



PENDEKATAN REGRESI KONTINUM FUNGSIONAL KEKAR DALAM PEMODELAN KALIBRASI

ISMAH



**PROGRAM STUDI STATISTIKA DAN SAINS DATA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI DISERTASI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa disertasi dengan judul “Pendekatan Regresi Kontinum Fungsional Kekar Dalam Pemodelan Kalibrasi” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir disertasi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus Tahun 2024

Ismah
G161180051

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



RINGKASAN

ISMAH. Pendekatan Regresi Kontinum Fungsional Kekar Dalam Pemodelan Kalibrasi. Dibimbing oleh ERFIANI, AJI HAMIM WIGENA, dan BAGUS SARTONO.

Kalibrasi peubah ganda merupakan proses menghubungkan sekumpulan data hasil pengukuran yang relatif mudah dan murah diperoleh dengan sekumpulan data hasil pengukuran yang sukar dan mahal memperolehnya. Tujuan kalibrasi peubah ganda berdasarkan sudut pandang statistika adalah menemukan model $E(Y) = f(\mathbf{X})$, untuk prediksi Y dengan akurasi dan presisi yang tinggi.

Salah satu penerapan kalibrasi dalam bidang kesehatan adalah pemodelan hubungan antara kadar glukosa darah yang diukur dengan teknik invasif dan non-invasif berupa spektrum yang menggambarkan sejauh mana cahaya inframerah atau elektromagnetik diserap oleh molekul-molekul glukosa dalam jaringan kulit. Data ini mencakup informasi tentang intensitas cahaya pada berbagai panjang gelombang. Pola spektrum diharapkan dapat menduga kadar glukosa darah.

Teknik yang digunakan untuk mengukur glukosa darah pada alat non-invasif adalah spektroskopi. Hasil keluaran spektroskopi memungkinkan adanya pergeseran pencaran, karena objek yang sama diukur beberapa kali tidak tepat menghasilkan spektrum yang sama. Ukuran objek memiliki pengaruh yang nyata terhadap spektrum. Spektrum yang keluar sering kali tidak sesuai dengan semestinya, hal ini terjadi karena pencaran cahaya yang disebabkan sifat fisik dan kimiawi dari objek yang diamati. Pengaruh yang muncul akibat sifat fisik dan kimiawi objek dapat menyebabkan penyimpangan cahaya, ketidaklinieran pencaran, dan ketidakkonsistenan respon. Hal inilah yang mendasari dibutuhkan suatu metode prapemrosesan, sehingga diperoleh data yang konsisten dan mulus (*smooth*). Berdasarkan hasil kajian literatur diperoleh metode *Wavelet* menunjukkan hasil pendugaan model yang lebih baik.

Data fungsional muncul sebagai hasil dari pengukuran atau pengamatan dari fenomena yang berubah atau bervariasi sepanjang waktu, ruang, atau dimensi lainnya. Penyebab utama munculnya data fungsional adalah sifat alami dari fenomena atau proses yang diamati seperti pada teknik spektroskopi. Regresi kontinum fungsional (RKF) merupakan pengembangan dari regresi kontinum yang dapat digunakan untuk mengatasi bila jumlah peubah bebas (p) lebih besar daripada jumlah amatan ($p \gg n$) dan data berbentuk fungsional. Pengukuran memungkinkan adanya nilai-nilai pencilan (*outlier*) dalam data yang diperoleh, sedangkan RKF berdasarkan kajian simulasi maupun empirik tidak resisten terhadap pencilan, sehingga perlu dikembangkan RKF yang kekar (RKFK).

Performa metode dalam mengatasi $p \gg n$ dan data fungsional, dalam penelitian ini dieksplorasi terhadap kelima metode yang diklasifikasi menjadi dua bagian, metode fungsional dan non fungsional. Metode fungsional yang diterapkan yaitu RKF yang diterapkan dengan mengkombinasikan *Wavelet* (RKF-*Wavelet*) dan Regresi fungsional (RF), sedangkan metode non fungsional yaitu regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), dan regresi kuadrat terkecil (RKT). Performa masing-masing metode dikaji menggunakan data simulasi, peubah bebas dibangkitkan dari fungsi sinus yang dikolaborasikan dengan distribusi seragam (*uniform*) dengan $p = 365$ dan jumlah pengamatan yang beragam

(n) yaitu 50, 100, dan 200. Kelima metode juga dikaji kekekarannya, dengan teknik simulasi data, yang dibangkitkan besar persentase bobot pencilan dari 2% hingga 10%.

Hasil kajian simulasi menunjukkan semakin bertambah jumlah data (n) performa dari kelima metode semakin baik, yang ditunjukkan dengan menurunnya nilai *root mean squares error of prediction* (RMSEP) dan *mean absolute error* (MAE), serta meningkatkan nilai korelasi antara peubah amatan dan peubah dugaan. Analisis secara menyeluruh disimpulkan bahwa RKF-*Wavelet* mempunyai ukuran kebaikan model yang lebih unggul dibandingkan dengan keempat metode lainnya.

Penerapan pendekatan RKF-*Wavelet* pada data pengukuran glukosa darah memberikan hasil RMSEP 86,57, korelasi 0,1718, MAE 58,11, dan *mean absolute percentage error* (MAPE) 41,48. Nilai yang diperoleh lebih kecil dibandingkan dengan RF, RKU, RKTP dan RKT.

RKFK yang dikembangkan dievaluasi menggunakan data simulasi dan diperoleh ukuran kebaikan model yang lebih baik dibandingkan kelima metode RKF, RF, RKU, RKTP, dan RKT. Penerapan RKFK pada data pengukuran glukosa darah, diperoleh hasil RMSEP 33,25, korelasi 0,92, MAE 16,36, dan MAPE 15,01, nilai-nilai tersebut lebih baik dibandingkan dengan kelima metode lainnya.

Performa RKFK tergantung pada fungsi pembobotan yang dipilih, sehingga dalam penelitian ini dilakukan analisis terhadap RKFK dengan empat fungsi pembobot yang umum digunakan adalah Huber, Hampel, Ramsay, dan Tukey (*Bisquare*). Analisis dilakukan dengan mengkaji pada data simulasi dan empirik, dan diperoleh hasil RKFK dengan fungsi pembobot Huber menunjukkan performa yang lebih baik dibandingkan dengan tiga fungsi pembobot lainnya.

Kata kunci: kalibrasi, regresi fungsional, regresi kontinum fungsional, regresi kontinum kekar





SUMMARY

ISMAH. Robust Functional Continuum Regression Approach in Calibration Modelling. Supervised by ERFIANI, AJI HAMIM WIGENA, and BAGUS SARTONO.

Multiple variable calibration is the process of correlating a set of measurement data that is relatively easy and cheap to obtain with a set of measurement data that is difficult and expensive to obtain. The goal of multiple variable calibration from a statistical point of view is to find a model $E(Y) = f(\mathbf{X})$, to predict Y with high accuracy and precision.

One application of calibration in healthcare is modeling the relationship between blood glucose levels measured by invasive and non-invasive techniques in the form of spectra that describe the extent to which infrared or electromagnetic light is absorbed by glucose molecules in skin tissue. This data includes information about the intensity of light at various wavelengths. The spectrum pattern is expected to predict blood glucose levels.

The technique used to measure blood glucose in non-invasive devices is spectroscopy. The output of spectroscopy allows for a shift in scatter, as the same object measured several times does not exactly produce the same spectrum. The size of the object has a noticeable influence on the spectrum. The spectra that come out are often not what they should be, this occurs due to the scattering of light caused by the physical and chemical properties of the object being observed. Influences that arise due to the physical and chemical properties of the object can cause light deviation, nonlinearity of scattering, and inconsistency of response. This is what underlies the need for a preprocessing method, so that consistent and smooth data is obtained. Based on the results of the literature review, the Wavelet method shows better model estimation results.

Functional data arises as a result of measurements or observations of phenomena that change or vary over time, space, or other dimensions. The main reason for the emergence of functional data is the nature of the observed phenomenon or process such as in spectroscopy techniques. Continuum functional regression (RKF) is an extension of continuum regression that can be used to cope when the number of independent variables (p) is greater than the number of observations ($p \gg n$) and the data is functional. Measurement allows the presence of outlier values in the data obtained, while RKF based on simulation and empirical studies is not resistant to outliers, so it is necessary to develop a robust RKF (RKFK).

The performance of the methods in dealing with $p \gg n$ and functional data, in this study explored the five methods classified into two parts, functional and non-functional methods. The functional methods applied are RKF by combining Wavelet (RKF-Wavelet) and functional regression (RF), while the non-functional methods are principal component regression (RKU), partial least squares regression (RKTP), and least squares regression (RKT). The performance of each method is studied using simulated data, the independent variable is generated from a sinus function collaborated with a uniform distribution with $p = 365$ and a varying number of observations (n) of 50, 100, and 200. The five methods are also studied

for robustness, with data simulation techniques, which are generated with a large percentage of outlier weights from 2% to 10%.

The results of the simulation study showed that as the number of data (n) increased, the performance of the five methods improved, as indicated by decreasing the root mean squares error of prediction (RMSEP) and mean absolute error (MAE) values, as well as increasing the correlation value between observed variables and predicted variables. Overall analysis concluded that RKF-Wavelet has a superior measure of model goodness compared to the other four methods.

The application of the RKF-Wavelet approach to blood glucose measurement data gives the results of RMSEP 86.57, correlation 0.1718, MAE 58.11, and mean absolute percentage error (MAPE) 41.48. The values obtained are smaller than those of RF, RKU, RKTP and RKT.

The developed RKFK was evaluated using simulated data and obtained a better measure of model goodness than the five methods RKF, RF, RKU, RKTP, and RKT. Application of RKFK on blood glucose measurement data, the results obtained RMSEP 33.25, correlation 0.92, MAE 16.36, and MAPE 15.01, these values are better than the other five methods.

The performance of RKFK depends on the chosen weighting function, so in this study, an analysis of RKFK with four commonly used weighting functions is conducted, namely Huber, Hampel, Ramsay, and Tukey (*Bisquare*). The analysis was carried out by examining simulated and empirical data, and the results obtained from RKFK with the Huber weighting function showed better performance compared to the other three weighting functions.

Keywords: calibration, functional regression, functional continuum regression, robust continuum regression.





@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PENDEKATAN REGRESI KONTINUM FUNGSIONAL KEKAR DALAM PEMODELAN KALIBRASI

ISMAH

Disertasi
sebagai salah satu syarat untuk memperoleh gelar
Doktor pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI STATISTIKA DAN SAINS DATA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

@Hak cipta milik IPB University

IPB University





Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji Luar Komisi Pembimbing pada Ujian Tertutup Disertasi:

- 1 Dr. Farit Mochamad Afendi, M.Si
- 2 Dr. Agus Mohamad Soleh, S.Si, M.T

Promotor Luar Komisi Pembimbing pada Sidang Promosi Terbuka Disertasi:

- 1 Dr. Farit Mochamad Afendi, M.Si
- 2 Prof. Dr. Ma'mun Murod, S.Sos., M.Si 



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

**Judul Disertasi: Pendekatan Regresi Kontinum Fungsional Kekar dalam
Pemodelan Kalibrasi**

Nama : Ismah
NIM : G161180051

Disetujui oleh

Pembimbing 1:
Dr. Ir. Erfiani, M.Si _____

Pembimbing 2:
Prof. Dr. Ir. Aji Hamim Wigena, M.Sc _____

Pembimbing 3:
Dr. Bagus Sartono, M.Si _____

Diketahui oleh

Ketua Program Studi:
Dr. Anang Kurnia, S.Si., M.Si _____
NIP. 197308241997021001

**Dekan Fakultas Matematika dan Ilmu Pengetahuan
Alam :**
Dr. Berry Juliandi, S.Si., M.Si _____
NIP. 197807232007011001

Tanggal Ujian: 26 Juni 2024

Tanggal Lulus:

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Oktober 2021 sampai bulan Juni 2024 ini ialah pemodelan kalibrasi, dengan judul “Pendekatan Regresi Kontinum Fungsional Kekar Dalam Pemodelan Kalibrasi”.

Terima kasih penulis ucapkan kepada para pembimbing, Dr. Erfiani, M.Si, Prof. Dr. Aji Hamim Wigena, M.Sc, dan Dr. Bagus Sartono, M.Si yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada Bapak Dr. Anang Kurnia, M.Si selaku Ketua Program Studi, dan penguji luar komisi pembimbing Dr. Farit Mochamad Afendi, M.Si dan Dr. Agus Mohamad Soleh, S.Si., M.T, yang telah memberikan saran dan masukan dalam ujian tertutup penulis.

Ungkapan terima kasih kepada Bapak Dr. Ir. Setiawan, MS dari Institut Teknologi Sepuluh November (ITS) yang telah berkenan memberikan pemahaman tentang regresi kontinum, Dr. Renan Prasta Jenie, STP. MT dari Universitas Binawan sekaligus alumni IPB, dan Umam Hidayaturrohman, S.Stat., M.Si yang menjadi anggota pengembang alat glukosa darah noninvasif, yang telah berkenan memberikan informasi tentang data glukosa darah non invasif. Dr. Faisal, M.Si dari Universitas Bina Nusantara yang telah membantu memahami mengenai teori dengan pendekatan analisis real.

Di samping itu, sebagai bentuk penghargaan penulis tujukan kepada Rektor Universitas Muhammadiyah Jakarta (Prof. Dr. Ma'mun Murod, M.Si) beserta jajaran Wakil Rektor, Dekan Fakultas Ilmu Pendidikan UMJ (Prof. Dr. Iswan, M.Si) beserta jajaran Wakil Dekan dan Kaprodi, Direktur Pascasarjana UMJ (Prof. Dr. Herwina Bahar, MA), rekan-rekan Dosen Prodi Pendidikan Matematika (Bu Mita, Bu Hastri, Bu Viarti, Bu Arlin, dan Bu Ririn), serta Bapak Ibu dosen Fakultas Ilmu Pendidikan UMJ.

Ungkapan terima kasih yang tak terhingga, atas do'a dan dukungan teruntuk ayahanda terhormat (alm. Muhammad Hasyim), ibunda terkasih (alm. Maisaroh), Ibu dan Bapak mertua (Hj. Endeh Sumarni, S.Ag, dan Drs. Abdul Rozak), suami tercinta (Muhamad Farid Fachrurozi, S.Si, M.Pd), anak-anak tersayang (Rafa Mumtaz Aljabbar, Lathifah Aulia Zahwa, dan Alea Nisa Jameela), serta seluruh keluarga besar. Teman-teman STK 2018 (Mba Tonah, Mba vera, Mba Rini, Mba Femmy, Mba Ita, Mba Yenni, Bu Viarti, Mas Hengki dan Mba Dila), dan keluarga besar Pascasarjana STK IPB, terima kasih untuk dukungan dan telah berkenan menjadi tempat diskusi serta sharing pengetahuan.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2024

Ismah

DAFTAR ISI

DAFTAR TABEL	xv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	5
1.4 Manfaat	6
1.5 Ruang Lingkup	6
1.6 Kebaruan (<i>novelty</i>)	7
II TINJAUAN PUSTAKA	10
2.1 Sistem Non-Invasif untuk Pengukuran Glukosa Darah	10
2.2 Transformasi <i>Wavelet</i>	12
2.3 Transformasi <i>Wavelet</i> Kontinu	14
2.4 Regresi Fungsional	15
2.5 Regresi Kontinum	16
2.6 Regresi Kontinum Fungsional	17
2.7 Fungsi Pembobot	18
III PENDEKATAN REGRESI KONTINUM FUNGSIONAL PADA DATA HASIL TRANSFORMASI WAVELET DALAM MODEL KALIBRASI PENGUKURAN GLUKOSA NON-INVASIF	21
3.1 Pendahuluan	21
3.2 Peringkasan Data	22
3.3 Eksplorasi Data	23
3.4 Peringkasan Data	26
3.5 Pemodelan Kalibrasi Glukosa Darah Non-Invasif	26
3.6 Kesimpulan	29
IV KAJIAN KEKEKARAN BERBAGAI METODE PEMODELAN KALIBRASI	30
4.1 Pendahuluan	30
4.2 Analisis Kekekaran Berbagai Metode Pemodelan Kalibrasi	31
4.3 Kesimpulan	35
V PENGEMBANGAN METODE REGRESI KONTINUM FUNGSIONAL KEKAR	36
5.1 Pendahuluan	36
5.2 Kajian Teoritis	37
5.3 Usulan Metodologi (Kebaruan)	40
5.4 Evaluasi Kinerja Metodologi yang Diusulkan	41
5.5 Ilustrasi Data Empirik	47
5.6 Kesimpulan	49
VI PERFORMA ANALISIS FUNGSI PEMBOBOT PADA REGRESI KONTINUM FUNGSIONAL KEKAR	50



6.1	Pendahuluan	50
6.2	Metode	51
6.3	Hasil dan Pembahasan	51
VII	PEMBAHASAN UMUM	57
VIII	SIMPULAN DAN SARAN	63
8.1	Simpulan	63
8.2	Saran	63
	DAFTAR PUSTAKA	65
	LAMPIRAN	70
	RIWAYAT HIDUP	89

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1	Rancangan Kondisi Pengukuran Glukosa Darah Non-Invasif	11
2	Fungsi Pembobot	19
3	Deskripsi Data Kadar Glukosa Darah	23
4	Korelasi antar peubah X1-X7	26
5	Deskripsi Hasil Analisis	27
6	Ukuran Keباikan Metode Menggunakan Data Simulasi	32
7	Waktu Komputasi	43
8	Analisis Ragam Ukuran Keباikan Model	44
9	Ukuran Keباikan Model: Data Simulasi	45
10	Ukuran Keباikan dengan Variasi α	47
11	Ukuran Keباikan Model	48
12	Ukuran Keباikan Model RKFK dengan Variasi Fungsi Pembobot	52
13	Ukuran Keباikan Model RKFK: Data Empirik	56

DAFTAR GAMBAR

1	Visualisasi Sampel Hasil Pengukuran Glukosa Non-Invasif	4
2	Visualisasi Sampel Hasil Transformasi dengan <i>Fourier</i>	4
3	Peta Konsep Penelitian	7
4	Bagan Perkembangan Penelitian Data Pengukuran Non-Invasif	9
5	Alur Kerja Alat Purwarupa Non-Invasif	10
6	Spektrum hasil pengukuran Glukosa Darah Non-Invasif	12
7	Visualisasi <i>mother Wavelet</i> Daubechies	14
8	Grafik Fungsi Pembobot	20
9	Visualisasi data kadar glukosa darah	24
10	Visualisasi Kadar Glukosa Darah Responden 17 dan 34	24
11	Visualisasi Kadar Glukosa Dari Responden 65 dan 67	25
12	Visualisasi Data Luasan	25
13	Visualisasi Data Hasil Pemulusan dengan Transformasi <i>Wavelet</i> pada Responden 1- Responden 5	26
14	Grafik data aktual dengan nilai dugaan	29
15	Visualisasi Ukuran Keباikan Model Data Simulasi	34
16	Eksplorasi Data Simulasi dan Empirik, (a) Peubah X data simulasi, (b) Peubah X data empirik, (c) Peubah Y data Simulasi, (d) Peubah Y data Empirik	42
17	Diagram pencar ukuran keباikan model	46
18	Diagram pencar antara data aktual dengan data dugaan model	48
19	Ukuran Keباikan Model: Data simulasi, (a) RMSEP, (b) korelasi, dan	53
20	Visualisasi Level data Glukosa Darah	54
21	Visualisasi Data Residu Intensitas (a) Antar Responden 1, 2, 3, dan 4, (b) Antar Ulangan 1 sampai 5 (U1, U2, U3, U4, U5) pada Responden 1	55
22	Visualisasi Data Hasil Ringkasan	55
23	Visualisasi Data Aktual dan Data Dugaan	56

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR LAMPIRAN

1	Lampiran 1 Sampel data Luasan Pengukuran glukosa dalam darah non invasif (X) sebanyak 8 responden, 5 ulangan dan 100 peubah.	71
2	Lampiran 2 Data peubah Y	75
3	Lampiran 3 Kode program r untuk membangkitkan data fungsional	76
4	Lampiran 4 Definisi, proposisi, dan pembuktiannya	76
5	Lampiran 5 Ukuran Keباikan Model Regresi Komponen Utama (RKU):	79
6	Lampiran 6 Ukuran Keباikan Model Regresi Kuadrat Terkecil Parsial (RKTP):	80
7	Lampiran 7 Ukuran Keباikan Model Regresi Kontinum Fungsional (RKF):	81
8	Lampiran 8 Ukuran Keباikan Model Regresi Kontinum Fungsional Kekar	83
9	Lampiran 9 Daftar Penelitian Data Non invasif Glukosa Darah IPB	84

I PENDAHULUAN

1.1 Latar Belakang

Kalibrasi peubah ganda merupakan proses menghubungkan sekumpulan data hasil pengukuran yang relatif mudah dan murah diperoleh dengan sekumpulan data hasil pengukuran yang sukar dan mahal memperolehnya (Brouckaert *et. al.* 2017). Tujuan kalibrasi peubah ganda berdasarkan sudut pandang statistika adalah menemukan model $E(Y) = f(X)$, untuk prediksi Y dengan akurasi dan presisi yang tinggi.

Beberapa metode yang telah diterapkan dalam model kalibrasi antara lain regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), regresi Ridge, pendekatan Bayes, Jaringan Syaraf Tiruan (JST), dan metode lainnya. Regresi kontinum (RK) adalah salah satu metode yang dapat diterapkan dalam memodelkan kalibrasi serta mengatasi permasalahan yang sering muncul dalam pemodelan kalibrasi.

RK pertama kali dikembangkan oleh Stone dan Brooks (1990). Formulasi RK kemudian dikembangkan oleh Björkström dan Sundberg (1999) yang disebut dengan istilah *Generalized view on continuum regression*. Kelemahan pada RK yaitu tidak resisten terhadap pencilan, menjadi alasan Serneels *et. al.* (2005) mengembangkan RK yang bersifat kekar yang disebut *robust continuum regression*, sedangkan penerapan regresi kontinum untuk data spektroskopi pernah dilakukan oleh Setiawan (2007).

RK merupakan metode yang dikembangkan dari regresi kuadrat terkecil (RKT), RKU, dan RKTP. RKU bekerja dengan membentuk peubah baru (laten) yang merupakan kombinasi linier dari matriks peubah bebas yang mempunyai ragam maksimal. Peubah baru yang terbentuk saling bebas dan memiliki dimensi yang lebih kecil. RKTP membentuk peubah baru dengan memaksimalkan koragam peubah bebas dengan peubah respon. RKTP mengkontruksi faktor yang dapat menerangkan keragaman peubah X dan pada saat yang bersamaan faktor tersebut memiliki hubungan dengan Y , sehingga akan diperoleh jumlah faktor yang lebih sedikit dibandingkan dengan RKU. RK mengkontruksi peubah baru dengan mengkombinasikan prinsip pada RKT, RKU dan RKTP yaitu dengan memaksimalkan keragaman peubah X serta koragam antara peubah X dan Y .

Proses pengukuran data dengan teknik spektroskopi akan diperoleh data spektrum yang termasuk dalam kategori data fungsional. Terbentuknya data fungsional akibat pengukuran atau pengamatan dari fenomena yang berubah atau bervariasi sepanjang waktu, ruang, atau dimensi lainnya. Data fungsional adalah jenis data statistik yang dihasilkan dari fungsi atau proses yang berubah sepanjang waktu atau ruang. Setiap observasi dalam data fungsional, bukanlah sekedar nilai tunggal, tetapi merupakan fungsi kontinu yang menggambarkan perubahan atau variasi dalam suatu fenomena. Data fungsional memiliki karakteristik khusus yang membuatnya sulit diolah dengan metode regresi konvensional yang dirancang untuk data tabular biasa. Data fungsional biasanya mengandung serangkaian fungsi kontinu yang merepresentasikan perubahan sepanjang waktu atau dimensi lainnya, sehingga memerlukan pendekatan analisis yang lebih tepat. Pada umumnya data fungsional memiliki dimensi yang besar, terdapat multikolinearitas atau dengan kata lain peubah bebas mengalami perubahan di sepanjang waktu tertentu sehingga

diduga terdapat interaksi antar peubah, memiliki hubungan tak linier serta bersifat dinamis.

Pendekatan analisis untuk data fungsional yang telah dikembangkan antara lain regresi fungsional, regresi komponen utama fungsional (RKUF), regresi kuadrat terkecil parsial fungsional (RKTPF), dan regresi kontinum fungsional (RKF) yang diperkenalkan oleh Zhou (2019). RKF sebagai pengembangan dari regresi kontinum yang dapat diterapkan pada data fungsional. RKF merupakan bentuk umum dari RKUF dan RKTPF. Hasil studi kasus dengan data simulasi dan numerik, diperoleh RKF lebih akurat pada kondisi data bersifat fungsional, dibandingkan dengan RKUF dan RKTPF baik dalam estimasi dan prediksi, selain itu RKF juga menunjukkan penduga yang konsisten (Zhou, 2019).

Salah satu penerapan model kalibrasi adalah menganalisis hubungan antara kadar glukosa darah hasil pengukuran secara invasif (Y), dengan kadar glukosa darah hasil pengukuran secara non-invasif (X). Pengukuran glukosa darah dengan teknik invasif, merupakan metode konvensional dengan cara pengambilan sampel darah dari pembuluh darah, kemudian dianalisis di laboratorium menggunakan alat pengukur gula darah atau meteran glukosa darah. Pengambilan sampel ini dapat dilakukan dengan jarum suntik yang menembus kulit, hal tersebut tentunya menimbulkan cedera fisik atau perasaan tidak nyaman pada pasien. Pengukuran glukosa darah dengan teknik non-invasif menggunakan prinsip spektroskopi dengan menggunakan cahaya (inframerah), untuk mengukur kadar glukosa tanpa penetrasi kulit. Teknik non-invasif melibatkan analisis cahaya yang melewati jaringan kulit dan berinteraksi dengan glukosa. Data spektrum yang dihasilkan digunakan untuk menghitung kadar glukosa dalam darah.

Metode non-invasif mengurangi rasa tidak nyaman dalam pengambilan sampel darah, namun, non-invasif memiliki sensitivitas yang rendah dan dipengaruhi faktor-faktor eksternal, sehingga alat pengukur gula darah non-invasif masih digunakan bersamaan dengan pengukuran gula darah konvensional untuk memverifikasi akurasi. Pemodelan kalibrasi akan diterapkan yang bertujuan untuk memastikan bahwa alat pengukur non-invasif dapat memberikan hasil yang dapat diandalkan dan relevan dalam mendeteksi penyakit diabetes.

Data hasil pengukuran non-invasif diketahui terdapat data pencilan, hal tersebut berdasarkan pada hasil penelitian Li *et. al.* (2010) yang menerapkan regresi parsial robust M pada pengukuran non-invasif. Rosni (2019) dalam penelitiannya diperoleh data pencilan pada data hasil pengukuran glukosa yang dilakukan oleh Tim Biomarking Institut Pertanian Bogor (IPB) tahun 2017. Berdasarkan hasil penelitian Rosni, maka dilakukan penelitian lanjutan dengan data yang sama oleh Herianti (2020) dengan mengembangkan pemodelan kalibrasi kekar, Rasiyanti (2020) menerapkan analisis regresi ridge robust, dan Novia (2021) menerapkan regresi kuantil komponen utama pada data hasil pengukuran glukosa darah non-invasif.

Hasil kajian Zhou (2019) menyatakan bahwa RKF tidak resisten terhadap data pencilan. RKF berpotensi untuk dikembangkan menjadi metode kekar sehingga resisten terhadap data pencilan. Seperti halnya dengan RK yang tidak resisten terhadap pencilan, sehingga Serneels *et. al.* (2005) mengembangkan regresi kontinum kekar.

Keberadaan pencilan dapat berakibat buruk terhadap hasil analisis. Salah satu upaya untuk menghilangkan data pencilan yaitu dengan memberikan bobot nol

pada data pencilan berdasarkan fungsi pembobot Tukey dan Hampel, sehingga pencilan akan hilang atau dengan kata lain akan menghilangkan data amatan. Upaya tersebut dianggap kurang efektif dalam mengatasi keberadaan pencilan, karena data amatan yang berupa hasil pengukuran diperoleh melalui proses yang cukup lama serta biaya yang mahal, sehingga keputusan untuk menghilangkan data perlu dipertimbangkan lagi. Alternatif lain yang dapat dilakukan untuk mengatasi keberadaan pencilan adalah memberikan nilai bobot pada data yang bersifat proporsional bergantung pada seberapa jauh jarak data pada nilai batasan (*cut off*), seperti fungsi pembobot Huber dan Ramsay. Deteksi pencilan dilakukan berdasarkan nilai galat dari data pengamatan dengan data dugaan, atau menggunakan jarak Mahalanobis. Jarak Mahalanobis adalah metrik yang digunakan untuk mengukur sejauh mana suatu titik dalam ruang fitur (data) dari sebuah pusat atau titik referensi tertentu. Jarak Mahalanobis digunakan dalam berbagai aplikasi analisis data dan pengenalan pencilan, seperti pada metode Robust Principal Component Analysis (RobPCA) (Hubert *et. al.* 2005).

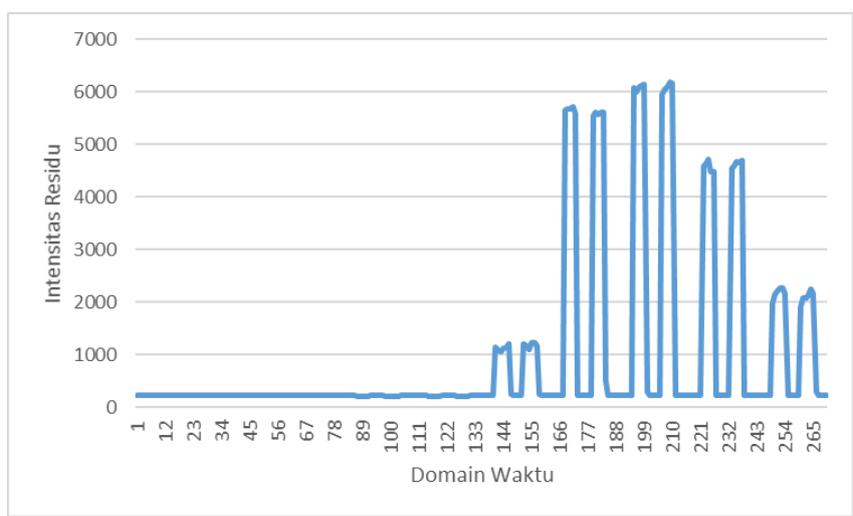
Teknik spektroskopi yang digunakan untuk mengukur glukosa darah pada alat non-invasif memungkinkan adanya pergeseran pencaran, karena objek yang sama diukur beberapa kali tidak tepat menghasilkan spektrum yang sama (Jayanti *et. al.* 2018). Hruschka (1987) menyatakan bahwa ukuran objek memiliki pengaruh yang nyata terhadap spektrum. Spektrum yang keluar sering kali tidak sesuai dengan semestinya, hal ini terjadi karena pencaran cahaya yang disebabkan sifat fisik dan kimiawi dari objek yang diamati (Arnita 2005). Pengaruh yang muncul akibat sifat fisik dan kimiawi objek dapat menyebabkan penyimpangan cahaya, ketidaklinieran pencaran, dan ketidakkonsistenan respon (Blanco *et. al.* 1998). Hal inilah yang mendasari dibutuhkan suatu metode prapemrosesan, sehingga diperoleh data yang konsisten atau mulus (*smooth*).

Alat non invasif glukosa darah yang dikembangkan tim Biomarking IPB pada Tahun 2017 dan 2019, menggunakan *Fourier* untuk mentransformasi data. Transformasi *Fourier* diperkenalkan pertama kali oleh matematikawan Prancis bernama J. *Fourier* tahun 1822, yang kemudian tahun 1965 dikembangkan algoritma *Fast Fourier Transform* (FFT) yaitu algoritma yang memiliki kecepatan lebih tinggi dalam komputasi (Polikar 1997). Metode transformasi *Fourier* digunakan untuk mendekomposisi sinyal kedalam fungsi eksponensial atau kedalam bentuk komponen sinusoidal. Perbandingan data asli (sebelum transformasi) hasil pengukuran non-invasif dengan data hasil transformasi *Fourier* dapat dilihat pada Gambar 1 dan Gambar 2. Tampak bahwa pada Gambar 2 visualisasi sampel hasil transformasi lebih *smooth* dan menunjukkan data fungsional dibandingkan dengan Gambar 1.

