Rata-rata lama sekolah di Jawa Timur yang kurang dari 9,03286 tahun meliputi Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Pasuruan, Kabupaten Nganjuk, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Tuban,

Kabupaten Lamongan, Kabupaten Bangkalan, Kabupaten Sampang, Kabupaten Pamekasan, dan Kabupaten Sumenep. Jika rata-rata lama sekolah mengalami kenaikan sebesar 1 tahun, maka Angka Harapan Hidup di wilayah-wilayah tersebut akan mengalami kenaikan sebesar 0,176008 tahun.

b. Rata-rata lama sekolah di Jawa Timur yang berada pada rentang 8,04625 tahun sampai 8,4525 tahun adalah Kabupaten Mojokerto dan Kabupaten Jombang. Jika rata-rata lama sekolah mengalami kenaikan sebesar 1 tahun, maka Angka Harapan Hidup di wilayah tersebut akan mengalami penurunan sebesar 1,8808 tahun.

c. Rata-rata lama sekolah di Jawa Timur yang berada pada rentang nilai 8,4525 tahun sampai 10,0775 tahun meliputi Kabupaten Gresik, Kota Blitar, Kota Probolinggo, Kota Pasuruan, dan Kota Batu. Jika rata-rata lama sekolah di wilayah-wilayah tersebut dinaikkan sebesar 1 tahun, maka Angka Harapan Hidup akan mengalami kenaikan sebesar 1,42629 tahun.

d. Rata-rata lama sekolah di Jawa Timur yang lebih dari 10,0775 tahun meliputi Kabupaten Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu. Jika rata-rata lama sekolah mengalami kenaikan sebesar 1%, maka Angka Harapan Hidup akan mengalami penurunan sebesar 0,6682 tahun.

Pemodelan data Angka Harapan Hidup di Provinsi Jawa Timur tahun 2013 menggunakan regresi semiparametrik dengan estimator campuran Spline *Truncated* dan Deret Fourier menghasilkan R^2 sebesar 99,62%, yang berarti bahwa variabel-variabel prediktor mampu menjelaskan variasi dalam variabel respon Angka Harapan Hidup sebesar 99,62 %

Berdasarkan hasil interpretasi-interpretasi tersebut, interpretasi model untuk data Angka Melek Huruf kurang dari 85,7431, model untuk data Angka Melek Huruf pada rentang 85,7431-87,5513% dan model untuk data Angka Melek Huruf lebih besar dari 94,7837% tidak sesuai dengan teori, dimana ketika Angka Melek Huruf mengalami kenaikan, maka Angka Harapan Hidup mengalami penurunan. Serta interpretasi model untuk data rata-rata lama sekolah pada rentang 8,04625-

8,4525 tahun dan model untuk data rata-rata lama sekolah lebih besar dari 10,0775 tahun juga tidak sesuai dengan teori, dimana ketika rata-rata lama sekolah mengalami kenaikan, maka Angka Harapan Hidup mengalami penurunan.

Angka Melek Huruf dan rata-rata lama sekolah merupakan komponen dari pendidikan. Dimana pendidikan merupakan salah satu faktor yang memberikan sumbangan yang besar terhadap perkembangan kehidupan sosial dan ekonomi suatu wilayah. Sehingga, seharusnya semakin tinggi tingkat pendidikan disuatu wilayah, maka Angka Harapan Hidupnya pun juga mengalami peningkatan.

Angka Melek Huruf yang kurang dari 85,7431 % adalah wilayah tapal kupa. Wilayah-wilayah yang masuk dalam kategori wilayah tapal kedua merupakan wilayah-wilayah yang memiliki Angka Melek Huruf yang rendah dibandingkan dengan wilayah lainnya di Jawa Timur. Namun, ketika Angka Melek Hurufnya mengalami kenaikan, justru Angka Harapan Hidupnya mengalami penurunan. Ketidaksesuain teori tersebut dimungkinkan karena mata pencaharian di wilayah-wilayah tersebut sekitar 60% umumnya adalah petani dan nelayan. Dimana profesi tersebut tidak membutuhkan spesifikasi pendidikan yang tinggi.

Angka Melek Huruf yang berada pada rentang 85,7431% - 87,5513%, meliputi Kabupaten Lumajang dan Kabupaten Ngawi, serta Angka Melek Huruf yang berada pada rentang lebih besar dari 94,7837% meliputi Kabupaten Tulungagung, Kabupaten Sidoarjo, Kabupaten Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Kota Mojokerto, dan Kota Surabaya. Rentang Angka Melek huruf tersebut, sudah terbilang cukup baik. Selain itu, wilayah-wilayah tersebut merupakan wilayah-wilayah yang memiliki Angka Melek Huruf yang cukup tinggi. Namun, ketika Angka Melek Huruf mengalami kenaikan di wilayah tersebut, Angka Harapan Hidupnya justru mengalami penurunan. Hal ini tidak sesuai dengan teori, dimungkinkan karena adanya ketidakmerataan Angka Melek Huruf di daerah-daerah tersebut.

Rata-rata lama sekolah yang berada pada rentang 8,4625-8,4525 tahun, yaitu Kabupaten Mojokerto dan Kabupaten Jombang. Mayoritas mata pencaharian di Kabupaten Jombang adalah buruh/karyawan swasta serta perdagangan dan jasa. Sedangkan, di Kabupaten Mojokerto mayoritas bermata pencaharian disektor industri. Tenaga kerja pada kedua wilayah ini, didominasi oleh lulusan dengan

tingkat pendidikan SLTP. Ketidaksesuaian teori dimungkinkan karena tenaga kerja pada kedua wilayah tersebut didominasi oleh tenaga kerja dengan spesifikasi pendidikan SLTP, sehingga sumbangsih yang diberikan oleh tenaga kerja dari kedua wilayah tersebut belum mampu memenuhi kebutuhan disektor industri dan menyebabkan terjadinya kesenjangan sosial dan ekonomi di wilayah tersebut.

Selain itu, jika dilihat berdasarkan Kabupaten/Kota yang rata-rata lama sekolahnya lebih besar dari 10,0775 tahun, terdapat beberapa Kabupaten/Kota yang berada pada rentang tersebut, yaitu Kabupaten Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu. Dimungkinkan ketidaksesuain teori antara rata-rata lama sekolah dengan Angka Harapan Hidup di wilayah-wilayah tersebut, disebabkan karena wilayah tersebut memiliki rata-rata lama sekolah yang tinggi namun tidak merata.

Kasus Angka Harapan Hidup merupakan kasus yang cukup kompleks, sebab Angka Harapan Hidup dijadikan sebagai indikator untuk mengukur tingkat keberhasilan program kesehatan dan program sosial ekonomi disuatu wilayah. Angka Harapan Hidup disuatu wilayah tidak hanya dipengaruhi oleh satu atau dua faktor saja. Oleh karena itu, kondisi ketidaksesuain antara Angka Harapan dengan Angka Melek Huruf dan rata-rata lama sekolah dibeberapa wilayah-wilayah tertentu, dimungkinkan masih ada faktor-faktor lain yang tidak tercakup dalam model ini. Oleh karena itu, hal ini bisa menjadi temuan untuk melakukan penelitian lebih lanjut mengenai wilayah-wilayah dimana terjadinya ketidaksesuaian tersebut.

Halaman ini sengaja dikosongkan

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Estimasi model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier diperoleh menggunakan metode *Penalized Least Square* (PLS), yaitu dengan menggabungkan komponen *goodness of fit* dan *penalty*.

$$\underset{f,g,h}{Min}\left\{ n^{-1}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})^{T}(\tilde{Y}-\mathbf{X}\tilde{\beta}-\mathbf{G}\tilde{\phi}-\mathbf{D}\tilde{a})+\tilde{a}^{T}\mathbf{U}\tilde{a}\right\}$$

Diperoleh Estimator kurva regresi parametrik linier :

$$\hat{\tilde{f}}(x)=\mathbf{X}\hat{\tilde{\beta}}=\mathbf{A}(K,\lambda,k)\tilde{Y},$$

dengan

$$\mathbf{A}(K,\lambda,k)=\mathbf{X}\Big[\Big(\big[I-PDRX\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[QX-QDRX\big]\Big)^{-1}$$
$$\Big(\big[P-PDR\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[Q-QRD\big]\Big)\Big].$$

Estimator kurva regresi Spline *Truncated* :

$$\hat{\tilde{g}}(t)=\mathbf{G}\hat{\tilde{\phi}}=\mathbf{B}(K,\lambda,k)\tilde{Y},$$

dengan

$$\mathbf{B}(K,\lambda,k)=\mathbf{G}\Bigg[\Bigg(\big[I-PDRX\big]\Big[\Big(\big[I-PDRX\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[QX-QDRX\big]\Big)^{-1}$$
$$\Big(\big[P-PDR\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[Q-QRD\big]\Big)\tilde{Y}\Big]+\big[PG-PDRG\big]\Bigg)^{-1}$$
$$\big[P-PDR\big]\Bigg].$$

Estimator kurva regresi Deret Fourier :

$$\hat{\tilde{h}}(z)=\mathbf{D}\hat{\tilde{a}}=\mathbf{C}(K,\lambda,k)\tilde{Y},$$

dengan

$$\mathbf{C}(K,\lambda,k)=\mathbf{D}\Bigg(R-RX\Big[\Big(\big[I-PDRX\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[QX-QDRX\big]\Big)^{-1}$$
$$\Big(\big[P-PDR\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[Q-QRD\big]\Big)\Big]$$
$$-RG\Bigg[\Bigg(\big[I-PDRX\big]\Big[\Big(\big[I-PDRX\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[QX-QDRX\big]\Big)^{-1}$$
$$\Big(\big[P-PDR\big]-\big[PG-PDRG\big]\big[I-QDRG\big]^{-1}\big[Q-QRD\big]\Big)\tilde{Y}\Big]+\big[PG-PDRG\big]\Bigg)^{-1}\big[P-PDR\big]\Bigg]\Bigg).$$

Sehingga, diperoleh Estimasi model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier, sebagai berikut :

$$\hat{y} = \hat{\hat{f}}(x) + \hat{g}(t) + \hat{\hat{h}}(z) = \left[\mathbf{A}(k,\lambda,K) + \mathbf{B}(k,\lambda,K) + \mathbf{C}(k,\lambda,K)\right]\tilde{Y} = \mathbf{F}(K,\lambda,k)\tilde{Y},$$

dengan, $\mathbf{F}(K,\lambda,k) = \mathbf{A}(K,\lambda,k) + \mathbf{B}(K,\lambda,k) + \mathbf{C}(K,\lambda,k).$

2. Pemodelan data Angka Harapan Hidup di Provinsi Jawa Timur tahun 2013 menggunakan model regresi semiparametrik campuran Spline *Truncated* dan Deret Fourier dengan variabel respon adalah Angka Harapan Hidup (y), dan variabel-variabel prediktornya adalah variabel Angka Kematian Bayi (x) yang dihampiri dengan parametrik linear, variabel Angka Melek Huruf (t_1), rata-rata lama sekolah (t_2) dihampiri dengan fungsi Spline *Truncated* linear, dan variabel persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan (t_2) serta variabel Tingkat Partisipasi Angkatan Kerja (z_2) dihampiri dengan fungsi Deret Fourier. Model terbaik diperoleh dari model yang memiliki nilai GCV paling minumum yaitu model dengan komponen Spline *Truncated* tiga titik knot dan komponen Deret Fourier dengan osilasi $k = 2$. Namun, karena perbedaan nilai GCV model tiga titik knot dengan osilasi $k = 1$, $k = 2$ dan $k = 3$ tidak signifikan berbeda, sehingga dengan alasan parsimoni model maka dipilih model terbaik yaitu model dengan tiga titik knot dan osilasi $k = 1$.

$$\begin{aligned}\hat{y}_i = {} & 148.9104 - 0,250025x_i - 0,0315882t_{1i} - 0,182079(t_{1i} - 85,7431)_+ + \\ & 0,277273(t_{1i} - 87,5513)_+ - 0,508491(t_{1i} - 94,7837)_+ + 0,176008t_{2i} + \\ & -2,05685(t_{2i} - 8,04625)_+ + 3,30714(t_{2i} - 8,4525)_+ - 2,09454(t_{2i} - 10,0775)_+ + \\ & 0,162176z_{1i} + 2,23833\cos z_{1i} + 0,0278582z_{2i} - 0,00201112\cos z_{2i}\end{aligned}$$

Pemodelan data Angka Harapan Hidup di Jawa Timur menggunakan estimator campuran Spline *Truncated* dan Deret Fourier menghasilkan nilai R^2 sebesar 99,62% yang berarti bahwa variabel-variabel prediktor mampu menjelaskan variabel respon Angka Harapan Hidup sebesar 99,62%. Dari hasil pemodelan diperoleh bahwa semakin tinggi Angka Kematian Bayi, maka semakin rendah Angka Harapan Hidup. Hal tersebut telah sesuai dengan teori Angka Harapan Hidup yang menyatakan bahwa Angka Harapan Hidup berbanding terbalik dengan Angka Kematian Bayi. Sedangkan untuk variabel Angka Melek Huruf

dan variabel rata-rata lama sekolah, terdapat interpretasi yang tidak sesuai dengan teori. Dimana terdapat beberapa Kabupaten/Kota di Provinsi Jawa Timur, ketika Angka Melek Huruf dan rata-rata lama sekolahnya mengalami kenaikan, Angka Harapan Hidupnya mengalami penurunan. Ketidaksesuain dengan teori tersebut terjadi dikarenakan adanya perbedaan karakteristik yang berdeda-beda dari setiap Kabupaten/Kota yang ada di Provinsi Jawa Timur. Selain itu, dimungkinkan oleh adanya ketidakmerataan Angka Melek Huruf dan rata-rata lama sekolah di Provinsi Jawa Timur.

5.2 Saran

Berdasarkan hasil analisis dan kesimpulan yang diperoleh, saran yang dapat diberikan adalah :

1. Metode yang digunakan dalam penelitian ini adalah regresi semiparametrik campuran Spline *Truncated* linier dan Deret Fourier dengan *trend*. Penelitian ini dapat pula dikembangkan untuk data longitudinal. Selain itu, penelitian ini terbatas pada estimasi titik, untuk penelitian selanjutnya dapat dilakukan estimasi interval dan uji statistik untuk melihat variabel-variabel prediktor yang memberi pengaruh signifikan terhadap variabel respon.
2. Bagi pengambil kebijakan dibidang pendidikan ; untuk meningkatkan Angka Melek Huruf dan rata-rata lama sekolah diperlukan penyuluhan-penyuluhan tentang kesadaran mengenai pentingnya pendidikan, serta melakukan pemerataan sarana dan prasarana pendidikan.

Halaman ini sengaja dikosongkan

DAFTAR PUSTAKA

Asrini, L., J. (2013). *Model Regresi Semiparametrik Deret Fourier*. Institut Teknologi Sepuluh Nopember Surabaya.

Badan Pusat Statistik Jawa Timur, (2012a), *Laporan Eksekutif Kesehatan Jawa Timur*, Badan Pusat Statistik Provinsi Jawa Timur.

Badan Pusat Statistik Jawa Timur, (2012b), *Laporan Eksekutif Pendidikan Jawa Timur*, Badan Pusat Statistik Provinsi Jawa Timur.

Badan Pusat Statistik Jawa Timur, (2013a), *Keadaan Angkatan Kerja di Jawa Timur*, Badan Pusat Statistik Provinsi Jawa Timur.

Badan Pusat Statistik Jawa Timur, (2013b), *Survei Sosial Ekonomi Nasional Jawa Timur*, Badan Pusat Statistik Provinsi Jawa Timur.

Bilodeau, M. (1992). Fourier Smoother and Additive Models. *The Canadian Journal of Statistics,* 3: 257-269.

Budiantara, I.N. (2000). Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI),* 6: 285-290.

Budiantara, I.N. (2009). Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang, *Pidato Pengukuhan Untuk Jabatan Guru Besar Dalam Bidang Ilmu: Matematika Statistika dan Probabilitas*, Institut Teknologi Sepuluh Nopember: Surabaya.

Draper, N.R. and Smith, H. (1992). *Applied Regression Analysis, Second Edition*. John Wiley and sons, Inc. New York.

Eubank, R.L., (1988), *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, New York.

Eubank, R. L., (1999), *Spline Smoothing and Nonparametric Regression*, Marcel Deker: New York.

Firdial, L. (2010). *Pemodelan Angka Harapan Hidup di Provinsi Jawa Timur dan Jawa Tengah dengan Metode Georaphically Weighted Regression (GWR),* Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Gujarati, D.N. (2004). *Basic Econometrics* (4th ed). New York : McGraw-Hill.

Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press: New York.

Hesikumalasari, (2016). *Pemodelan Regresi Semiprametrik menggunakan Estimator Campuran Spline Truncated dan Kernel.* Institut Teknologi Sepuluh Nopember Surabaya.

Purnomo, A.,A.,S., I. (2016). *Estimator Campuran Kernel dan Spline Trincated Linier dalam Regresi Nonparametrik.* Institut Teknologi Sepuluh Nopember Surabaya.

Rakhmawati, D.,P. (2011). *Analisis Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup Di Provinsi Jawa Barat,* Universitas Gadjah Mada.

Riskiyanti, R. (2010). *Analisis Regresi Multivariat Berdasarkan Faktor-faktor yang Mempengaruhi Derajat Kesehatan di Provinsi Jawa Timur*, Institut Teknologi Sepuluh Nopember Surabaya.

Rory, (2016). *Regresi Campuran Nonparametrik Spline Linier Truncated dan Fungsi Kernel untuk Pemodelan Data Kemiskinan di Papua.* Institut Teknologi Sepuluh Nopember Surabaya.

Rismal, (2016). *Estimasi Model Campuran Spline Truncated dan Kernel dalam Regresi Nonparametrik Multivariabel.* Institut Teknologi Sepuluh Nopember Surabaya.

Rencher, Alvin dan Schaalje, G. (2008), *Linear Models in Statistics*, 2nd Edition, John Willey and Sons Inc., New Jersey.

Singh, K.G.,Siahpush,M. (2006). Widening Socioenomic Inequalities in US life expectancy, 1980-2000, *International Journal of Epidemiology,* 35, 969-979.

Siminoff, J.S. (1996), *Smoothing Methods in Statistics*, Springer, New York.

Sudiarsa, I. W., Budiantara, I. N., Suhartono, & Purnami, S. W. (2015). "Combined Estimator Fourier Series and Spline Truncated Multivariable Nonparametric Regression". *Applied Mathematical Sciences,* Vol. 9, No. 100, hal. 4997-5010.

Sugiarti, A.,P. (2013). *Analisis Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik Spline.* Institut Teknologi Sepuluh Nopember Surabaya.

Syisliawati. (2016). *Estimasi Interval untuk kurva regresi campuran spline truncated dan kernel dalam regresi nonparametrik*. Institut Teknologi Sepuluh Nopember Surabaya.

Tripena, A., & Budiantara, I. N. (2007). "Fourier Estimator in Nonparametric Regression". *International Conference on Natural Sciences and Applied Natural Scienes.* Ahmad Dahlan University, Yogyakarta.

Wahba, G. (1990). *Spline Models for Observational Data.* Philadelphia: Society for Industrial and Applied Mathematics.