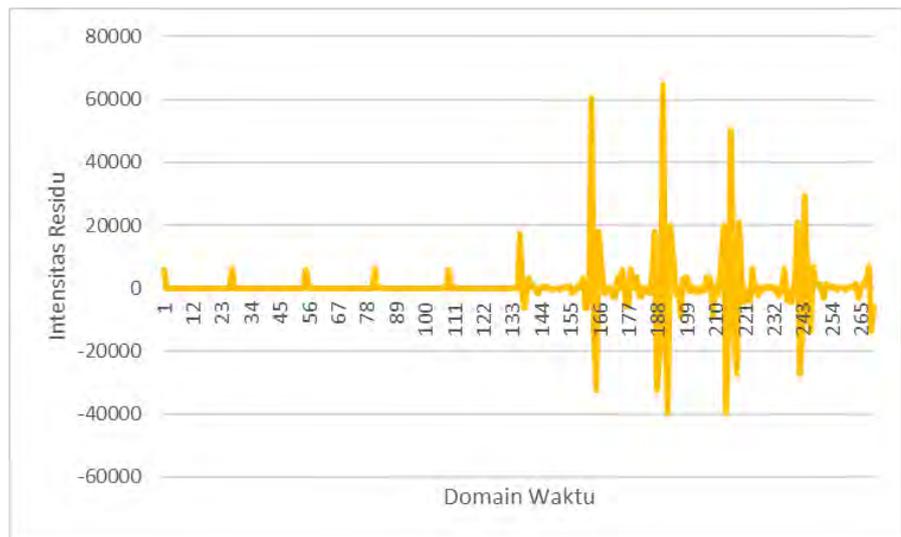
Metode yang populer untuk pemulusan data selain *Fourier*, adalah *Wavelet*, namun berdasarkan hasil kajian literatur diperoleh metode *Wavelet* menunjukkan hasil pendugaan model yang lebih baik. Transformasi *Wavelet* merepresentasikan suatu kurva, seperti spektrum, sebagai kombinasi linear kurva-kurva lain yang relatif lebih sederhana yang disebut fungsi basis atau fungsi *Wavelet* (Fearn 1999). Fungsi basis tersebut diperoleh dengan dilatasi dan translasi dua jenis fungsi *Wavelet* yang disebut *father Wavelet* dan *mother Wavelet* (Nason dan Silverman 1995). Dalam analisis *Fourier* fungsi basis yang digunakan adalah fungsi cosinus, sehingga metode *Wavelet* dapat dipandang sebagai perluasan dari analisis *Fourier*. *Wavelet* berasal dari fungsi skala, dari fungsi ini dapat dibuat sebuah *mother*



Wavelet. *Wavelet- Wavelet* lainnya akan muncul dari hasil penskalaan, dilatasi dan translasi (pergeseran) *mother Wavelet*.



Gambar 1 Visualisasi Sampel Hasil Pengukuran Glukosa Non-Invasif



Gambar 2 Visualisasi Sampel Hasil Transfomasi dengan *Fourier*

Pada penerapan transformasi *Wavelet*, perlu dipilih *mother Wavelet* untuk membangun fungsi basis. Penggunaan *mother Wavelet* yang berbeda akan memberikan hasil analisis yang berbeda. Berdasarkan hasil kajian Jang *et. al.* (2021) fungsi *mother Wavelet* yang optimal adalah Daubechies, sehingga dalam penelitian ini akan digunakan fungsi *mother Wavelet* Daubechies dalam mentransformasi data.

Berdasarkan kajian terhadap penelitian sebelumnya serta permasalahan yang dihadapi, maka akan dilakukan penelitian pemodelan kalibrasi menggunakan pendekatan regresi kontinum fungsional kekar (RKFK), sebagai perkembangan dari RKF pada data hasil pengukuran darah non-invasif setelah dilakukan prapemrosesan dengan mentransformasi data menggunakan *Wavelet*. Tahapan dalam penelitian ini terdiri dari tiga kajian. **Pertama**, pemodelan kalibrasi dengan kelima pendekatan yang terdiri dari metode fungsional yang meliputi RF dan RKF,

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

serta non-fungsional RKT, RKU, dan RKTP pada data hasil transformasi *Wavelet*, hasil kajian diharapkan akan memperoleh metode transformasi dan metode regresi untuk model kalibrasi yang tepat. **Kedua**, melakukan kajian kekekaran dari kelima pendekatan yang dilakukan pada kajian pertama, menggunakan data simulasi data pada berbagai bobot pencilan (2% hingga 10%) dan jumlah pengamatan (50, 100, dan 200). **Ketiga**, pengembangan metode RKFK untuk data hasil transformasi *Wavelet*, dalam kajian ini akan digunakan data simulasi dan data empirik hasil pengukuran gula darah dengan teknik invasif dan non-invasif. Pengembangan metode RKFK akan dieksplorasi menggunakan empat jenis fungsi pembobot yaitu Huber, Hampel, Ramsay, dan Tukey (*Bisquare*), untuk mendapatkan performa RKFK dengan fungsi pembobot yang terbaik.

1.2 Rumusan Masalah

Transformasi data dalam metode spektroskopi dilakukan sebagai proses *smoothing* atau penghalusan data. Tujuan penghalusan data adalah mengurangi *noise* atau fluktuasi acak dalam data spektroskopi sehingga informasi yang signifikan dapat lebih mudah diidentifikasi. Seperti halnya dalam pengukuran kadar glukosa darah non-invasif, metode transformasi yang umum diterapkan pada alat pengukuran adalah *Fourier*. Berdasarkan kajian yang telah dilakukan metode *Wavelet* yang merupakan perkembangan dari metode *Fourier* adalah metode transformasi yang lebih baik dibandingkan dengan *Fourier*. Oleh karena itu, kajian terkait dengan metode transformasi *Wavelet* menjadi perlu dilakukan untuk mengetahui keunggulannya dalam mentransformasi data, dan diperoleh pemodelan kalibrasi yang lebih baik.

Setelah melalui tahapan transformasi, untuk mendapatkan pendugaan terhadap data pengukuran kadar glukosa dalam darah dari dua alat yaitu invasif yang diasumsikan sebagai peubah Y, dan non-invasif yang diasumsikan sebagai peubah X dengan akurasi yang tinggi, maka perlu dibuat pemodelan kalibrasi. Permasalahan yang sering muncul dalam kalibrasi adalah data yang memiliki bersifat fungsional, dan terdapat pencilan. Permasalahan tersebut mengakibatkan regresi dengan metode kuadrat terkecil tidak dapat diterapkan. Penelitian-penelitian yang telah dilakukan pada umumnya hanya mengatasi satu atau beberapa dari permasalahan yang ada pada kalibrasi. Penelitian ini akan mengembangkan model kalibrasi dengan metode yang dimodifikasi untuk mengatasi permasalahan data fungsional dan pencilan, yaitu regresi kontinum fungsional kekar. Tahap validasi model akan dilakukan untuk mengetahui kehandalan dari metode dalam pemodelan.

1.3 Tujuan

Berdasarkan permasalahan yang dihadapi pada pemodelan kalibrasi, pada disertasi ini penulis memfokuskan tujuan penelitian ini sebagai berikut:

- Mengkaji transformasi data dengan *Wavelet* menggunakan *mother Wavelet Daubechies*
- Mengkaji pemodelan kalibrasi dengan pendekatan fungsional dan non fungsional
- Mengkaji kekekaran metode fungsional dan non fungsional
- Mengembangkan metode regresi kontinum fungsional kekar

- e. Mengembangkan model kalibrasi data hasil pengukuran gula darah invasif dan non-invasif menggunakan regresi kontinum fungsional kekar dengan transformasi *Wavelet*.

1.4 Manfaat

Penelitian ini diharapkan memberikan manfaat diantaranya :

1. *Wavelet* dapat diterapkan sebagai metode transformasi data hasil pengukuran, sehingga diperoleh data dengan selang yang lebih lebar yang mengandung informasi yang lebih banyak.
2. Metode hasil pengembangan dapat digunakan sebagai alternatif dalam pemodelan kalibrasi, dan tidak sensitif terhadap permasalahan yang ada pada pemodelan kalibrasi.
3. Model yang dikembangkan akan menjadi alternatif untuk menduga kadar glukosa dalam darah dengan akurasi yang tinggi dan waktu yang relatif lebih cepat, serta tidak menimbulkan rasa sakit atau cedera pada tubuh manusia.

1.5 Ruang Lingkup

Ruang lingkup penelitian ini meliputi pengembangan metode regresi kontinum fungsional kekar (RKFK) yang akan diterapkan pada pemodelan kalibrasi. Pemodelan kalibrasi yang digunakan untuk menganalisis hubungan antara data hasil pengukuran glukosa darah dengan Teknik non-invasif (X) dan dengan teknik invasif (Y). Teknik non-invasif yang dilakukan menggunakan prinsip spektroskopi.

Data hasil pengukuran dengan metode spektroskopi pada umumnya akan ditransformasi dengan tujuan mendapatkan informasi yang lebih berguna atau relevan. Transformasi data dalam spektroskopi mencakup prapemrosesan data dengan cara menghilangkan *noise* dari data dan meratakan spektrum untuk memudahkan analisis selanjutnya. Metode transformasi yang digunakan dalam penelitian ini adalah *Wavelet*. *Wavelet* bekerja efektif bergantung pada fungsi basis yang digunakan, berdasarkan hasil kajian Jang *et. al.* (2021) fungsi basis yang optimal adalah Daubechies, sehingga dalam penelitian ini akan digunakan fungsi *mother Wavelet* Daubechies dalam mentransformasi data.

Hasil pengukuran dengan teknik non-invasif menggunakan metode spektroskopi berupa nilai residu intensitas. Jumlah data residu intensitas yang dimiliki setiap responden berbeda-beda sehingga diperlukan peringkasan data, agar dapat dilakukan analisis data. Peringkasan data dilakukan dengan menentukan nilai luas pada area yang telah terbagi menjadi 20 bagian. Teknik membagi area berdasarkan lama waktu yang dihasilkan pada saat pengukuran gula darah untuk setiap responden.

RKFK yang akan dikembangkan mengacu pada fungsi pembobot, dalam penelitian ini akan dilakukan analisis RKFK dengan keempat jenis fungsi pembobot yaitu Huber, Hampel, Ramsay, dan Tukey (*Bisquare*). Ketangguhan dan sifat efisiensi dari RKFK ditunjukkan melalui studi simulasi, dan penerapan pada data empirik hasil pengukuran glukosa darah dengan teknik non-invasif.

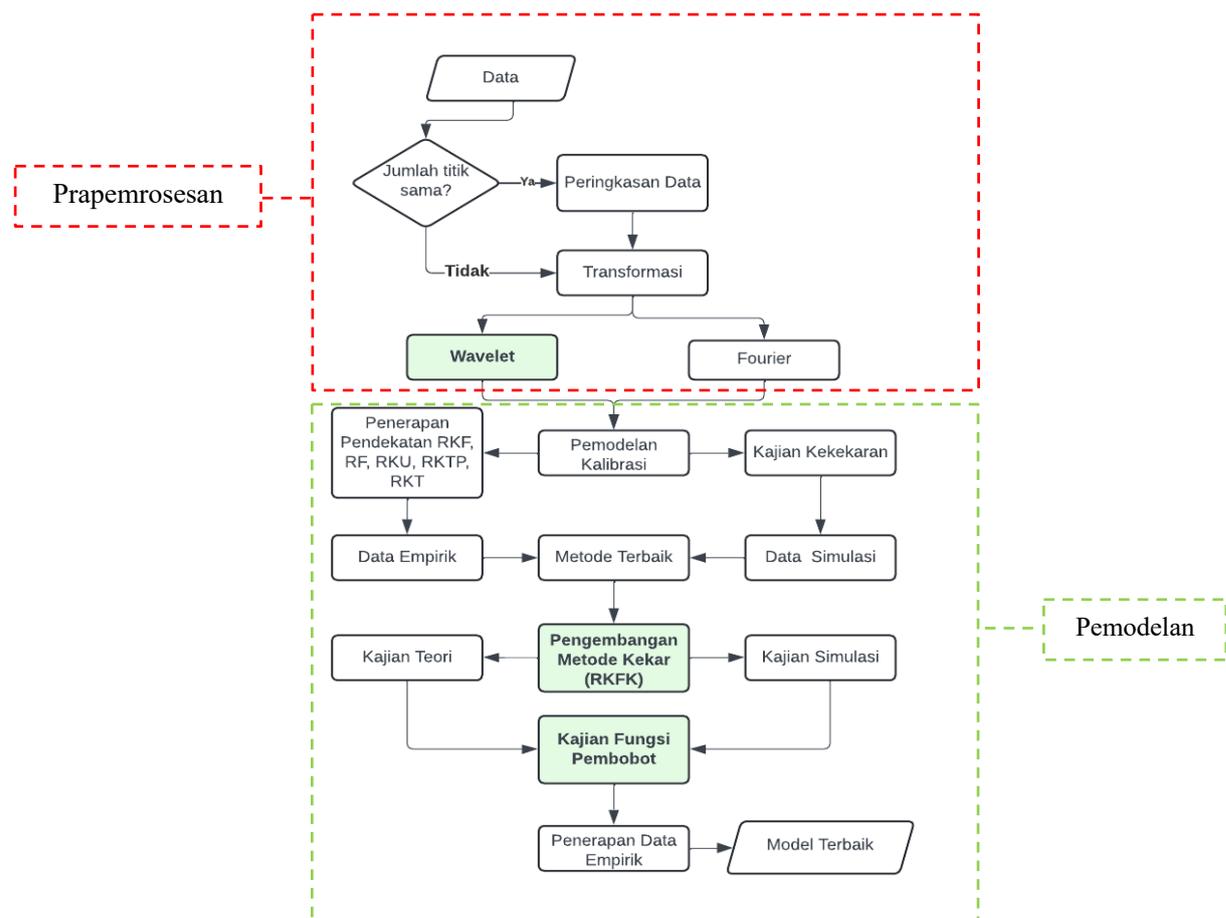
Evaluasi terhadap model dilakukan dengan mengukur seberapa konsisten model dalam menggeneralisasi pada bagian-bagian yang berbeda dari data (*resampling*). Metode *resampling* yang populer digunakan adalah *Jackknife* dan *Bootstrap*, namun dalam penelitian ini metode *Jackknife* digunakan untuk mengevaluasi model. Hal tersebut berdasarkan hasil penelitian Ariani *et. al.* (2017)

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

dan Rodliyah (2016) yang menyebutkan metode *Jackknife* lebih baik dibandingkan dengan *Bootstrap* dalam menentukan nilai dugaan dan selang kepercayaan parameter regresi.

1.6 Kebaruan (*novelty*)

Aspek kebaruan dalam penelitian ini dapat dilihat pada dua bagian dalam tahapan penelitian ini yaitu prapemrosesan dan metode analisis data untuk pemodelan kalibrasi, yang dapat dilihat pada Gambar 3.



Gambar 3 Peta Konsep Penelitian

Bagian pertama, prapemrosesan yaitu melakukan pengecekan jumlah data dalam setiap responden, apabila jumlah data berbeda maka akan dilakukan peringkasan data kemudian dilanjutkan dengan tahap transformasi, apabila jumlah sama sama maka dilakukan proses transformasi. Tujuan transformasi yaitu melakukan pemulusan data (*smoothing*) yang bertujuan untuk menghilangkan fluktuasi acak atau variasi kecil dalam data sehingga pola atau tren yang lebih umum dapat lebih mudah diamati, serta membantu mengurangi efek *noise* atau variasi acak dalam data.

Pemulusan data dilakukan menggunakan metode transformasi *Wavelet*, yang merupakan pengembangan dari *Fourier*, terlihat pada fungsi basis yang digunakan

dalam *Fourier* adalah cosinus, sedangkan pada *Wavelet* fungsi basis bergantung pada penetapan fungsi *mother* dan *father Wavelet*. Penerapan secara empirik memperoleh hasil bahwa metode transformasi *Wavelet* memberikan hasil pendugaan yang relatif baik, dalam hal teknik pemulusan data mengacu pada hasil penelitian Barsanti dan Gilmore (2011).

Mother Wavelet memiliki kemampuan memberikan efek *smoothing* pada data karena karakteristiknya yang dapat mereduksi frekuensi tinggi atau detail dalam sinyal. Beberapa fungsi *mother Wavelet*, terutama yang bersifat *smoothing* atau memiliki kernel yang melibatkan fungsi Gaussian, dapat digunakan untuk tujuan ini. Beberapa *mother Wavelet* yang biasanya cocok untuk *smoothing* meliputi: *Gaussian Wavelet*, *Coiflet Wavelet*, *Daubechies Wavelet*, *Symlet Wavelet*, dan *Haar Wavelet*. Beberapa kajian telah menunjukkan *Daubechies* merupakan *mother Wavelet* yang lebih optimal digunakan untuk *smoothing* (Barsanti dan Gilmore 2011).

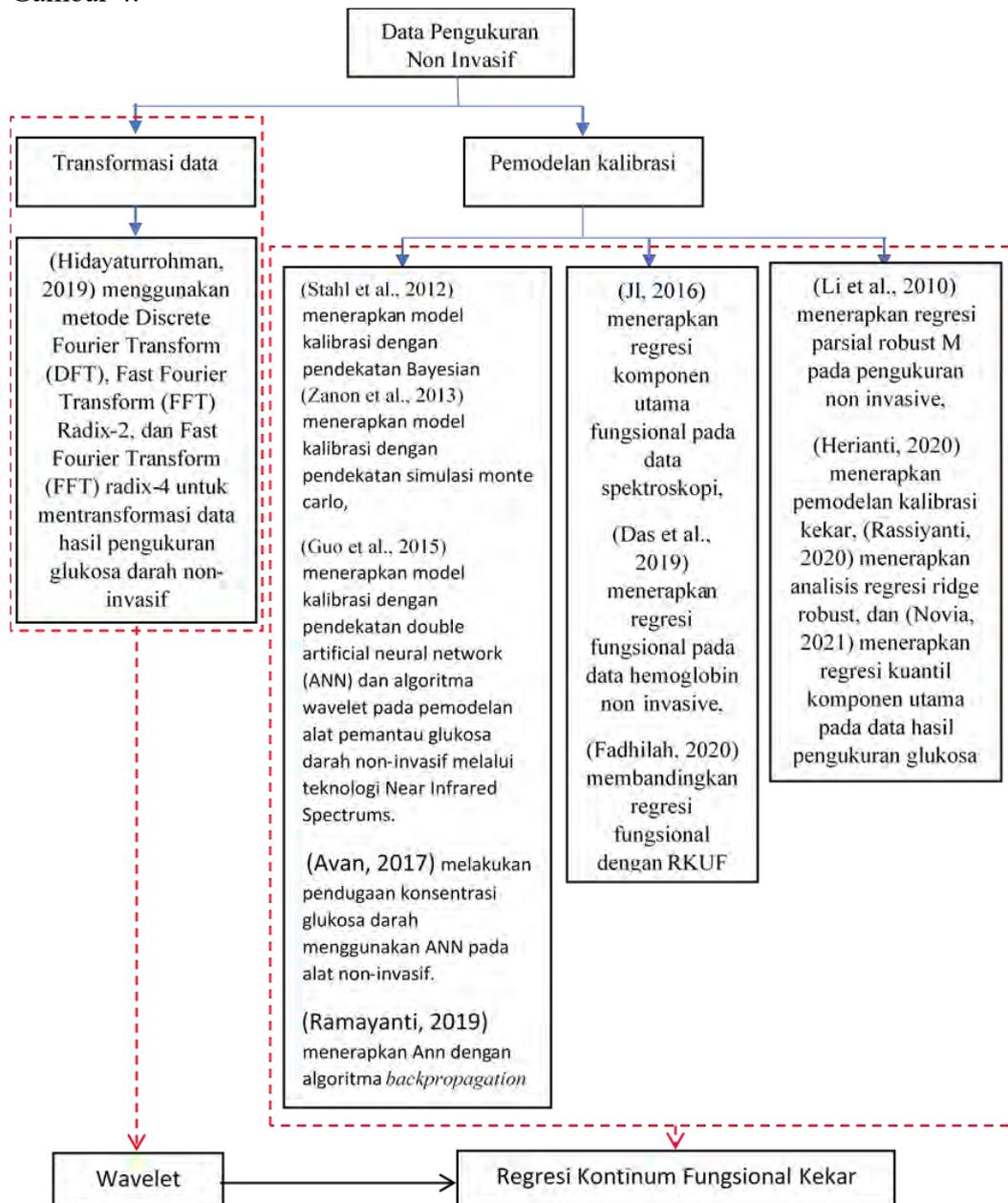
Proses *smoothing* menggunakan *Wavelet* dapat melibatkan penghilangan koefisien tinggi atau merata-ratakan sejumlah koefisien tertentu. Hal ini bergantung pada *Wavelet* yang digunakan dan parameter pengaturan transformasi *Wavelet*. Proses ini dapat membantu menghilangkan *noise* tinggi frekuensi atau detail kecil yang mungkin tidak relevan dalam beberapa aplikasi, seperti dalam pengolahan sinyal atau citra.

Bagian kedua, keterbaruan pada pengembangan metode analisis data untuk pemodelan kalibrasi, memberikan versi yang kekar dari formulasi regresi kontinum fungsional. Kekekaran dilakukan dengan membentuk matrik pembobot, yang berfungsi untuk mengurangi dampak pencilan atau nilai ekstrem terhadap estimasi model. Besaran nilai bobot bergantung pada nilai sisaan. Apabila nilai sisaan kurang dari atau sama dengan nilai k (konstanta *tuning*) maka nilai bobot sama dengan satu, dan apabila nilai sisaan lebih dari nilai k maka nilai bobot lebih kecil dari satu (proporsional). Nilai yang lebih kecil dari k menghasilkan kekekaran yang lebih besar terhadap pencilan, tetapi dengan mengorbankan efisiensi yang lebih rendah ketika kesalahan terdistribusi secara normal. Pemilihan nilai k memerlukan keseimbangan antara kekekaran terhadap pencilan dan efisiensi terhadap data yang tidak terpengaruh pencilan.

Penerapan berbagai metode untuk memodelkan kalibrasi data pengukuran glukosa darah non-invasif telah banyak dilakukan. Perkembangan metode terus dilakukan menyesuaikan kebutuhan dan jenis data, serta permasalahan yang muncul dalam pemodelan kalibrasi, sehingga metode konvensional dianggap tidak relevan untuk diterapkan. Pengukuran glukosa darah dengan teknik non-invasif memberikan data fungsional yang mencerminkan perubahan kadar glukosa sepanjang waktu. Regresi kontinum fungsional (RKF) merupakan pengembangan dari regresi kontinum yang dapat diaplikasikan pada data fungsional.

RKF dinyatakan Zhou (2019) sebagai metode yang efektif dalam mengatasi permasalahan multikolinearitas dengan jumlah peubah bebas lebih besar dari jumlah pengamatan ($p \gg n$) dan data bersifat fungsional. Pada beberapa kondisi kesehatan, respon tubuh dapat menjadi tidak stabil atau fluktuatif terhadap stimulus alat pengukur, sehingga menghasilkan nilai yang mungkin dianggap sebagai pencilan, namun RKF tidak resisten terhadap keberadaan pencilan. Oleh karena itu diperlukan pengembangan metode analisis data untuk pemodelan kalibrasi, dengan versi yang kekar dari formulasi regresi kontinum fungsional.

Performa RKF dan sifat kekekarannya dianalisis dengan metode simulasi dan kajian empirik, dan dilakukan perbandingan terhadap metode fungsional lainnya yaitu regresi fungsional (RF), serta ketiga metode non fungsional yaitu regresi komponen utama (RKU, regresi kuadrat terkecil parsial (RKTP), dan regresi kuadrat terkecil (RKT). Hasil kajian dengan metode simulasi dan empirik menjadi acuan dalam mengembangkan metode regresi kontinum fungsional kekar (RKFK). Perkembangan hasil penelitian terhadap data hasil pengukuran glukosa serta kontribusi penelitian ini dalam mengembangkan model kalibrasi, ditunjukkan pada Gambar 4.



Gambar 4 Bagan Perkembangan Penelitian Data Pengukuran Non-Invasif

II TINJAUAN PUSTAKA

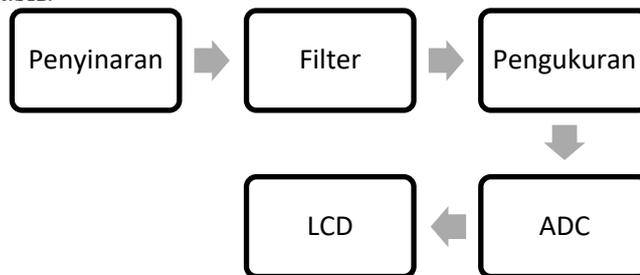
2.1 Sistem Non-Invasif untuk Pengukuran Glukosa Darah

Pengukuran glukosa dalam darah dapat dilakukan secara invasif dan non-invasif. Invasif adalah teknik fundamental dengan mendeteksi kadar glukosa darah berdasarkan darah kapiler yang diperoleh melalui tusuk jari (glukometer) (Ciudin *et. al.*, 2012). Sedangkan non-invasif adalah tindakan medis yang dilakukan tanpa memasukkan alat ke dalam tubuh, sehingga tidak mencederai anggota tubuh. Salah satu metode yang digunakan dalam pengukuran glukosa darah non-invasif yaitu spektroskopi (Dantu *et. al.*, 2014).

Pengukuran glukosa dengan alat invasif akan menimbulkan cedera pada tubuh yang mengakibatkan rasa sakit hingga trauma pada pasien. Teknik invasif dalam pengukuran glukosa melalui proses yang membutuhkan waktu cukup lama, setelah sampel darah diambil dari pasien, darah akan dibawa ke laboratorium untuk dilakukan pengecekan. Proses dengan waktu yang lama, tidak cukup efektif karena kadar glukosa dalam darah seseorang bersifat dinamis, sehingga hasil pengukuran yang telah dilakukan mungkin tidak lagi akurat dengan kondisi pasien pada saat menerima hasil. Kadar glukosa yang bersifat dinamis, mengakibatkan pasien harus sering mengontrol kondisi glukosa pada tubuhnya, hal ini apabila dilakukan dengan teknik invasif tidak efektif dan efisien, karena akan menimbulkan trauma dan kerusakan pada sel tubuh.

Biomarking IPB telah mengembangkan alat purwarupa non-invasif untuk mengukur kadar glukosa darah dengan metode spektroskopi. Metode spektroskopi menganalisis jumlah cahaya yang diserap jari pada panjang gelombang tertentu untuk setiap tingkat glukosa darah (Dantu *et. al.*, 2014). Spektroskopi sering digunakan untuk mengidentifikasi suatu substansi melalui penyerapan dan pemancaran suatu spektrum (Braeuer, 2015). Besarnya penyerapan ini bergantung pada konsentrasi dari glukosa darah.

Sistem kerja purwarupa alat pendeteksi glukosa darah secara non-invasif dimulai dengan pasien/sampel disensor menggunakan cahaya inframerah, dengan menyinari bagian tubuh tertentu dengan sinar inframerah, yang kemudian ditangkap oleh sensor dalam bentuk nilai analog. Konversi sinyal analog ke bentuk sinyal digital menggunakan *Analog to Digital Converter* (ADC) tipe ADS 1115 berbentuk bilangan diskret 0-32767, dan kemudian server akan mengeluarkan luaran berupa nilai intensitas tegangan (residu) dalam domain waktu yang ditampilkan dalam *Liquid crystal display* (LCD). Gambar 5 merupakan alur kerja pengukuran darah non-invasif.



Gambar 5 Alur Kerja Alat Purwarupa Non-Invasif

Alat yang dikembangkan oleh Biomarking IPB dirangkai menggunakan 2 jenis lampu inframerah, yaitu lampu yang pertama dengan panjang gelombang 1550 nm untuk mengetahui kadar air dalam tubuh dan lampu yang kedua dengan panjang gelombang 1600 nm untuk mengetahui kadar glukosa darah dalam tubuh. Rancangan kondisi dengan modulasi atau tingkat pencahayaan lampu yang berbeda dalam pengukuran dapat dilihat pada Tabel 1 (Novia, 2021).

Tabel 1 Rancangan Kondisi Pengukuran Glukosa Darah Non-Invasif

Modulasi	Tingkat Pencahayaan	Periode			
		Kedua Lampu Mati	Lampu Pertama Hidup	Lampu Kedua Hidup	Kedua Lampu Hidup
0	0/1023=0%	1	11	21	31
10	111/1023=10.85%	2	12	22	32
20	222/1023=21.70%	3	13	23	33
30	333/1023=32.55%	4	14	24	34
40	444/1023=43.44%	5	15	25	35
50	555/1023=54.25%	6	16	26	36
60	666/1023=65.10%	7	17	27	37
70	777/1023=75.95%	8	18	28	38
80	888/1023=86.80%	9	19	29	39
90	999/1023=97.65%	10	20	30	40

Pengukuran dilakukan dengan 4 kondisi rancangan lampu yang berbeda yaitu kedua lampu mati, lampu pertama hidup, lampu kedua hidup, dan kedua lampu mati. Setiap kondisi terdiri dari 10 periode dan lama waktu pengukuran untuk masing-masing kondisi adalah 5 detik, sehingga total lama waktu pengukuran adalah 20 detik. Spektrum kadar glukosa darah terdiri dari sepuluh modulasi yaitu 0% hingga 90%.

Modulasi 10 adalah penamaan untuk tingkat pencahayaan 10,85% yang diperoleh dari pembagian banyaknya gelombang cahaya yaitu 111 dibagi dengan gelombang cahaya maksimum yang terbaca oleh alat yaitu 1023. Modulasi 20 adalah penamaan untuk tingkat pencahayaan 21,70% yang diperoleh dari pembagian banyaknya gelombang cahaya yaitu 222 dibagi dengan gelombang cahaya maksimum yang terbaca oleh alat yaitu 1023. Perhitungan ini juga berlaku pada penamaan modulasi lainnya. Masing-masing modulasi diulang sebanyak dua kali untuk setiap pengukuran.

Kondisi yang terjadi pada modulasi yang membentuk puncak adalah ketika lampu kedua menyala dan kedua lampu menyala, sedangkan untuk kondisi kedua lampu mati dan lampu pertama hidup tidak membentuk puncak atau spektrum tampak *flat* (datar). Penelitian ini menggunakan kondisi pada saat lampu pertama menyala (1600 nm). Pemilihan kondisi tersebut agar tidak ada pengaruh dari lampu dengan panjang gelombang 1550 nm, karena lampu tersebut berfungsi untuk mengetahui kadar air, sedangkan lampu dengan 1600 nm untuk mengetahui kadar glukosa darah dalam tubuh.

Hasil spektrum pengukuran kadar glukosa non-invasif dengan satu jari tangan dari satu responden diulang sebanyak lima kali ulangan. Pola spektrum yang dihasilkan merupakan intensitas dari cahaya inframerah yang dilewatkan terhadap

time domain. Hasil pengukuran spektrum salah satu responden ditunjukkan pada Gambar 6 berikut.



Gambar 6 Spektrum hasil pengukuran Glukosa Darah Non-Invasif

Setiap modulasi terdiri atas dua kali pengukuran, sehingga terdapat sepuluh puncak yang terbentuk pada setiap kondisi seperti yang tampak pada Gambar 6. Data puncak spektrum kadar glukosa darah umumnya memiliki data residu intensitas paling sedikit lima data dan paling banyak tujuh data. Data awal dan data akhir pada setiap puncak dibuang karena dikhawatirkan data tersebut merupakan noise dari alat non-invasif saat memulai dan mengakhiri pengukuran.

2.2 Transformasi *Wavelet*

Wavelet pertama kali diperkenalkan oleh Alfred Haar pada tahun 1909. Jean Morlet dan Alex Grossmann di awal tahun 1980-an menamai *Wavelet* yang berasal dari bahasa Perancis, *ondelette* artinya gelombang kecil. Kata *onde* diterjemahkan ke dalam Bahasa Inggris menjadi *wave*, yang kemudian digabung dengan *let* dari kata asalnya sehingga terbentuk kata *Wavelet*. *Wavelet* merupakan alat analisis yang biasa digunakan untuk menyajikan data atau fungsi atau operator ke dalam komponen-komponen frekuensi yang berlainan, dan kemudian mengkaji setiap komponen dengan suatu resolusi yang sesuai dengan skalanya (Sutarno 2010).

Wavelet adalah metode transformasi yang memetakan $L^2(R) \rightarrow L^2(R^2)$, dan merupakan pengembangan dari metode *Fourier Transform* dan *Short Time Fourier Transform* (STFT) (Akansu dan Haddad, 2001). Seperti halnya STFT, *Wavelet* mentransformasi sinyal dalam domain waktu menjadi sinyal dalam domain waktu dan frekuensi (yang dalam hal ini dibentuk menjadi domain translasi dan skala). Translasi adalah sebuah bentuk transformasi dari domain waktu. Skala adalah bentuk transformasi dari frekuensi, dengan nilai *scale* berbanding terbalik dengan nilai frekuensi.

Transformasi *Wavelet* dalam konteks kemometrika, mengacu pada teknik pengolahan sinyal atau data yang menggunakan analisis *Wavelet* untuk mengubah data menjadi domain *Wavelet*. Transformasi *Wavelet* adalah metode yang kuat dan fleksibel dalam menganalisis data multivariat dan dapat digunakan untuk mengungkapkan struktur, pola, atau informasi yang terkandung dalam data.

Wavelet mempunyai 3 peranan penting (Ogden 1997). Pertama, *denoising* (pengurangan derau): Transformasi *Wavelet* dapat digunakan untuk mengurangi derau dalam data kemometrika. Derau sering kali menjadi tantangan dalam analisis kemometrika karena dapat mempengaruhi estimasi parameter dan menghasilkan kesalahan yang signifikan. Dengan menerapkan transformasi *Wavelet*, sinyal yang

tidak relevan atau derau dapat dihilangkan atau diredam, sehingga meningkatkan kualitas analisis data. Kedua, kompresi data: Transformasi *Wavelet* juga digunakan untuk kompresi data dalam analisis kemometrika. Dalam analisis multivariat, dataset sering kali memiliki dimensi yang tinggi, yang memerlukan penyimpanan dan pengolahan yang besar. Dengan menggunakan transformasi *Wavelet*, data dapat dikompresi menjadi representasi yang lebih ringkas dengan mempertahankan sebagian besar informasi penting. Hal ini dapat mengurangi ukuran data, mengurangi kompleksitas komputasi, dan memfasilitasi penyimpanan dan pengiriman data yang lebih efisien. Ketiga, fitur ekstraksi: Transformasi *Wavelet* juga digunakan dalam ekstraksi fitur atau representasi yang lebih baik dari data kemometrika. Dalam analisis multivariat, fitur yang tepat dapat memainkan peran penting dalam menggambarkan perbedaan dan pola dalam data. Transformasi *Wavelet* dapat mengungkapkan fitur-fitur penting dan saling keterkaitan antar peubah, yang dapat digunakan untuk membangun model atau analisis lebih lanjut.