LAMPIRAN

Lampiran 1. Data Angka Harapan Hidup di Jawa Timur Tahun 2013 dan Faktor-Faktor yang Diduga Berpengaruh

Kabupaten/Kota	y	x	t_1	t_2	z_1	z_2
Pacitan	72,18	22,12	91,67	7,01	3,54	79,44
Ponorogo	70,85	25,83	89,37	7,49	3,27	71,81
Trenggalek	72,33	20,8	93,07	7,33	3,43	77,46
Tulungagung	72,02	21,4	94,92	7,79	4,06	71,52
Blitar	71,8	23,12	92,12	7,41	3,72	71,99
Kediri	70,65	26,83	92,97	7,75	3,33	68,39
Malang	69,7	29,46	91,22	7,08	3,55	68,74
Lumajang	67,95	36,92	86,63	6,52	2,97	65,63
Jember	63,64	55,42	83,79	6,8	3,31	65,01
Banyuwangi	68,58	32,56	88,44	7,25	3,6	72,84
Bondowoso	63,95	52,28	81,22	5,94	3,1	67,48
Situbondo	63,95	53,82	78,62	6,28	2,95	68,62
Probolinggo	62,1	62,45	80,95	6,31	3,73	72,81
Pasuruan	64,81	49,74	91,71	6,88	2,99	70,72
Sidoarjo	71,43	23,36	97,91	10,23	4,81	67,37
Mojokerto	71,13	23,99	94,47	8,22	4,15	67,87
Jombang	70,64	27,05	94,45	8,06	3,18	64,18
Nganjuk	69,82	30,46	91,16	7,62	3,4	69,64
Madiun	69,68	30,64	90,04	7,47	4,08	69,86
Magetan	71,96	22,29	91,42	7,86	3,38	71,5
Ngawi	70,97	25,83	85,99	7,06	4,32	73,17
Bojonegoro	67,81	38,24	85,13	6,72	3,97	72,99
Tuban	68,71	32,86	86	6,82	2,87	70,01

Keterangan :

y **:** Angka Harapan Hidup

x **:** Angka Kematian Bayi

t_1 **:** Angka Melek Huruf

t_2 **:** Rata-rata Lama Sekolah

z_1 **:** Persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan

z_2 **:** Tingkat Partisipasi Angkatan Kerja

Lampiran 1. Data Angka Harapan Hidup di Jawa Timur Tahun 2013 dan Faktor-Faktor yang Diduga Berpengaruh (Lanjutan)

Kabupaten/Kota	y	x	t_1	t_2	z_1	z_2
Lamongan	68,98	33,25	89,09	7,79	2,91	70,5
Gresik	71,57	22,65	96,38	9	3,78	67,6
Bangkalan	64,02	53,69	82,93	5,75	2,05	70,61
Sampang	64,52	51,72	69,47	4,39	1,14	72,37
Pamekasan	65,19	49	84,48	6,42	1,58	77,97
Sumenep	65,49	47,48	78,75	5,73	3,83	75,59
Kediri	71,36	23,3	97,86	10,29	4,18	64,18
Blitar	73	18,71	97,48	9,87	3	66,53
Malang	71,14	22,84	98,38	10,89	3,77	65,99
Probolinggo	71,16	23,13	92,66	8,79	1,9	63,7
Pasuruan	66,75	38,38	97,12	9,07	2,84	69,13
Mojokerto	72,48	21,38	97,58	10,12	4,42	70,18
Madiun	71,89	22,62	98,15	10,54	4,1	66,39
Surabaya	72,13	22,48	98,4	10,12	3,64	67,86
Batu	70,32	27,91	93,37	8,76	2,21	70,58

Keterangan :

y **:** Angka Harapan Hidup

x **:** Angka Kematian Bayi

t_1 **:** Angka Melek Huruf

t_2 **:** Rata-rata Lama Sekolah

z_1 **:** Persentase bayi berumur 0-11 bulan yang diberi ASI 4-6 bulan

z_2 **:** Tingkat Partisipasi Angkatan Kerja

Lampiran 2. Syntax R Model Semiparametrik Campuran

```
options(digits=8)
name="D:/122d_k=1/" #read directiory
data=read.table(paste0(name,"ahh122d.csv"),header=T,sep=';')
kf=1  # nilai k
mf=2 #banyaknya variabel fourier
mt=2 #banyaknya variabel spline

GCV1=function(data,kf,name,mf,mt)
{
  #fungsi GCV untuk 1 knot
  #data adalah data yang digunakan
  #kf adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  #mf adalah banyanya variabel fourier
  #mt adalah banyanya variabel spline

  options(digits=6)
  #mf=2 #banyaknya variabel fourier
  #mt=2 #banyaknya variabel spline
  mp=ncol(data)-(1+mf+mt)#banyaknya variabel parametrik
  m1=mt

  #load package Matrix dan pracma
  library(Matrix)
  library(pracma)
  library(gtools)

  data=as.matrix(data)
  y=as.matrix(data[,1]) #inisiasi vektor y
  x=as.matrix(data[,-1])
  n=nrow(y)
  m=ncol(x)
  dataA=as.matrix(x[,(m+1-mt):m]) #var non Par
  datapar=as.matrix(x[,(mf+1):(mf+mp)]) #var parametrik
  nk=27 #banyak patahan (knot)
  #nl=100  #banyak patahan lambda
#menginisasi nilai knot
  knot1=matrix(ncol=m1,nrow=nk)
  for (i in (1:m1))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=nk)
      knot1[j,i]=a[j]
    }
  }
  knot1=as.matrix(knot1[2:(nk-1),])
  aa=rep(1,n)
  data1=matrix(ncol=m1,nrow=n)#in akan menjadi X
```

```
#persamaan deret fourier
  xfou=as.matrix(x[,1:mf])
  W=matrix(nrow=n)
  D=matrix(0,ncol=1,nrow=1)
  for (i in (1:mf))
  {
    der=fungsik(kf,xfou[,i])
    W=cbind(W,der$W)
    D=blkdiag(D,der$D)
  }
  W=as.matrix(W[,-1])
  D=as.matrix(D[-1,-1])
  a2=nk-2
  nprint=(min(25,a2))

  #menginisiasi nilai titik titik lambda
  temp1=c()
  for ( i in 1:7)
  {
    if (mf==1) temp0=seq((10^(-(1+i))),10^(-(i)),by=(1*10^(-(1+i))))
else temp0=seq((10^(-(1+i))),10^(-(i)),by=(3*10^(-(1+i))))
    temp1=c(temp1,temp0)
  }
  if (mf==1)
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.02),temp1)))) else
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.1),temp1))))
  #lambda1=as.matrix(c(seq(10^(-9),10^(-6),by=0.00000002),seq(10^(-
6),10^-3,by=0.000002),seq(10^(-3),10^-2,by=0.0002)))
  options(expressions=1e5)
  lambda2=permutations(length(lambda11),mf,lambda11)
  options(expressions=5000)

  lambda1=lambda2
  nl=nrow(lambda1)
#membuat matrix kosong untuk nilai GCV,MSE,Rsq,SSE,SST,SSR
  GCV=matrix(nrow=a2,ncol=nl)
  MSE=matrix(nrow=a2,ncol=nl)
  Rsq=matrix(nrow=a2,ncol=nl)
  SSE=matrix(nrow=a2,ncol=nl)
  SST=matrix(nrow=a2,ncol=nl)
  SSR=matrix(nrow=a2,ncol=nl)
  lambda=matrix(ncol=nl);colnames(GCV)<-c(1:nl);colnames(MSE)<-
c(1:nl);colnames(Rsq)<-c(1:nl)
  for(i in 1:a2)
  {
    lambda=rbind(lambda,t(lambda2))
  }
  lambda=lambda[-1,]


  for (i in 1:a2)
  {
```

```
    #menghitung knot
    for (j in 1:m1)
    {
      for (k in 1:n)
      {
        #i=1;k=1;j=1
        if (dataA[k,(j)]<knot1[i,j]) data1[k,j]=0 else
data1[k,j]=dataA[k,(j)]-knot1[i,j] #menghitung knot
      }
    }
    satu=rep(1,n)
    X=cbind(satu,datapar,dataA,data1) # nilai matrix x (sudah
termasuk nilai knot)
    I1=diag(ncol(X))
    I2=diag(nrow(X))
    for (j in 1:nl)
    {
#menghitung nilai estimasi
      lambdaA=matrix(ncol=1)
        for (k in 1:mf)
        {

lambdaA=cbind(lambdaA,matrix(c(rep(lambda2[j,k],(ncol(D)/mf))),nrow=
1))
        }
        lambdaA=lambdaA[,-1]
      Sk=pinv(t(W)%*%W+n*lambdaA*D)
      B1=pinv(I1-(pinv(t(X)%*%X)%*%t(X)%*%W%*%Sk%*%t(W)%*%X))
      B2=pinv(t(X)%*%X)%*%t(X)%*%(I2-W%*%Sk%*%t(W))
      B3=B1%*%B2
      A1=Sk%*%t(W)%*%(I2-X%*%B3)
      C=W%*%A1+X%*%B3
      yhat=C%*%y  # mencari nilai yhat
      SSE[i,j]=sum((y-yhat)^2)      #mencari SSE
      SSR[i,j]=sum((yhat-mean(y))^2)      #mencari SSR
      SST[i,j]=SSR[i,j]+SSE[i,j]    #mencari SST
      Rsq[i,j]=(SSR[i,j]/(SST[i,j]))*100  #mencari R square
      MSE[i,j]=(t(y)%*%t(I2-C)%*%(I2-C)%*%y)/n  #mencari MSE
      GCVbwh=sum(diag(I2-C))/n
      GCV[i,j]=MSE[i,j]/GCVbwh      ##mencari GCV
    }
  }
  GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
#menyiapkan output sort GCV terkecil
  R11=matrix(ncol=1);       colnames(R11)<-"Rsq"
  MSE11=matrix(ncol=1);     colnames(MSE11)<-"MSE"
  GCV11=matrix(ncol=1);     colnames(GCV11)<-"GCV"
  ke11=matrix(ncol=1);colnames(ke11)<-"Knot_ke"
  kelam11=matrix(ncol=1);colnames(kelam11)<-"lambda_ke"
  for ( i in 1:nl)
  {
    ke11=rbind(ke11,matrix(c(1:a2),ncol=1))
    R11=rbind(R11,matrix(Rsq[,i],ncol=1))
    MSE11=rbind(MSE11,matrix(MSE[,i],ncol=1))
    GCV11=rbind(GCV11,matrix(GCV[,i],ncol=1))
    kelam11=rbind(kelam11,matrix(rep(i,a2),ncol=1))
  }
  ke11=matrix(ke11[-1,],ncol=1)
  R11=matrix(R11[-1,],ncol=1)
  MSE11=matrix(MSE11[-1,],ncol=1)
  GCV11=matrix(GCV11[-1,],ncol=1)
  kelam11=matrix(kelam11[-1,],ncol=1)

  knot11=matrix(nrow=nrow(ke11),ncol=ncol(knot1));
  L11=matrix(nrow=nrow(ke11),ncol=ncol(lambda2));
  for ( i in 1:nrow(ke11))
  {
    knot11[i,]=knot1[ke11[i],]
    L11[i,]=lambda2[kelam11[i],]
  }
gab11=cbind(GCV11,R11,MSE11,L11,ke11,knot11)
  gab11copy=gab11
  dataprint=matrix(nrow=nprint,ncol=ncol(gab11copy)) #gcv, Rsq, mse, 
lambda, knot
  for (i in 1:nprint)
  {
    s1=min(gab11copy[,1])
    pos=which(s1==gab11copy[,1])
    if (length(pos)==1) pos=pos else pos=randsample(pos,1)
    dataprint[i,]=gab11copy[pos,]
    posdellambda=0
    for (j in 1:mf)
    {

posdellambda=c(posdellambda,which(gab11copy[pos,(3+j)]==gab11copy[,(
3+j)]))
    }
    posdellambda=posdellambda[-1]
    posdelknot  =which(gab11copy[pos,4+mf]==gab11copy[,4+mf])
    gab11copy=gab11copy[-c(posdellambda,posdelknot),]
  }
  dataprint=dataprint[order(dataprint[,1],-
dataprint[,2],dataprint[,3],dataprint[,4+mf]),]
  colnames(dataprint)<-
```

```
c("GCV","Rsq","MSE",paste0('lambda',c(1:mf)),"knot_ke",
                         paste0('knot',c(1:mt),1))
#save hasil menggunakan file csv
  write.csv(gab11,file=paste(name,"output data All 1 knot
k=",kf,".csv"))
  write.csv(dataprint,file=paste(name,"output data yang di print 1
knot k=",kf,".csv"))

  #print output
  cat("==============================================","\n")
  cat("HASIL GCV terkecil dengan fourier-Spline 1 knot k=",kf,"\n")
  cat("==============================================","\n")
  cat("GCV     = ",dataprint[1,1],"\n")
  cat("Rsquare = ",dataprint[1,2],"\n")
  cat("MSE     = ",dataprint[1,3],"\n")
  cat("Lambda  = ",dataprint[1,c(4:(4+mf-1))],"\n")
  cat("Knot    = ",dataprint[1,(4+mf+1):ncol(dataprint)],"\n")
  cat("Nilai GCV 25 ","\n")
  print(dataprint)

  #membuat plot antara GCV vs Lambda pada posisi lambda optimum
  pos=which(min(gab11[,1])==gab11[,1])
  if (length(pos)==1) pos=pos else pos=randsample(pos,1)
  #posknot  =which(gab11[pos,5]==gab11[,5])
  posknot=which(dataprint[1,4+mf]==gab11[,4+mf])
  logl=matrix(ncol=mf,nrow=nrow(gab11))
  for ( j in 1:(mf))
  {
    temp=0
    temp=which(dataprint[1,3+j]==gab11[,3+j])
    logl[1:length(temp),j]=temp
  }
  mypath <- file.path(paste(name,"plot_1 knot K=",kf,".jpg", sep =
""))
  jpeg(file=mypath,width=1104,height=694)
  par(mfrow=c((ceiling(mf/2)),2))
  for ( i in 1:mf)
  {
    posl=c(na.omit(c(logl[,-i])))
    gabpos=c(posl,posknot)
    if (mf==1) posfix=gabpos else
posfix=gabpos[duplicated(c(posl,posknot))]
    while ((anyDuplicated(posfix))!=0)
    {
      posfix=posfix[duplicated(posfix)]
    }
pos2=which(pos==posfix)
    posbawah=pos2-(min(10,pos2))+1
    posatas=pos2+(min(10,length(posfix)-pos2))
    xyplot=gab11[posfix[posbawah:posatas],c(1,3+i)]
    xyplot=xyplot[order(xyplot[,2]),]
    gcvplot=xyplot[,1];
    lambdaplot=xyplot[,2]
```

```
plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='b
',main=paste("Plot_ 1 knot, k=",kf,"lambda ke ",i))
  }
#save plot GCV vs Lambda pada posisi lambda optimum
  dev.off()
  par(mfrow=c(1,1))

  #lambdaplot=gab11[posknot,4]
  #gcvplot=gab11[posknot,1]
  #par(mfrow=c(3,2))

#plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='
b',main=paste("knot1 optimum k=",kf))

  #membuat matrix x untuk estimasi selanjutnya
  lambda1=dataprint[1,c(4,(4+mf-1))]
  knotgcv=matrix(dataprint[1,(4+mf+1):ncol(dataprint)],nrow=1)
  for (j in 1:m1)
  {
    for (k in 1:n)
    {
      if (dataA[k,(j)]<knotgcv[j]) data1[k,j]=0 else
data1[k,j]=dataA[k,(j)]-knotgcv[j]
    }
  }
  X=cbind(satu,datapar,dataA,data1)

list(D=D,knot=knotgcv,mingcv=dataprint[1,1],Rsqgcv=dataprint[1,2],W=
W,X=X,lambda=lambda1)
}
GCV2=function(data,kf,name,mf,mt)
{
  #fungsi GCV untuk 2 knot
  #data adalah data yang digunakan
  #kf adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  #mf adalah banyanya variabel fourier
  #mt adalah banyanya variabel spline

  options(digits=6)
  #mf=2 #banyaknya variabel fourier
  #mt=2 #banyaknya variabel spline
  mp=ncol(data)-(1+mf+mt)
  m1=mt
  library(Matrix)
  library(pracma)
  library(gtools)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
  data=as.matrix(data)
  y=as.matrix(data[,1]) #inisiasi vektor y
  x=as.matrix(data[,-1])
  n=nrow(y)
  m=ncol(x)
  dataA=as.matrix(x[,(m+1-mt):m]) #var non Par
  datapar=as.matrix(x[,(mf+1):(mf+mp)]) #var parametrik
  nk=22 #banyak patahan (knot)
  #nl=100  # banyak patahan lambda
#menginisasi nilai knot
  knot1=matrix(ncol=m1,nrow=nk)
  for (i in (1:m1))
  {
    for (j in (1:(nk)))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=nk)
      knot1[j,i]=a[j]
    }
  }
  knot=as.matrix(knot1[2:((nk)-1),])
  a2=nrow(knot)#a2=nk-2
  z=(a2*(a2-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m1))
  {
    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*m1+1)]
  a3=nrow(knot2)
  nprint=(min(25,a3))
  aa=rep(1,n)
   #persamaan deret fourier
  xfou=as.matrix(x[,1:mf])
  W=matrix(nrow=n)
  D=matrix(0,ncol=1,nrow=1)
  for (i in (1:mf))
{
    der=fungsik(kf,xfou[,i])
    W=cbind(W,der$W)
    D=blkdiag(D,der$D)
  }
W=as.matrix(W[,-1])
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
D=as.matrix(D[-1,-1])
  data1=matrix(ncol=(2*m1),nrow=n)#in akan menjadi X
#menginisiasi nilai titik titik lambda
  temp1=c()
  for ( i in 1:7)
  {
    if (mf==1) temp0=seq((10^(-(1+i))),10^(-(i)),by=(1*10^(-(1+i))))
else temp0=seq((10^(-(1+i))),10^(-(i)),by=(3*10^(-(1+i))))
    temp1=c(temp1,temp0)
  }
  if (mf==1)
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.02),temp1)))) else
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.1),temp1))))
  #lambda1=as.matrix(c(seq(10^(-9),10^(-6),by=0.00000002),seq(10^(-