Transformasi *Wavelet* merupakan suatu teknik pemrosesan sinyal dengan Analisis Multi Resolusi (AMR). Ide dasar metode transformasi *Wavelet* adalah merepresentasikan suatu kurva sebagai kombinasi linear kurva-kurva lain yang relatif sederhana yang disebut fungsi basis atau fungsi *Wavelet* (Fearn 1998).

Berdasarkan bilangan dilatasi dan translasi transformasi *Wavelet* terbagi menjadi 2, yaitu Transformasi *Wavelet* Diskrit (TWD) dan Transformasi *Wavelet* Kontinu (TWK). Perbedaan diantara keduanya adalah pada nilai koefisien skala (a) dan koefisien translasi (b). Termasuk TWK apabila suatu fungsi yang didekomposisi ke dalam fungsi-fungsi *Wavelet* diambil bilangan dilatasi dan translasi yang kontinu. Termasuk TWD apabila bilangan dilatasi dan translasi berupa bilangan bulat bukan negatif.

Pada penerapan transformasi *Wavelet*, perlu dipilih *mother Wavelet* untuk membangun fungsi basis. Penggunaan *mother Wavelet* yang berbeda akan memberikan hasil analisis yang berbeda. Beberapa contoh *mother Wavelet* adalah Haar, Daubechies, Symlets, Coiflets, BiorSplines, ReverseBior, Meyer, DMeyer, Gaussian, Mexican_hat, Morlet, Complex Gaussian, Shannon, Frequency B-Spline, Complex Morlet, Riyad, dan lain sebagainya. Dalam penelitian ini akan dikaji hasil analisis Transformasi *Wavelet* dengan pemilihan *mother Wavelet* yang cukup populer berdasarkan hasil penelitian Suma'inna dan Gumilar (2013), dan *mother Wavelet* yang optimal berdasarkan hasil penelitian Jang *et. al.* (2021) yaitu Daubechies.

Fungsi dalam *Wavelet* terdiri dari *father Wavelet* (φ) dan *mother Wavelet* (ψ). Sebuah fungsi dikatakan *Wavelet* apabila memenuhi syarat berikut ini:

$$1. \int_{-\infty}^{\infty} \varphi(x) dx = \int_{-\infty}^{\infty} \psi^2(x) dx = 1 \quad (\text{father Wavelet})$$

$$2. \int_{-\infty}^{\infty} \psi(x) dx = 0 \quad (\text{mother Wavelet})$$

Fungsi basis diperoleh dari hasil dilatasi dan translasi fungsi $\psi(x)$. Bentuk umum fungsi basis sebagai berikut:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right); \quad a > 0, -\infty < b < \infty \quad (1)$$

dengan $a = 2^{-j}, b = k2^{-j}, j, k \in Z = \{\dots, -1, 0, 1, \dots\}$, maka akan diperoleh sekumpulan fungsi basis yang saling ortogonal. Bentuk fungsi ortogonal tersebut sebagai berikut:

$$\psi_{j,k}(x) = \frac{1}{\sqrt{|2^{-j}|}} \psi(2^j x - k) = 2^{-\frac{j}{2}} \psi(2^j x - k); \quad j, k \in Z \quad (2)$$

2.3 Transformasi *Wavelet* Kontinu

Transformasi *Wavelet* Kontinu (TWK) didefinisikan dalam bentuk dilatasi dan translasi dari fungsi *mother Wavelet* $\psi(t)$ (Akansu dan Haddad 2001):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (3)$$

dari persamaan (3) sinyal yang ditransformasi merupakan fungsi dari dua parameter a dan b yaitu parameter skala (dilatasi) dan parameter translasi, sedangkan t adalah waktu. TWK memetakan sebuah fungsi $f(t)$ kedalam ruang skala waktu sebagai berikut

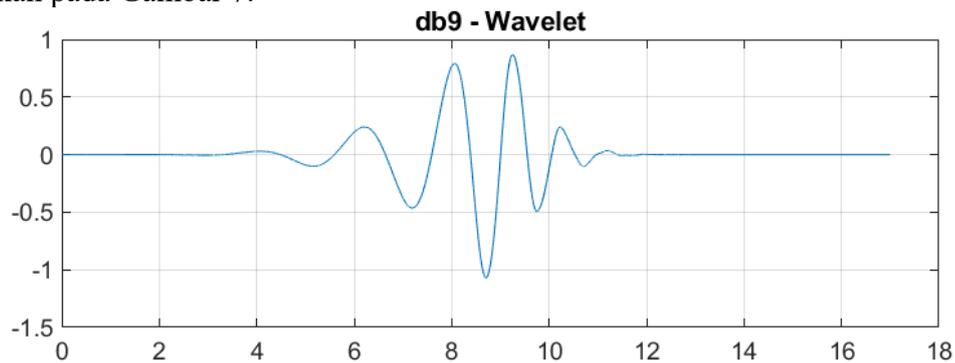
$$W_f(a, b) = \int \psi_{a,b}(t) f(t) dt = \langle \psi_{a,b}(t), f(t) \rangle \quad (4)$$

$f(t)$ adalah sinyal yang dianalisis dan $W_f(a, b)$ adalah hasil transformasi.

Kebalikan dari TWK adalah sebagai berikut:

$$f(t) = \iint W_f(a, b) \psi_{a,b}(t) db da \quad (5)$$

Pemilihan *mother Wavelet* diperlukan untuk membentuk fungsi basis, hal inilah yang membedakan transformasi *Wavelet* dengan *Fourier* yang hanya melibatkan fungsi cosinus. Visualisasi dari fungsi *mother Wavelet* Daubechies ditunjukkan pada Gambar 7.



Gambar 7 Visualisasi *mother Wavelet* Daubechies

Fungsi *mother Wavelet* pada *Wavelet* Daubechies diberikan pada persamaan 6 (Rowe dan Abbott 1995),

$$\psi(t) = \sqrt{2} \sum_{k=0}^{N-1} (-1)^k c_{N-1-k} \varphi(2t - k) \quad (6)$$

dan *father Wavelet* $\varphi(t) = \sqrt{2} \sum_{k=0}^{N-1} c_k \varphi(2t - k)$, dengan fungsi skala dinormalisasi dengan $\int \varphi(t) dt = 1$.

Cara kerja TWK berdasarkan Polikar (1997), yang pertama adalah menentukan *mother Wavelet* tertentu, dengan melibatkan jendela dalam proses yang diperoleh dari proses dilatasi dan translasi (menggeser) *mother Wavelet*. Selanjutnya komputasi diawali dengan $a=1$ dan TWK dihitung untuk semua nilai a . Namun, tergantung pada sinyal, transformasi lengkap pada umumnya tidak diperlukan karena perhitungan transformasi untuk interval terbatas memadai.

Wavelet ditempatkan pada awal sinyal pada titik yang sesuai dengan $t=0$. Fungsi *Wavelet* pada skala "1" dikalikan dengan sinyal dan kemudian diintegrasikan pada semua waktu. Hasil integrasi ini kemudian dikalikan dengan nilai konstanta $\frac{1}{\sqrt{a}}$. Tujuan perkalian ini adalah untuk normalisasi energi sehingga

sinyal yang berubah akan memiliki energi yang sama pada setiap skala. Hasil akhir adalah nilai transformasi, yaitu nilai TWK pada waktu nol dan skala $s = 1$. Dengan kata lain, nilai transformasi tersebut adalah nilai yang sesuai dengan titik $a=1$ dan $b=0$ di bidang skala waktu.

Wavelet pada skala $a=1$ kemudian bergeser ke arah kanan sebesar b ke lokasi $t=b$, dan persamaan 6 dihitung untuk mendapatkan transformasi nilai pada $t=b$, $a=1$ dalam bidang waktu-frekuensi. Prosedur ini diulang sampai *Wavelet* mencapai akhir dari sinyal. Satu baris titik-titik di atas bidang skala-waktu untuk $a=1$ telah didapatkan. Selanjutnya nilai a dinaikkan sedikit demi sedikit, dan prosedur penghitungan yang sama diulang.

2.4 Regresi Fungsional

Perkembangan teknologi, berdampak pada permintaan terhadap Analisis Data Fungsional (ADF) semakin meningkat (Zhou, 2019). ADF dikenalkan oleh Ramsay (1982) yang berhubungan dengan analisis dan teori data yang berupa fungsi, gambar dan bentuk, ataupun objek yang lebih umum (Wang *et. al.*, 2016).

Misalkan terdapat pengamatan berpasangan $(x_1, y_1), \dots, (x_n, y_n)$ yang berdistribusi bebas dan identik sebagai (\mathbf{X}, y) dengan y adalah skalar dan $\mathbf{X} = \mathbf{X}(\cdot)$ adalah proses acak yang biasanya disebut sebagai waktu yang diperoleh dari \mathbb{T} yang merupakan ukuran Lebesgue. $V_X: L^2(\mathbb{T}) \rightarrow L^2(\mathbb{T})$ merupakan kovarian X , untuk semua $f \in L^2(\mathbb{T})$, $L^2(\mathbb{T})$ merupakan ruang L^2 pada selang waktu \mathbb{T} . Kovarian X diformulasikan seperti berikut ini,

$$V_X(f)(\cdot) = \int_{\mathbb{T}} f(s)v_x(s, \cdot)ds \quad (7)$$

dengan $v_x(s, t) = cov\{\mathbf{X}(s), \mathbf{X}(t)\}$. Diasumsikan $v_x \in L^2(\mathbb{T}^2)$ yang berimplikasi banyak nilai tidak negatif, sehingga nilai eigen secara berurutan menjadi menurun (*decreasing*), $\lambda_1 \geq \lambda_2 \geq \dots$, sehingga penjumlahan nilai eigen kurang dari tak hingga, $\lambda_1 + \lambda_2 + \dots < \infty$.

Misalkan terdapat fungsi skalar dalam model regresi yang menghubungkan peubah respon y dengan integral *inner product* X dengan fungsi yang tidak diketahui dan tidak acak β , $\beta = \beta(\cdot)$ diasumsikan orthogonal ke ruang nol dari V_X , yang secara matematis bentuk hubungan tersebut dapat dituliskan sebagai berikut,

$$Y = E(y) + \int_{\mathbb{T}} \beta\{X - E(X)\} + \sigma \varepsilon \quad (8)$$

dengan σ bernilai positif dan ε menyebar normal dengan rata-rata nol dan ragam unit, sedangkan $\int_{\mathbb{T}} f$ adalah ringkasan dari integral Lebesgue $\int_{\mathbb{T}} f(t)dt$.

Dimensi infinit ruang struktur fungsional membuat tantangan dalam menganalisis data, dimensi ruang parameter melebihi jumlah pengamatan, sehingga reduksi dimensi sangat diperlukan dalam membentuk model yang tepat. Untuk mengestimasi fungsi koefisien β dan mengestimasi nilai harapan bersyarat dengan formula sebagai berikut

$$\eta(X_0) = E(y|X = X_0) = E(y) + \beta\{X_0 - E(X)\} \quad (9)$$

dengan X_0 berdistribusi seperti X .

Konsep dari analisis fungsional yaitu memperoleh penduga parameter β berdasarkan fungsi basis, sehingga pendekatan bakunya adalah dengan menyatakan β dalam bentuk kombinasi linier dari fungsi w_1, \dots, w_p yang merupakan fungsi basis w_1, w_2, \dots dalam $L^2(\mathbb{T})$. Zhou (2019) menyatakan pendekatan fungsi koefisien model β dan nilai harapan bersyarat $\eta(X_0)$ sebagai berikut,

$$\begin{aligned}\boldsymbol{\beta}_p &= \arg \min_{\theta \in \text{span}(w_1, \dots, w_p)} E[\mathbf{y} - E(\mathbf{y}) - \int_{\mathbb{T}} \theta \{\mathbf{X} - E(\mathbf{X})\}]^2 \\ &= \arg \min_{\theta \in \text{span}(w_1, \dots, w_p)} E[\int_{\mathbb{T}} (\boldsymbol{\beta} - \theta) \{\mathbf{X} - E(\mathbf{X})\}]^2\end{aligned}\quad (10)$$

dan

$$\eta_p(\mathbf{X}_0) = E(\mathbf{y}) + \int_{\mathbb{T}} \boldsymbol{\beta}_p \{\mathbf{X}_0 - E(\mathbf{X})\} \quad (11)$$

$\text{span}(w_1, \dots, w_p)$ adalah ruang linier yang dibentang oleh w_1, \dots, w_p .

Penduga dari persamaan (10) dan (11) diperoleh sebagai berikut

$$\hat{\boldsymbol{\beta}}_p = \arg \min_{\theta \in \text{span}(w_1, \dots, w_p)} \frac{1}{n} \sum_{i=1}^n \{\mathbf{y}_i - \bar{\mathbf{y}} - \int_{\mathbb{T}} \theta (\mathbf{X}_i - \bar{\mathbf{X}})\}^2 \quad (12)$$

$$\hat{\eta}_p(\mathbf{X}_0) = \bar{\mathbf{y}} + \int_{\mathbb{T}} \hat{\boldsymbol{\beta}}_p (\mathbf{X}_0 - \bar{\mathbf{X}}) \quad (13)$$

Dalam proses pendugaan, meskipun memungkinkan pemilihan \hat{w}_j (basis polynomial, basis *Fourier*, basis B-spline, dan lainnya) tidak bergantung pada data, namun akan lebih baik pemilihan \hat{w}_j bergantung pada data.

2.5 Regresi Kontinum

Salah satu masalah regresi yaitu terdapat korelasi antar peubah bebas atau yang disebut dengan multikolinieritas, maka penerapan metode kuadrat terkecil (MKT) akan menghasilkan ragam yang besar. Beberapa metode yang disebut sebagai regularisasi (penyusutan) ragam yang bias, merupakan alternatif dari MKT. Metode regularisasi yang dikenal diantaranya regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), regresi ridge (RR) yang dikenalkan oleh Brown (1993), serta Frank dan Friedman (1993), dan regresi kontinum yang diperkenalkan oleh Stone dan Brooks (1990), yang mempertimbangkan spektrum sebagai peubah regresi, dan masing-masing memiliki nilai parameter α . Konsep dari teknik regresi kontinum adalah menggabungkan MKT, RKTP dan RKU sebagai kasus khusus, bergantung dengan nilai α tertentu, dan juga metode RR yang dimodifikasi.

Misalkan \mathbf{X} matriks data yang sudah dipusatkan (*centered*) berukuran $n \times p$ dan disebut peubah bebas, sedangkan \mathbf{y} adalah vektor peubah respon berukuran $n \times 1$ yang sudah dipusatkan, $\boldsymbol{\beta}$ adalah vektor parameter regresi berukuran $1 \times p$, serta $\boldsymbol{\varepsilon}$ adalah vektor galat berukuran $1 \times n$. Regresi Kontinum dikembangkan berdasarkan model regresi linear klasik sebagai berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Pada model regresi linear terboboti formula matematis dapat ditulis dengan maksimumkan kuadrat korelasi antara peubah bebas dengan peubah respon sebagai berikut,

$$r_w^2 = \frac{(\sum_{i=1}^n y_i w^T x_i)^2}{(\sum_{i=1}^n y_i^2) \sum_{i=1}^n (w^T x_i)^2} = \frac{(w^T s)^2}{\|y_i\|^2 w^T s w} \quad (14)$$

dengan x_i adalah vektor pengamatan peubah bebas ke- i ($i= 1, 2, \dots, n$) yang berukuran $(p \times 1)$, $s = \mathbf{X}^T \mathbf{y}$ dan $\mathbf{S} = \mathbf{X}^T \mathbf{X}$.

Pada regresi kontinum peubah baru diformulasikan dalam model peubah laten sebagai berikut (Serneels *et. al.*, 2005):

$$\mathbf{y} = \mathbf{T}_h \boldsymbol{\xi} + \boldsymbol{\varepsilon} \quad (15)$$

dengan $\mathbf{T}_h = \mathbf{X}\mathbf{W}_h$ dan $\mathbf{W}_h = (w_1, w_2, \dots, w_h)$ matriks berukuran $p \times h$ ($h < p$) dan disebut matriks pembobot.

Stone dan Brooks (1990) memformulasikan matriks pembobot tersebut sebagai berikut:

$$\mathbf{w}_{i,\alpha} = \arg \max \left\{ \text{cov}(\mathbf{X}\mathbf{w}, \mathbf{y})^2 \text{var}(\mathbf{X}\mathbf{w})^{\frac{\alpha}{\alpha-1}-1} \right\} \quad (16)$$

dengan $\|\mathbf{w}_i\| = 1$ dan $\text{cov}(\mathbf{X}\mathbf{w}_i, \mathbf{y}\mathbf{w}_j) = 0$ untuk $j < i$, dan $\alpha \in [0,1)$. Persamaan (16) merupakan generalisasi dari MKT ketika $\alpha = 0$, RKU ketika $\alpha \approx 1$ dan RKTP ketika $\alpha = \frac{1}{2}$.

Karena masalah ketika banyaknya peubah bebas lebih kecil dari banyak pengamatan ($p \ll n$) telah teratasi dengan ditentukan peubah laten (15), sehingga pendugaan parameter dapat diperoleh dengan MKT dan di formulasikan sebagai berikut.

$$\hat{\boldsymbol{\beta}}_{\alpha,h} = (\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T \mathbf{y} \quad (17)$$

dan nilai harapannya (Setiawan, 2007),

$$\begin{aligned} E[\hat{\boldsymbol{\beta}}_{\alpha,h} | \mathbf{W}] &= E[(\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T \mathbf{y} | \mathbf{w}] \\ &= E[(\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T (\mathbf{T}_h \boldsymbol{\xi} + \boldsymbol{\varepsilon}) | \mathbf{w}] \\ &= \boldsymbol{\xi} + E[(\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T \boldsymbol{\varepsilon} | \mathbf{w}] \\ &= \boldsymbol{\xi} + E[(\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \mathbf{W}_h^T \mathbf{X}^T \boldsymbol{\varepsilon} | \mathbf{w}] \\ &= \boldsymbol{\xi} + E[(\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \mathbf{W}_h^T \mathbf{X}^T E(\boldsymbol{\varepsilon})] \\ &= \boldsymbol{\xi} + E[(\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \mathbf{W}_h^T \mathbf{X}^T \boldsymbol{\varepsilon} | \mathbf{w}] \\ &= \boldsymbol{\xi} \end{aligned}$$

sedangkan ragamnya

$$\begin{aligned} \text{var}[\hat{\boldsymbol{\beta}}_{\alpha,h} | \mathbf{W}] &= \text{var}[(\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T \mathbf{y} | \mathbf{W}] \\ &= \text{var}[(\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \mathbf{W}_h^T \mathbf{X}^T \mathbf{y} | \mathbf{W}] \\ &= (\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \mathbf{W}_h^T \mathbf{X}^T \text{var}(\mathbf{y}) \mathbf{X} \mathbf{W}_h (\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \\ &= (\mathbf{W}_h^T \mathbf{X}^T \mathbf{X} \mathbf{W}_h)^{-1} \sigma^2 \end{aligned}$$

Ketika model regresi dikalikan dengan sebuah matriks pembobot \mathbf{W}_h menjadi $\hat{\mathbf{y}}_{\alpha,h} = \mathbf{X} \mathbf{W}_h \hat{\boldsymbol{\beta}}_{\alpha,h}$, pemberian bobot tersebut dilakukan untuk mengatasi keberadaan data pencilan sehingga model regresi menjadi kekar atau resisten terhadap pencilan, maka untuk menduga koefisien regresi $\boldsymbol{\beta}$ pada model dapat diperoleh dari formulasi $\hat{\boldsymbol{\beta}}_{\alpha,h} = \mathbf{W}_h (\mathbf{T}_h^T \mathbf{T}_h)^{-1} \mathbf{T}_h^T \mathbf{y}$. Metode tersebut yang kemudian disebut dengan regresi kontinum kekar yang dikembangkan oleh Serneels *et. al.* (2005).

2.6 Regresi Kontinum Fungsional

Fungsi basis kontinum dinyatakan dengan $\{w_{1,\alpha}, \dots, w_{p,\alpha}\}$, dan telah ditentukan $\alpha \in [0,1)$, $w_{1,\alpha}, \dots, w_{p,\alpha}$ diperoleh dengan memaksimumkan matriks pembobot $\mathbf{T}_\alpha(\mathbf{w})$, seperti persamaan berikut,

$$w_{i,\alpha} = \arg \max \mathbf{T}_\alpha(\mathbf{w}) \quad (18)$$

dengan kendala $\|\mathbf{w}\| = 1$ dan $\int_{\mathbb{T}} \mathbf{w} V_X(w_{1,\alpha}) = \dots = \int_{\mathbb{T}} \mathbf{w} V_X(w_{i-1,\alpha}) = 0$ (19)

$\mathbf{T}_\alpha(\mathbf{w})$ diformulasikan sebagai berikut,

$$\mathbf{T}_\alpha = \mathbf{T}_\alpha(\mathbf{w}) = \left\{ \int_{\mathbb{T}} \mathbf{w} V_X(\mathbf{w}) \right\}^{\frac{\alpha}{1-\alpha}-1} \text{cov}^2(\mathbf{y}, \int_{\mathbb{T}} \mathbf{X} \mathbf{w}) \quad (20)$$

Masalah optimisasi (18) yang dibatasi oleh (19) serupa dengan persamaan (16). Persamaan (20) adalah bentuk umum dari RKF, dalam kasus khusus akan menjadi

RKU fungsional dan RKTP fungsional, tergantung dari nilai parameter yang ditetapkan seperti halnya pada RK. Fungsi koefisien model RKF $\boldsymbol{\beta}$ dan nilai harapan bersyarat $\eta(\mathbf{X}_0)$ analogi dari persamaan (10) dan (11), sehingga diperoleh persamaan sebagai berikut

$$\begin{aligned}\boldsymbol{\beta}_{p,\alpha} &= \arg \min_{\theta \in \text{span}(w_{1,\alpha}, \dots, w_{p,\alpha})} E[\int_{\mathbb{T}} (\boldsymbol{\beta} - \theta) \{ \mathbf{X} - E(\mathbf{X}) \}]^2 \\ &= \sum_{i=1}^p \{ \int_{\mathbb{T}} \boldsymbol{\beta} V_X(w_{i,\alpha}) \} \{ \int_{\mathbb{T}} w_{i,\alpha} V_X(w_{i,\alpha}) \}^{-\frac{1}{2}} w_{i,\alpha}\end{aligned}\quad (21)$$

dan

$$\begin{aligned}\eta_{p,\alpha}(X_0) &= E(\mathbf{y}) + \int_{\mathbb{T}} \boldsymbol{\beta}_{p,\alpha} \{ \mathbf{X}_0 - E(\mathbf{X}) \} \\ &= E(\mathbf{y}) + \sum_{i=1}^p \{ \int_{\mathbb{T}} \boldsymbol{\beta} V_X(w_{i,\alpha}) \} \{ \int_{\mathbb{T}} w_{i,\alpha} V_X(w_{i,\alpha}) \}^{-\frac{1}{2}} \int_{\mathbb{T}} w_{i,\alpha} \{ \mathbf{X}_0 - E(\mathbf{X}) \}\end{aligned}\quad (22)$$

Definisi $w_{i,\alpha}$ pada persamaan (18) dapat menjadi acuan untuk menentukan dugaan fungsi basis kontinu $\hat{w}_{i,\alpha}$ secara empirik. $\hat{w}_{i,\alpha}$ diperoleh dengan memaksimalkan $\hat{T}_\alpha(w)$ yang ditunjukkan pada persamaan (23) beserta fungsi kendala, hal ini berdasarkan pada proposisi 1 (Lampiran 4).

$$\begin{aligned}\text{maximize}_w \hat{T}_\alpha(w) &= \left\{ \frac{1}{n} \sum_{j=1}^n (\mathbf{y}_j - \bar{\mathbf{y}}) \int_{\mathbb{T}} w(\mathbf{X}_j - \bar{\mathbf{X}}) \right\}^2 \{ \int_{\mathbb{T}} w \hat{V}_X(w) \}^{\frac{\alpha}{1-\alpha}-1} \\ \text{kendala } \|w\| &= 1 \text{ dan } \int_{\mathbb{T}} w \hat{V}_X(\hat{w}_{1,\alpha}) = \dots = \int_{\mathbb{T}} w \hat{V}_X(\hat{w}_{i-1,\alpha}) = 0 \\ \text{dengan } \hat{V}_X(f)(\cdot) &= \int_{\mathbb{T}} f(s) \hat{v}_x(s, \cdot) ds\end{aligned}\quad (23)$$

Selanjutnya parameter (21) dan (22) secara berturut-turut dapat diduga oleh parameter (24) dan (25) berikut

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{p,\alpha} &= \arg \min_{\theta \in \text{span}(\hat{w}_{1,\alpha}, \dots, \hat{w}_{p,\alpha})} \frac{1}{n} \sum_{j=1}^p \{ \mathbf{y}_j - \bar{\mathbf{y}} - \int_{\mathbb{T}} \theta(\mathbf{X}_j - \bar{\mathbf{X}}) \}^2 \\ &= \sum_{i=1}^p \{ \int_{\mathbb{T}} \boldsymbol{\beta} \hat{V}_X(\hat{w}_{i,\alpha}) \} \{ \int_{\mathbb{T}} \hat{w}_{i,\alpha} \hat{V}_X(\hat{w}_{i,\alpha}) \}^{-\frac{1}{2}} \hat{w}_{i,\alpha} \\ &= \sum_{i=1}^p \text{cov}(\mathbf{y}, \int_{\mathbb{T}} \mathbf{X} \hat{w}_{i,\alpha}) \widehat{\text{var}}^{-\frac{1}{2}}(\int_{\mathbb{T}} \mathbf{X} \hat{w}_{i,\alpha}) \hat{w}_{i,\alpha}\end{aligned}\quad (24)$$

dan

$$\hat{\eta}_{p,\alpha}(\mathbf{X}_0) = \bar{\mathbf{y}} + \int_{\mathbb{T}} \hat{\boldsymbol{\beta}}_{p,\alpha} (\mathbf{X}_0 - \bar{\mathbf{X}}) \quad (25)$$

2.7 Fungsi Pembobot

Pencilan adalah pengamatan yang sangat berbeda dari data lainnya (Montgomery *et. al.*, 2012). Pencilan dapat terjadi akibat adanya kesalahan dalam pendataan atau jenis kesalahan lainnya. Data pencilan dapat berakibat buruk terhadap hasil pendugaan yang salah dan nilai dugaan kuadrat rata-rata sisaan menjadi sangat besar, tetapi pencilan mungkin juga tidak berakibat buruk melainkan menjadi informasi yang sangat berguna mengenai proses yang sedang diteliti. Salah satu teknik yang dapat digunakan untuk mengatasi pencilan adalah dengan memberikan nilai bobot yang berbeda-beda pada setiap amatan.

Banyak metode untuk menentukan fungsi pembobot agar diperoleh penduga yang kekar. Penduga yang kekar adalah penduga yang resisten terhadap keberadaan pencilan. Jenis fungsi pembobot yang populer diantaranya yaitu Huber, Hampel,

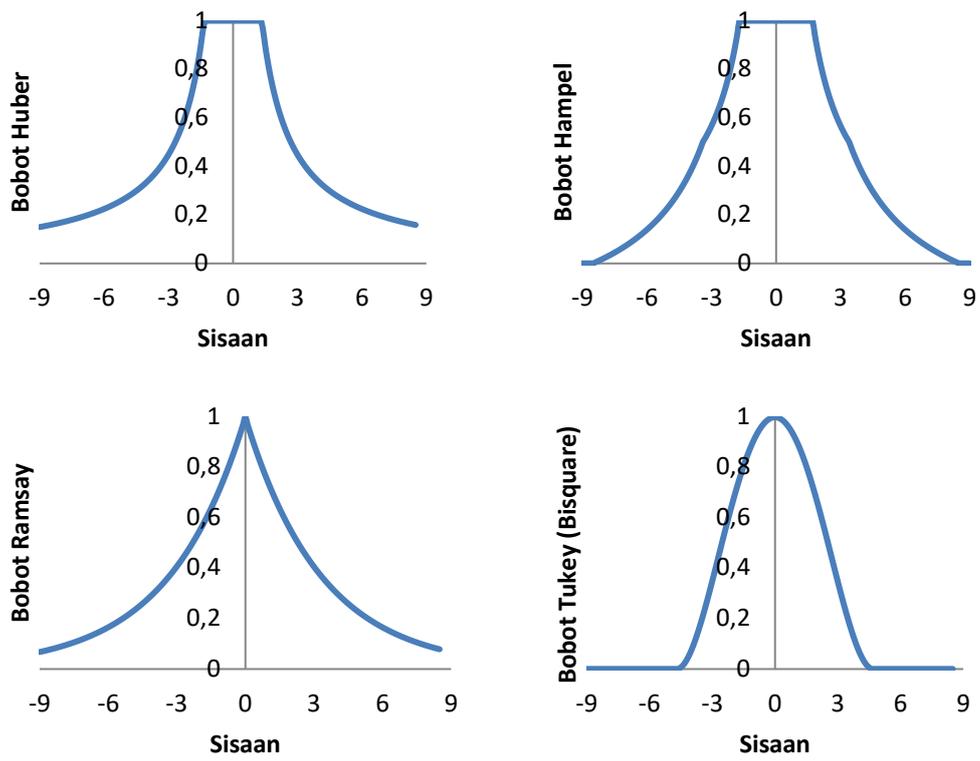
Ramsay dan Tukey (*Bisquare*). Fungsi pembobot bergantung pada nilai sisaan , keempat fungsi pembobot dengan kriteria pembobotan masing-masing dapat dilihat pada Tabel 2 dan grafik nilai bobot pada Gambar 8 (Fox 2002).

Tabel 2 Fungsi Pembobot

No	Metode	Fungsi Pembobot
1	Huber	$w(e_i) = \begin{cases} 1 & \text{for } e_i \leq k \\ \frac{k}{ e_i } & \text{for } e_i > k \end{cases}$ $k = 1,345$
2	Hampel	$w(e_i) = \begin{cases} 1 & \text{for } e_i \leq k_1 \\ \frac{k_1}{ e_i } & \text{for } k_1 < e_i < k_2 \\ \frac{k_1 - e_i }{k_1} & \text{for } k_2 < e_i < k_3 \\ 0 & \text{for } e_i > k_3 \end{cases}$ $k_1 = 1,7; k_2 = 3,4; k_3 = 8,5$
3	Ramsay	$w(e_i) = \exp(-k e_i) \quad \text{for } e_i < \infty$ $k = 0,3$
4	Tukey (<i>Bisquare</i>)	$w_T(e) = \begin{cases} \left[1 - \left(\frac{e}{k}\right)^2\right]^2 & \text{for } e \leq k \\ 0 & \text{for } e > k \end{cases}$ $k = 4,685$

Berdasarkan Fox (2002) k adalah parameter tuning, dengan $k = 1,345$ untuk Huber dan $k = 4,685$ untuk Tukey, konstanta k_1 , k_2 , k_3 dalam Hampel, dan k pada metode lainnya disebut parameter penyetel (*tunning*) dengan $0 < k_1 < k_2 < k_3 < \infty$ dan $0 < k < \infty$, dan nilai sisaan adalah $e_i = y_i - \hat{y}_i$. Fungsi pembobot Hampel dan Tukey (*Bisquare*) bernilai 0 untuk kriteria $|e| > k$, pemberian bobot 0 pada data pencilan artinya bahwa data tersebut akan dieliminasi, sehingga jumlah data (pengamatan) akan berkurang, hal ini sangat disayangkan karena tidak semua pencilan berdampak buruk seperti yang dinyatakan oleh Montgomery (2012), selain itu keputusan mengeliminasi pada data hasil pengukuran sangat disayangkan karena proses memperoleh data hasil pengukuran membutuhkan biaya yang besar dengan proses yang lama mulai dari persiapan hingga keluar hasil (Ismah *et. al.*, 2009).

Besaran nilai bobot pada masing-masing jenis dapat dilihat pada Gambar 8. Tampak grafik pada pembobot Huber dan Ramsay tidak menyentuh sumbu X (sisaan) yang artinya bahwa semakin besar nilai sisaan maka nilai bobot mendekati nol (tidak sama dengan nol), hal ini berbeda dengan pembobot Hampel dan Tukey (*Bisquare*) terlihat grafik menyentuh sumbu X yang artinya semakin besar nilai sisaan maka nilai bobot akan sama dengan nol.



Gambar 8. Grafik Fungsi Pembobot

III PENDEKATAN REGRESI KONTINUM FUNGSIONAL PADA DATA HASIL TRANSFORMASI *WAVELET* DALAM MODEL KALIBRASI PENGUKURAN GLUKOSA NON-INVASIF

3.1 Pendahuluan

Data dalam bidang statistika merupakan bahan mentah yang diolah, sehingga diperoleh satu atau lebih informasi penting secara kuantitatif ataupun kualitatif tentang suatu keadaan. Pendekatan statistika tradisional pada umumnya digunakan untuk data yang berbentuk tunggal atau bukan fungsi, namun untuk data yang berbentuk fungsi atau kontinu, pendekatan statistika tradisional dianggap kurang tepat karena tidak dapat memberikan informasi yang relevan dengan bentuk fungsi.