6),10^-3,by=0.000002),seq(10^(-3),10^-2,by=0.0002)))
  options(expressions=1e5)
  lambda2=permutations(length(lambda11),mf,lambda11)
  options(expressions=5000)

  lambda1=lambda2
  nl=nrow(lambda1)
#membuat matrix kosong untuk nilai GCV,MSE,Rsq,SSE,SST,SSR
  GCV=matrix(nrow=a3,ncol=nl)
  MSE=matrix(nrow=a3,ncol=nl)
  Rsq=matrix(nrow=a3,ncol=nl)
  SSE=matrix(nrow=a3,ncol=nl)
  SST=matrix(nrow=a3,ncol=nl)
  SSR=matrix(nrow=a3,ncol=nl)
  lambda=matrix(ncol=nl);colnames(GCV)<-c(1:nl);colnames(MSE)<-
c(1:nl);colnames(Rsq)<-c(1:nl)
  for(i in 1:a3)
  {
    lambda=rbind(lambda,t(lambda2))
  }
  lambda=lambda[-1,]


  for (i in 1:a3)
  {
    #menghitung knot
    for (j in 1:(2*m1))
    {
      if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
      for (k in 1:n)
      {
        if (dataA[k,b]<knot2[i,j]) data1[k,j]=0 else
data1[k,j]=dataA[k,b]-knot2[i,j] #menghitung knot
      }
    }
satu=rep(1,n)
    X=cbind(satu,datapar,dataA,data1) # nilai matrix x (sudah
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
termasuk nilai knot)
    I1=diag(ncol(X))
I2=diag(nrow(X))
    for (j in 1:nl)
    {
      #menghitung nilai estimasi
      lambdaA=matrix(ncol=1)
      for (k in 1:mf)
      {

lambdaA=cbind(lambdaA,matrix(c(rep(lambda2[j,k],(ncol(D)/mf))),nrow=
1))
      }
      lambdaA=lambdaA[,-1]
      Sk=pinv(t(W)%*%W+n*lambdaA*D)
      B1=pinv(I1-(pinv(t(X)%*%X)%*%t(X)%*%W%*%Sk%*%t(W)%*%X))
      B2=pinv(t(X)%*%X)%*%t(X)%*%(I2-W%*%Sk%*%t(W))
      B3=B1%*%B2
      A1=Sk%*%t(W)%*%(I2-X%*%B3)
      C=W%*%A1+X%*%B3
      yhat=C%*%y  # mencari nilai yhat
      SSE[i,j]=sum((y-yhat)^2)      #mencari SSE
      SSR[i,j]=sum((yhat-mean(y))^2)      #mencari SSR
      SST[i,j]=SSR[i,j]+SSE[i,j]    #mencari SST
      Rsq[i,j]=(SSR[i,j]/(SST[i,j]))*100  #mencari R square
      MSE[i,j]=(t(y)%*%t(I2-C)%*%(I2-C)%*%y)/n  #mencari MSE
      GCVbwh=sum(diag(I2-C))/n
      GCV[i,j]=MSE[i,j]/GCVbwh      ##mencari GCV
    }
  }
GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)

  #menyiapkan output sort GCV terkecil
  R11=matrix(ncol=1);       colnames(R11)<-"Rsq"
  MSE11=matrix(ncol=1);     colnames(MSE11)<-"MSE"
  GCV11=matrix(ncol=1);     colnames(GCV11)<-"GCV"
  ke11=matrix(ncol=1);colnames(ke11)<-"Knot_ke"
  kelam11=matrix(ncol=1);colnames(kelam11)<-"lambda_ke"
  for ( i in 1:nl)
  {
    ke11=rbind(ke11,matrix(c(1:a3),ncol=1))
    R11=rbind(R11,matrix(Rsq[,i],ncol=1))
    MSE11=rbind(MSE11,matrix(MSE[,i],ncol=1))
    GCV11=rbind(GCV11,matrix(GCV[,i],ncol=1))
    kelam11=rbind(kelam11,matrix(rep(i,a3),ncol=1))
  }
  ke11=matrix(ke11[-1,],ncol=1)
  R11=matrix(R11[-1,],ncol=1)
  MSE11=matrix(MSE11[-1,],ncol=1)
  GCV11=matrix(GCV11[-1,],ncol=1)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
kelam11=matrix(kelam11[-1,],ncol=1)
knot11=matrix(nrow=nrow(ke11),ncol=ncol(knot2));
  L11=matrix(nrow=nrow(ke11),ncol=ncol(lambda2));
  for ( i in 1:nrow(ke11))
  {
    knot11[i,]=knot2[ke11[i],]
    L11[i,]=lambda2[kelam11[i],]
  }
gab11=cbind(GCV11,R11,MSE11,L11,ke11,knot11)
  gab11copy=gab11
  dataprint=matrix(nrow=nprint,ncol=ncol(gab11copy)) #gcv, Rsq, mse,
lambda, knot
  for (i in 1:nprint)
  {
    s1=min(gab11copy[,1])
    pos=which(s1==gab11copy[,1])
    if (length(pos)==1) pos=pos else pos=randsample(pos,1)
    dataprint[i,]=gab11copy[pos,]
    posdellambda=0
    for (j in 1:mf)
    {

posdellambda=c(posdellambda,which(gab11copy[pos,(3+j)]==gab11copy[,(
3+j)]))
    }
    posdellambda=posdellambda[-1]
    posdelknot  =which(gab11copy[pos,4+mf]==gab11copy[,4+mf])
    gab11copy=gab11copy[-c(posdellambda,posdelknot),]
  }
dataprint=dataprint[order(dataprint[,1],-
dataprint[,2],dataprint[,3],dataprint[,5]),]
  colnames(dataprint)<-
c("GCV","Rsq","MSE",paste0('lambda',c(1:mf)),"knot_ke",
                         paste0('_',c(1:(2*mt))))
#save hasil menggunakan file csv
  write.csv(gab11,file=paste(name,"output data All 2 knot
k=",kf,".csv"))
  write.csv(dataprint,file=paste(name,"output data yang di print 2
knot k=",kf,".csv"))
print output
  cat("==============================================","\n")
  cat("HASIL GCV terkecil dengan fourier-Spline 2 knot k=",kf,"\n")
  cat("==============================================","\n")
  cat("GCV     = ",dataprint[1,1],"\n")
  cat("Rsquare = ",dataprint[1,2],"\n")
  cat("MSE     = ",dataprint[1,3],"\n")
  cat("Lambda  = ",dataprint[1,c(4:(4+mf-1))],"\n")
  cat("Knot    = ",dataprint[1,(4+mf+1):ncol(dataprint)],"\n")
  cat("Nilai GCV 25 ","\n")
  print(dataprint)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
#membuat plot antara GCV vs Lambda pada posisi lambda optimum
  pos=which(min(gab11[,1])==gab11[,1])
  if (length(pos)==1) pos=pos else pos=randsample(pos,1)
#posknot  =which(gab11[pos,5]==gab11[,5])
  posknot=which(dataprint[1,4+mf]==gab11[,4+mf])
  logl=matrix(ncol=mf,nrow=nrow(gab11))
  for ( j in 1:(mf))
{
    temp=0
    temp=which(dataprint[1,3+j]==gab11[,3+j])
    logl[1:length(temp),j]=temp
  }
  mypath <- file.path(paste(name,"plot_2 knot K=",kf,".jpg", sep =
""))
  jpeg(file=mypath,width=1104,height=694)
  par(mfrow=c((ceiling(mf/2)),2))
  for ( i in 1:mf)
  {
    posl=c(na.omit(c(logl[,-i])))
    gabpos=c(posl,posknot)
    if (mf==1) posfix=gabpos else
posfix=gabpos[duplicated(c(posl,posknot))]
    while ((anyDuplicated(posfix))!=0)
    {
      posfix=posfix[duplicated(posfix)]
    }
pos2=which(pos==posfix)
    posbawah=pos2-(min(10,pos2))+1
    posatas=pos2+(min(10,length(posfix)-pos2))
    xyplot=gab11[posfix[posbawah:posatas],c(1,3+i)]
    xyplot=xyplot[order(xyplot[,2]),]
    gcvplot=xyplot[,1];
    lambdaplot=xyplot[,2]

plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='b
',main=paste("Plot_ 1 knot, k=",kf,"lambda ke ",i))
  }
#save plot GCV vs Lambda pada posisi lambda optimum
  dev.off()
  par(mfrow=c(1,1))
#test
  #lambdaplot=gab11[posknot,4]
  #gcvplot=gab11[posknot,1]
  #par(mfrow=c(3,2))

#plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='
b',main=paste("knot1 optimum k=",kf))
#membuat matrix x untuk estimasi selanjutnya
  lambda1=dataprint[1,c(4,(4+mf-1))]
  knotgcv=matrix(dataprint[1,(4+mf+1):ncol(dataprint)],nrow=1)
  for (j in 1:(2*m1))
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
#membuat matrix x untuk estimasi selanjutnya
  lambda1=dataprint[1,c(4,(4+mf-1))]
  knotgcv=matrix(dataprint[1,(4+mf+1):ncol(dataprint)],nrow=1)
  for (j in 1:(2*m1))
GCV3=function(data,kf,name,mf,mt)
{
  #fungsi GCV untuk 3 knot
  #data adalah data yang digunakan
  #kf adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  #mf adalah banyanya variabel fourier
#mt adalah banyanya variabel spline

  options(digits=6)
  #mf=2 #banyaknya variabel fourier
  #mt=2 #banyaknya variabel spline
  mp=ncol(data)-(1+mf+mt)#banyaknya variabel parametrik
  m1=mt

  #load package Matrix dan pracma
  library(Matrix)
  library(pracma)
  library(gtools)

  data=as.matrix(data)
  y=as.matrix(data[,1])#inisiasi vektor y
  x=as.matrix(data[,-1])
  n=nrow(y)
  m=ncol(x)
  dataA=as.matrix(x[,(m+1-mt):m]) #var non Par
  datapar=as.matrix(x[,(mf+1):(mf+mp)]) #var parametrik
  nk=17 #banyak patahan (knot)

  #menginisasi nilai knot
  knot1=matrix(ncol=m1,nrow=nk)
  for (i in (1:m1))
  {
    for (j in (1:(nk)))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=nk)
      knot1[j,i]=a[j]
    }
  }
knot=as.matrix(knot1[2:((nk)-1),])
  a2=nrow(knot)#a2=nk-2
  z=(a2*(a2-1)*(a2-2)/6)
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m1))
  {
    knot2=rbind(rep(NA,3))
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
    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*m1+1)]
  a3=nrow(knot1)
  nprint=(min(25,a3))
  aa=rep(1,n)
  #persamaan deret fourier
  xfou=as.matrix(x[,1:mf])
  W=matrix(nrow=n)
  D=matrix(0,ncol=1,nrow=1)
  for (i in (1:mf))
  {
    der=fungsik(kf,xfou[,i])
    W=cbind(W,der$W)
    D=blkdiag(D,der$D)
  }
  W=as.matrix(W[,-1])
  D=as.matrix(D[-1,-1])
  data1=matrix(ncol=3*m1,nrow=n)#in akan menjadi X
  #menginisiasi nilai titik titik lambda
  temp1=c()
  for ( i in 1:7)
  {
    if (mf==1) temp0=seq((10^(-(1+i))),10^(-(i)),by=(1*10^(-(1+i))))
else temp0=seq((10^(-(1+i))),10^(-(i)),by=(3*10^(-(1+i))))
    temp1=c(temp1,temp0)
  }
  if (mf==1)
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.02),temp1)))) else
lambda11=as.matrix(sort(unique(c(seq(0.2,1,by=0.1),temp1))))
  #lambda1=as.matrix(c(seq(10^(-9),10^(-6),by=0.00000002),seq(10^(-
6),10^-3,by=0.000002),seq(10^(-3),10^-2,by=0.0002)))
  options(expressions=1e5)
  lambda2=permutations(length(lambda11),mf,lambda11)
  options(expressions=5000)
lambda1=lambda2
  nl=nrow(lambda1)

  #membuat matrix kosong untuk nilai GCV,MSE,Rsq,SSE,SST,SSR
  GCV=matrix(nrow=a3,ncol=nl)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
MSE=matrix(nrow=a3,ncol=nl)
  Rsq=matrix(nrow=a3,ncol=nl)
SSE=matrix(nrow=a3,ncol=nl)
  SST=matrix(nrow=a3,ncol=nl)
  SSR=matrix(nrow=a3,ncol=nl)
  lambda=matrix(ncol=nl);colnames(GCV)<-c(1:nl);colnames(MSE)<-
c(1:nl);colnames(Rsq)<-c(1:nl)
  for(i in 1:a3)
  {
    lambda=rbind(lambda,t(lambda2))
  }
lambda=lambda[-1,]
  for (i in 1:a3)
  {
    #menghitung knot
    for (j in 1:ncol(knot1))
    {
      b=ceiling(j/3)
      for (k in 1:n)
      {
        if (dataA[k,b]<knot1[i,j]) data1[k,j]=0 else
data1[k,j]=dataA[k,b]-knot1[i,j] #menghitung knot
      }
    }
satu=rep(1,n)
    X=cbind(satu,datapar,dataA,data1)# nilai matrix x (sudah
termasuk nilai knot)
    I1=diag(ncol(X))
    I2=diag(nrow(X))
    for (j in 1:nl)
    {
#menghitung nilai estimasi
      lambdaA=matrix(ncol=1)
      for (k in 1:mf)
      {

lambdaA=cbind(lambdaA,matrix(c(rep(lambda2[j,k],(ncol(D)/mf))),nrow=
1))
      }
      lambdaA=lambdaA[,-1]
      Sk=pinv(t(W)%*%W+n*lambdaA*D)
      B1=pinv(I1-(pinv(t(X)%*%X)%*%t(X)%*%W%*%Sk%*%t(W)%*%X))
      B2=pinv(t(X)%*%X)%*%t(X)%*%(I2-W%*%Sk%*%t(W))
      B3=B1%*%B2
      A1=Sk%*%t(W)%*%(I2-X%*%B3)
      C=W%*%A1+X%*%B3
      yhat=C%*%y# mencari nilai yhat
SSE[i,j]=sum((y-yhat)^2)#mencari SSE
      SSR[i,j]=sum((yhat-mean(y))^2)#mencari SSR
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
      SST[i,j]=SSR[i,j]+SSE[i,j]#mencari SST
      Rsq[i,j]=(SSR[i,j]/(SST[i,j]))*100#mencari R square
      MSE[i,j]=(t(y)%*%t(I2-C)%*%(I2-C)%*%y)/n#mencari MSE
      GCVbwh=sum(diag(I2-C))/n
      GCV[i,j]=MSE[i,j]/GCVbwh  ##mencari GCV
    }
  }
GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)

  #menyiapkan output sort GCV terkecil
  R11=matrix(ncol=1);       colnames(R11)<-"Rsq"
  MSE11=matrix(ncol=1);     colnames(MSE11)<-"MSE"
  GCV11=matrix(ncol=1);     colnames(GCV11)<-"GCV"
  ke11=matrix(ncol=1);colnames(ke11)<-"Knot_ke"
  for ( i in 1:nl)
  {
    ke11=rbind(ke11,matrix(c(1:a3),ncol=1))
    R11=rbind(R11,matrix(Rsq[,i],ncol=1))
    MSE11=rbind(MSE11,matrix(MSE[,i],ncol=1))
    GCV11=rbind(GCV11,matrix(GCV[,i],ncol=1))
  }
  ke11=matrix(ke11[-1,],ncol=1)
  R11=matrix(R11[-1,],ncol=1)
  MSE11=matrix(MSE11[-1,],ncol=1)
  GCV11=matrix(GCV11[-1,],ncol=1)
  kelam11=matrix(ncol=1);colnames(kelam11)<-"lambda_ke"
  for ( i in 1:nl)
  {
kelam11=rbind(kelam11,matrix(rep(i,a3),ncol=1))
  }
  kelam11=matrix(kelam11[-1,],ncol=1)

  knot11=matrix(nrow=nrow(ke11),ncol=ncol(knot1));
  L11=matrix(nrow=nrow(ke11),ncol=ncol(lambda2));
  for ( i in 1:nrow(ke11))
  {
    knot11[i,]=knot1[ke11[i],]
    L11[i,]=lambda2[kelam11[i],]
  }
  gab11=cbind(GCV11,R11,MSE11,L11,ke11,knot11)
  gab11copy=gab11
  dataprint=matrix(nrow=nprint,ncol=ncol(gab11copy)) #gcv, Rsq, mse, 
lambda, knot
  for (i in 1:nprint)
  {
    s1=min(gab11copy[,1])
    pos=which(s1==gab11copy[,1])
    if (length(pos)==1) pos=pos else pos=randsample(pos,1)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
    dataprint[i,]=gab11copy[pos,]
    posdellambda=0
    for (j in 1:mf)
    {

posdellambda=c(posdellambda,which(gab11copy[pos,(3+j)]==gab11copy[,(
3+j)]))
    }
    posdellambda=posdellambda[-1]
    posdelknot  =which(gab11copy[pos,4+mf]==gab11copy[,4+mf])
    gab11copy=gab11copy[-c(posdellambda,posdelknot),]
  }
dataprint=dataprint[order(dataprint[,1],-
dataprint[,2],dataprint[,3],dataprint[,5]),]
  colnames(dataprint)<-
c("GCV","Rsq","MSE",paste0('lambda',c(1:mf)),"knot_ke",
                         paste0('_',c(1:(3*mt))))
  #save hasil menggunakan file csv
  write.csv(gab11,file=paste(name,"output data All 3 knot
k=",kf,".csv"))
  write.csv(dataprint,file=paste(name,"output data yang di print 3
knot k=",kf,".csv"))
#print output
  cat("==============================================","\n")
  cat("HASIL GCV terkecil dengan fourier-Spline 3 knot k=",kf,"\n")
  cat("==============================================","\n")
  cat("GCV     = ",dataprint[1,1],"\n")
  cat("Rsquare = ",dataprint[1,2],"\n")
  cat("MSE     = ",dataprint[1,3],"\n")
  cat("Lambda  = ",dataprint[1,c(4:(4+mf-1))],"\n")
  cat("Knot    = ",dataprint[1,(4+mf+1):ncol(dataprint)],"\n")
  cat("Nilai GCV 25 ","\n")
  print(dataprint)
#membuat plot antara GCV vs Lambda pada posisi lambda optimum
  pos=which(min(gab11[,1])==gab11[,1])
  if (length(pos)==1) pos=pos else pos=randsample(pos,1)
  #posknot  =which(gab11[pos,5]==gab11[,5])
  posknot=which(dataprint[1,4+mf]==gab11[,4+mf])
  logl=matrix(ncol=mf,nrow=nrow(gab11))
  for ( j in 1:(mf))
  {
temp=0
    temp=which(dataprint[1,3+j]==gab11[,3+j])
    logl[1:length(temp),j]=temp
  }
  mypath <- file.path(paste(name,"plot_3 knot K=",kf,".jpg", sep =
""))
  jpeg(file=mypath,width=1104,height=694)
  par(mfrow=c((ceiling(mf/2)),2))
  for ( i in 1:mf)
  {
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
    posl=c(na.omit(c(logl[,-i])))
    gabpos=c(posl,posknot)
    if (mf==1) posfix=gabpos else
posfix=gabpos[duplicated(c(posl,posknot))]
    while ((anyDuplicated(posfix))!=0)
    {
      posfix=posfix[duplicated(posfix)]
    }
pos2=which(pos==posfix)
    posbawah=pos2-(min(10,pos2))+1
    posatas=pos2+(min(10,length(posfix)-pos2))
    xyplot=gab11[posfix[posbawah:posatas],c(1,3+i)]
    xyplot=xyplot[order(xyplot[,2]),]
    gcvplot=xyplot[,1];
    lambdaplot=xyplot[,2]

plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='b
',main=paste("Plot_ 1 knot, k=",kf,"lambda ke ",i))
  }
  #save plot GCV vs Lambda pada posisi lambda optimum
  dev.off()
  par(mfrow=c(1,1))


  #test
  #lambdaplot=gab11[posknot,4]
  #gcvplot=gab11[posknot,1]
  #par(mfrow=c(3,2))

#plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",type='
b',main=paste("knot1 optimum k=",kf))

  #membuat matrix x untuk estimasi selanjutnya
  lambda1=dataprint[1,c(4,(4+mf-1))]
  knotgcv=matrix(dataprint[1,(4+mf+1):ncol(dataprint)],nrow=1)
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:n)
    {
      if (dataA[k,b]<knotgcv[j]) data1[k,j]=0 else
data1[k,j]=dataA[k,b]-knotgcv[j]
    }
  }
  X=cbind(satu,datapar,dataA,data1)

list(D=D,knot=knotgcv,mingcv=dataprint[1,1],Rsqgcv=dataprint[1,2],W=
W,X=X,lambda=lambda1)
}
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
GCV_kom=function(data,kf,name,xknot,mf,mt)
  {
      #fungsi GCV untuk kombinasi knot
  #data adalah data yang digunakan
  #kf adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  # xknot adalah titik knot pada GCV 1,2 dan 3
  #mf adalah banyanya variabel fourier
  #mt adalah banyanya variabel spline


      #kf= number of k
    options(digits=6)
    #mf=2 #banyaknya variabel fourier
    #mt=2 #banyaknya variabel spline
    mp=ncol(data)-(1+mf+mt)#banyaknya variabel parametrik
      m1=mt

    data=as.matrix(data)
    y=as.matrix(data[,1])     #inisiasi vektor y
    x=as.matrix(data[,-1])
    n=nrow(y)
    m=ncol(x)
    dataA=as.matrix(x[,(m+1-mt):m]) #var non Par
    datapar=as.matrix(x[,(mf+1):(mf+mp)]) #var parametrik
    F=diag(n)
    xknot=as.matrix(xknot)
    m=0

      #mencari titik titik kombinasi
    a=loop(m1)$a

      #menginisiasi nilai titik titik lambda
      temp1=c()
      for ( i in 1:12)
      {
        temp0=seq((10^(-(1+i))),10^(-(i)),by=(2*10^(-(2+i))))
        temp1=c(temp1,temp0)
      }
      lambda1=as.matrix(sort(unique(c(seq(0.1,1,by=0.01),temp1))))
#    lambda1=as.matrix(c(seq(10^(-9),10^(-
6),by=0.00000002),seq(10^(-6),10^-3,by=0.000002),seq(10^(-3),10^-
2,by=0.0002)))
    nl=length(lambda1)
    a3=((3^m1)-3)
#membuat matrix kosong untuk nilai GCV,MSE,Rsq,SSE,SST,SSR
    Rsq=matrix(nrow=((3^m1)-3),ncol=nl)
    GCV=matrix(ncol=nl,nrow=((3^m1)-3))
    MSE=matrix(ncol=nl,nrow=((3^m1)-3))
    SSE=matrix(ncol=nl,nrow=((3^m1)-3))
    SSR=matrix(ncol=nl,nrow=((3^m1)-3))
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
SST=matrix(ncol=nl,nrow=((3^m1)-3))
lambda=matrix(ncol=nl);colnames(GCV)<-lambda1;colnames(MSE)<-
lambda1;colnames(Rsq)<-lambda1
    for(i in 1:((3^m1)-3))
    {
      lambda=rbind(lambda,t(lambda1))
    }
    lambda=lambda[-1,]
    kkkkom=matrix(nrow=((3^m1)-3),ncol=3*m1)
#persamaan deret fourier
    xfou=as.matrix(x[,1:mf])
    W=matrix(nrow=n)
    D=matrix(0,ncol=1,nrow=1)
    for (i in (1:mf))
    {
      der=fungsik(kf,xfou[,i])
      W=cbind(W,der$W)
      D=blkdiag(D,der$D)
    }
    W=as.matrix(W[,-1])
    D=as.matrix(D[-1,-1])
    ma=0
for (i in 1:((3^m1)-3))
    {
      im=0
      for(i2 in 1:m1)
      {
        #menghitung nilai knot kondisional mengikuti kombinasi
            for (h in 1:1)
          if (a[i,i2]==1)
          {
            gab=as.matrix(xknot[i2,1])
            gen=as.matrix(dataA[,i2])
            aa=matrix(nrow=n,ncol=1)
            for (j in 1:1)
              for (w in 1:n)
                if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]     #menghitung knot
          }
        else
if (a[i,i2]==2)
          {
            gab=as.matrix(xknot[i2,2:3]) ;gab=t(gab)
            gen=as.matrix(cbind(dataA[,i2],dataA[,i2]))
            aa=matrix(nrow=n,ncol=2)
            for (j in 1:2)
              for (w in 1:n)
                if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]

          }
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
        else
        {
          gab=as.matrix(xknot[i2,4:6]);gab=t(gab)
          gen=as.matrix(cbind(dataA[,i2],dataA[,i2],dataA[,i2]))
          aa=matrix(nrow=n,ncol=3)
          for (j in 1:3)
            for (w in 1:n)
              if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]

        }

        if (im==0) {ma=aa;im=1} else {ma=cbind(ma,aa)}
      }
satu=rep(1,n)
      X=cbind(satu,datapar,dataA,na.omit(ma))
      #X=cbind(satu,dataA,dataA^2,na.omit(ma))
      I1=diag(ncol(X))
      I2=diag(nrow(X))
      for (j in 1:nl)
      {
          #menghitung nilai estimasi
        Sk=pinv(t(W)%*%W+n*lambda[i,j]*D)
        B1=pinv(I1-(pinv(t(X)%*%X)%*%t(X)%*%W%*%Sk%*%t(W)%*%X))
        B2=pinv(t(X)%*%X)%*%t(X)%*%(I2-W%*%Sk%*%t(W))
        B3=B1%*%B2
        A1=Sk%*%t(W)%*%(I2-X%*%B3)
        C=W%*%A1+X%*%B3
        yhat=C%*%y# mencari nilai yhat
        SSE[i,j]=sum((y-yhat)^2)#mencari SSE
        SSR[i,j]=sum((yhat-mean(y))^2)#mencari SSR
        SST[i,j]=SSR[i,j]+SSE[i,j]#mencari SST
        Rsq[i,j]=(SSR[i,j]/(SST[i,j]))*100#mencari R square
        MSE[i,j]=(t(y)%*%t(I2-C)%*%(I2-C)%*%y)/n#mencari MSE
        GCVbwh=sum(diag(I2-C))/n
        GCV[i,j]=MSE[i,j]/GCVbwh##mencari GCV
      }
#menggabungkan knot kombinasi dengan titik kombinasi
        kkk=matrix(nrow=1,ncol=3*m1)
      for (i3 in 1:m1)
      {
        if (a[i,i3]==1)
        {s=xknot[i3,1];
         ns1=length(s);
         kkk[1,(3*i3-2):((3*i3-2)+ns1-1)]=s
        }else
          if (a[i,i3]==2)
{s=xknot[i3,2:3]
           ns1=length(s);
           kkk[1,(3*i3-2):((3*i3-2)+ns1-1)]=s
          }else
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
          {s=xknot[i3,4:6]
           ns1=length(s);
           kkk[1,(3*i3-2):((3*i3-2)+ns1-1)]=s}
      }
      kkk=as.matrix(kkk)
      ckkk=length(kkk)
      kkk=t(kkk)
      kkkkom[i,1:ckkk]=kkk
    }
#menyiapkan output sort GCV terkecil
  R11=matrix(Rsq,ncol=1);       colnames(R11)<-"Rsq"
  L11=matrix(lambda,ncol=1);    colnames(L11)<-"Lambda"
  MSE11=matrix(MSE,ncol=1);     colnames(MSE11)<-"MSE"
  GCV11=matrix(GCV,ncol=1);     colnames(GCV11)<-"GCV"
  ke11=matrix(rep(1:a3,nl),ncol=1);  colnames(ke11)<-"Kombinasi Knot
ke"
  knot11=matrix(nrow=(a3*nl),ncol=ncol(kkkkom))
  n1=nrow(GCV)
  nprint=(min(25,n1))
  for (i in 1:nl)
  { knot11[(i*a3-(a3-1)):(i*a3),]=(kkkkom)   }
  gab11=cbind(GCV11,R11,MSE11,L11,ke11,knot11)
  gab11copy=gab11
  dataprint=matrix(nrow=nprint,ncol=ncol(gab11copy)) #gcv, Rsq, mse,
lambda, knot
  for (i in 1:nprint)
  {
s1=min(gab11copy[,1])
    pos=which(s1==gab11copy[,1])
    if (length(pos)==1) pos=pos else pos=randsample(pos,1)
    dataprint[i,]=gab11copy[pos,]
    posdellambda=which(gab11copy[pos,4]==gab11copy[,4])
    posdelknot  =which(gab11copy[pos,5]==gab11copy[,5])
    gab11copy=gab11copy[-c(posdellambda,posdelknot),]
  }
  dataprint=dataprint[order(dataprint[,1],-
dataprint[,2],dataprint[,3],dataprint[,5]),]
  dataprint1=as.matrix(dataprint[,1]);colnames(dataprint1)<-"GCV"
  dataprint2=as.matrix(dataprint[,2]);colnames(dataprint2)<-"Rsq"
  dataprint3=as.matrix(dataprint[,3]);colnames(dataprint3)<-"MSE"
  dataprint4=as.matrix(dataprint[,4]);colnames(dataprint4)<-"Lambda"