Data yang berbentuk fungsi disebut dengan data fungsional, Grenander and Karhunen tahun 1950an memperkenalkan analisis data fungsional sebagai pendekatan alternatif untuk menganalisis data fungsional. Ramsay dan Silverman (2005) mengembangkan analisis data fungsional melalui pendekatan regresi fungsional yang kemudian berkembang pesat dalam bidang statistik serta didorong oleh banyaknya aplikasi yang menghasilkan jenis data fungsional. Setiap elemen sampel data fungsional dianggap sebagai fungsi acak. Fungsi-fungsi yang didefinisikan pada umumnya adalah waktu, panjang gelombang, lokasi spasial, probabilitas, dan lainnya.

Data fungsional memiliki struktur data dengan dimensi yang besar, dan menjadi sumber informasi yang luas, namun sangat memungkinkan terdapat permasalahan dalam menganalisis data fungsional. Permasalahan yang mungkin akan dihadapi adalah teori perkembangan dan komputasi. Regresi komponen utama fungsional (RKUF) adalah salah satu metode yang dikembangkan untuk menganalisis data fungsional. Penerapan RKUF telah banyak digunakan, seperti halnya digunakan dalam model regresi linier fungsional untuk data longitudinal (Yao *et. al.*, 2005) dan juga digunakan sebagai teknik pemulusan (Ramsay dan Silverman, 2005). Pendekatan lain yang dikembangkan dalam pemodelan data fungsional selain RKUF adalah regresi kuadrat terkecil parsial fungsional (RKTPF). RKTPF digunakan dalam pemodelan regresi ketika himpunan peubah bebas bertipe fungsional dengan data berupa fungsi atau kurva dari beberapa parameter t yang merupakan waktu atau panjang gelombang (Preda *et. al.*, 2007).

Hasil pengukuran gula darah dengan teknik noninvasif berbentuk spektrum panjang gelombang, dan merupakan salah satu jenis data fungsional. Teknik noninvasif merupakan cara pengukuran gula darah tanpa memberikan rasa sakit pada pasien, karena teknik noninvasif diterapkan dengan menembakkan sinar inframerah ke ujung jari pasien yang kemudian sinar tersebut akan menangkap cahaya yang terlewati dan seterusnya akan diinterpretasi sebagai ukuran gula darah. Teknik pengukuran gula darah lainnya adalah invasif, yaitu dengan cara mengambil sampel darah dengan cara menusukkan jarum kedalam tubuh pasien, hal ini tentu memberikan rasa sakit pada pasien serta terjadi kerusakan sel.

Alat noninvasif yang dikembangkan oleh Biomarking Institut Pertanian Bogor perlu dianalisa akurasi hasil keluaran alat serta kecepatan pengukuran. Metode untuk menganalisa akurasi dan kecepatan alat dapat dilakukan dengan pemodelan

kalibrasi, yaitu menghubungkan hasil pengukuran gula darah dengan teknik noninvasif dan hasil pengukuran gula darah dengan teknik invasif. Tujuan kalibrasi berdasarkan sudut pandang statistika adalah menemukan model $E(Y) = f(\mathbf{X})$, untuk prediksi Y dengan akurasi dan presisi yang tinggi.

Analisis pemodelan kalibrasi pada umumnya akan ditemui masalah dengan jumlah peubah bebas (p) lebih besar dari jumlah responden (n), selain itu jenis data fungsional yang akan diterapkan seperti hasil pengukuran gula darah non-invasif mengakibatkan penerapan pendekatan metode kuadrat terkecil (MKT) dianggap tidak efektif. Pengembangan metode Regresi kontinum fungsional (RKF) yang merupakan pengembangan dari regresi kontinum adalah metode alternatif yang dapat diterapkan pada data fungsional (Zhou, 2019). Algoritma pada RKF hampir sama dengan regresi kontinum yaitu menetapkan peubah laten $\mathbf{y} = \mathbf{T}_h \boldsymbol{\xi} + \boldsymbol{\varepsilon}$ dengan $\mathbf{T}_h = \mathbf{X}\mathbf{W}_h$ dan $\mathbf{W}_h = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_h)$ matriks berukuran $p \times h$ ($h < p$) dan disebut matriks pembobot (Serneels *et. al.*, 2005). Formulasi matematika matriks pembobot dalam regresi kontinum sebagai berikut (Stone dan Brooks, 1990):

$$\mathbf{w}_{i,\alpha} = \arg \max \left\{ \text{cov}(\mathbf{X}\mathbf{w}, \mathbf{y})^2 \text{var}(\mathbf{X}\mathbf{w})^{\alpha-1} \right\}$$

dengan $\|\mathbf{w}_i\| = 1$ dan $\text{cov}(\mathbf{X}\mathbf{w}_i, \mathbf{y}\mathbf{w}_j) = 0$ untuk $j < i$, dan $\alpha \in [0, 1)$.

Regresi kontinum adalah bentuk umum dari regresi komponen utama (RKU) dan Regresi kuadrat terkecil parsial (RKTP), persamaan (1) akan menjadi MKT ketika $\alpha = 0$, RKU ketika $\alpha = 1$ dan RKTP ketika $\alpha = \frac{1}{2}$.

Data hasil pengukuran gula darah non-invasif sangat mungkin terdapat *noise* yang ditimbulkan akibat dari kecenderungan kesalahan manusia, atau berasal dari instrument atau alat yang digunakan untuk mengumpulkan data. Teknik *smoothing* merupakan salah satu cara untuk menghilangkan atau mengurangi *noise*, hal ini perlu dilakukan karena *noise* dapat mempengaruhi hasil prediksi atau memberikan informasi yang tidak akurat. Teknik *smoothing* yang akan dilakukan dalam penelitian ini menggunakan metode *Wavelet*.

Konsep dasar *Wavelet* yaitu merepresentasikan fungsi asal menjadi fungsi basis atau fungsi *Wavelet* yang dihasilkan dari hasil dilatasi dan translasi fungsi *mother Wavelet* ($\boldsymbol{\psi}$) dan *Father Wavelet* ($\boldsymbol{\varphi}$). *Wavelet* merupakan pengembangan dari *Fourier*, terlihat pada fungsi basis yang digunakan dalam *Fourier* adalah cosinus, sedangkan pada *Wavelet* fungsi basis bergantung pada penetapan fungsi *mother* dan *father Wavelet*. Hasil penerapan secara empirik memperoleh hasil bahwa metode transformasi *Wavelet* memberikan hasil pendugaan yang relatif lebih baik dibandingkan dengan *Fourier* (Barsanti dan Gilmore, 2011).

3.2 Peringkasan Data

Data yang digunakan dalam penelitian ini adalah data primer hasil pengukuran glukosa darah menggunakan alat invasif dan non-invasif. Pengukuran dilakukan pada tahun 2019. Sampel yang digunakan sebanyak 74 orang yang merupakan masyarakat umum yang berusia 21–87 tahun yang dilakukan pada tanggal 13 Juli 2019 di Kebon Pedes, Kota Bogor, Jawa Barat. Pengukuran dilakukan di laboratorium Bio Kimia, Departemen Gizi Masyarakat IPB. Data penelitian merupakan sebagian luaran dari penelitian “Pengembangan dan Uji Klinis Purwarupa Alat Pendeteksian Kadar Glukosa Darah Non-Invasif Institut Pertanian Bogor” tahun 2019.

Hasil pengukuran dengan teknik non-invasif menggunakan metode spektroskopi berupa nilai residu intensitas. Jumlah data residu intensitas yang dimiliki setiap responden berbeda-beda sehingga diperlukan peringkasan data, agar dapat dilakukan analisis data. Peringkasan dilakukan dengan teknik luas pernah dilakukan oleh Rassiyanthi (2020) dan Novia (2021). Teknik luas yang digunakan melibatkan keseluruhan data untuk memperoleh data tunggal yang merupakan nilai luas dari kurva yang terbentuk, sehingga tidak kehilangan informasi pada data (Novia, 2021). Secara keseluruhan data spektrum residu intensitas terdiri atas 5 ulangan, setiap ulangan terdiri atas 10 modulasi, setiap modulasi terdapat 2 data puncak sehingga diperoleh 100 peubah residu intensitas, yang kemudian akan disebut sebagai peubah bebas. Hasil pengukuran dengan teknik invasif berupa kadar glukosa darah dengan nilai tunggal yang akan menjadi peubah respon.

Luas yang akan ditentukan menggunakan pendekatan Aturan Trapezoida. Aturan Trapezoida adalah suatu metode pendekatan integral numerik dengan polinom orde satu. Dalam metode ini, kurva yang berbentuk lengkung di dekatkan dengan garis lurus sedemikian sehingga, bentuk di bawah kurvanya seperti trapesium. Formulasi matematis luasan dengan aturan trapezoida adalah sebagai berikut:

$$\int_a^b f(x) \approx \frac{h}{n} [f(x_0) + f(x_n) + 2 \sum_{i=1}^{n-1} f(x_i)]$$

dengan $h = b - a$, $a = x_0$ adalah batas bawah area yang akan ditentukan luasannya, dalam hal ini berupa *time domain* (ns), $b = x_n$ adalah batas atas area yang akan ditentukan luasannya, $f(x)$ adalah fungsi yang membentuk area luasan, dalam hal ini berupa residu intensitas cahaya (Cd), dan n adalah banyak rentang dalam *time domain*.

3.3 Eksplorasi Data

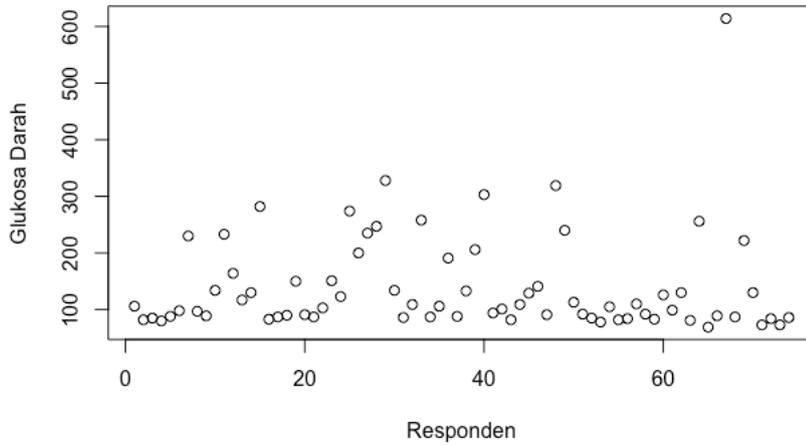
3.3.1 Kadar Glukosa Darah

Data kadar glukosa darah dalam penelitian ini diketahui nilai terendah sebesar 69 mg/dL dan tertinggi sebesar 614 mg/dL, dengan rata-rata sebesar 140.7 mg/dL. *American Diabetes Association* (ADA) membagi kadar glukosa darah kedalam tiga kategori yaitu normal (< 100 mg/dL), pra-diabetes (100 mg/dL - 125 mg/dL) dan diabetes (> 125 mg/dL) (*American Diabetes Association* 2014), sehingga berdasarkan kategori tersebut maka diperoleh data sebanyak 35 responden termasuk dalam kategori normal, 11 responden termasuk dalam kategori pradiabetes dan terdapat 28 responden termasuk dalam kategori diabetes.

Deskripsi data dapat dilihat pada tabel 3, sedangkan visualisasi data ditunjukkan pada Gambar 9. Tampak pada Gambar 9, kadar glukosa darah responden cenderung diatas nilai 100 mg/dL.

Tabel 3 Deskripsi Data Kadar Glukosa Darah

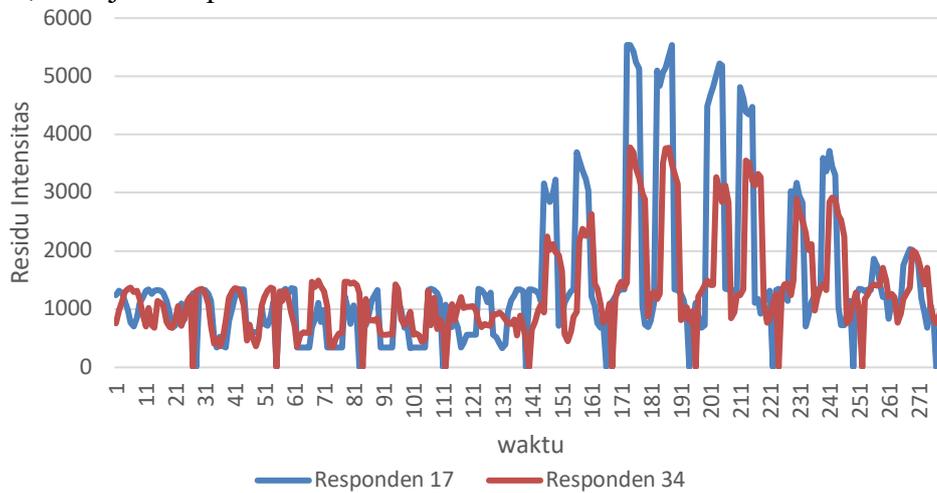
N	Minimum	Maksimum	Rata-Rata	Simpangan Baku
74	69	614	140.7	87.7



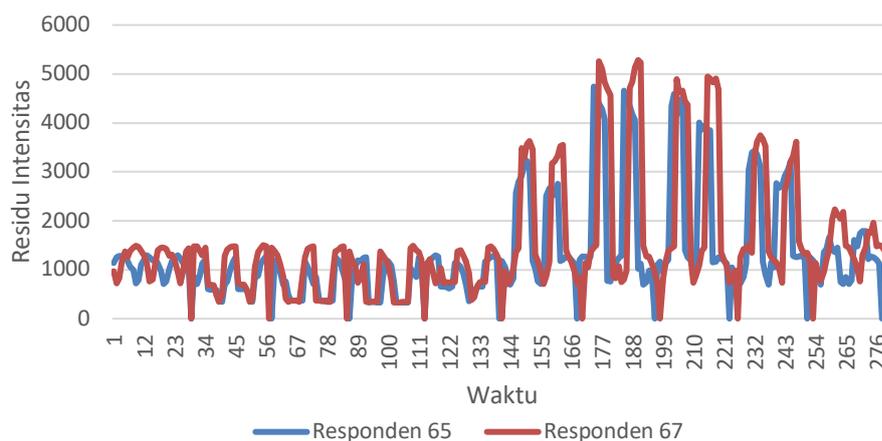
Gambar 9 Visualisasi data kadar glukosa darah

3.3.2 Data Residu Intensitas

Visualisasi data nilai residu intensitas responden 17 dan 34 pada ulangan kedua dengan kadar glukosa darah yang sama yaitu 87 mg/dl, ditunjukkan pada Gambar 10. Visualisasi data nilai residu intensitas responden 65 dan 67 pada ulangan kedua dengan kadar glukosa darah yang berbeda yaitu 69 mg/dl dan 614 mg/dl, ditunjukkan pada Gambar 11.



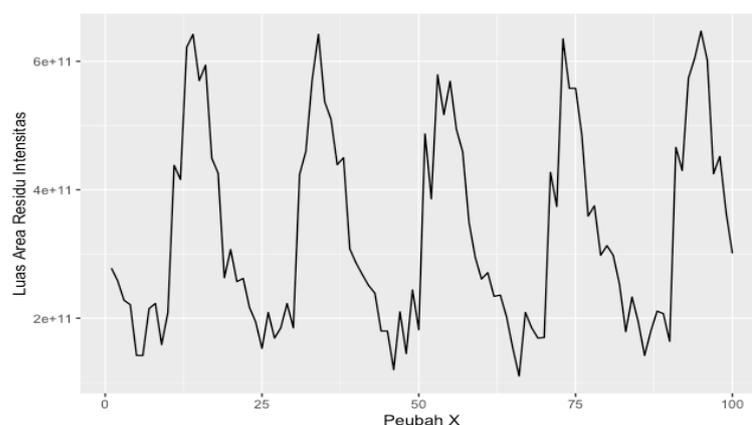
Gambar 10 Visualisasi Kadar Glukosa Darah Responden 17 dan 34



Gambar 11 Visualisasi Kadar Glukosa Dari Responden 65 dan 67

Tampak pada Gambar 10 pola yang dihasilkan dari kedua responden lebih fluktuatif dan beragam, sedangkan pada Gambar 11 pola yang terbentuk hampir serupa antar responden, terlihat pola data saling beririsan. Pada grafik tersebut terlihat jelas bahwa posisi antar pola grafik kedua responden terdapat pergeseran. Hal tersebut dapat disimpulkan jumlah titik dan waktu yang terbaca berbeda-beda antar responden dan antar ulangan, sehingga perlu dilakukan peringkasan (reduksi) data agar diperoleh jumlah titik yang sama antar responden dan ulangan, yang kemudian dapat dilakukan analisis data.

Teknik peringkasan yang dilakukan dengan menentukan nilai luasan dari kurva menggunakan metode numerik aturan trapezoida. Hasil peringkasan diperoleh jumlah variabel setiap responden untuk satu kali ulangan adalah 20 peubah, sehingga diperoleh 100 peubah untuk 5 kali ulangan. Visualisasi data hasil peringkasan pada responden pertama ditunjukkan pada Gambar 12 berikut.



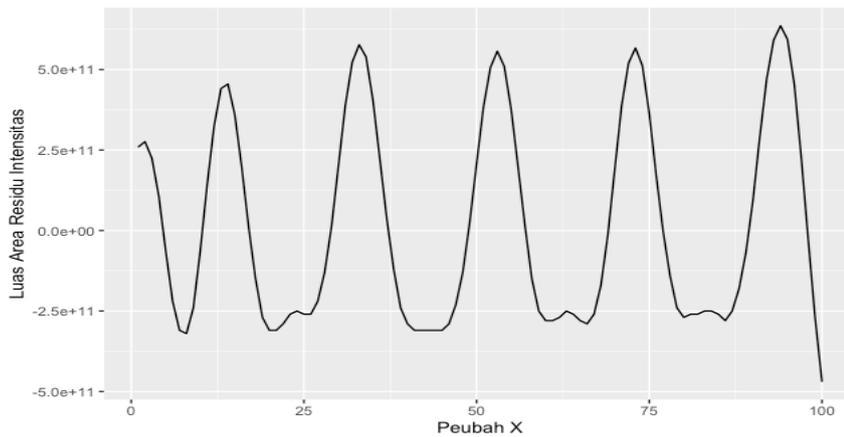
Gambar 12 Visualisasi Data Luasan

Tampak pada Gambar 12 data hasil peringkasan memiliki pola yang konsisten di setiap ulangan, data yang membentuk puncak merupakan data ulangan ke-1 hingga ke-5, hasil ini tampak berbeda dari Gambar 10 dan 11 sebelum dilakukan peringkasan.

3.4 Peringkasan Data

Hasil pengukuran data umumnya mengandung *noise* (gangguan) yang mengakibatkan penurunan kualitas data, serta dapat mempengaruhi hasil analisis statistic, sehingga perlu dilakukan transformasi data melalui teknik pemulusan sebelum data dianalisis. *Wavelet* digunakan untuk mentransformasi data penelitian, dan hasil transformasi diperoleh pola yang dapat dilihat pada Gambar 13, sebagai visualisasi data dari sampel responden pertama hingga kelima.

Tampak pada Gambar 13, pola data lebih mulus apabila dibandingkan dengan Gambar 12, dan cenderung memiliki pola yang sama antar responden.



Gambar 13 Visualisasi Data Hasil Pemulusan dengan Transformasi *Wavelet* pada Responden 1- Responden 5

3.5 Pemodelan Kalibrasi Glukosa Darah Non-Invasif

Data hasil pengukuran kalibrasi pada umumnya memiliki permasalahan multikolinearitas pada peubah bebas. Deteksi multikolinearitas tampak dari nilai korelasi antar peubah bebas yaitu residu intensitas (**X**) yang cukup tinggi, seperti yang terlihat pada Tabel 4 berikut ini.

Tabel 4 Korelasi antar peubah X1-X7

Korelasi	X1	X2	X3	X4	X5	X6	X7
X1	1	0,98	0,92	0,69	0,06	-0,37	-0,53
X2	0,98	1	0,98	0,80	0,20	-0,26	-0,46
X3	0,92	0,98	1	0,90	0,38	-0,10	-0,33
X4	0,69	0,80	0,90	1	0,72	0,30	0,05
X5	0,06	0,20	0,38	0,72	1	0,87	0,70
X6	-0,37	-0,26	-0,10	0,30	0,87	1	0,96
X7	-0,53	-0,46	-0,33	0,05	0,70	0,96	1

Tabel 4 menggambarkan nilai korelasi antar peubah bebas yang diambil sampel yaitu X1 sampai dengan X7. Berdasarkan Tabel 4 tampak bahwa nilai korelasi antar peubah sangat tinggi seperti nilai korelasi antar X1 dan X2 diperoleh 0.98 atau antar X1 dan X3 diperoleh 0.92, sehingga terindikasi permasalahan multikolinearitas, yang mengakibatkan regresi kuadrat terkecil tidak tepat digunakan sebagai metode analisis data.

Pemodelan kalibrasi glukosa darah non-invasif dalam kajian ini menggunakan dua pendekatan fungsional dan non fungsional. Pendekatan fungsional yang diterapkan adalah regresi kontinum fungsional (RKF), dan regresi fungsional (RF), sedangkan pendekatan non fungsional terdiri dari regresi kuadrat terkecil (RKT), regresi komponen utama (RKU), dan regresi kuadrat terkecil parsial terkecil (RKTP).

Berdasarkan data korelasi antar peubah bebas pada Tabel 4 diketahui terdapat multikolinieritas, sehingga alternatif solusi agar RKT dapat diterapkan yaitu menggunakan pendekatan *stepwise forward selection*, yaitu suatu pendekatan seleksi peubah dengan memilih sekumpulan peubah bebas yang paling relevan untuk menduga peubah respon. Kriteria evaluasi yang digunakan dalam menyeleksi peubah adalah *Akaike Information Criterion* (AIC).

Hasil analisis RKF diperoleh parameter α yang optimum adalah 0.70 dan parameter p adalah 5 (Lampiran 7), dan hasil analisis RKU diperoleh jumlah komponen yang optimum adalah 8 (Lampiran 5), sedangkan hasil analisis RKTP diperoleh jumlah komponen adalah 2 (Lampiran 6). Analisis data dengan RKF diawali dengan melakukan transformasi data kedalam fungsi basis *B-Spline*.

Evaluasi model ditentukan berdasarkan ukuran kebaikan model dengan *root mean square error of prediction* (RMSEP), korelasi antar data aktual dan data dugaan, mean absolute error (MAE), serta nilai relatif untuk mengetahui besaran kesalahan nilai dugaan (\hat{y}_i) terhadap nilai aktual (y_i) yaitu *mean absolute percentage error* (MAPE). Berikut formulasi untuk menghitung keempat kriteria ukuran kebaikan model,

$$\text{RMSEP} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

$$\text{korelasi} = \frac{\sum (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{(n-1)s_y s_{\hat{y}}}$$

s_y dan $s_{\hat{y}}$ merupakan simpangan baku data pengamatan dan dugaan

$$\text{MAE} = \frac{\sum |y_i - \hat{y}_i|}{n}$$

$$\text{MAPE} = \frac{\sum \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} 100\%$$

Hasil analisis kelima metode dapat dilihat pada Tabel 5.

Tabel 5 Deskripsi Hasil Analisis

Metode	Ukuran			
	RMSEP	Korelasi	MAE	MAPE
RKF	86,57	0,1718	58,11	41,48
RKU	86,86	0.1740	58,06	41,47
RKTP	89,05	0.1397	61,35	44,99
RKT	101,21	0.0594	95,49	55,85
RF	121,9	0,3346	95,49	83,10

Tampak pada Tabel 5 diketahui rata-rata ukuran kebaikan model dari hasil analisis. Berdasarkan nilai RMSEP diketahui metode terbaik adalah RKF yang memperoleh nilai RMSEP sebesar 86,57, metode terbaik kedua adalah metode RKU diperoleh nilai RMSEP 86,86, selanjutnya RKTP sebesar 89,05, sedangkan

nilai RMSEP yang diperoleh RKT dan RF memiliki selisih yang cukup tinggi dibandingkan dengan RKF, RKU, dan RKTP, yaitu sebesar 101,21 dan 121,9.

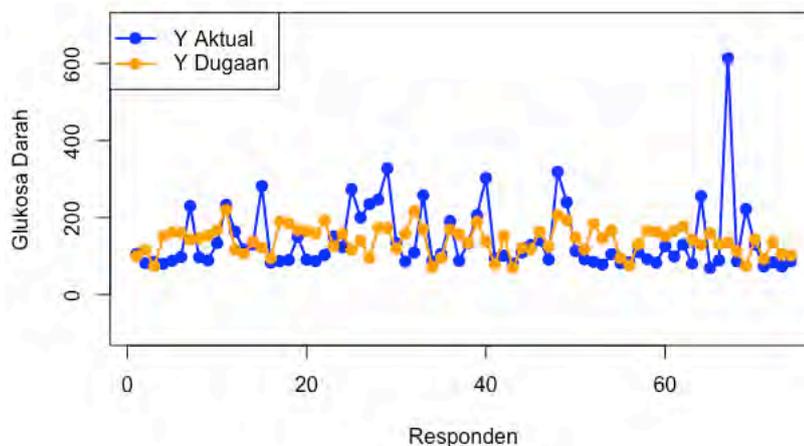
Berdasarkan nilai MAE yang diperoleh, pendekatan RKU lebih baik dari yang lainnya, karena memperoleh nilai MAE terkecil yaitu 58,06, namun nilai MAE yang diperoleh tidak berbeda jauh dari RKF yaitu 58,11, sehingga selisihnya adalah 0,05. RKTP memperoleh nilai MAE 61,35 lebih tinggi dibandingkan RKU dan RKF, dan lebih baik dibandingkan dengan RKT dan RF yang masing-masing diperoleh nilai MAE sebesar 95,49. Nilai MAE dianggap ukuran yang lebih baik dibandingkan dengan RMSEP karena nilai mutlak selisih sisaan dalam formulasinya menjadikan MAE tidak dipengaruhi oleh pencilan, sedangkan RMSEP dipengaruhi oleh pencilan karena dalam formulasinya kuadrat selisih sisaan.

Berdasarkan nilai korelasi antara data aktual dengan data dugaan, tampak metode RF memperoleh nilai korelasi tertinggi sebesar 33,46%, sedangkan RKF dengan RKU memperoleh nilai korelasi yang tidak jauh berbeda yaitu 17,18% dan 17,40%, selanjutnya RKTP diperoleh nilai korelasi sebesar 13,97%, dan RKT memperoleh nilai korelasi yang sangat berbeda jauh dengan metode lainnya yaitu sebesar 5,94%.

Hasil MAPE yang diperoleh seperti halnya yang diperoleh MAE, RKU lebih baik karena diperoleh nilai MAPE terkecil yaitu 41,47 dibandingkan dengan yang lainnya, namun nilai MAPE RKU hanya selisih 0,01 lebih baik dibandingkan dengan RKF yaitu 41,48. Selanjutnya RKTP diperoleh nilai MAPE sebesar 44,99 lebih tinggi dibandingkan RKU dan RKF, namun lebih kecil dibandingkan dengan RKT dan RF yang masing-masing diperoleh nilai MAPE sebesar 55,85 dan 83,10.

Hasil analisis yang diperoleh menunjukkan performa RKF hampir sama dengan RKU. Hal ini diduga dipengaruhi oleh karakteristik data yang kurang sesuai dengan asumsi model fungsional, atau hubungan antar peubah bebas dan peubah respon kompleks, sehingga pendekatan fungsional tidak dapat menangkap hubungan antar peubah bebas dan peubah respon dengan baik.

Kebaikan model RKF dapat dilihat pada pola dan korelasi antara nilai dugaan dengan data aktual seperti yang ditunjukkan pada Gambar 14. Tampak pada Gambar 14 pola data dugaan tidak sepenuhnya mengikuti pola data aktual, hal ini yang sesuai dengan besaran korelasi yang diperoleh yaitu 17,18%.



Gambar 14 Grafik data aktual dengan nilai dugaan

3.6 Kesimpulan

Pada kajian ini menunjukkan bahwa metode RKF memiliki performa yang cukup baik dalam pemodelan kalibrasi, meskipun ukuran kebaikan model yang diperoleh tidak terlalu berbeda jauh dengan metode RKU dan RKTP. Hasil RMSEP, MAE dan MAPE yang diperoleh oleh masing-masing metode cukup besar, serta korelasi cukup kecil. Hal ini menunjukkan fenomena data aktual yang diperoleh dengan teknik invasif tidak cukup baik digambarkan oleh data yang diperoleh dengan teknik non invasif.

Banyak faktor yang mempengaruhi fenomena data baik eksternal ataupun internal alat ukur. Berdasarkan data yang diperoleh maka perlu dilakukan evaluasi terhadap alat ukur non invasif yang telah dikembangkan, sehingga data yang diperoleh dengan teknik non invasif dapat mendekati nilai akurasi dan prediksi dari alat ukur invasif sebagai metode konvensional, dan metode non invasif dapat menjadi acuan yang digunakan di bidang kesehatan dan kedokteran.

Dari hasil analisis dapat dikatakan bahwa metode reduksi seperti RKF, RKU, dan RKTP merupakan metode yang unggul apabila digunakan ketika terjadi multikolinearitas dengan jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan. Berdasarkan hasil analisis yang diperoleh metode RF dan RKT sebagai metode yang non reduksi menunjukkan performa kurang efektif digunakan untuk kasus dengan jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan. Secara menyeluruh dapat dikatakan bahwa pendekatan RKF lebih baik karena sifat fungsional yang dimiliki tidak menghilangkan pola data. Dengan demikian dapat dikatakan bahwa RKF merupakan pendekatan yang efektif dalam menganalisis model kalibrasi.

IV KAJIAN KEKEKARAN BERBAGAI METODE PEMODELAN KALIBRASI

4.1 Pendahuluan

Pencilan (*outlier*) adalah observasi yang secara signifikan berbeda atau jauh dari pola umum data. Pencilan dapat merupakan nilai yang sangat rendah atau sangat tinggi dibandingkan dengan nilai-nilai lain dalam data. Pencilan sering kali dianggap sebagai data yang tidak biasa, ekstrem, atau anomali dalam suatu distribusi atau dataset.

Pencilan dapat muncul karena beberapa alasan, termasuk kesalahan pengukuran, kesalahan pengumpulan data, variasi alami dalam populasi, atau kejadian yang jarang atau tidak umum. Pencilan dapat menjadi sumber ketidakakuratan dalam analisis statistik dan dapat mempengaruhi hasil secara signifikan jika tidak dikelola dengan baik. Pencilan dapat memiliki efek yang signifikan pada analisis data, termasuk pengaruh terhadap deskripsi statistik, pendugaan parameter, dan pemodelan statistik. Pencilan yang tidak diatasi dapat menyebabkan kesalahan interpretasi atau pengambilan kesimpulan yang salah.

Dalam konteks model kalibrasi, pencilan mengacu pada data yang tidak konsisten atau tidak cocok dengan model yang dikembangkan. Data ini biasanya merupakan data observasi yang ekstrem atau data yang tidak sesuai dengan pola umum dalam dataset. Keberadaan pencilan dapat mempengaruhi kualitas dan keandalan model kalibrasi. Dalam pemodelan kalibrasi, pencilan dapat memiliki dampak yang signifikan terhadap hasil estimasi atau prediksi. Pencilan dapat menyebabkan bias dalam parameter estimasi dan meningkatkan kesalahan prediksi.

Cara singkat penanganan pencilan adalah penghapusan data untuk mencegah pengaruhnya pada estimasi atau prediksi. Namun, penghapusan data harus dilakukan dengan hati-hati dan hanya jika pencilan diyakini sebagai kesalahan pengukuran. Dalam beberapa kasus, pencilan mungkin mengandung informasi penting atau menunjukkan keadaan yang sebenarnya, oleh karena itu, penting untuk mempertimbangkan asal-usul dan penyebab pencilan sebelum mengambil tindakan terhadap pencilan, serta pemilihan metode kekar yang akan digunakan untuk membentuk model kalibrasi.

Pengembangan metode kekar untuk mengatasi pencilan telah banyak dilakukan, seperti metode *M-estimation* yang dikembangkan oleh Beaton dan Tukey (1974), Huber (1981) mengembangkan fungsi pembobot Huber, Harrell (2012) mengembangkan metode estimasi kekar menggunakan regresi kuadrat terkecil, Rousseeuw dan Leroy (2005) mengembangkan metode *M-estimation* yang lebih kekar dengan fungsi pembobot yang memuat fungsi pembobot *Bisquare* dan Hampel, Donoho dan Johnstone (1995) mengembangkan metode estimasi kekar menggunakan transformasi *Wavelet*.

Berdasarkan hasil kajian terapan yang telah dilakukan pada BAB sebelumnya, diketahui performa regresi kontinum fungsional (RKF) tidak cukup baik dalam pemodelan kalibrasi glukosa darah noninvasif. Hal tersebut dapat dikaji lebih dalam terhadap kekekar RKF beserta keempat metode lainnya regresi fungsional (RF), serta tiga metode non fungsional yaitu regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), dan regresi kuadrat terkecil (RKT).