  dataprint5=as.matrix(dataprint[,5]);colnames(dataprint5)<-
"knot_ke"

dataprint=cbind(dataprint1,dataprint2,dataprint3,dataprint4,dataprin
t5,dataprint[,6:ncol(dataprint)])
#save hasil menggunakan file csv
  write.csv(gab11,file=paste(name,"output data All kombinasi knot
k=",kf,".csv"))
  write.csv(dataprint,file=paste(name,"output data yang di print
kombinasi knot k=",kf,".csv"))
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
#print output
cat("==============================================","\n")
      cat("HASIL GCV terkecil dengan fourier-Spline kombinasi knot
k=",kf,"\n")
      cat("=============================================","\n")
      cat("GCV     = ",dataprint[1,1],"\n")
      cat("Rsquare = ",dataprint[1,2],"\n")
      cat("MSE     = ",dataprint[1,3],"\n")
      cat("Lambda  = ",dataprint[1,4],"\n")
      cat("Kombinasi Knot   = ",a[dataprint[1,5],],"\n")
      cat("Knot    = ",dataprint[1,6:ncol(dataprint)],"\n")
      cat("Nilai GCV ","\n")
      print(dataprint)

      #membuat plot antara GCV vs Lambda pada posisi lambda optimum
      pos=which((min(gab11[,1])==gab11[,1])&(dataprint[1,5]==gab11[,
5]))
      #pos=which((min(gab11[,1])==gab11[,1]))
      posknot  =which(dataprint[1,5]==gab11[,5])
      pos2=which(pos==posknot)
      posbawah=pos2-(min(10,pos2))+1
      posatas=pos2+(min(10,length(posknot)-pos2))
      lambdaplot=gab11[posknot[posbawah:posatas],4]
      gcvplot=gab11[posknot[posbawah:posatas],1]

      #save plot GCV vs Lambda pada posisi lambda optimum
      mypath <- file.path(paste(name,"plot_knot kombinasi
K=",kf,".jpg", sep = ""))
      jpeg(file=mypath,width=1104,height=1104)
      plot((lambdaplot),gcvplot,col="red",ylab="GCV",xlab="lambda",t
ype='b',main=paste("Plot_ knot kombinasi, k=",kf))
      dev.off()
      #membuat matrix x untuk estimasi selanjutnya
      com=as.matrix(a[dataprint[1,5],])
      com=t(com)
      im=0
      for(i2 in 1:m1)
      {
        for (h in 1:1)
            if (com[,i2]==1)
            {
              gab=as.matrix(xknot[i2,1])
              gen=as.matrix(dataA[,i2])
              aa=matrix(nrow=n,ncol=1)
  for (j in 1:1)
                  for (w in 1:n)
                    if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]

            }
        else
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
      if (com[,i2]==2)
            {
              gab=as.matrix(xknot[i2,2:3]) ;gab=t(gab)
              gen=as.matrix(cbind(dataA[,i2],dataA[,i2]))
              aa=matrix(nrow=n,ncol=2)
              for (j in 1:2)
                  for (w in 1:n)
                    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(xknot[i2,4:6]);gab=t(gab)
            gen=as.matrix(cbind(dataA[,i2],dataA[,i2],dataA[,i2]))
            aa=matrix(nrow=n,ncol=3)
            for (j in 1:3)
              for (w in 1:n)
                  if (gen[w,j]<gab[h,j]) aa[w,j]=0 else
aa[w,j]=gen[w,j]-gab[h,j]

        }
        if (im==0) {ma=aa;im=1} else {ma=cbind(ma,aa)}
      }
      X=cbind(satu,datapar,dataA,na.omit(ma))
list(D=D,knot=dataprint[1,6:ncol(dataprint)],mingcv=s1,Rsqgcv=datapr
int[1,2],lambda=dataprint[1,4],W=W,X=X,comknot=a[dataprint[1,5],])
  }

  GCV=function(data,kf,name,mf,mt)
  {
    #fungsi utama untuk memanggil fungsi yang lain
  #data adalah data yang digunakan
  #kf adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  #mf adalah banyanya variabel fourier
  #mt adalah banyanya variabel spline
      library(gtools)
library(pracma)
    cat("\n","Nilai GCV untuk K=",kf,"\n","\n")
    G1=GCV1(data,kf,name,mf,mt)      #running GCV 1 knot
    G2=GCV2(data,kf,name,mf,mt)      #running GCV 2 knot
    G3=GCV3(data,kf,name,mf,mt)      #running GCV 3 knot
#if(mt!=1) { #kondisi ketika hanya ada 1 variabel spline knot
kombinasi tidak akan dirunning
      #knot1k1=cbind(t(matrix((G1$knot),nrow=1)),t(matrix((G2$knot),
nrow=2)),t(matrix((G3$knot),nrow=3)))
    ##xknot=as.matrix(knot1k1)
    #Gk=GCV_kom(data,kf,name,xknot,mf,mt) #running knot kombinasi
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
#minGCV=min(as.matrix(cbind(G1$mingcv,G2$mingcv,G3$mingcv,Gk$mingcv)
))
    #} else
{minGCV=min(as.matrix(cbind(G1$mingcv,G2$mingcv,G3$mingcv,10^8)))}

      minGCV=min(as.matrix(cbind(G1$mingcv,G2$mingcv,G3$mingcv,10^8)
))
      #mencari nilai GCV minimum dari semua titik knot ( 1 knot, 2
knot, 3 knot )
      k=kf;
    if (minGCV==G1$mingcv) {
      W=G1$W ; X=G1$X;lambda=G1$lambda  ;knot1=G1$knot;D=G1$D;ink='1
knot';comknot=1
    }else if (minGCV==G2$mingcv) {
      W=G2$W ; X=G2$X;lambda=G2$lambda  ;knot1=G2$knot;D=G2$D;ink='2
knot';comknot=2
    }else if (minGCV==G3$mingcv) {
      W=G3$W ; X=G3$X;lambda=G3$lambda  ;knot1=G3$knot;D=G3$D;ink='3
knot';comknot=3
    } else {
      W=Gk$W ; X=Gk$X;lambda=Gk$lambda
;knot1=Gk$knot;D=Gk$D;ink='knot Kombinasi';comknot=Gk$comknot}
    #print output
    cat("============================================","\n")
    cat("Nilai GCV terkecil dengan fourier-Spline ","\n")
    cat("============================================","\n")
    cat("GCV        = ",minGCV,"\n")
    cat("knot       = ",ink,"\n")
      cat("Titik Knot = ",knot1,"\n")
    cat("Lambda     = ",lambda,"\n")
    cat("k          = ",k,"\n")
    cat("\n")
    estim=estimasi(data,X,W,D,lambda,k,name,comknot,mf,mt) #running
estimasi setelah titik knot dan lambada optimum diketahui
    C=estim$C ;  res=estim$res ;  Rsq=estim$Rsq ;  yhat=estim$yhat
  }
estimasi=function(data,X,W,D,lambda,k,name,comknot,mf,mt)
  {   #fungi estimasi
  #data adalah data yang digunakan
  # X, W, D, lambda adalah nilai X, W, D, dan lambda pada GCV
terkecil
  #k adalah nilai k yang digunakan
  #name adalah nama folder untuk menyimpan output
  #comknot adalah kombinasi knot yang optimum
  #mf adalah banyanya variabel fourier
  #mt adalah banyanya variabel spline

    #mf=2 #banyaknya variabel fourier
    #mt=2 #banyaknya variabel spline
    mp=ncol(data)-(1+mf+mt)
    m1=mt
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
      #inisiasi nama baris pada beta
    name1=''
    for ( i in 1:mf)
    {name0=c(paste0('b',i),paste0(paste0('a',rep(i,((ncol(W)/mf)-
1))),c(0:k)))
     name1=c(name1,name0)}
    name1=name1[-1]
    name2=paste0('beta',c(0:mp))
    name3=paste0('psi',c(1:mt))
    if (length(comknot)==1) comknot=rep(comknot,mt) else
comknot=comknot
    name14=''
    for ( i in 1:length(comknot))
    {
name04=paste0('alfa',i,c(1:comknot[i]))
      name14=c(name14,name04)
    }
    name4=name14[-1]
    namer=c(name1,name2,name3,name4)


    data=as.matrix(data)
    y=as.matrix(data[,1]);colnames(y)<-"y"  #inisiasi vektor y
    n=nrow(y)
    X=as.matrix(X)
W=as.matrix(W)
    D=as.matrix(D)
    I1=diag(ncol(X))
    I2=diag(nrow(X))

      #menghitung nilai estimasi
          lambdaA=matrix(ncol=1)
        for (k in 1:mf)
        {

lambdaA=cbind(lambdaA,matrix(c(rep(lambda[k],(ncol(D)/mf))),nrow=1))
        }
        lambdaA=lambdaA[,-1]
            Sk=pinv(t(W)%*%W+n*lambdaA*D)
    #Sk=pinv(t(W)%*%W+n*lambda*D)
    B1=pinv(I1-(pinv(t(X)%*%X)%*%t(X)%*%W%*%Sk%*%t(W)%*%X))
    B2=pinv(t(X)%*%X)%*%t(X)%*%(I2-W%*%Sk%*%t(W))
    B3=B1%*%B2
    A1=Sk%*%t(W)%*%(I2-X%*%B3)
estim_alfa=A1%*%y #estimasi deret fourier
    estim_beta=B3%*%y #estimasi spline
    C=W%*%A1+X%*%B3
    yhat=C%*%y;colnames(yhat)<-"yhat"     # mencari nilai yhat
    SSE=sum((y-yhat)^2)#mencari SSE
    SSR=sum((yhat-mean(y))^2)#mencari SSR
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
    SST=SSR+SSE#mencari SST
    Rsq=(SSR/(SST))*100#mencari R square
    MSE=(t(y)%*%t(I2-C)%*%(I2-C)%*%y)/n#mencari MSE
    GCVbwh=sum(diag(I2-C))/n
    GCV=MSE/GCVbwh##mencari GCV
    res=y-yhat;colnames(res)<-"residual"  #mencari residual
    cat("\n")
    cat("\n")
#menggabungkan estimasi deret fourier dan spline

B=matrix(rbind(matrix(estim_alfa,ncol=1),matrix(estim_beta,ncol=1)))
    n1=length(B)
    colnames(B)<-"       estimasi_parameter"
    MSR=SSR/(n1-1) #mencari nilali MSR
#uji F (uji serentak)
    Fhit=MSR/MSE
    pvalue=pf(Fhit,(n1-1),(n-n1),lower.tail=FALSE)

      #uji t (uji individu)
    predic=cbind(W,X)
    thit=rep(NA,n1)
    pval=rep(NA,n1)
    MSE=as.numeric(MSE)
    SE=sqrt(diag(MSE*(pinv(t(predic)%*%predic))))#mencari nilai SE
koefisien
    for (i in 1:n1)
    {
      thit[i]=B[i,1]/SE[i]
      pval[i]=2*(pt(abs(thit[i]),(n-n1),lower.tail=FALSE))
    }
    thit=as.matrix(thit)
    colnames(thit)<-"t_hitung"
    colnames(B)<-"parameter_beta"
#print output
    tg1=cbind(B,thit,pval);rownames(tg1)<-namer
    cat("========================================","\n")
    cat("Estimasi Parameter","\n")
    cat("========================================","\n")
    print(tg1)
    cat("\n")
    cat("Analysis of Variance","\n")
    cat("======================================","\n")
    cat("Source     ","df","\t","SS
","\t","\t","MS","\t","\t","Fhit","\n")
    cat("Regression ",(n1-1),"\t",SSR,"\t",MSR,"\t",Fhit,"\n")
    cat("Error      ",n-n1,"\t",SSE,"\t",MSE,"\n")
cat("Total    ",n-1,"\t",SST,"\n")
    cat("======================================","\n")
    cat("s=",sqrt(MSE),"  Rsq=",Rsq,"\n")
    cat("pvalue(F)=",pvalue,"\n")
    predic=cbind(W,X)
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
    predic1=cbind(y,yhat,res)
#save hasil estimasi
    write.csv(tg1,file=paste(name,"estimasi_parameter k=",k,"
.csv"))
    write.csv(predic,file=paste(name,"estimasi_matriks_W_dan_X
k=",k," .csv"))
    write.csv(predic1,file=paste(name,"estimasi_y,yhatdan residual
k=",k," .csv"))
    #write.csv(C,file=paste(name,"estimasi_C k=",k," .csv"))
list(matx=predic,res=res,yhat=yhat,C=C,Rsq=Rsq,B=B,X=X,W=W,MSE=MSE)
  }

fungsik=function(k,xfour)
{#fungsi untuk mencari D dan W pada fourier
  fou1=as.matrix(xfour)
  n=length(fou1)
  fou2=rep(0.5,length(xfour))
  fouk=matrix(ncol=k,nrow=nrow(fou1))
  for (i in 1:k)
  { fouk[,i]=cos(i*(fou1*(2*pi)/n)) }
#  { fouk[,i]=cos(i*fou1) }
  W=cbind(fou1,fou2,fouk)
cc=mat.or.vec(nc=1,nr=k+2)
  cc[1]=0;cc[2]=0
  for (i in 1:k)
  {   cc[i+2]=i^4   }
  D=diag(cc)
  list(D=D,W=W)
}

loop=function(n)
{
   #fungsi untuk mencari kombinasi nilai knot
   #n adalah banyaknya variabel yang akan dikombinasikan

  #mencari kombinasi yang mungkin.
  a=matrix(ncol=n,nrow=3^(n))
  for (i in 1:n)
  {
    a[,i]=rep(c((rep(1,3^(n-i))),(rep(2,3^(n-i))),(rep(3,3^(n-
i)))),3^(i-1))
  }
  aa=matrix(ncol=1,nrow=3)

  #menghilangkan kombinasi yang sama
  for(i in 1:3)
  {
    for(j in 1:(3^(n)))
    {
    if (all(a[j,]==i)) aa[i,]=j
    }
```

Lampiran 2. Syntax R Model Semiparametrik Campuran (Lanjutan)

```
  }
  a=a[-aa,]
  list(a=a)
}

a=GCV(data,kf,name,mf,mt)
```

BIOGRAFI PENULIS

Khaerun Nisa, Lahir di Bulukumba, Sulawesi Selatan pada tanggal 5 Desember 1991. Penulis merupakan anak dari pasangan Bapak Muh.Ramli dan Ibu Hasmi. Penulis memulai jenjang pendidikan formal di Sekolah Dasar Negeri 122 Batupanyu, Kabupaten Bulukumba pada tahun 1998-2004. Kemudian penulis melanjutkan pendidikan di Sekolah Menengah Pertama Negeri 2 Herlang tahun 2004-2007. Pada tahun yang sama penulis melanjutkan pendidikan ke Sekolah Menengah Atas Negeri 1 Herlang pada tahun 2007-2010. Penulis kemudian melanjutkan pendidikan Sarjana (S1) di Universitas Negeri Makassar (UNM) pada Program Studi Matematika dan selesai pada tahun 2014. Pada tahun 2015 Penulis melanjutkan studi Pasca Sarjana (S2) di Institut Teknologi Sepuluh Nopember Program Studi Statistika. Segala kritik dan saran yang berkaitan dengan tesis ini dapat dikirim melalui email : khaerunnisaramli@yahoo.co.id.