Kekekaraan masing-masing metode diamati berdasarkan ukuran kebaikan model dengan nilai *Root Mean Square Error Prediction* (RMSEP), korelasi antar peubah respon pengamatan dengan peubah respon dugaan, *mean absolute error* (MAE), dan *mean absolute percentage error* (MAPE). Ukuran MAPE memiliki kekurangan diantaranya, ketika nilai dugaan (\hat{y}_i) atau nilai aktual (y_i) adalah bilangan negatif maka akan diperoleh nilai MAPE lebih besar dari 100% (Kim dan Kim, 2016). Berdasarkan hal tersebut maka dalam penelitian ini, ukuran MAPE direvitalisasi menjadi akar dari MAPE untuk kajian simulasi.

4.2 Analisis Kekekaran Berbagai Metode Pemodelan Kalibrasi

Analisis kekekaran untuk kelima metode yaitu RKF, RKU, RKTP, RKT, dan RF menggunakan metode simulasi. Metode simulasi dapat memberikan efisiensi dan kemudahan dalam menganalisis suatu model matematika. Metode simulasi yang digunakan dalam penelitian ini menggunakan data bangkitan sebanyak sembilan puluh satu peubah bebas, dan satu peubah respon dengan jumlah pengamatan (n) yang bervariasi yaitu 50, 100, dan 200 dan proporsi pencilan 2% hingga 10%.

Data dibangkitkan sebanyak 365 peubah bebas yang merupakan fungsi $\sin\left(\frac{k\pi}{45}\right) +$ distribusi uniform $(365, -0.1, 0.1)$, dengan $k = [0, 364]$, dan 1 peubah respon $y_i = \sum_{j=1}^{365} x_{i,j}$, $i = 1, 2, \dots, n$, masing-masing peubah dibangkitkan sebanyak 300 ulangan. Data simulasi dianalisis menggunakan kelima metode RKF, RKU, RKTP, RKT, dan RF dengan ulangan sebanyak 100 kali, sehingga diperoleh dari masing-masing metode nilai rata-rata RMSEP, korelasi antar peubah respon pengamatan dengan peubah respon dugaan, dan rata-rata MAE.

Tabel 6 berikut ini menunjukkan nilai rata-rata RMSEP, korelasi dan MAE yang diperoleh dari masing-masing metode berdasarkan proporsi pencilan dan jumlah pengamatan yang bervariasi.

Tabel 6 Ukuran Kebajikan Metode Menggunakan Data Simulasi

Observasi	Pencilan	RMSEP					Korelasi					MAE					MAPE (%)				
		RKF	RKU	RKTP	RKT	RF	RKF	RKU	RKTP	RKT	RF	RKF	RKU	RKTP	RKT	RF	RKF	RKU	RKTP	RKT	RF
N=50	2%	2,201	2,415	2,632	4,463	2,331	0,068	0,01	0,012	-0,008	0,042	1,153	1,486	1,761	3,49	1,359	23,343	23,125	23,357	26,559	23,367
	3%	2,684	2,914	3,166	5,191	2,832	0,027	-0,003	0,005	0,016	0,008	1,538	1,899	2,209	4,122	1,754	24,4	24,422	24,392	28,943	24,391
	4%	3,041	3,302	3,587	5,938	3,229	0,017	0,003	0,015	0,01	-0,005	1,859	2,237	2,575	4,752	2,105	25,604	26,009	25,611	30,391	25,613
	5%	3,4	3,639	3,92	6,483	3,568	-0,005	0,009	0,027	0,025	-0,009	2,216	2,563	2,897	5,249	2,427	26,527	27,156	26,537	32,03	26,545
	6%	3,704	3,953	4,263	6,918	3,872	-0,035	-0,003	0,016	0,044	-0,027	2,521	2,852	3,205	5,584	2,715	27,282	28,197	27,306	29,739	27,312
	7%	3,959	4,22	4,528	7,597	4,133	-0,045	-0,008	0,019	0,011	-0,043	2,82	3,132	3,475	6,166	2,987	27,083	27,81	27,097	32,719	27,16
	8%	4,14	4,409	4,738	7,85	4,352	-0,038	0,016	0,037	0,044	-0,04	3,035	3,352	3,693	6,373	3,228	28,068	28,899	28,063	34,576	28,06
	9%	4,363	4,604	5,031	8,407	4,544	-0,033	0,019	0,024	0,011	-0,027	3,304	3,569	3,976	6,856	3,466	28,828	29,637	28,796	34,003	28,804
	10%	4,558	4,805	5,255	8,89	4,745	-0,058	0,007	0,014	0,041	-0,038	3,544	3,783	4,194	7,243	3,691	29,539	30,568	29,534	34,57	29,538
	N=100	2%	1,471	1,606	1,96	2,891	1,522	0,256	0,09	0,091	0,049	0,21	0,546	0,826	1,273	2,082	0,661	15,773	15,972	15,773	18,126
3%		1,802	1,941	2,38	3,422	1,857	0,176	0,053	0,06	0,037	0,138	0,752	1,037	1,593	2,465	0,869	16,638	17,079	16,638	18,628	16,644
4%		2,065	2,214	2,718	3,869	2,129	0,144	0,041	0,053	0,024	0,105	0,943	1,224	1,863	2,821	1,061	18,271	19,006	18,271	22,834	18,272
5%		2,306	2,448	3,029	4,258	2,369	0,113	0,036	0,044	0,025	0,081	1,133	1,399	2,128	3,134	1,239	18,736	19,546	18,736	21,896	18,741
6%		2,516	2,667	3,307	4,598	2,581	0,087	0,021	0,034	0,001	0,06	1,309	1,566	2,347	3,394	1,407	19,583	21,409	19,583	24,021	19,58
7%		2,713	2,852	3,555	4,961	2,773	0,053	0,019	0,026	0,004	0,035	1,485	1,729	2,577	3,707	1,574	20,211	21,044	20,211	24,545	20,213
8%		2,889	3,015	3,765	5,256	2,944	0,044	0,029	0,031	0,005	0,034	1,654	1,881	2,761	3,926	1,736	21,585	22,934	21,585	26,427	21,597
9%		3,029	3,181	3,972	5,484	3,097	0,05	0,027	0,03	0,014	0,035	1,805	2,038	2,94	4,109	1,898	22,288	24,088	22,288	27,728	22,292
10%		3,182	3,333	4,193	5,848	3,252	0,023	0,019	0,017	-0,008	0,015	1,958	2,187	3,135	4,357	2,047	23,935	25,776	23,935	29,127	23,934
N=200		2%	1,016	1,136	1,258	1,296	1,037	0,444	0,198	0,255	0,195	0,414	0,256	0,537	0,704	0,749	0,326	13,619	13,625	13,619	16,15
	3%	1,243	1,358	1,537	1,548	1,268	0,361	0,143	0,186	0,143	0,328	0,36	0,632	0,892	0,902	0,432	14,22	14,218	14,22	17,406	14,221
	4%	1,431	1,544	1,766	1,753	1,459	0,311	0,115	0,151	0,118	0,276	0,465	0,721	1,051	1,033	0,533	14,962	14,878	14,88	18,488	14,883
	5%	1,601	1,71	1,97	1,94	1,628	0,27	0,096	0,127	0,095	0,238	0,563	0,81	1,199	1,155	0,627	15,042	15,403	15,43	18,693	15,387
	6%	1,75	1,862	2,156	2,099	1,778	0,239	0,073	0,105	0,087	0,207	0,655	0,895	1,333	1,267	0,718	15,255	15,919	15,563	19,57	15,791
	7%	1,889	1,998	2,32	2,247	1,915	0,205	0,059	0,093	0,072	0,178	0,754	0,984	1,463	1,373	0,808	15,543	16,579	16,48	20,695	16,437
	8%	2,015	2,121	2,469	2,384	2,042	0,19	0,061	0,089	0,075	0,164	0,846	1,07	1,584	1,483	0,9	15,872	17,048	17,203	21,201	16,874
	9%	2,128	2,237	2,615	2,537	2,158	0,179	0,057	0,079	0,058	0,151	0,936	1,153	1,699	1,598	0,992	16,141	17,683	17,828	21,874	17,331
	10%	2,244	2,349	2,759	2,645	2,27	0,156	0,046	0,064	0,052	0,133	1,027	1,233	1,81	1,672	1,076	16,376	18,24	19,036	22,274	17,844

Pembangkitan data yang dirancang berbasis fungsional, sehingga dapat diasumsikan bahwa data memiliki tingkat korelasi antar peubah bebas yang tinggi, atau dapat dikatakan terjadi multikolinearitas. Berdasarkan asumsi tersebut maka penetapan metode RKU dan RKTP, yang merupakan metode reduksi peubah (komponen) sebagai evaluasi terhadap kinerja RKF dianggap sudah tepat. Pendekatan RF dan RKT sebagai metode non reduksi juga menjadi evaluasi bagi metode reduksi (RKF, RKU, dan RKTP). Tampak bahwa RKT sebagai metode non reduksi memperoleh nilai RMSEP dan MAE paling besar. Hal ini berlaku konsisten ketika jumlah data ataupun persentase pencilan meningkat.

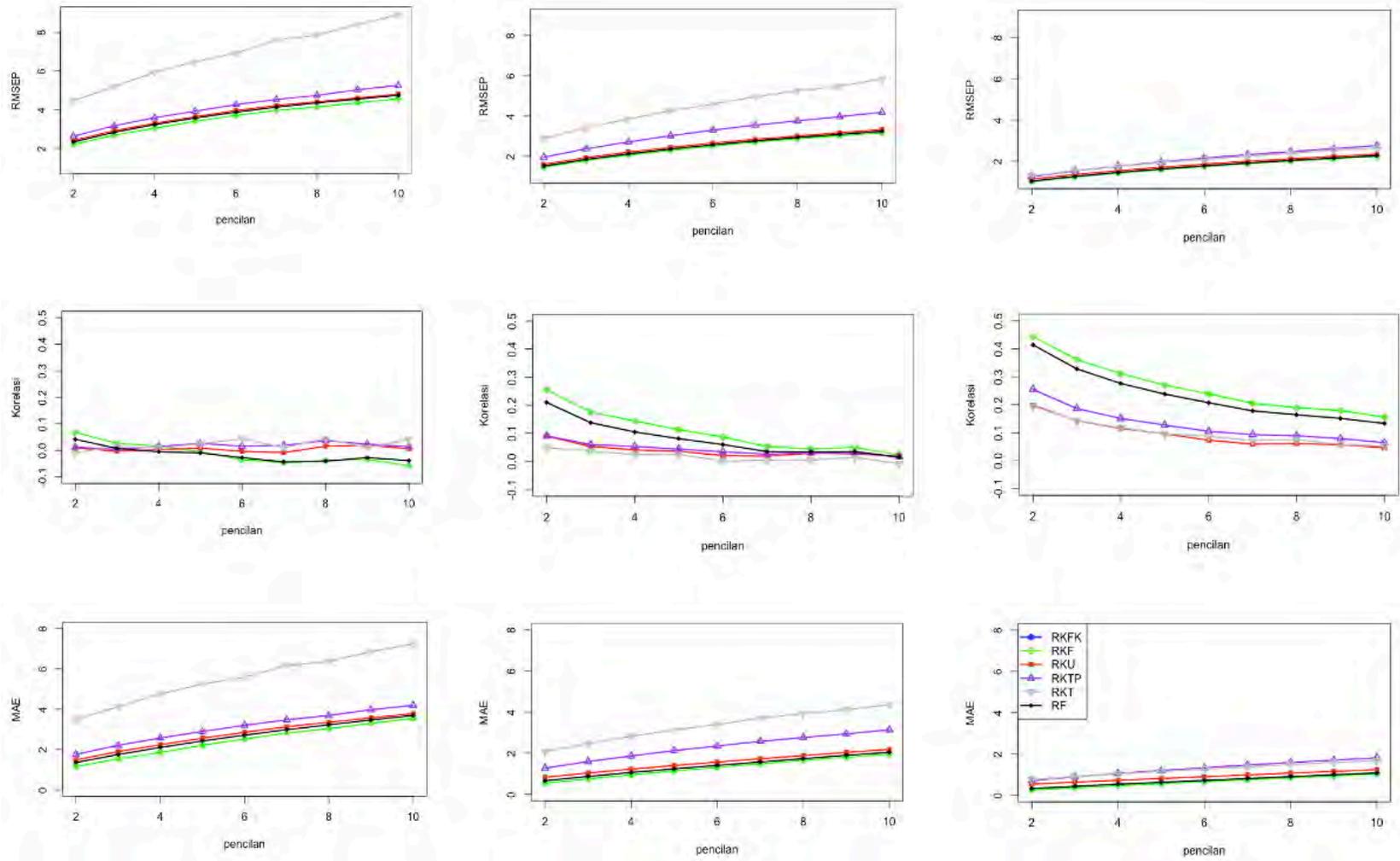
Tampak pada Tabel 6 nilai rata-rata RMSEP dan MAE yang diperoleh dari masing-masing metode semakin membesar seiring dengan bertambahnya bobot pencilan yang diberikan. Berbeda dengan RMSEP dan MAE, nilai korelasi semakin menurun seiring dengan bertambahnya bobot pencilan. Hal ini dapat diamati pada Gambar 15. Demikian sehingga dapat dikatakan kelima metode tidak resisten terhadap keberadaan pencilan.

Berdasarkan hasil analisis terlihat nilai rata-rata RMSEP yang diperoleh metode RKF lebih kecil dari keempat metode lainnya pada kondisi $n = 50, 100,$ dan 200 serta bobot pencilan dari 2% hingga 10%, namun nilai rata-rata RMSEP metode RKF yang diperoleh tidak terlalu berbeda jauh dengan metode RF yang juga merupakan metode fungsional, kondisi ini juga serupa dengan nilai rata-rata MAE yang diperoleh metode RKF yang lebih kecil dari metode lainnya.

Berdasarkan nilai rata-rata korelasi antara peubah respon pengamatan dengan peubah respon pendugaan, metode RKF memperoleh nilai korelasi yang lebih baik dibandingkan keempat metode lainnya. Hasil analisis ini dapat diamati pada visualiasi nilai ukuran RMSEP, korelasi dan MAE untuk kelima metode berturut-turut pada Gambar 15.

Nilai relatif sebagai ukuran kesalahan persentase mutlak, diukur menggunakan kuadrat rata-rata MAPE. Nilai relatif yang telah diperoleh dari masing-masing metode terlihat, terjadi penurunan pada saat jumlah data (n), dan terjadi peningkatan pada saat tingkat persentase bobot dinaikkan secara konsisten pada kondisi $n = 50, 100,$ dan 200 . Nilai relatif antar metode memiliki perbedaan, namun selisih perbedaan antar metode tidak cukup jauh berbeda, kecuali metode RKT. Hal ini dapat dikatakan bahwa metode yang bersifat reduksi dan/atau bersifat fungsional memiliki nilai persentase relatif yang lebih baik dalam mengatasi permasalahan jumlah peubah bebas lebih besar dari jumlah amatan, serta pencilan.

Tampak jelas pada Gambar 15 terjadi penurunan nilai RMSEP dan MAE pada masing-masing metode seiring dengan bertambahnya jumlah data (n). Hal tersebut berbanding terbalik dengan nilai korelasi yang diperoleh. Semakin meningkat seiring dengan bertambahnya jumlah data. Grafik RKF yang ditunjukkan pada garis berwarna orange konsisten selalu berada paling bawah untuk nilai rata-rata RMSEP dan MAE pada variasi jumlah data (n) 50, 100, dan 200, sedangkan untuk nilai korelasi grafik RKF selalu berada paling atas.



Gambar 15 Visualisasi Ukuran Kebajikan Model Data Simulasi

4.3 Kesimpulan

Berdasarkan hasil analisis kekekaran metode RKF, RKU, RKTP, RKT, dan RF, dengan teknik simulasi, yang dirancang dengan jumlah pengamatan, dan juga proporsi pencilan yang bervariasi, dapat disimpulkan bahwa semakin banyak jumlah pengamatan nilai rata-rata RMSEP dan MAE semakin menurun untuk kelima metode, sedangkan nilai korelasi antar peubah respon pengamatan dengan peubah respon dugaan terus meningkat. Data simulasi yang dirancang bersifat fungsional, dengan asumsi terjadi multikolinearitas, dapat disimpulkan pendekatan RKT sebagai metode non reduksi memiliki performa yang kurang baik. RF juga merupakan metode non reduksi, namun sifat fungsional yang dimiliki RF menjadikan metode ini memiliki performa yang lebih baik dibandingkan dengan RKT.

Kelima metode menunjukkan ketidakkekaran berdasarkan nilai RMSEP dan MAE yang terus meningkat seiring bertambahnya nilai proporsi pencilan, dan nilai korelasi konsisten menurun. Peningkatan jumlah data mempengaruhi ukuran kebaikan model, nilai RMSEP dan MAE konsisten menurun, dan nilai korelasi konsisten membesar dengan meningkatnya jumlah data.

Perbandingan antara metode fungsional dengan non fungsional dari hasil analisis, diketahui bahwa metode fungsional RKF dan RF menghasilkan nilai RMSEP dan mae yang lebih kecil dibandingkan dengan metode yang bukan fungsional yaitu RKU, RKTP dan RKT, selain itu nilai korelasi yang diperoleh oleh metode fungsional lebih besar dibandingkan dengan ketiga metode yang bukan fungsional, hal ini terjadi secara konsisten meskipun jumlah pengamatan dan nilai proporsi pencilan semakin membesar. Hasil analisis pada kajian simulasi ini berbeda dengan hasil kajian terapan yang dilakukan pada BAB 3 sebelumnya. Hal ini karena data simulasi peubah bebas dan respon yang dirancang berasal dari fungsi kontinu.

Hasil kajian simulasi terhadap kekekaran dari kelima metode, secara menyeluruh dapat disimpulkan RKF lebih baik dibandingkan dengan keempat metode lainnya. Metode fungsional seperti RKF dan RF lebih efektif digunakan ketika data memenuhi asumsi fungsional, dibandingkan dengan metode non fungsional.

V PENGEMBANGAN METODE REGRESI KONTINUM FUNGSIONAL KEKAR

5.1 Pendahuluan

Regresi kontinum (RK) yang dikembangkan oleh Stone dan Brooks (1990) merupakan gabungan dari metode regresi kuadrat terkecil (RKT), regresi kuadrat terkecil parsial (RKTP) dan regresi komponen utama (RKU). Prinsip kerja RK didasari oleh penentuan jumlah peubah laten dan penentuan atau diseleksi parameter tuning α dalam interval $[0, 1)$ melalui validasi silang yang memungkinkan seseorang untuk memutuskan nilai α mana yang terbaik untuk data yang ada (Qannari dan Hanafi, 2005). RK mengkonstruksi peubah baru dengan mengkombinasikan prinsip pada RKT, RKTP dan RKU yaitu dengan memaksimalkan keragaman peubah X serta koragam antara peubah X dan Y. Regresi kontinum pengembangan dari RKT, RKTP dan RKU jika α sama dengan 0, 0,5 atau mendekati 1 (Serneels *et. al.*, 2005).

Eksplorasi terhadap tiga metode, yaitu RK, RKU, RKTP yang dilakukan terhadap data simulasi pada berbagai struktur korelasi matriks peubah bebas menunjukkan bahwa RK mempunyai kinerja yang lebih unggul dibanding metode lainnya (Setiawan dan Notodiputro, 2009). RK terus mengalami perkembangan yang dilakukan oleh banyak peneliti, seperti Björkström dan Sundberg (1999) memberikan gambaran umum mengenai pendekatan RK serta menunjukkan perbedaan yang berhubungan dengan regresi kuadrat terkecil (RKT). Serneels *et. al.*, (2005) mengembangkan *robust continuum regression* yang resisten terhadap pencilan, *continuum power regression* oleh Jong *et. al.* (2001), pendekatan regresi kontinum sederhana oleh Qannari dan Hanafi (2005), dan regresi kontinum fungsional oleh Zhou (2019).

Regresi kontinum fungsional (RKF) merupakan pengembangan dari RK yang dapat mengakomodir data fungsional. Prinsip RKF mengacu pada metode analisis data fungsional (ADF) sebagai cabang statistik yang menangani data yang dapat dikonseptualisasikan sebagai fungsi dari peubah kontinu. Data dalam ADF umumnya berbentuk kurva (atau permukaan) yang halus dalam dimensi waktu atau ruang (Wang *et. al.*, 2016).

Data fungsional memiliki karakteristik khusus yang membuatnya sulit diolah dengan metode regresi konvensional yang dirancang untuk data tabular biasa. Pada umumnya data fungsional memiliki dimensi yang besar, terdapat multikolinearitas atau dengan kata lain peubah bebas mengalami perubahan di sepanjang waktu tertentu sehingga diduga terdapat interaksi antar peubah, memiliki hubungan tak linier serta bersifat dinamis. Pendekatan analisis untuk data fungsional yang umum digunakan antara lain regresi fungsional, regresi komponen utama fungsional (RKUF), dan regresi kuadrat terkecil parsial fungsional (RKTPF). Hasil kajian dengan data simulasi dan empirik diperoleh bahwa RKF lebih akurat daripada RKUF dan RKTPF baik dalam estimasi ataupun prediksi, dan juga menunjukkan penduga yang konsisten (Zhou, 2019).

Data hasil pengukuran dengan dimensi yang besar sangat berpeluang memiliki data pencilan, karena data memiliki banyak variasi. Data pencilan adalah observasi yang berbeda secara signifikan dari pola umum data (Hassan *et. al.*, 2022).

Keberadaan pencilan dapat mempengaruhi hasil analisis data atau penduga parameter dalam pemodelan. Mengatasi masalah pencilan dalam analisis regresi penting untuk mendapatkan model yang lebih akurat dan dapat diandalkan, tetapi pencilan mungkin juga tidak berakibat buruk melainkan menjadi informasi yang sangat berguna mengenai proses yang sedang diteliti (Montgomery *et. al.*, 2012).

Salah satu teknik yang dapat digunakan untuk mengatasi pencilan adalah dengan memberikan nilai bobot yang berbeda-beda pada setiap amatan (Aguinis *et. al.*, 2013). Fungsi pembobot yang umum digunakan adalah Huber, Hampel, Ramsay, dan Tukey (*Bisquare*). Fungsi pembobot Hampel dan Tukey bernilai 0 untuk kriteria nilai residual lebih besar dari konstanta yang ditetapkan, pemberian bobot 0 pada data pencilan artinya bahwa data tersebut akan dieliminasi, sehingga jumlah data (pengamatan) akan berkurang. Hal ini sangat disayangkan karena tidak semua pencilan berdampak buruk seperti yang dinyatakan oleh Montgomery *et. al.*, (2012). Selain itu keputusan mengeliminasi pada data hasil pengukuran sangat disayangkan karena proses memperoleh data hasil pengukuran lebih sulit dibandingkan dengan data survei karena membutuhkan biaya yang besar dengan proses yang lama mulai dari persiapan hingga keluar hasil (Ismah *et. al.*, 2009), oleh karena itu keputusan menggunakan fungsi pembobot Huber dalam mengatasi pencilan dianggap tepat.

RKF dinyatakan sebagai metode yang kurang resisten terhadap pencilan. Hal ini berdasarkan hasil penelitian Zhou (2019), dan Ismah *et. al.* (2024) melalui kajian simulasi dan empirik. Hasil eksplorasi terhadap metode RK hingga permasalahan data fungsional dan pencilan, maka perlu dikembangkan metode RK yang dapat mengatasi kedua permasalahan tersebut.

Kebaruan dalam penelitian ini adalah menerapkan fungsi pembobot Huber untuk mendeteksi dan mengatasi pencilan. Teknis pembobotan pada data observasi dilakukan dengan membangkitkan masing-masing data observasi sebanyak nilai bobot yang telah dinormalisasi dengan cara membagi masing-masing pembobot dengan nilai pembobot yang paling kecil. Data observasi yang telah dibangkitkan akan dilakukan pemodelan menggunakan pendekatan RKF. Teknik ini yang kemudian disebut regresi kontinum fungsional kekar (RKFK).

5.2 Kajian Teoritis

Dalam kajian ini akan diuraikan ekivalensi dari basis kontinum fungsional, berdasarkan proposisi berikut ini.

Proposisi 2. Misalkan $X^{(1,\alpha)} = X - E(X)$ dan $Y^{(1,\alpha)} = Y - E(Y)$. Untuk setiap bilangan bulat $j \geq 2$, diberikan $w_{k,\alpha}$ yang memenuhi $\int_{\mathbb{T}} w_{k,\alpha} V_X(w_{k,\alpha}) > 0$ untuk semua $k \in \{1, \dots, j-1\}$, sehingga dapat dituliskan

$$X^{(j,\alpha)} = X - E(X) - \sum_{k=1}^{j-1} \int_{\mathbb{T}} w_{k,\alpha} \{X - E(X)\} \left\{ \int_{\mathbb{T}} w_{k,\alpha} V_X(w_{k,\alpha}) \right\}^{-\frac{1}{2}} V_X(w_{k,\alpha})$$

dan

$$Y^{(j,\alpha)} = Y - n_{j-1,\alpha}(X) = \int_{\mathbb{T}} \beta X^{(j,\alpha)}.$$

kemudian, $w_{j,\alpha}$ dalam persamaan (17) dapat ditemukan dengan memaksimalkan $T_{j,\alpha}^*$. Pada *unit sphere*, yaitu:

$$W_{j,\alpha} = \arg \max_{w: \|w\|=1} T_{j,\alpha}^*(w),$$

dengan

$$\begin{aligned} T_{j,\alpha}^*(w) &= \text{cov}^2 \left\{ Y^{(j,\alpha)}, \int_{\mathbb{T}} X^{(j,\alpha)} w \right\} \left\{ \int_{\mathbb{T}} w V_{X^{(j,\alpha)}}(w) \right\}^{\frac{\alpha}{1-\alpha}-1} \\ &= \left\{ \int_{\mathbb{T}} \beta V_{X^{(j,\alpha)}}(w) \right\}^2 \left\{ \int_{\mathbb{T}} w V_{X^{(j,\alpha)}}(w) \right\}^{\alpha/(1-\alpha)-1} \end{aligned}$$

Proposisi 3. Tetapkan bilangan bulat $i \in \{1, \dots, n\}$. Misalkan $\hat{X}_i^{(1,\alpha)} = X_i - \bar{X}$ dan $\hat{Y}_i^{(1,\alpha)} = Y_i - \bar{Y}$. Untuk setiap bilangan bulat $j \geq 2$, diberikan $\hat{w}_{k,\alpha}$ dengan $\int_{\mathbb{T}} \hat{w}_{k,\alpha} \hat{V}_X(\hat{w}_{k,\alpha}) > 0$ untuk semua $k \in \{1, \dots, j-1\}$, sehingga dapat dituliskan

$$\hat{X}^{(j,\alpha)} = X_i - \bar{X} - \sum_{k=1}^{j-1} \int_{\mathbb{T}} \hat{w}_{k,\alpha} \{X_i - \bar{X}\} \left\{ \int_{\mathbb{T}} \hat{w}_{k,\alpha} \hat{V}_X(\hat{w}_{k,\alpha}) \right\}^{-\frac{1}{2}} \hat{V}_X(\hat{w}_{k,\alpha})$$

dan

$$\hat{Y}^{(j,\alpha)} = Y_i - \hat{n}_{j-1,\alpha}(X_i) = \int_{\mathbb{T}} \beta \hat{X}_i^{(j,\alpha)}.$$

kemudian

$$\hat{W}_{j,\alpha} = \arg \max_{w: \|w\|=1} \hat{T}_{j,\alpha}^*(w), \quad (25)$$

dengan

$$\begin{aligned} \hat{T}_{j,\alpha}^*(w) &= \widehat{\text{cov}}^2 \left\{ \hat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \hat{X}^{(j,\alpha)} w \right\} \left\{ \int_{\mathbb{T}} w \hat{V}_X(w) \right\}^{\alpha/(1-\alpha)-1} \\ &= \left\{ \int_{\mathbb{T}} \beta \hat{V}_X(w) \right\}^2 \left\{ \int_{\mathbb{T}} w \hat{V}_X(w) \right\}^{\alpha/(1-\alpha)-1}. \end{aligned}$$

dengan $\hat{V}_{\hat{X}^{(j,\alpha)}} = \hat{V}_{\hat{X}^{(j,\alpha)}}(s, t) = \sum_{i=1}^n \hat{X}_i^{(j,\alpha)}(s) \hat{X}_1^{(j,\alpha)}(t) / n$.

Persamaan (17) dan (25), masing-masing basis kontinum fungsional didefinisikan sebagai himpunan pemaksimal dari masalah optimasi. Proposisi 4 berikut ini diperoleh bentuk alternatif tetapi lebih eksplisit dari solusi yang dikembangkan dengan menyesuaikan fungsi β .

Proposisi 4. Diberikan $\alpha \in [0, 1)$ dan $w_{1,\alpha}, \dots, w_{j-1,\alpha}$. Misal $\lambda_k^{(j,\alpha)}$ menunjukkan nilai eigen ke- k teratas dari $V_{X^{(j,\alpha)}}$ dengan fungsi eigen yang sesuai $\phi_k^{(j,\alpha)}$. Misalkan $\lambda_1^{(j,\alpha)}$ memiliki nilai $m \geq 1$, yaitu $\lambda_1^{(j,\alpha)} = \dots = \lambda_m^{(j,\alpha)} > \lambda_{m+1}^{(j,\alpha)}$. Jika $V_{X^{(j,\alpha)}}(\beta)$ tidak ortogonal dengan rentang $\{\phi_1^{(j,\alpha)}, \dots, \phi_m^{(j,\alpha)}\}$, maka ada $\delta^{(j,\alpha)} \in (-1, 0) \cup (0, \infty)$ sehingga $w_{j,\alpha}$ adalah panjang satuan dan

$$w_{j,\alpha} \propto \sum_{k=1}^{\infty} \frac{\lambda_k^{(j,\alpha)} \int_{\mathbb{T}} \beta \phi_k^{(j,\alpha)}}{\lambda_k^{(j,\alpha)} + \lambda_1^{(j,\alpha)} / \delta^{(j,\alpha)}} \phi_k^{(j,\alpha)}, \quad (26)$$

dengan tiga nilai batas dari $\delta^{(j,\alpha)}$, yaitu, $-1, 0$ dan ∞ , sesuai dengan fungsional RKU ($\delta^{(j,\alpha)} \rightarrow -1$), fungsional RKTP ($\delta^{(j,\alpha)} \rightarrow 0$) dan fungsional MKT ($\delta^{(j,\alpha)} \rightarrow \infty$), secara berurutan.

Penentuan nilai maksimum pada persamaan (22) sukar untuk dihitung secara analitik karena masih mengandung integral. Bentuk integral atau persamaan yang mengandung integral atau tidak “*closed form*” belum dapat memberikan suatu penyelesaian, kecuali dengan pendekatan numerik menggunakan simulasi dengan ulangan dalam jumlah yang banyak. Zhou (2019) memberikan bentuk yang sederhana untuk dapat diimplementasikan seperti yang ditunjukkan dalam proposisi 5.

Proposisi 5. Perbaiki $\widehat{w}_{1,\alpha}, \dots, \widehat{w}_{j-1,\alpha}$, misalkan $\widehat{\lambda}_k^{(j,\alpha)}$ adalah nilai eigen kth terbesar dari $\widehat{V}_{\widehat{X}^{(j,\alpha)}}$ dengan fungsi eigen yang sesuai $\widehat{\phi}_k^{(j,\alpha)}$. Misalkan $\widehat{\lambda}_1^{(j,\alpha)} = \dots = \widehat{\lambda}_m^{(j,\alpha)} > \widehat{\lambda}_{m+1}^{(j,\alpha)}$. Jika $\widehat{V}_{\widehat{X}^{(j,\alpha)}}(\beta) = \sum_{i=1}^n \widehat{X}_i^{(j,\alpha)} \widehat{Y}_i^{(j,\alpha)} / n$ tidak ortogonal terhadap $\text{span}\{\widehat{\phi}_1^{(j,\alpha)}, \dots, \widehat{\phi}_m^{(j,\alpha)}\}$, maka terdapat $\widehat{\delta}^{(j,\alpha)} \in (-1, 0) \cup (0, \infty)$ sehingga

$$\widehat{w}_{j,\alpha} = \left[\sum_{k=1}^{\infty} \frac{\widehat{cov}^2 \left\{ \widehat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_k^{(j,\alpha)} \right\}}{\left\{ \widehat{\lambda}_k^{(j,\alpha)} + \widehat{\lambda}_1^{(j,\alpha)} / \widehat{\delta}^{(j,\alpha)} \right\}^2} \right]^{-1/2} \sum_{k=1}^{\infty} \frac{\widehat{cov} \left\{ \widehat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_k^{(j,\alpha)} \right\}}{\widehat{\lambda}_k^{(j,\alpha)} + \widehat{\lambda}_1^{(j,\alpha)} / \widehat{\delta}^{(j,\alpha)}} \widehat{\phi}_k^{(j,\alpha)}.$$

Nilai $\widehat{\delta}^{(j,\alpha)}$ yang belum diketahui dalam persamaan (12) dapat diperoleh sebagai berikut

$$\widehat{\delta}^{(p,\alpha)} = \underset{\delta \in (-1,0) \cup (0,\infty)}{\operatorname{argmax}} Q_{j,\alpha}(\delta) = \underset{\delta \in (-1,0) \cup (0,\infty)}{\operatorname{argmin}} -\ln Q_{p,\alpha}(\delta)$$

dengan

$$Q_{p,\alpha}(\delta) = \left\{ \sum_{j=1}^{\infty} \frac{\widehat{cov}^2 \left\{ \widehat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_j^{(j,\alpha)} \right\}}{\widehat{\lambda}_k^{(j,\alpha)} + \frac{\widehat{\lambda}_1^{(j,\alpha)}}{\delta}} \right\}^2 \left\{ \sum_{j=1}^{\infty} \frac{\widehat{cov}^2 \left\{ \widehat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_j^{(j,\alpha)} \right\}}{\left(\widehat{\lambda}_k^{(j,\alpha)} + \frac{\widehat{\lambda}_1^{(j,\alpha)}}{\delta} \right)^2} \right\}^{\frac{\alpha}{(1-\alpha)}}$$

$$\left\{ \sum_{j=1}^{\infty} \frac{\widehat{cov}^2 \left\{ \widehat{Y}^{(j,\alpha)}, \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_j^{(j,\alpha)} \right\} \widehat{var} \left\{ \int_{\mathbb{T}} \widehat{X}^{(j,\alpha)} \widehat{\phi}_j^{(j,\alpha)} \right\}}{\left(\widehat{\lambda}_k^{(j,\alpha)} + \frac{\widehat{\lambda}_1^{(j,\alpha)}}{\delta} \right)^2} \right\}^{\frac{\alpha}{(1-\alpha)-1}}$$

Sebagai upaya untuk mengurangi beban komputasi dan meningkatkan efisiensi maka untuk menghitung $\widehat{X}_i^{(j,\alpha)}$ dan $\widehat{\beta}_{j,\alpha}$ menggunakan cara rekursif yaitu teknik pengulangan yang tidak memerlukan iterasi.

$$\widehat{X}_i^{(p,\alpha)} = \widehat{X}_i^{(j-1,\alpha)} - \widehat{var}^{-\frac{1}{2}} \left(\int_{\mathbb{T}} X \widehat{w}_{j-1,\alpha} \right) \left\{ \int_{\mathbb{T}} \widehat{X}_i^{(1,\alpha)} \widehat{w}_{j-1,\alpha} \right\} \widehat{V}_X(\widehat{w}_{j-1,\alpha})$$

dan

$$\widehat{\beta}_{j,\alpha} = \widehat{\beta}_{j-1,\alpha} + \widehat{cov}(Y, \int_{\mathbb{T}} X \widehat{w}_{j,\alpha}) \widehat{var}^{-\frac{1}{2}} \left(\int_{\mathbb{T}} X \widehat{w}_{j,\alpha} \right) \widehat{w}_{j,\alpha}$$

dimulai dengan $\widehat{X}^{(1,\alpha)} = X_i - \bar{X}$, $\widehat{Y}^{(1,\alpha)} = Y_i - \bar{Y}$ dan $\widehat{\beta}_{0,\alpha} = 0$.

Hasil yang diperoleh RKF bergantung pada penentuan dua nilai parameter yaitu α sebagai parameter kontinum, dan p sebagai jumlah fungsi basis yang dimasukkan dalam model. Penentuan kedua parameter yang optimal dapat melalui kajian simulasi berdasarkan pada nilai RMSEP yang terkecil.

5.3 Usulan Metodologi (Kebaruan)

Konsep dalam penelitian ini adalah mengembangkan pendekatan regresi kontinum fungsional yang resisten terhadap pencilan (kekar) sehingga kemudian dinamakan sebagai regresi kontinum fungsional kekar (RKFK). Prinsip yang dikembangkan adalah dengan melibatkan fungsi pembobot dalam proses analisis, seperti halnya pada regresi kuadrat terkecil terboboti.

Pemberian bobot secara proporsional, dapat mengurangi pengaruh bias dalam pendugaan. Bobot yang diperoleh berdasarkan kriteria pada masing-masing metode pada Tabel 2 akan ditransformasi dengan cara menambahkan nilai bobot dengan konstanta, dan membagi setiap bobot dengan nilai minimum dari sekumpulan bobot. Teknik transformasi dilakukan agar terhindar dari terjadinya reduksi data (bobot nol) seperti yang berlaku pada fungsi pembobot Hampel dan Tukey (*Bisquare*).

Metodologi yang diusulkan dalam penelitian ini menggandakan masing-masing amatan sebanyak jumlah bobot yang diperoleh setelah ditransformasi. Apabila bobot berbentuk data kontinu (desimal), maka lakukan pembulatan kebawah, agar menghindari jumlah data membesar dan berakibat pada lamanya proses komputasi. Tahapan selanjutnya data yang telah digandakan akan dilakukan proses regresi dengan RKF, sehingga diperoleh nilai sisaan. Proses akan berulang hingga diperoleh nilai konvergen yang diperoleh ketika $|e_n - e_{n-1}| \leq 0,00001$. Adapun algoritma RKF yang dikembangkan oleh Zhou (2019) adalah sebagai berikut :

```

for  $(p, \alpha)$  in finite set do           #  $p$  dan  $\alpha \in [0,1)$  adalah parameter tuning
  for  $i$  from 1 to  $n$  do
    if  $p = 1$  then
       $\hat{X}_i^{(p,\alpha)} \leftarrow X_i - \bar{X}$       #  $X_i$  data pengamatan ke-  $i$  (peubah bebas),  $\bar{X}$  rata-rata,  $\hat{X}_i$ 
                                         data dugaan  $X_i$ 
       $\hat{Y}_i^{(p,\alpha)} \leftarrow Y_i - \bar{Y}$     #  $Y_i$  data pengamatan ke-  $i$  (peubah respon),  $\bar{Y}$  rata-rata,  $\hat{Y}_i$ 
                                         data dugaan  $Y_i$ 
       $\hat{\beta}_{p-1,\alpha} \leftarrow 0$          #  $\hat{\beta}$  penduga koefisien model
    else
       $\hat{X}_i^{(p,\alpha)} \leftarrow \hat{X}_i^{(p-1,\alpha)} - c_2 \cdot c_3 \cdot \hat{V}_X(\hat{\omega}_{p-1,\alpha})$  #  $\hat{V}_X$  penduga kovarians X,  $\hat{\omega}$  penduga
                                         fungsi basis
       $\hat{Y}_i^{(p,\alpha)} \leftarrow \hat{Y}_i^{(p-1,\alpha)} - \hat{\eta}_{p-1,\alpha}(X_i)$  #  $\hat{\eta}$  penduga nilai harapan bersyarat
    end if
  end for
   $\hat{\lambda}_j^{(p,\alpha)}, \hat{\phi}_j^{(p,\alpha)} \leftarrow$  nilai eigen ke- $j$  dan  $\hat{V}_{\hat{X}^{(p,\alpha)}}$  fungsi eigen
   $a_j \leftarrow c\widehat{ov} \left\{ \hat{Y}^{(p,\alpha)}, \int_{\mathbb{T}} \hat{X}^{(p,\alpha)} \hat{\phi}_j^{(p,\alpha)} \right\}$ 
   $b_j \leftarrow v\widehat{ar} \left\{ \int_{\mathbb{T}} \hat{X}^{(p,\alpha)} \hat{\phi}_j^{(p,\alpha)} \right\}$ 
   $Q_{p,\alpha}(\delta) \leftarrow \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta}} \right\}^2 \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\left( \hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta} \right)^2} \right\}^{\frac{\alpha}{(1-\alpha)}}$ 
   $\hat{\delta}^{(p,\alpha)} \leftarrow \underset{\delta \in (-1,0) \cup (0,\infty)}{\operatorname{argmin}} - \ln Q_{p,\alpha}(\delta)$ 

```

$$\hat{\omega}_{p,\alpha} \leftarrow \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\left(\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta^{(p,\alpha)}} \right)^2} \right\}^{\frac{1}{2}} \sum_{j=1}^{\infty} \frac{a_j}{\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta^{(p,\alpha)}}} \hat{\phi}_j^{(p,\alpha)}$$

$$c_1 \leftarrow \widehat{\text{cov}}(Y, \int_{\mathbb{T}} X \hat{\omega}_{p,\alpha})$$

$$c_2 \leftarrow \widehat{\text{var}}^{-\frac{1}{2}} \left(\int_{\mathbb{T}} X \hat{\omega}_{p,\alpha} \right)$$

$$c_3 \leftarrow \int_{\mathbb{T}} \hat{X}_i^{(1,\alpha)} \hat{\omega}_{p,\alpha}$$

$$\hat{\beta}_{p,\alpha} \leftarrow \hat{\beta}_{p-1,\alpha} + c_1 c_2 \hat{\omega}_{p,\alpha}$$

for i from 1 to n **do**

$$\hat{\eta}_{p,\alpha}(X_i) \leftarrow \bar{Y} + \int_{\mathbb{T}} \hat{X}_i^{(1,\alpha)} \hat{\beta}_{p,\alpha}$$

end for

end for
optimal(p, α) \leftarrow Min RMSEP

Algoritma RKFK yang diusulkan dalam penelitian adalah sebagai berikut:

1. Regresikan x dan y menggunakan metode RKF
2. Hitung nilai sisaan (e_i) untuk masing-masing pengamatan berdasarkan model yang diperoleh pada langkah 1
3. Hitung nilai $\hat{\sigma}$ dengan $\hat{\sigma} = 1,4826 (\text{median}|e_i - \text{median}(e_i)|)$ (Rousseeuw dan Croux, 1993)
4. Hitung nilai pembobot untuk masing-masing pengamatan (w_i) berdasarkan nilai sisaan yang diperoleh pada Langkah 2, formulasi bobot menggunakan empat jenis fungsi pembobot yang tertera pada Tabel 2.
5. Transformasi nilai bobot $w^*(e_i) = \frac{w(e_i)+0,1}{\min(w(e_i)+0,1)}$
6. Bangkitkan data x dan y sebanyak masing-masing bobot yang diperoleh pada Langkah 5
7. Regresikan data bangkitan pada Langkah 6 menggunakan metode RKF.
8. Ulangi Langkah 2 hingga 7 sampai mendapatkan nilai konvergen. Nilai konvergen diperoleh pada saat $|e_n - e_{n-1}| \leq 0,00001$.

5.4 Evaluasi Kinerja Metodologi yang Diusulkan

Bagian ini membahas performa RKFK dalam pemodelan. Kinerja metodologi RKFK akan dievaluasi dengan metode simulasi. RKFK akan diinvestigasi berdasarkan ukuran kebaikan model yaitu korelasi nilai pengamatan (y_i) dengan nilai dugaan (\hat{y}_i) untuk setiap pengamatan sejumlah n , *root mean square error of prediction* (RMSEP), *mean absolute of error* (MAE) dan *mean absolute percentage Error* (MAPE) menggunakan software R.

$$\text{korelasi} = \frac{\sum (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{(n-1)s_y s_{\hat{y}}}$$

s_y dan $s_{\hat{y}}$ merupakan simpangan baku data pengamatan dan dugaan

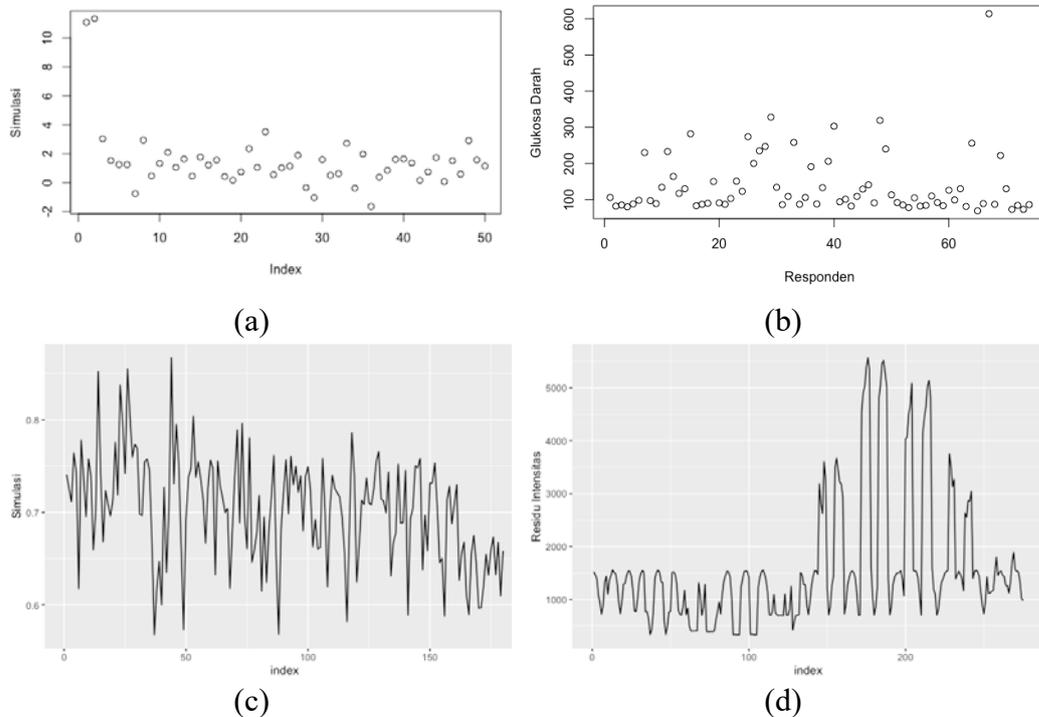
$$\text{RMSEP} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

$$\text{MAE} = \frac{\sum |y_i - \hat{y}_i|}{n}$$

$$\text{MAPE} = \frac{\sum \frac{|y_i - \hat{y}_i|}{y_i}}{n} 100\%$$

Kinerja RKFK dievaluasi dengan cara membandingkan ketiga ukuran kebaikan model tersebut yang diperoleh RKFK dengan metode fungsional lainnya seperti regresi kontinum fungsional (RKF), dan regresi fungsional (RF), serta metode non fungsional yang terdiri dari regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), dan regresi kuadrat terkecil (RKT).

Data simulasi diasumsikan terdapat 365 peubah bebas yang merupakan fungsi $\sin\left(\frac{i\pi}{45}\right) + \text{distribusi uniform}(365, -0.1, 0.1)$, $i = [0, 364]$, dan 1 peubah respon $y_i = \sum_{j=1}^{365} x_{i,j}$, $j = 1, 2, \dots, n$, masing-masing peubah memiliki 300 ulangan, dan data bangkitan yang disiapkan adalah $n = 50$, $n = 100$, dan $n = 200$ serta pemberian penciran yang bervariasi (2% - 10%). Pola data simulasi menyesuaikan pola data empirik yang dapat dilihat pada Gambar 16.



Gambar 16 Eksplorasi Data Simulasi dan Empirik, (a) Peubah X data simulasi, (b) Peubah X data empirik, (c) Peubah Y data Simulasi, (d) Peubah Y data Empirik

Proses analisis menggunakan RKFK membutuhkan waktu yang cukup lama, karena membutuhkan proses regresi sebanyak dua kali untuk menghasilkan nilai dugaan (\hat{y}). Nilai dugaan dapat digunakan untuk memperoleh nilai sisaan, yang akan dijadikan acuan untuk menetapkan besaran bobot untuk setiap amatan. Proses tersebut tentu membutuhkan waktu yang cukup lama. Selain itu iterasi yang dibutuhkan untuk menentukan nilai konvergen seperti pada langkah 8 dalam algoritma RKFK, juga mempengaruhi lama komputasi.

Berdasarkan permasalahan waktu komputasi, maka dalam penelitian ini dikaji aspek waktu komputasi yang dibutuhkan untuk masing-masing metode. Konsep analisis komputasi yaitu dengan menentukan waktu awal dan berakhirnya program komputasi. Selisih waktu awal dan akhir menjadi lama waktu komputasi yang dibutuhkan masing-masing metode. Analisis waktu komputasi untuk keenam metode dilakukan dengan pengulangan sebanyak 10 kali, dengan variasi pencilan 2%, 5%, dan 10%, serta variasi jumlah data (n) 50, 100, dan 200. Hasil analisis waktu komputasi keenam metode dapat dilihat pada Tabel 7.

Tabel 7 Waktu Komputasi

Metode	Waktu Komputasi (dalam menit)								
	n=50			n=100			n=200		
	2%	5%	10%	2%	5%	10%	2%	5%	10%
RKFK	2,0	2,1	1,6	6,3	6,8	7,2	8,8	9,3	9,6
RKF	1,5	1,5	1,5	3,8	3,7	3,7	23,8	23,4	22,6
RKU	1,7	1,5	1,5	4,2	3,8	2,3	24,8	23,9	23,1
RKTP	1,8	1,8	1,8	5,7	5,5	5,5	35,3	34,8	32,8
RKT	2,3	2,7	2,4	6,9	4,9	5,6	11,5	11,1	12,3
RF	1,7	2,0	1,8	6,6	5,3	5,9	37,5	35,6	32,2

Tampak waktu komputasi tercepat adalah metode RKF, dan yang paling lama adalah RKT. RKFK membutuhkan waktu yang cukup lama, namun selisih waktu yang dibutuhkan pada saat jumlah data (n) bertambah, tidak terlalu jauh seperti pada metode lainnya. Ketika $n = 200$ nampak selisih waktu yang dibutuhkan RKFK adalah kurang lebih 2 hingga 3 menit dari ketika $n = 100$. Kondisi tersebut berbeda dengan metode lainnya, pada saat $n = 200$ waktu yang terdeteksi sangat lama membutuhkan waktu lebih dari 11 hingga 37 menit, atau selisih kurang lebih 6 hingga 30 menit ketika $n = 100$. Dengan demikian dapat dikatakan bahwa RKFK masih optimal digunakan ketika jumlah data semakin banyak.

Hasil evaluasi terhadap keenam metode dengan ukuran kebaikan model, ditunjukkan pada Tabel 9. Tampak pada Tabel 9 nilai rata-rata RMSEP dan MAE yang diperoleh dari masing-masing metode tampak semakin membesar seiring dengan bertambahnya bobot pencilan yang diberikan, sedangkan nilai korelasi semakin menurun, sehingga dapat dikatakan semua metode terpengaruh oleh keberadaan pencilan. Berdasarkan hasil analisis terlihat performa RKFK lebih unggul dibandingkan dengan kelima metode lainnya, nilai rata-rata RMSEP yang diperoleh lebih kecil dari kelima metode lainnya pada kondisi $n = 50, 100, \text{ dan } 200$ serta bobot pencilan dari 2% hingga 10%. Berdasarkan nilai RMSEP dan MAE yang diperoleh dengan metode fungsional RKFK, RKF, dan RF terlihat lebih baik dibanding dengan metode non fungsional yaitu RKU, RKTP dan RKT.

Berdasarkan nilai rata-rata korelasi antara data pengamatan dan nilai dugaan, metode RKFK memperoleh nilai korelasi jauh lebih tinggi dibandingkan kelima metode lainnya. Semakin besar jumlah data terlihat nilai korelasi semakin membesar, sebaliknya semakin besar persentase pencilan dalam data maka nilai korelasi semakin menurun.

Hasil analisis dapat diamati pada diagram pencar ukuran kebaikan model pada Gambar 17 berdasarkan nilai rata-rata RMSEP, korelasi dan MAE untuk keenam

metode. Terlihat jelas grafik nilai rata-rata RMSEP dengan pendekatan RKFK (berwarna biru) berada paling bawah diantara kelima metode lainnya. Hal tersebut serupa dengan nilai rata-rata MAE yang diperoleh RKFK, terlihat grafik RKFK konsisten berada paling bawah meskipun jumlah data semakin besar dan nilai persentase pencilan semakin besar. Sedangkan pada grafik nilai rata-rata korelasi RKFK berada paling atas secara konsisten meskipun jumlah data semakin besar dan nilai persentase pencilan semakin besar.

Kajian simulasi menghasilkan dua faktor yaitu pencilan dan jumlah amatan. Kedua faktor tersebut perlu dikaji untuk mengetahui pengaruh dari masing-masing faktor, dan interaksi kedua faktor terhadap ukuran kebaikan model. Hasil analisis ragam dua arah dengan interaksi dapat dilihat pada Tabel 8.

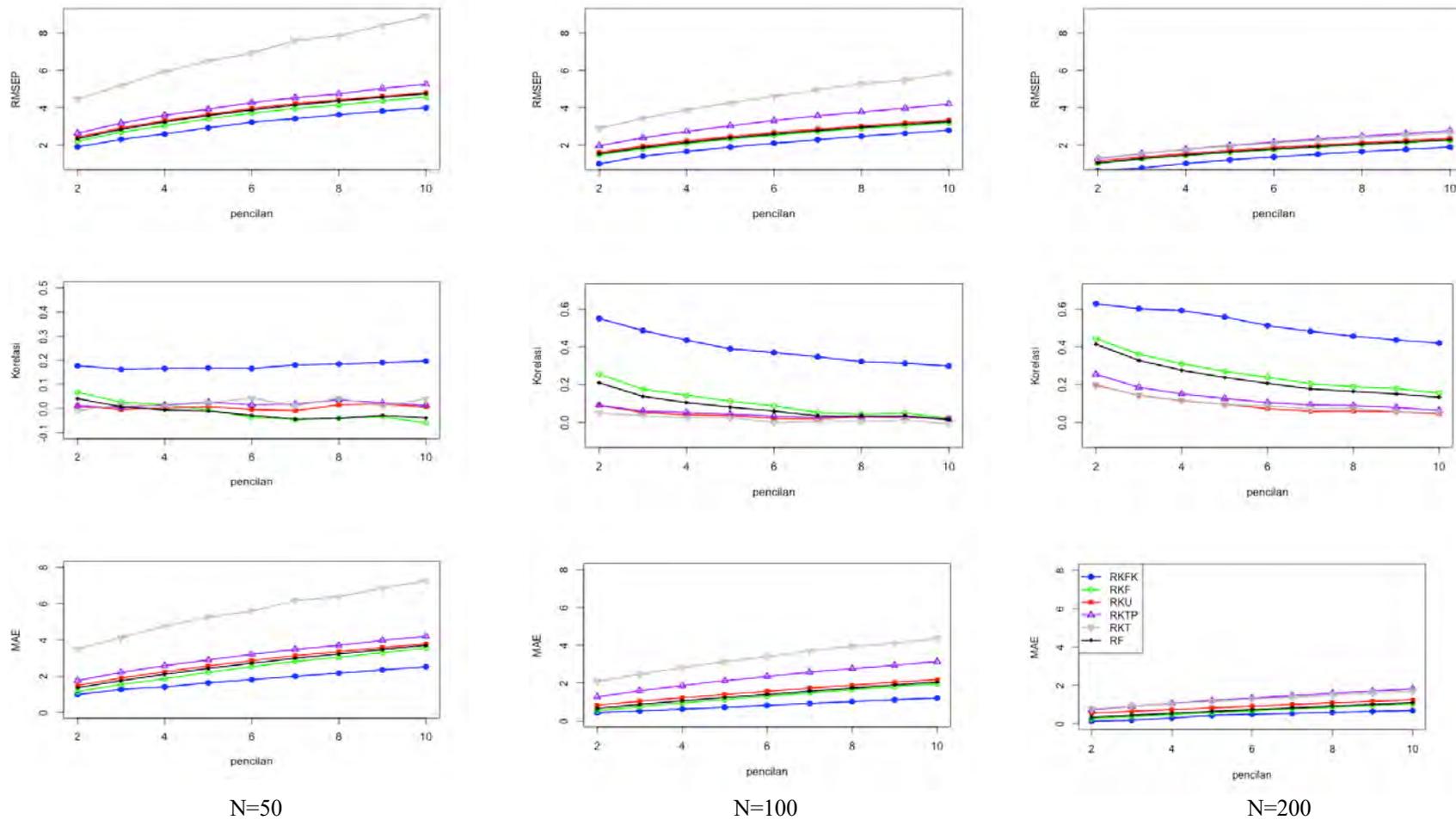
Tabel 8 Analisis Ragam Ukuran Kebaikan Model

Sumber Keragaman	Derajat Bebas (db)	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)	F Hitung	Pr(>F)
Pencilan	8	66,17	8,27	8,907	9,44E-10
Jumlah Amatan	2	157,62	78,81	84,866	< 2e-16
Pencilan x Jumlah Amatan	16	5,39	0,34	0,363	0,9888
Galat	135	125,37	0,93		

Berdasarkan nilai p yang diperoleh pada kolom Pr(>F) di Tabel 8, faktor pencilan memberikan pengaruh signifikan terhadap ukuran kebaikan model dengan mengendalikan faktor jumlah amatan (n) pada tingkat signifikansi 5%. Faktor jumlah amatan juga memberikan pengaruh signifikan terhadap ukuran kebaikan model, dengan mengendalikan faktor pencilan pada tingkat signifikansi 5%. Interaksi antara pencilan dan jumlah amatan tidak memberikan pengaruh terhadap ukuran kebaikan model, sehingga tidak perlu dilakukan uji lanjutan.

Tabel 9 Ukuran Keباikan Model: Data Simulasi

Obs	Pencilan	RMSEP						Korelasi						MAE						MAPE					
		RKFK	RKF	RKU	RKTP	RKT	RF	RKFK	RKF	RKU	RKTP	RKT	RF	RKFK	RKF	RKU	RKTP	RKT	RF	RKFK	RKF	RKU	RKTP	RKT	RF
N=50	2%	1,894	2,201	2,415	2,632	4,463	2,331	0,178	0,068	0,01	0,012	-0,008	0,042	0,9917	1,153	1,486	1,761	3,49	1,359	19,427	23,343	23,125	23,357	26,559	23,36
	3%	2,295	2,684	2,914	3,166	5,191	2,832	0,163	0,027	-0,003	0,005	0,016	0,008	1,2722	1,538	1,899	2,209	4,122	1,754	19,689	24,4	24,422	24,392	28,943	24,39
	4%	2,592	3,041	3,302	3,587	5,938	3,229	0,166	0,017	0,003	0,015	0,01	-0,005	1,4027	1,859	2,237	2,575	4,752	2,105	20,105	25,604	26,009	25,611	30,391	25,61
	5%	2,918	3,4	3,639	3,92	6,483	3,568	0,168	-0,005	0,009	0,027	0,025	-0,009	1,6323	2,216	2,563	2,897	5,249	2,427	20,374	26,527	27,156	26,537	32,03	26,54
	6%	3,218	3,704	3,953	4,263	6,918	3,872	0,166	-0,035	-0,003	0,016	0,044	-0,027	1,8143	2,521	2,852	3,205	5,584	2,715	20,543	27,282	28,197	27,306	29,739	27,31
	7%	3,409	3,959	4,22	4,528	7,597	4,133	0,181	-0,045	-0,008	0,019	0,011	-0,043	1,9902	2,82	3,132	3,475	6,166	2,987	19,997	27,083	27,81	27,097	32,719	27,1
	8%	3,618	4,14	4,409	4,738	7,85	4,352	0,185	-0,038	0,016	0,037	0,044	-0,04	2,1696	3,035	3,352	3,693	6,373	3,228	20,474	28,068	28,899	28,063	34,576	28,06
	9%	3,816	4,363	4,604	5,031	8,407	4,544	0,191	-0,033	0,019	0,024	0,011	-0,027	2,3471	3,304	3,569	3,976	6,856	3,466	20,904	28,828	29,637	28,796	34,003	28,8
	10%	3,991	4,558	4,805	5,255	8,89	4,745	0,197	-0,058	0,007	0,014	0,041	-0,038	2,5134	3,544	3,783	4,194	7,243	3,691	21,566	29,539	30,568	29,534	34,57	29,53
	N=100	2%	0,988	1,471	1,606	1,96	2,891	1,522	0,550	0,256	0,09	0,091	0,049	0,21	0,4292	0,546	0,826	1,273	2,082	0,661	13,045	15,773	15,972	15,773	18,126
3%		1,398	1,802	1,941	2,38	3,422	1,857	0,486	0,176	0,053	0,06	0,037	0,138	0,5275	0,752	1,037	1,593	2,465	0,869	13,261	16,638	17,079	16,638	18,628	16,64
4%		1,653	2,065	2,214	2,718	3,869	2,129	0,435	0,144	0,041	0,053	0,024	0,105	0,6253	0,943	1,224	1,863	2,821	1,061	13,555	18,271	19,006	18,271	22,834	18,27
5%		1,888	2,306	2,448	3,029	4,258	2,369	0,390	0,113	0,036	0,044	0,025	0,081	0,7194	1,133	1,399	2,128	3,134	1,239	13,541	18,736	19,546	18,736	21,896	18,74
6%		2,089	2,516	2,667	3,307	4,598	2,581	0,370	0,087	0,021	0,034	0,001	0,06	0,8205	1,309	1,566	2,347	3,394	1,407	13,722	19,583	21,409	19,583	24,021	19,58
7%		2,279	2,713	2,852	3,555	4,961	2,773	0,347	0,053	0,019	0,026	0,004	0,035	0,9181	1,485	1,729	2,577	3,707	1,574	13,81	20,211	21,044	20,211	24,545	20,21
8%		2,466	2,889	3,015	3,765	5,256	2,944	0,322	0,044	0,029	0,031	0,005	0,034	1,0204	1,654	1,881	2,761	3,926	1,736	13,812	21,585	22,934	21,585	26,427	21,59
9%		2,624	3,029	3,181	3,972	5,484	3,097	0,313	0,05	0,027	0,03	0,014	0,035	1,1138	1,805	2,038	2,94	4,109	1,898	14,179	22,288	24,088	22,288	27,728	22,29
10%		2,784	3,182	3,333	4,193	5,848	3,252	0,298	0,023	0,019	0,017	-0,008	0,015	1,2119	1,958	2,187	3,135	4,357	2,047	14,273	23,935	25,776	23,935	29,127	23,93
N=200		2%	0,623	1,016	1,136	1,258	1,296	1,037	0,628	0,444	0,198	0,255	0,195	0,414	0,1153	0,256	0,537	0,704	0,749	0,326	11,434	13,619	13,625	13,619	16,15
	3%	0,762	1,243	1,358	1,537	1,548	1,268	0,602	0,361	0,143	0,186	0,143	0,328	0,1722	0,36	0,632	0,892	0,902	0,432	11,54	14,22	14,218	14,22	17,406	14,22
	4%	1,007	1,431	1,544	1,766	1,753	1,459	0,592	0,311	0,115	0,151	0,118	0,276	0,2836	0,465	0,721	1,051	1,033	0,533	11,617	14,962	14,878	14,88	18,488	14,88
	5%	1,201	1,601	1,71	1,97	1,94	1,628	0,558	0,27	0,096	0,127	0,095	0,238	0,4289	0,563	0,81	1,199	1,155	0,627	11,682	15,042	15,403	15,43	18,693	15,38
	6%	1,359	1,75	1,862	2,156	2,099	1,778	0,512	0,239	0,073	0,105	0,087	0,207	0,4788	0,655	0,895	1,333	1,267	0,718	11,755	15,255	15,919	15,563	19,57	15,79
	7%	1,505	1,889	1,998	2,32	2,247	1,915	0,482	0,205	0,059	0,093	0,072	0,178	0,5280	0,754	0,984	1,463	1,373	0,808	11,899	15,543	16,579	16,48	20,695	16,43
	8%	1,641	2,015	2,121	2,469	2,384	2,042	0,456	0,19	0,061	0,089	0,075	0,164	0,5771	0,846	1,07	1,584	1,483	0,9	11,917	15,872	17,048	17,203	21,201	16,87
	9%	1,767	2,128	2,237	2,615	2,537	2,158	0,436	0,179	0,057	0,079	0,058	0,151	0,6263	0,936	1,153	1,699	1,598	0,992	12,057	16,141	17,683	17,828	21,874	17,33
	10%	1,884	2,244	2,349	2,759	2,645	2,27	0,420	0,156	0,046	0,064	0,052	0,133	0,6754	1,027	1,233	1,81	1,672	1,076	12,159	16,376	18,24	19,036	22,274	17,84



Gambar 17 Diagram pencar ukuran kebaikan model

5.5 Ilustrasi Data Empirik

Data yang digunakan dalam penelitian ini adalah data glukosa darah yang diukur menggunakan teknik invasif oleh Laboratorium PRODIA IPB Indonesia (Y) dan teknik non invasif menggunakan Purwarupa Alat Pendeteksian Gula Darah berbasis spektroskopi yang dikembangkan oleh Tim Fisika dan Statistika IPB tahun 2019 (X). Sampel data yang diperoleh sebanyak 74 orang yang merupakan masyarakat umum yang tinggal di lingkungan IPB Bogor Indonesia, yang berusia 21-87 Tahun.

Validasi model dilakukan menggunakan metode *Jacknife*. Metode *Jacknife* melakukan *resampling* dengan cara satu observasi pada suatu dataset dihilangkan pada setiap iterasi, dan penduga dihitung ulang untuk setiap subset dari dataset yang dihasilkan. Nilai α optimum diperoleh dari hasil analisis berdasarkan nilai RMSEP terkecil seperti yang ditunjukkan pada Tabel 10 dengan $\alpha \in [0,1)$, sedangkan nilai optimum untuk parameter p diperoleh 3 (Lampiran 8). Berdasarkan ukuran kebaikan pada Tabel 10, terlihat bahwa nilai optimum α dengan $p = 3$ adalah ketika 0,1. Nilai RMSEP terkecil diperoleh yaitu 33,25, nilai korelasi yang diperoleh adalah yang terbesar yaitu 0,9281, dan juga nilai MAE yang terkecil diperoleh 16,36, sedangkan ukuran rata-rata relatif sisaan (MAPE) yang terkecil yaitu 15,01.

Tabel 10 Ukuran Kebaikan dengan Variasi α

Alfa	RMSEP	Korelasi	MAE	MAPE (%)
0,1	33,25	0,9281	16,36	15,01
0,2	80,08	0,4304	47,25	30,9
0,3	80,17	0,4299	47,16	30,38
0,4	80,17	0,4299	47,16	30,38
0,5	80,17	0,4299	47,16	30,38
0,6	80,87	0,4138	47,63	30,44
0,7	82,31	0,3881	48,68	30,44
0,8	83,45	0,3539	50,11	31,45
0,9	85,41	0,2883	51,68	32,24
1	80,62	0,416	47,78	30,83

Evaluasi terhadap kinerja RKFK dilakukan perbandingan ukuran kebaikan model dengan pendekatan fungsional lainnya yaitu regresi kontinum fungsional (RKF), regresi fungsional (RF), serta pendekatan non fungsional yaitu regresi komponen utama (RKU), regresi kuadrat terkecil parsial (RKTP), dan regresi kuadrat terkecil (RKT). Hasil analisis dengan keenam pendekatan dapat dilihat pada Tabel 11. Hasil analisis RKU diperoleh jumlah komponen yang optimum adalah 8 (Lampiran 5), sedangkan hasil analisis RKTP diperoleh jumlah komponen yang optimum adalah 2 (Lampiran 6). Pendekatan RKT yang dilakukan menggunakan metode *stepwise forward selection* untuk mengatasi permasalahan multikolinearitas dan jumlah peubah yang lebih besar dari jumlah pengamatan.

Model yang memiliki kinerja terbaik adalah model dengan nilai RMSEP paling kecil dan nilai koefisien korelasi yang paling besar. Tampak pada Tabel 11 diketahui RKFK memiliki kinerja yang lebih baik dibandingkan dengan keempat

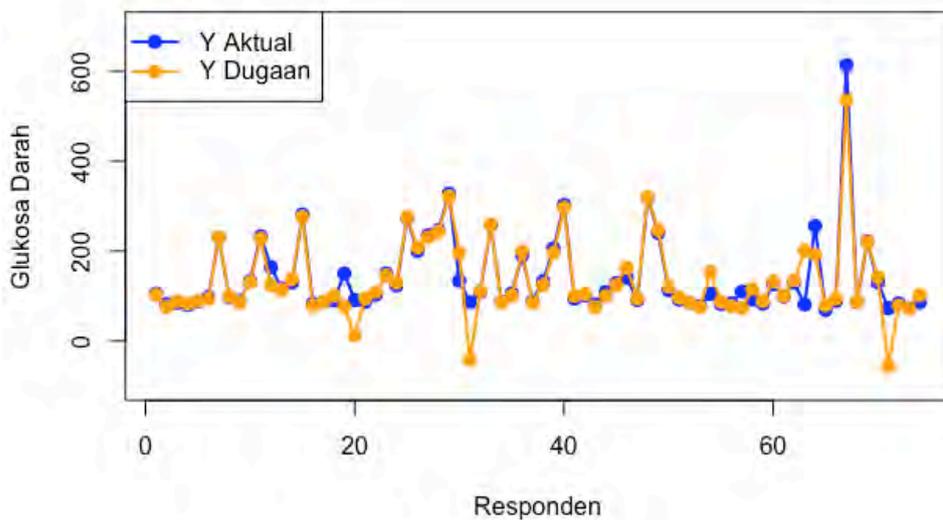
metode lainnya, karena memperoleh nilai RMSEP yang paling kecil yaitu 33,25, serta nilai korelasi yang cukup baik yaitu 0,9281, sedangkan hasil MAE yang diperoleh serta nilai MAPE juga terlihat paling kecil, masing-masing diperoleh nilai 16,36 dan 15,01.

Hasil analisis dengan metode RKT dan RF diperoleh nilai RMSEP dan MAE, serta ukuran rata-rata relatif sisaan (MAPE) tampak jauh lebih besar dibandingkan dengan metode yang lainnya. Hal ini dapat dikatakan bahwa RKT dan RF belum optimal digunakan untuk memodelkan data dengan kasus multikolinearitas, dengan jumlah peubah lebih besar dari jumlah pengamatan.

Tabel 11 Ukuran Kebaikan Model

Metode	Ukuran			
	RMSEP	Korelasi	MAE	MAPE (%)
RKFK	33,25	0,9281	16,36	15,01
RKF	86,57	0,1718	58,11	41,48
RKU	86,86	0,17400	58,06	41,47
RKTP	89,05	0,13979	61,35	44,99
RKT	101,21	0,05947	95,49	55,85
RF	121,9	0,3346	95,49	83,10

Pola dan korelasi data aktual dengan data dugaan menggunakan pendekatan RKFK ditunjukkan pada Gambar 18. Tampak pola data dugaan (grafik berwarna merah) menyerupai data aktual (grafik berwarna biru), sehingga dapat dikatakan korelasi antara kedua data tersebut cukup tinggi, sesuai dengan hasil analisis ukuran korelasi yang diperoleh pada Tabel 10.



Gambar 18 Diagram pencar antara data aktual dengan data dugaan model RKFK

5.6 Kesimpulan

Dari hasil analisis dapat dikatakan bahwa metode reduksi seperti regresi kontinum fungsional, regresi komponen utama, dan regresi kuadrat terkecil parsial merupakan metode yang unggul apabila digunakan ketika terjadi multikolinearitas atau jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan.

Dalam kasus analisis data fungsional penerapan regresi fungsional lebih baik karena tidak menghilangkan pola data, dengan demikian dapat dikatakan bahwa regresi kontinum fungsional merupakan pendekatan yang efektif dalam menganalisis model kalibrasi yang umumnya memiliki data fungsional, dan terdapat beberapa permasalahan yang diantaranya terjadi multikolinearitas atau jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan. Selain dua permasalahan tersebut data hasil pengukuran pada umumnya juga memiliki pencilan yang dapat mempengaruhi hasil pendugaan parameter. Sangat disayangkan apabila diambil keputusan dengan menghilangkan data pencilan dari data hasil pengukuran seperti pada data glukosa darah yang digunakan dalam penelitian ini, karena waktu yang diperlukan untuk menghasilkan data cukup lama serta memerlukan biaya yang cukup mahal.

Berdasarkan hasil analisis menggunakan data simulasi dan empirik terlihat performa RKFK lebih unggul. Regresi kontinum fungsional kekar sebagai metode kebaruan yang dikembangkan dari regresi kontinum fungsional, menjadi metode yang efektif digunakan untuk menggambarkan hubungan antara data spektrum cahaya non invasif dengan kadar glukosa darah invansif.

VI PERFORMA ANALISIS FUNGSI PEMBOBOT PADA REGRESI KONTINUM FUNGSIONAL KEKAR

6.1 Pendahuluan

Data fungsional sering kali tidak sesuai untuk dianalisis dengan metode konvensional karena sifatnya yang kompleks dan dinamis (Ramsay dan Hooker, 2017). Data fungsional umumnya mencakup peubah berbentuk fungsi yang berevolusi dari waktu ke waktu atau dalam konteks domain tertentu yang berkaitan erat dengan kontinuitas, seperti kurva, spektrum, atau gambar. Fungsi-fungsi tersebut sering kali didefinisikan dalam bentuk waktu, tetapi juga dapat berupa lokasi spasial, panjang gelombang, probabilitas, dan banyak lagi (Horvath dan Kokoszka, 2012). Secara intrinsik, data fungsional memiliki dimensi tak terbatas. Dimensi yang tinggi secara intrinsik dari data ini membawa tantangan pada teori dan komputasi, dan tantangan ini bervariasi sesuai dengan bagaimana data fungsional diambil sampelnya (Wang *et. al.*, 2016). Namun, struktur data berdimensi tinggi atau tak terbatas merupakan sumber informasi yang kaya dan ada banyak tantangan menarik untuk penelitian dan analisis data.

Penanganan data fungsional membutuhkan pendekatan yang lebih canggih agar diperoleh model yang dapat menghasilkan prediksi yang lebih baik. Beberapa metode yang umum digunakan untuk analisis data fungsional antara lain (Kokoszka dan Reimherr, 2017): Analisis Fungsional yang melibatkan penggunaan teknik matematika dan statistik yang dirancang khusus untuk menangani data fungsional, contohnya adalah analisis harmonik, analisis *Wavelet*, regresi komponen utama fungsional (RKUF), atau pendekatan regresi kuadrat terkecil parsial fungsional (RKTPF). Fungsi Splines dan Basis merupakan pendekatan yang melibatkan pendekatan data fungsional dengan fungsi basis seperti splines atau *Wavelet*, data fungsional dipecah menjadi komponen-komponen yang lebih sederhana untuk analisis lebih lanjut. Model tak linear dan *Machine Learning* untuk data yang sangat kompleks. Model tak linear dan teknik *machine learning* dapat digunakan untuk mengekstrak pola yang tersembunyi dalam data fungsional, misalnya, jaringan syaraf tiruan, mesin vektor pendukung, atau pohon keputusan yang diterapkan pada data fungsional (Somvanshi *et. al.*, 2016). Analisis Spasial dan Geostatistik digunakan untuk data fungsional yang terkait dengan dimensi spasial, seperti data citra satelit atau medis, dengan tujuan untuk memahami pola spasial dan temporal pada data tersebut (Caprarelli dan Fletcher, 2014).

Metode analisis harus dipilih sesuai dengan struktur dan sifat data fungsional tertentu. Kombinasi dari berbagai teknik analisis dan pendekatan pemodelan sering kali diperlukan untuk mendapatkan pemahaman yang komprehensif tentang data fungsional. Regresi kontinum fungsional kekar (RKFK) merupakan pengembangan dari regresi kontinum yang dicetuskan oleh Stone dan Brooks (1990). RKFK dapat diterapkan pada data fungsional dan bersifat kekar, yaitu tahan terhadap pencilan.

Pencilan adalah pengamatan yang secara signifikan berbeda dari pola umum pada data, dan dapat mempengaruhi hasil analisis (Modi dan Oza, 2016). Pencilan tidak selalu menjadi gangguan, terkadang pencilan dapat memberikan informasi penting atau wawasan yang berharga, karena adanya variabilitas alamiah dari data (de Cheveigne dan Dorothee, 2018). Salah satu alternatif yang dapat digunakan untuk mengatasi pencilan adalah dengan memberikan nilai bobot yang lebih rendah

pada observasi yang dianggap sebagai pencilan, sehingga mengurangi pengaruhnya terhadap hasil analisis. Fungsi pembobotan membantu mengurangi bias yang mungkin terjadi dengan mengurangi pengaruh pencilan secara proporsional, daripada menghilangkannya sama sekali (Yazici, 2016). Ketangguhan RKFK tergantung pada fungsi pembobotan yang dipilih. Fungsi pembobot yang umum digunakan adalah Huber, Hampel, Ramsay, dan Tukey (*Bisquare*).

Berdasarkan fungsi pembobot tersebut, penelitian ini akan menguji kinerja RKFK terhadap keberadaan pencilan dengan empat fungsi pembobot yaitu Huber, Hampel, Ramsay, dan Tukey (*Bisquare*). Bobot yang diperoleh berdasarkan kriteria pada masing-masing metode pada Tabel 2 akan ditransformasi dengan cara menambahkan nilai bobot dengan konstanta, dan membagi setiap bobot dengan nilai minimum dari sekumpulan bobot, dengan adanya transformasi maka akan terhindar dari terjadinya reduksi data (bobot nol) seperti yang berlaku pada fungsi pembobot Hampel dan Tukey (*Bisquare*).

6.2 Metode

Pada bagian ini akan dibahas mengenai kinerja RKFK dalam melakukan pemodelan dengan data simulasi dan akan diaplikasikan pada data empiris untuk mengilustrasikan hasil teoritisnya. Kinerja RKFK dengan kedua data tersebut akan diselidiki berdasarkan nilai goodness of fit, yaitu korelasi antara nilai observasi (y_i) dengan nilai prediksi (\hat{y}_i) untuk setiap n observasi, *root means square error of prediction* (RMSEP), *mean absolute error* (MAE), dan *mean absolute percentage error* (MAPE) dengan menggunakan perangkat lunak R.

Empat jenis fungsi pembobot Huber, Hampel, Ramsay dan Tukey (*Bisquare*) digunakan untuk tujuan analisis. Masing-masing kriteria fungsi pembobot ditunjukkan pada Tabel 2. Tahapan dalam analisis data mengacu pada algoritma RKFK yang dibahas pada Bab 5.

6.3 Hasil dan Pembahasan

6.3.1 Kajian Simulasi

Data simulasi dibangkitkan sebanyak 365 peubah bebas yang berasal dari fungsi $\sin(i\pi/45) +$ distribusi uniform $(365, -0.1, 0.1)$, $i = [0, 364]$, dan 1 peubah respon $y_i = \sum_{j=1}^{365} x_{i,j}$, $j = 1, 2, \dots, n$, masing-masing diulang sebanyak 50, dan data disiapkan untuk $n=50$, $n=100$, dan $n=200$, serta persentase pencilan dengan variasi 3%, 6%, dan 10%.

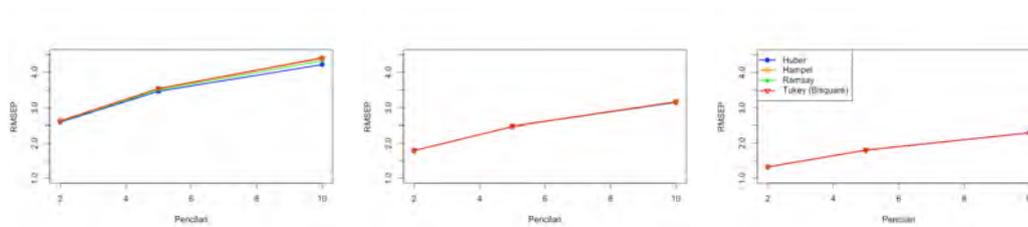
Berdasarkan hasil analisis data simulasi pada Tabel 12, RKFK menunjukkan performa yang baik ketika jumlah data meningkat, dan menurun ketika persentase pencilan meningkat. Hal ini secara konsisten dialami oleh semua fungsi pembobot.

RKFK dengan fungsi pembobot Huber memperoleh nilai RMSEP terkecil dibandingkan dengan metode lainnya pada saat $n=50$, namun menjadi terbesar pada saat n bertambah ($n=100$, dan $n=200$). Di sisi lain, nilai MAE RKFK dengan fungsi pembobot Huber paling besar ketika n kecil ($n=50$), dan menjadi paling kecil ketika $n=100$, dan $n=200$. Sementara itu, nilai korelasi yang diperoleh RKFK dengan fungsi pembobot Huber secara konsisten lebih besar daripada metode lainnya untuk semua n .

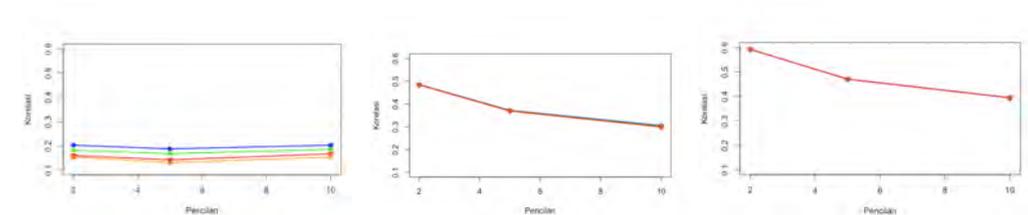
Tabel 12 Ukuran Keباikan Model RKFK dengan Variasi Fungsi Pembobot

Ukuran Keباikan Model	Fungsi Pembobot	n=50			n=100			n=200		
		3%	6%	10%	3%	6%	10%	3%	6%	10%
RMSEP	Huber	2,5791	3,4568	4,2168	1,7912	2,4650	3,1401	1,3224	1,8006	2,2881
	Hampel	2,6185	3,5394	4,4205	1,7877	2,4667	3,1577	1,3155	1,7943	2,2842
	Ramsay	2,5991	3,4978	4,3251	1,7893	2,4658	3,1496	1,3185	1,7973	2,2861
	Tukey (<i>Bisquare</i>)	2,6160	3,5320	4,3993	1,7897	2,4694	3,1603	1,3172	1,7967	2,2869
Korelasi	Huber	0,2034	0,1883	0,2034	0,4859	0,3727	0,3056	0,5926	0,4709	0,3954
	Hampel	0,1535	0,1313	0,1538	0,4838	0,3692	0,2992	0,5926	0,4705	0,3944
	Ramsay	0,1822	0,1685	0,1859	0,4849	0,3711	0,3025	0,5927	0,4708	0,3950
	Tukey (<i>Bisquare</i>)	0,1607	0,1427	0,1675	0,4838	0,3693	0,2994	0,5925	0,4705	0,3944
MAE	Huber	1,3698	1,9227	2,6394	0,6366	0,9371	1,3429	0,4841	0,6348	0,8363
	Hampel	1,3666	1,9097	2,6035	0,6463	0,9406	1,3340	0,4969	0,6452	0,8431
	Ramsay	1,3612	1,9055	2,5998	0,6406	0,9381	1,3367	0,4897	0,6392	0,8392
	Tukey (<i>Bisquare</i>)	1,3622	1,9053	2,5989	0,6424	0,9369	1,3310	0,4929	0,6413	0,8395

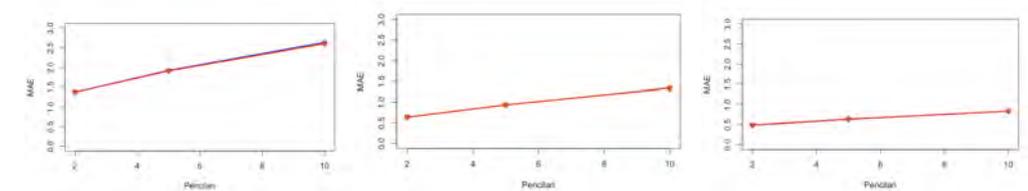
Secara keseluruhan nilai ukuran kebaikan model RKFk yang diperoleh untuk semua fungsi pembobot tidak jauh berbeda seperti yang ditunjukkan pada Gambar 19. Grafik cenderung tumpang tindih, sehingga pada penelitian ini kinerja RKFk dalam menangani pencilan dapat diaplikasikan dengan berbagai macam fungsi pembobot. Hal yang baru dalam penelitian ini adalah tidak menghilangkan atau memberikan bobot nol pada data observasi yang teridentifikasi sebagai pencilan.



(a) RMSEP



(b) Korelasi



(c) MAE

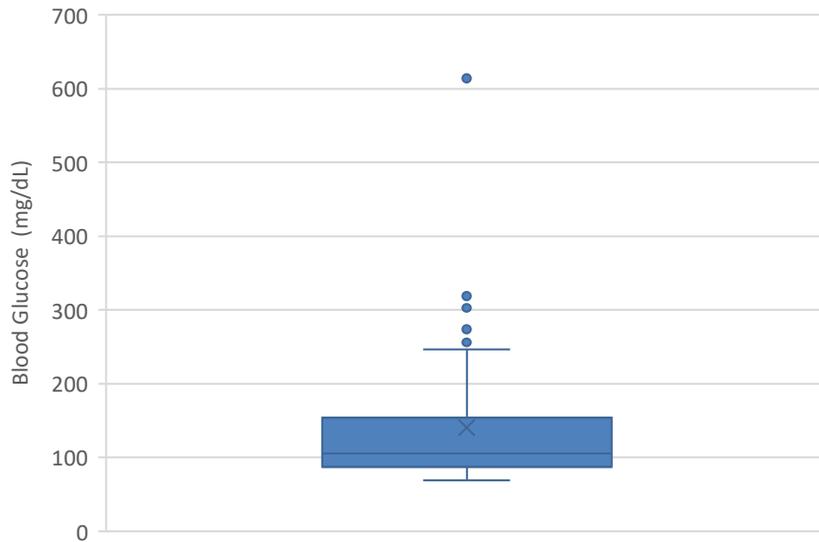
Gambar 19 Ukuran Kebaikan Model: Data simulasi, (a) RMSEP, (b) korelasi, dan (c) MAE

6.3.2 Kajian Empirik

Data yang digunakan dalam penelitian ini adalah data glukosa darah yang diukur dengan teknik invasif oleh Laboratorium PRODIA IPB Indonesia (Y) dan teknik invasif dengan menggunakan Prototipe Alat Deteksi Glukosa Darah Non-Invasif Berbasis Spektroskopi yang dikembangkan oleh Tim Fisika dan Statistika IPB pada tahun 2019 (X). Sampel data yang diperoleh sebanyak 74 orang yang merupakan masyarakat umum yang tinggal di lingkungan IPB Bogor Indonesia, berusia 21-87 tahun.

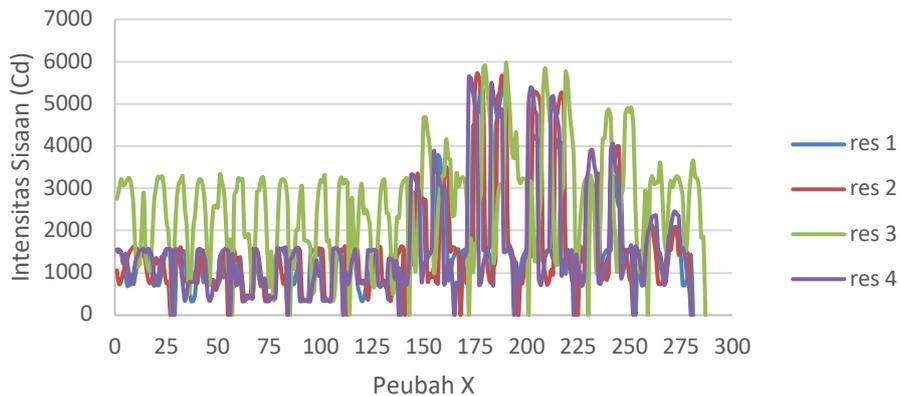
Data kadar glukosa darah pada penelitian ini menunjukkan nilai terendah 69 mg/dL dan tertinggi 614 mg/dL, dengan rata-rata 140,7 mg/dL. American Diabetes Association (ADA) membagi kadar glukosa darah menjadi tiga kategori, yaitu normal (< 100 mg/dL), prediabetes (100 mg/dL - 125 mg/dL) dan diabetes (> 125 mg/dL) (American Diabetes Association 2014), sehingga

berdasarkan kategori tersebut, didapatkan data sebanyak 35 responden termasuk dalam kategori normal, 11 responden termasuk dalam kategori prediabetes dan terdapat 28 responden yang termasuk dalam kategori diabetes. Visualisasi data ditunjukkan pada Gambar 20, kadar gula darah responden cenderung mendekati nilai 100 mg/dL, selain berada di atas nilai 100 mg/dL. Terdapat delapan kadar gula darah responden yang pencilan, antara lain 256 mg/dL, 258 mg/dL, 274 mg/dL, 282, mg/dL, 303 mg/dL, 319 mg/dL, 328 mg/dL, dan 614 mg/dL.

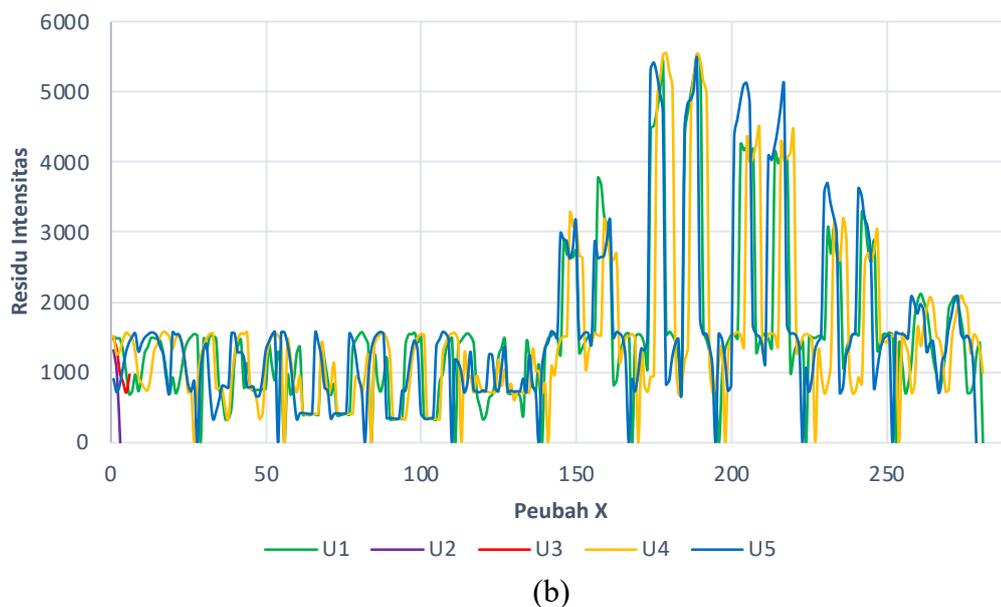


Gambar 20 Visualisasi Level data Glukosa Darah

Visualisasi data nilai residual intensitas dapat dilihat pada Gambar 21. Sampel yang ditampilkan adalah responden satu sampai dengan empat, pada ulangan kedua.



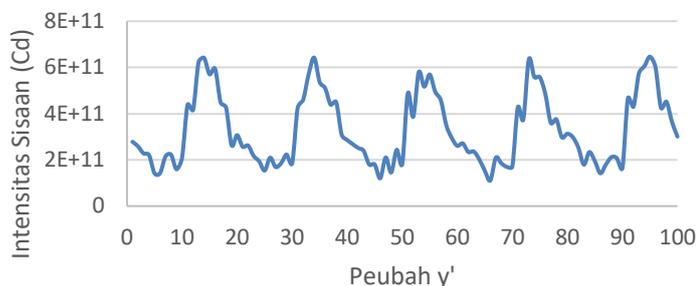
(a)



(b)
Gambar 21 Visualisasi Data Residu Intensitas (a) Antar Responden 1, 2, 3, dan 4, (b) Antar Ulangan 1 sampai 5 (U1, U2, U3, U4, U5) pada Responden 1

Terlihat jelas pada Gambar 21 bahwa pola yang terbentuk hampir mirip antar responden, terlihat bahwa pola data saling tumpang tindih. Jumlah data yang dihasilkan bervariasi antar responden dalam satu ulangan, sehingga perlu dilakukan peringkasan (reduksi) data untuk mendapatkan jumlah yang sama antar responden dan replikasi agar analisis data dapat dilakukan. Teknik peringkasan dilakukan dengan menentukan nilai luas kurva dengan menggunakan metode numerik aturan trapesium.

Hasil dari peringkasan didapatkan jumlah peubah untuk setiap responden untuk satu ulangan adalah 20 peubah sehingga didapatkan 100 peubah untuk 5 replikasi. Visualisasi data hasil peringkasan pada responden pertama ditunjukkan pada Gambar 22. Hasil peringkasan memiliki pola yang konsisten pada setiap replikasi, data yang membentuk puncak adalah data replikasi ke-1 hingga ke-5, hasil ini terlihat berbeda dengan Gambar 21 sebelum dilakukan peringkasan



Gambar 22 Visualisasi Data Hasil Ringkasan

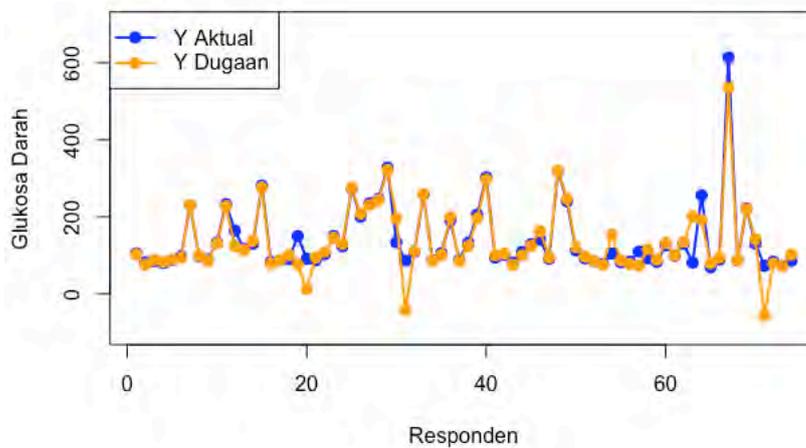
Hasil analisis data menggunakan pendekatan RKFK dengan fungsi pembobot Huber diperoleh ukuran kebaikan model pada Tabel 13. Nilai RMSEP yang diperoleh sebesar 33,25, korelasi sebesar 0,92, nilai MAE sebesar 19,60, dan nilai MAPE sebesar 15,01%. Analisis data menggunakan pendekatan RKFK

dengan fungsi pembobot Huber, telah dievaluasi pada kajian sebelumnya BAB 5 pada Tabel 11.

Tabel 13 Ukuran Kebaikan Model RKFK: Data Empirik

RMSEP	Korelasi	MAE	MAPE (%)
33,25	0,92	19,60	15,01

Kebaikan model dapat dievaluasi melalui pola data yang terbentuk. Data estimasi yang diperoleh dengan metode RKFK memiliki pola yang sama dengan data aktualnya. Hal ini dilihat dari tingkat korelasi yang diperoleh, dapat dilihat pada Gambar 23.



Gambar 23 Visualisasi Data Aktual dan Data Dugaan

RKFK adalah metode perbaikan dari RKF yang dikembangkan oleh Zhou (2019) yang tahan terhadap pencilan. Kekekaran RKFK terhadap pencilan bergantung pada fungsi pembobotan yang digunakan. Penelitian empat fungsi pembobot Huber, Hampel, Ramsay, dan Tukey (*Bisquare*) yang diterapkan pada RKF, sehingga RKF menjadi tahan terhadap pencilan dan disebut sebagai RKFK.

Hasil yang diperoleh fungsi pembobot Huber menunjukkan kinerja RKFK yang lebih baik dibandingkan dengan yang lain berdasarkan ukuran kebaikan model yang diperoleh, yaitu RMSEP, korelasi antara data aktual dengan model, dan MAE. Hasil penelitian ini sejalan dengan penelitian Nawaz *et. al.* (2022), namun kebaruan dari penelitian ini tidak menghilangkan (mereduksi) data observasi yang teridentifikasi sebagai pencilan, karena pada umumnya data pengukuran relatif sulit diperoleh karena membutuhkan biaya yang besar dan waktu yang lama, sehingga sangat disayangkan apabila data tersebut direduksi, selain itu pencilan terkadang bukan merupakan gangguan melainkan fenomena data dengan variabilitas yang tinggi. Dengan demikian pencilan memberikan informasi yang penting dan mendalam.

VII PEMBAHASAN UMUM

Pengembangan alat ukur glukosa darah non-invasif merupakan inovasi bidang penelitian yang terus berkembang, dan banyak upaya telah dilakukan untuk menciptakan teknologi yang dapat mengukur kadar glukosa dalam darah tanpa perlu menyuntikkan jarum atau mengambil sampel darah, karena hal tersebut dapat membuat cedera pasien dan merusak sel tubuh apabila dilakukan dalam jangka waktu yang lama. Beberapa pendekatan yang telah dilakukan dan sedang dikembangkan adalah spektroskopi. Spektroskopi merupakan metode yang menggunakan cahaya untuk mengukur kadar glukosa dalam darah. Ada beberapa jenis spektroskopi yang digunakan, termasuk spektroskopi inframerah dekat dan spektroskopi Raman. Teknologi ini terus dikembangkan untuk membuatnya lebih akurat dan dapat diandalkan. Beberapa perguruan tinggi atau lembaga yang bergerak dalam penelitian yang telah berhasil mengembangkan alat ukur glukosa darah non invasif, diantaranya adalah Departemen Fisika Institut Pertanian Bogor. Meskipun demikian, hingga saat ini teknik invasif tetap menjadi acuan dalam menentukan diagnosa dokter. Hal tersebut karena teknik invasif memiliki tingkat akurasi yang lebih baik dibandingkan dengan non invasif.

Evaluasi terhadap pengembangan alat ukur perlu dilakukan sehingga diperoleh hasil kinerja yang optimal. Pemodelan kalibrasi yang dilakukan dalam penelitian ini bertujuan untuk menentukan metode dengan performa yang unggul dalam menghubungkan data hasil pengukuran glukosa darah non invasif dengan data hasil pengukuran invasif. Metode dengan performa yang baik dapat menghasilkan model kalibrasi yang dapat digunakan untuk memprediksi ukuran glukosa darah seseorang melalui pola data yang dihasilkan dengan teknik non invasif dengan akurasi yang tinggi, sehingga pengukuran dengan teknik invasif tidak lagi menjadi acuan utama dalam menentukan diagnosa dokter.

Fenomena data hasil pengukuran dengan teknik spektroskopi adalah bersifat spektrum, dan data jenis spectrum dikatakan sebagai data fungsional. Analisis data fungsional (ADF) menjadi alternatif dalam menganalisis data fungsional dengan metode regresi modern yang telah disesuaikan dengan kondisi data fungsional. Dengan demikian pendekatan non fungsional menjadi tidak efektif lagi apabila digunakan untuk menganalisis data fungsional. Berbagai metode populer telah dikembangkan disesuaikan dengan kebutuhan data fungsional. Metode yang telah dikembangkan diantaranya regresi fungsional (RF), regresi komponen utama fungsional (RKU), regresi kuadrat terkecil parsial fungsional (RKTPF), regresi ridge fungsional (RRF), dan yang terkini yaitu regresi kontinum fungsional (RKF) yang dikembangkan oleh Zhou (2019).

RKF merupakan metode yang dikembangkan dari regresi kontinum yang dikenal oleh Stone dan Brooks (1990). Seperti halnya regresi kontinum, regresi kontinum fungsional juga merupakan metode fleksibilitas yang didasari dari konsep

RF, RKUF, dan RKTPF. Pengembangan model kalibrasi pada data glukosa darah non invasif telah dilakukan kajian pertama dalam penelitian ini, dengan menggunakan dua pendekatan yaitu fungsional dan non fungsional. Pendekatan fungsional yang digunakan adalah RKF dan RF, sedangkan pendekatan non fungsional yang digunakan adalah RKU, RKTP dan regresi kuadrat terkecil (RKT).

Data glukosa darah yang akan digunakan dalam penelitian ini melalui tahap prapemrosesan dengan cara mentransformasi data menggunakan metode *Wavelet* dengan *mother Wavelet* yaitu Daubechies. Tujuan dari transformasi adalah untuk menghilangkan pengaruh gangguan (*noise*) yang pada umumnya terjadi diluar kontrol atau karena kondisi alat. *Noise* dapat mempengaruhi hasil analisa, sehingga proses ini perlu dilakukan untuk meminimalisir pengaruhnya. Hasil eksplorasi data diperoleh jumlah data dalam spektrum yang dihasilkan oleh alat non invasif bervariasi pada setiap objek pengamatan, sehingga perlu dilakukan teknik peringkasan (reduksi) data agar jumlah data antar pengamatan sama, dan dapat dilakukan analisa lanjutan.

Teknik peringkasan yang dilakukan menggunakan luasan dengan pendekatan konsep trapezoida. masing-masing amatan akan dibentuk 20 luasan dengan membagi rata berdasarkan pada waktu yang dibutuhkan untuk mengukur satu kali ulangan pada amatan. Satu amatan diukur sebanyak 5 kali ulangan sehingga akan diperoleh 100 luasan dari masing-masing amatan.

Hasil yang diperoleh dari kajian pertama bahwa metode RKF memiliki performa yang cukup baik dalam pemodelan kalibrasi, meskipun ukuran kebaikan model yang diperoleh tidak terlalu berbeda jauh dengan metode RKU dan RKTP. Hasil *Root Mean Square Error of Prediction* (RMSEP), *Mean Absolute Error* (MAE), dan *mean absolute percentage error* (MAPE) yang diperoleh oleh masing-masing metode cukup besar, serta korelasi cukup kecil. Hal ini menunjukkan fenomena data aktual yang diperoleh dengan teknik invasif tidak cukup baik digambarkan oleh data yang diperoleh dengan teknik non invasif. Banyak faktor yang mempengaruhi fenomena data baik eksternal ataupun internal alat ukur. Berdasarkan data yang diperoleh maka perlu dilakukan evaluasi terhadap alat ukur non invasif yang telah dikembangkan sehingga data yang diperoleh dapat dengan teknik invasif dapat mendekati nilai akurasi dan prediksi dari alat ukur invasif sebagai metode konvensional, dan metode invasif sampai saat ini tetap menjadi rujukan utama yang digunakan di bidang kesehatan dan kedokteran.

Dari hasil analisis dapat dikatakan bahwa metode reduksi seperti RKF, RKU, dan RKTP merupakan metode yang unggul apabila digunakan ketika terjadi multikolinearitas dengan jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan. Berdasarkan hasil analisis yang diperoleh metode RF dan RKT sebagai metode yang non reduksi menunjukkan performa kurang efektif digunakan untuk kasus dengan jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan. Secara menyeluruh dapat dikatakan bahwa pendekatan RKF lebih baik karena sifat fungsional yang dimiliki tidak menghilangkan pola data. Dengan demikian dapat dikatakan bahwa regresi kontinum fungsional merupakan pendekatan yang efektif dalam menganalisis model kalibrasi.

Berdasarkan hasil penelitian sebelumnya dengan data yang sama pada penelitian ini diketahui bahwa terdapat data pencilan. Data pencilan adalah titik data yang secara signifikan berbeda dari pola umum dalam kumpulan data. Pencilan bisa menjadi hasil dari kesalahan pengukuran, anomali dalam proses, atau bahkan

representasi dari fenomena yang nyata tetapi tidak biasa. pencilan sering kali diidentifikasi dalam analisis data, karena dapat memengaruhi hasil analisis statistik dan pemodelan. Strategi untuk menangani pencilan bisa mencakup menghapusnya dari analisis, atau menyesuaikan teknik statistik yang resisten terhadap pencilan.

Berdasarkan pada hasil kajian pertama RKF merupakan pendekatan dengan performa yang cukup baik dalam pemodelan kalibrasi data glukosa darah non invasif, namun RKF dinyatakan tidak resisten terhadap pencilan berdasarkan kajian simulasi dan empirik yang dilakukan oleh Zhou (2019). Penelitian ini akan melakukan kajian kedua mengenai kekekaran (resistensi) dari kelima metode yang digunakan pada kajian pertama. Kekekaran kelima metode dikaji menggunakan data simulasi. Metode simulasi dapat memberikan efisiensi dan kemudahan dalam menganalisis suatu model matematika.

Metode simulasi yang digunakan dalam penelitian ini menggunakan data bangkitan sebanyak sembilan puluh satu peubah bebas, dan satu peubah respon dengan jumlah pengamatan (n) yang bervariasi yaitu 50, 100, dan 200 dan proporsi pencilan 2% hingga 10%. Fungsi yang dibangkitkan sebagai peubah bebas yaitu $\sin\left(\frac{i\pi}{45}\right) + \text{distribusi uniform}(365, -0.1, 0.1)$, $i = [0,364]$. Data bangkitan diulang sebanyak 100 kali.

Berdasarkan hasil analisis kekekaran metode RKF, RKU, RKTP, RKT, dan RF menggunakan data simulasi, yang dirancang dengan jumlah pengamatan yang bervariasi, dan juga proporsi pencilan yang bervariasi, dapat disimpulkan bahwa semakin banyak jumlah pengamatan nilai RMSEP dan MAE semakin menurun untuk kelima metode, sedangkan nilai korelasi antar peubah respon pengamatan dengan peubah respon dugaan terus meningkat. Kelima metode menunjukkan ketidakkekaran berdasarkan nilai RMSEP, dan MAE yang terus meningkat seiring bertambahnya nilai proporsi pencilan, dan nilai korelasi konsisten menurun.

Perbandingan antara metode fungsional dengan non fungsional dari hasil analisis, diketahui bahwa metode fungsional RKF dan RF menghasilkan nilai RMSEP dan MAE yang lebih kecil dibandingkan dengan metode yang bukan fungsional yaitu RKU, RKTP dan RKT. Selain itu nilai korelasi yang diperoleh oleh metode fungsional lebih besar dibandingkan dengan ketiga metode yang bukan fungsional, hal ini terjadi secara konsisten meskipun jumlah pengamatan dan nilai proporsi pencilan berubah. Perbandingan secara menyeluruh dapat disimpulkan bahwa kekekaran metode RKF masih lebih baik dibandingkan dengan RKU, RKTP, RKT, maupun RF.

Mengacu pada hasil kajian kedua maka dalam penelitian dilakukan kebaruan dengan mengembangkan RKF menjadi kekar, yang disebut dengan regresi kontinum fungsional kekar (RKFK). Konsep kekar yang digunakan adalah dengan memberikan bobot pada amatan yang terdeteksi pencilan berdasarkan nilai sisaan yang diperoleh. Kekekaran yang menjadi kebaruan dalam penelitian ini yaitu dengan cara membangkitkan data amatan sebanyak jumlah bobot yang diperoleh dan telah ditransformasi hingga diperoleh bilangan bulat positif (*counting*).

Besaran bobot yang diperoleh antara 0 sampai dengan 1, semakin besar sisaan yang diperoleh semakin kecil nilai bobot yang diberikan (mendekati 0), demikian sebaliknya semakin kecil sisaan yang diperoleh maka semakin besar nilai bobot yang diberikan (mendekati 1). Kriteria pemberian bobot bergantung pada fungsi pembobot yang digunakan. Kajian dalam penelitian ini menggunakan empat jenis fungsi pembobot yaitu Huber, Hampel, Ramsay, dan Tukey (*Bisquare*).

Berdasarkan teknik kekekaran yang telah dikonsepsikan maka algoritma dalam RKFK adalah sebagai berikut:

1. Regresikan x dan y menggunakan metode RKF
2. Hitung nilai sisaan (\mathbf{e}_i) untuk masing-masing pengamatan berdasarkan model yang diperoleh pada langkah 1
3. Hitung nilai $\hat{\sigma}$, dengan $\hat{\sigma} = \mathbf{1,4826}(\mathbf{median}|e_i - \mathbf{median}(e_i)|)$
4. Hitung nilai pembobot untuk masing-masing pengamatan (\mathbf{w}_i) berdasarkan nilai sisaan yang diperoleh pada Langkah 2, formulasi bobot menggunakan fungsi Huber.
5. Transformasi nilai bobot $\mathbf{w}^*(\mathbf{e}_i) = \frac{\mathbf{w}(\mathbf{e}_i+0,1)}{\mathbf{min}(\mathbf{w}(\mathbf{e}_i)+0,1)}$
6. Bangkitkan data x dan y sebanyak masing-masing bobot yang diperoleh pada Langkah 5
7. Regresikan data bangkitan pada Langkah 6 menggunakan metode RKF.
8. Ulangi Langkah 2 hingga 7 sampai mendapatkan nilai konvergen. Nilai konvergen diperoleh pada saat $|\mathbf{e}_n - \mathbf{e}_{n-1}| \leq \mathbf{0,00001}$.

Algoritma regresi kontinum fungsional yang digunakan, yang telah dikembangkan oleh Zhou (2019) sebagai berikut:

```

for  $(p, \alpha)$  in finite set do           #  $p$  dan  $\alpha \in [0,1)$  adalah parameter tuning
  for  $i$  from 1 to  $n$  do
    if  $p = 1$  then
       $\hat{X}_i^{(p,\alpha)} \leftarrow X_i - \bar{X}$    #  $X_i$  data pengamatan ke-  $i$  (peubah bebas),  $\bar{X}$  rata-rata,  $\hat{X}_i$ 
data dugaan  $X_i$ 
       $\hat{Y}_i^{(p,\alpha)} \leftarrow Y_i - \bar{Y}$    #  $Y_i$  data pengamatan ke-  $i$  (peubah respon),  $\bar{Y}$  rata-rata,  $\hat{Y}_i$ 
data dugaan  $Y_i$ 
       $\hat{\beta}_{p-1,\alpha} \leftarrow 0$          #  $\hat{\beta}$  penduga koefisien model
    else
       $\hat{X}_i^{(p,\alpha)} \leftarrow \hat{X}_i^{(p-1,\alpha)} - c_2 \cdot c_3 \cdot \hat{V}_X(\hat{\omega}_{p-1,\alpha})$  #  $\hat{V}_X$  penduga kovarians  $X$ ,  $\hat{\omega}$  penduga
fungsi basis
       $\hat{Y}_i^{(p,\alpha)} \leftarrow \hat{Y}_i^{(p-1,\alpha)} - \hat{\eta}_{p-1,\alpha}(X_i)$    #  $\hat{\eta}$  penduga nilai harapan bersyarat
    end if
  end for
   $\hat{\lambda}_j^{(p,\alpha)}, \hat{\phi}_j^{(p,\alpha)} \leftarrow$  nilai eigen ke- $j$  dan  $\hat{V}_{\hat{X}^{(p,\alpha)}}$  fungsi eigen
   $a_j \leftarrow c\hat{o}v \left\{ \hat{Y}^{(p,\alpha)}, \int_{\mathbb{T}} \hat{X}^{(p,\alpha)} \hat{\phi}_j^{(p,\alpha)} \right\}$ 
   $b_j \leftarrow v\hat{a}r \left\{ \int_{\mathbb{T}} \hat{X}^{(p,\alpha)} \hat{\phi}_j^{(p,\alpha)} \right\}$ 
   $Q_{p,\alpha}(\delta)$ 

$$\leftarrow \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta}} \right\}^2 \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\left( \hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta} \right)^2} \right\}^{\frac{\alpha}{(1-\alpha)}} \left\{ \sum_{j=1}^{\infty} \frac{a_j^2 b_j}{\left( \hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta} \right)^2} \right\}^{\frac{\alpha}{(1-\alpha)} - 1}$$

   $\hat{\delta}^{(p,\alpha)} \leftarrow \mathop{\text{argmin}}_{\delta \in (-1,0) \cup (0,\infty)} - \ln Q_{p,\alpha}(\delta)$ 

```

$$\hat{\omega}_{p,\alpha} \leftarrow \left\{ \sum_{j=1}^{\infty} \frac{a_j^2}{\left(\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta^{(p,\alpha)}} \right)^2} \right\}^{\frac{1}{2}} \sum_{j=1}^{\infty} \frac{a_j}{\hat{\lambda}_j^{(p,\alpha)} + \frac{\hat{\lambda}_1^{(p,\alpha)}}{\delta^{(p,\alpha)}}} \hat{\phi}_j^{(p,\alpha)}$$

$$c_1 \leftarrow \widehat{\text{cov}}(Y, \int_{\mathbb{T}} X \hat{\omega}_{p,\alpha})$$

$$c_2 \leftarrow \widehat{\text{var}}^{-\frac{1}{2}} \left(\int_{\mathbb{T}} X \hat{\omega}_{p,\alpha} \right)$$

$$c_3 \leftarrow \int_{\mathbb{T}} \hat{X}_i^{(1,\alpha)} \hat{\omega}_{p,\alpha}$$

$$\hat{\beta}_{p,\alpha} \leftarrow \hat{\beta}_{p-1,\alpha} + c_1 c_2 \hat{\omega}_{p,\alpha}$$

for i from 1 to n **do**

$$\hat{\eta}_{p,\alpha}(X_i) \leftarrow \bar{Y} + \int_{\mathbb{T}} \hat{X}_i^{(1,\alpha)} \hat{\beta}_{p,\alpha}$$

end for

$$\text{GCV}(p, \alpha) \leftarrow (n - p - 1)^{-2} \sum_{i=1}^n \{Y_i - \hat{\eta}_{p,\alpha}(X_i)\}^2 \quad \# \text{ generalized cross-validation (GCV)}$$

end for

$$\text{optimal}(p, \alpha) \leftarrow \underset{(p,\alpha)}{\text{argmin}} \text{GCV}(p, \alpha)$$

Proses analisis menggunakan RKFK membutuhkan waktu yang cukup lama, karena membutuhkan proses regresi sebanyak dua kali untuk menghasilkan nilai dugaan (\hat{y}). Berdasarkan permasalahan waktu komputasi, maka dalam penelitian ini dikaji aspek waktu komputasi dibutuhkan untuk masing-masing metode. Konsep analisis komputasi yaitu dengan menentukan waktu awal dan berakhirnya program komputasi. Selisih waktu awal dan akhir menjadi lama waktu komputasi yang dibutuhkan masing-masing metode. Berdasarkan hasil kajian terhadap waktu komputasi yang dibutuhkan, diketahui RKFK optimum digunakan ketika jumlah data semakin banyak.

Berdasarkan hasil analisis pada kajian ketiga dan keempat dapat dikatakan bahwa metode reduksi seperti regresi kontinum fungsional, regresi komponen utama, dan regresi kuadrat terkecil parsial merupakan metode yang unggul apabila digunakan ketika terjadi multikolinearitas atau jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan.

Dalam kasus analisis data fungsional penerapan regresi fungsional lebih baik karena tidak menghilangkan pola data. Dengan demikian dapat dikatakan bahwa pendekatan fungsional merupakan pendekatan yang efektif dalam menganalisis model kalibrasi, dengan teknik spektroskopi yang umumnya memiliki data fungsional, dan memiliki jumlah peubah bebas lebih besar dibandingkan dengan jumlah pengamatan.

Permasalahan lain pada data hasil pengukuran pada umumnya memiliki pencilan yang dapat mempengaruhi hasil pendugaan parameter, sangat disayangkan apabila diambil keputusan dengan menghilangkan data pencilan dari data hasil pengukuran seperti pada data glukosa darah yang digunakan dalam penelitian ini, karena waktu yang diperlukan untuk menghasilkan data cukup lama serta memerlukan biaya yang cukup mahal. Permasalahan ini menjadi dasar dilakukan kajian RKFK dengan keempat fungsi pembobot yaitu Huber, Hampel, Ramsya, dan Tukey (*Bisquare*).

Fungsi bobot Hampel dan Tukey (*Bisquare*) memiliki kriteria pemberian bobot 0 (nol), apabila nilai sisaan lebih besar dari nilai batasan (*cut off*) yang telah ditetapkan oleh masing-masing metode. Pendekatan algoritma RKFK mengakibatkan pemberian bobot 0 pada fungsi bobot Hampel dan Tukey (*Bisquare*) dapat dihindarkan, karena terdapat penambahan nilai konstanta lebih besar dari nol untuk setiap bobot (langkah 5 pada algoritma RKFK). Berdasarkan hasil analisis menggunakan data simulasi dan empirik terlihat performa RKFK dengan fungsi pembobot Huber lebih unggul.

Metode yang dikembangkan menggunakan regresi kontinum fungsional kekar merupakan pendekatan yang inovatif dalam penelitian ini. Regresi kontinum fungsional kekar memungkinkan penggabungan informasi dari spektrum cahaya dalam bentuk yang kontinu, sehingga dapat menangkap variasi kompleks dalam data spektrum glukosa darah yang dihasilkan oleh teknik spektroskopi non-invasif. Pendekatan ini dirancang untuk menangani tantangan yang dihadapi oleh data spektrum yang bervariasi akibat sifat fisik dan kimiawi dari objek yang diamati. Penggunaan regresi kontinum fungsional kekar, diharapkan dapat diperoleh model kalibrasi yang lebih kekar dan akurat, memungkinkan prediksi yang lebih tepat terhadap kadar glukosa darah meskipun dalam situasi di mana terdapat pergeseran atau ketidakteraturan dalam respons spektrum. Dengan demikian, metode ini dapat meningkatkan kinerja alat pengukur glukosa darah non-invasif secara signifikan.

Penelitian ini memberikan kontribusi terhadap perkembangan ilmu pengetahuan, khususnya dalam analisis data glukosa darah non-invasif yang telah dilakukan kajian sejak tahun 2017 hingga sekarang. Berbagai kajian yang telah dilakukan diantaranya adalah klasifikasi, pemodelan, transformasi data, hingga menentukan titik optimum. Penelitian yang menggunakan data glukosa darah non-invasif ini, telah dirangkum pada Lampiran 9.

VIII SIMPULAN DAN SARAN

8.1 Simpulan

Penelitian ini mengembangkan metode kebaruan yaitu regresi kontinum fungsional kekar dalam pemodelan kalibrasi alat pengukur glukosa darah, yang memungkinkan integrasi data spektrum cahaya secara lebih efektif. Pendekatan ini diharapkan dapat mengatasi tantangan variasi data yang kompleks akibat sifat fisik dan kimiawi dari objek yang diamati, dengan tujuan utama meningkatkan akurasi prediksi kadar glukosa darah dalam pengukuran non-invasif.

Hasil penelitian terhadap data simulasi dan empirik, berdasarkan faktor jumlah data amatan dan jumlah pencilan, diketahui bahwa performa regresi kontinum fungsional kekar lebih baik dibandingkan dengan metode lainnya, yaitu regresi fungsional, regresi kontinum fungsional, regresi kuadrat terkecil, regresi komponen utama, dan regresi kuadrat terkecil parsial. Setiap metode ditentukan nilai ukuran kebaikan model yaitu *root mean square error of prediction* (RMSEP), *mean absolute error* (MAE), korelasi antar data aktual dengan data dugaan, dan *mean absolute percentage error* (MAPE). Pendekatan regresi kontinum fungsional kekar menunjukkan nilai RMSEP, MAE, dan nilai relatif MAPE yang lebih kecil dibandingkan dengan metode lainnya. Sedangkan nilai korelasi yang diperoleh regresi kontinum fungsional kekar diperoleh lebih besar dibandingkan dengan metode lainnya.

Hasil analisis ragam diperoleh faktor jumlah amatan dan jumlah pencilan, berpengaruh terhadap nilai ukuran kebaikan model pada kajian simulasi. Hasil analisis komputasi waktu yang dibutuhkan oleh masing-masing pendekatan, dapat disimpulkan bahwa regresi kontinum fungsional kekar untuk jumlah amatan kecil membutuhkan waktu yang lebih lama dibandingkan pendekatan lainnya, namun untuk jumlah amatan yang besar waktu komputasi yang diperoleh lebih stabil, atau tidak berbeda jauh ketika jumlah amatan kecil. Hal tersebut berbeda dengan pendekatan lainnya.

Kajian terhadap pendekatan fungsi pembobot Huber, Hampel, Ramsay dan Tukey (*Bisquare*) pada regresi kontinum fungsional kekar, disimpulkan pendekatan Huber menunjukkan performa lebih baik. Faktor yang berpengaruh terhadap perbedaan dari masing fungsi pembobot adalah jumlah pencilan, sedangkan untuk jumlah amatan tidak berpengaruh secara signifikan.

8.2 Saran

Penelitian ini masih memiliki potensi besar untuk dapat dikembangkan dan diterapkan dengan berbagai fenomena data yang ada. Penelitian ini masih terbatas pada penerapan RKFK pada model kalibrasi data glukosa darah non invasif, sehingga sangat berpotensi diterapkan pada pemodelan kalibrasi dengan kasus data pada bidang yang lain, seperti di bidang perikanan, pertanian, farmasi, atau lainnya.

Pengembangan metode kekar juga dapat dilakukan pada metode regresi yang lainnya seperti regresi komponen utama fungsional kekar, regresi kuadrat terkecil parsial fungsional kekar, regresi ridge fungsional kekar, dan metode-metode lainnya. Konsep kekekaran juga masih banyak yang dapat dieksplorasi secara numerik maupun analitik.

Penetapan parameter tuning yang optimum dalam penelitian masih dilakukan secara manual. Potensi lain yang dapat dikembangkan dari penelitian ini, yaitu penetapan parameter *tunning* pada metode regresi kontinum fungsional kekar, dapat dilakukan menggunakan pendekatan *response surface*.

Reduksi data atau pembobotan nol seperti pada fungsi pembobot Hampel dan Tukey (*Bisquare*) mungkin akan menunjukkan performa yang lebih baik seperti pada penelitian Pratiwi *et. al.* (2018), namun tentunya akan kehilangan informasi data secara utuh. Penelitian ini berpotensi untuk dikembangkan dengan fungsi pembobot lain seperti Andrew, Welsh, atau yang lainnya, serta penelitian mengenai ukuran kebaikan model yang lain seperti bias atau *mean square of error* (MSE), dan ukuran relatif seperti *Mean Relative Absolute Error* (MRAE).

DAFTAR PUSTAKA

- Aguinis, H., Gottfredson, R. K., & Joo, H. 2013. Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2):270-301.
- Akansu AN, Haddad RA. 2001. *Multiresolution Signal Decomposition*. Second.
- Ariani, D., Nasution, Y. N., & Yuniarti, D. 2017. Perbandingan Metode Bootstrap Dan Jackknife Resampling Dalam Menentukan Nilai Estimasi Dan Interval Konfidensi Parameter Regresi. *EKSPONENSIAL*, 8(1), 43-50.
- Arnita. 2005. Koreksi pencaran dalam model kalibrasi peubah ganda pada data senyawa aktif gingerol serbuk rimpang jahe (*zingiber officinaleroscue*).
- Avan. 2017. Pendugaan Konsentrasi Glukosa dalam Darah Menggunakan Jaringan Syaraf Tiruan pada Alat Non-Invasive.
- Barsanti, R. J., & Gilmore, J. 2011. Comparing noise removal in the Wavelet and Fourier domains. In 2011 IEEE 43rd Southeastern Symposium on System Theory. 163-167.
- Beaton, A. E., & Tukey, J. W. 1974. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16(2):147-185.
- Benko M. 2004. *Functional Principal Components Analysis, Implementation and Applications*. :68.
- Berrendero JR, Justel A, Svarc M. 2011. Principal components for multivariate functional data. *Comput. Stat. Data Anal.* 55(9):2619–2634. doi:10.1016/j.csda.2011.03.011.
- Björkström, A., & Sundberg, R. 1999. A generalized view on continuum regression. *Scandinavian Journal of Statistics*, 26(1):17-30.
- Blanco, M., J. Coello, H. Iturriaga, S. MasPOCH, and C. de la Pezuela. 1998. Near-infrared spectroscopy in the pharmaceutical industry. *ApplSpectrosc* 123, 135–150.
- Brauer, A. (2015). Absorption spectroscopy. In *Supercritical Fluid Science and Technology*. 7: 347-366. Elsevier.
- Brouckaert D, Uyttersprot J-S, Broeckx W, De Beer T. 2017. Calibration transfer of a Raman spectroscopic quantification method for the assessment of liquid detergent compositions between two at-line instruments installed at two liquid detergent production plants. *Anal. Chim. Acta.* 984:1–18. doi:10.1016/j.aca.2017.07.044.
- Brown PJ. 1993. *Measurement, Regression, and Calibration*. Clarendon Press.
- Caprarelli, G., & Fletcher, S. 2014. A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. *Parasitology*, 141(5): 581-601.
- Chau FT, Liang YZ, Gao J, Shao XG. 2004. *Chemometrics: From Basic to Wavelet Transform*. John Wiley & Sons.
- Ciudin, A., Hernández, C., & Simó, R. 2012. Non-invasive methods of glucose measurement: current status and future perspectives. *Current diabetes reviews*, 8(1): 48-54.
- Dantu, V., Vempati, J., & Srivilliputhur, S. 2014. Non-invasive blood glucose monitor based on spectroscopy using a smartphone. In 2014 36th annual

- international conference of the IEEE engineering in medicine and biology society (pp. 3695-3698).
- Das D, Pasupathy KS, Haddad NN, Hallbeck MS, Zielinski MD, Sir MY. 2019. Improving Accuracy of Noninvasive Hemoglobin Monitors: A Functional Regression Model for Streaming SpHb Data. *IEEE Trans. Biomed. Eng.* 66(3):759–767.doi:10.1109/TBME.2018.2856091.
- de Cheveigné, A., & Arzounian, D. 2018. Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. *NeuroImage*, 172: 903-912.
- Donoho, D. L., & Johnstone, I. M. 1995. Adapting to unknown smoothness via Wavelet shrinkage. *Journal of the american statistical association*, 90(432):1200-1224.
- Fadhilah N. 2020. Perbandingan Metode Regresi Fungsional Dan Regresi Komponen Utama Fungsional Untuk Menduga Kadar Glukosa Darah Pada Alat Non-Invasif. :12.
- Fearn T. 1998. Wavelets. *SAGE J.* 9(5):10–12.doi:https://doi.org/10.1255/nirn.485.
- Fox, J., & Weisberg, S. 2002. Robust regression. An R and S-Plus companion to applied regression, 91, 6.
- Frank Ildiko E, Friedman JH. 1993. A Statistical View of Some Chemometrics Regression Tools. *Technometrics.* 35(2):109–135.doi:10.1080/00401706.1993.10485033.
- Fredriksson H. 2012. On the Short-Time Fourier Transform and Gabor Frames generated by B-splines.
- Guo DX, Shang YZ, Peng R, Yong SS, Wang XA. 2015. Noninvasive Blood Glucose Measurement Based on NIR Spectrums and Double ANN Analysis. *J. Biosci. Med.* 03(06):42–48.doi:10.4236/jbm.2015.36007.
- Hassan, A. F., Barakat, S., & Rezk, A. 2022. Towards a deep learning-based outlier detection approach in the context of streaming data. *Journal of Big Data*, 9(1):120.
- Herianti. 2020. Pemodelan Kalibrasi Kekar (Pengukuran Glukosa Darah Non-Invasif).
- Hidayaturrohman, U., Erfiani, E., & Afendi, F. M. 2020. Implementasi Transformasi Fourier Untuk Transformasi Domain Waktu Ke Domain Frekuensi Pada Luaran Purwarupa Alat Pendeteksian Gula Darah Secara Non-Invasif. *Indonesian Journal of Statistics and Its Applications*, 4(2), 234-244.
- Horváth, L., & Kokoszka, P. 2012. Inference for functional data with applications (Vol. 200). Springer Science & Business Media.
- Huber PJ. 1981. *Robust Statistics*. John Wiley & Sons.
- Hubert, M., Rousseeuw, P. J., & Vanden Branden, K. 2005. ROBPCA: a new approach to robust principal component analysis. *Technometrics*, 47(1):64-79.
- Hruschka, W. 1987. Data analysis: Wavelength selection methods. *Data analysis: Wavelength selection methods*, 35–56
- Ismah, Wigena AH, Djuraidah A. 2009. Pendekatan Regresi Kuadrat Terkecil Partial Robust Dalam Model Kalibrasi. :9.
- Ismah Ismah, Erfiani, Aji Hamim Wigena, Bagus Sartono. 2024. Functional Continuum Regression Approach to Wavelet Transformation Data in a Non-

- Invasive Glucose Measurement Calibration Model. *Mathematics and Statistics*, 12(1): 69 - 79. DOI: 10.13189/ms.2024.120110
- Jang YI, Sim JY, Yang J-R, Kwon NK. 2021. The Optimal Selection of Mother Wavelet Function and Decomposition Level for Denoising of DCG Signal. *Sensors*. 21(5):1851.doi:10.3390/s21051851.
- Jayanti, P. G. K., Anisa, R., & Aidi, M. N. 2018. Penerapan Teknik Prapemrosesan Smoothing Spline pada Data Hasil Pengukuran Alat Pemantau Kadar Glukosa Darah Non-Invasif. *Xplore: Journal of Statistics*, 2(2):15-23.
- Jl AR. 2016. *Pemodelan Regresi Komponen Utama Fungsional Pada Data Spektroskopi*. :61.
- Jong, S. D., Wise, B. M., & Ricker, N. L. 2001. Canonical partial least squares and continuum power regression. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 15(2): 85-100.
- Kokoszka, P., & Reimherr, M. 2017. *Introduction to functional data analysis*. Chapman and Hall/CRC.
- Kim, S., & Kim, H. 2016. A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669-679.
- Li Q, Yan H, Li L, Wu J, Zhang G. 2010. Application of Partial Robust M-Regression in Noninvasive Measurement of Human Blood Glucose Concentration with Near-Infrared Spectroscopy .pdf. *Guang Pu Xue Yu Guang Pu Fen Xi*. 30(2):2115-9.
- Modi, K., & Oza, B. 2016. Outlier analysis approaches in data mining. *International Journal of Innovative Research in Technology*, 3(7):6-12.
- Montgomery DC, Peck EA, Vining GG. 2012. *Introduction to Linear Regression Analysis*. :679.
- Nason, G. P., & Silverman, B. W. 1995. The stationary Wavelet transform and some statistical applications. In *Wavelets and statistics*. 281-299. New York, NY: Springer New York.
- Ningsih DP. 2017. Perbandingan Mother Wavelet Untuk Transformasi Wavelet Kontinu Pada Pemodelan Kalibrasi Peubah Ganda (Studi Kasus Data Kadar Gingerol Dan Kurkuminoid). :49.
- Novia SA. 2021. Model Kalibrasi Untuk Prediksi Kadar Gula Darah Non-Invasif Menggunakan Regresi Kuantil Komponen Utama.
- Ogden RT. 1997. *Essential Wavelets for Statistical Applications and Data Analysis*. Boston, MA: Birkhäuser Boston. [diunduh 2021 Mei 26]. Tersedia pada: <http://link.springer.com/10.1007/978-1-4612-0709-2>
- Polikar R. 1997. *The Wavelet Tutorial*.
- Preda, C., Saporta, G., & Lévéder, C. 2007. PLS classification of functional data. *Computational Statistics*, 22(2):223-235.
- Qannari, E. M., & Hanafi, M. 2005. A simple continuum regression approach. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 19(5-7):387-392.
- Ramayanti R. 2019. Simulasi Pendugaan Konsentrasi Glukosa Dalam Darah Pada Alat Non-Invasive. :52.
- Ramsay, J. O. 1982. When the data are functions. *Psychometrika*, 47:379-396.
- Ramsay, J. O., & Silverman, B. W. 2005. Smoothing functional data with a roughness penalty. *Functional Data Analysis*, 81-109.

- Ramsay, J., & Hooker, G. 2017. *Dynamic data analysis*. Springer New York, New York, NY. doi, 10, 978-1.
- Rassiyanti L. 2020. Analisis Regresi Ridge Robust Pada Pengukuran Kadar Glukosa Darah Non-Invasif. :14.
- Reiss PT, Ogden RT. 2007. Functional Principal Component Regression and Functional Partial Least Squares. *J. Am. Stat. Assoc.* 102(479):984–996.
- Rodliyah, I. 2016. Perbandingan Metode Bootstrap dan Jackknife dalam Mengestimasi Parameter Regresi Linier Berganda. *JMPM: Jurnal Matematika dan Pendidikan Matematika*, 1(1), 76-86.
- Rosni. 2019. Perbandingan Metode Regresi Komponen Utama, Regresi Kuadrat Terkecil Parsial dan Support Vector Regression dalam Menduga Kadar Glukosa Darah Non-Invasif.
- Rousseeuw, P. J., & Leroy, A. M. 2005. *Robust regression and outlier detection*. John Wiley & sons.
- Rousseeuw, P. J., & Croux, C. 1993. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424), 1273–1283. <https://doi.org/10.1080/01621459.1993.10476408>
- Rowe ACH, Abbott PC. 1995. Daubechies Wavelets and Mathematica. *Comput. Phys.* 9(6):635.doi:10.1063/1.168556.
- Serneels S, Croux C, Filzmoser P, Van Espen PJ. 2005. Partial robust M-regression. *Chemom. Intell. Lab. Syst.* 79(1–2):55–64.doi:10.1016/j.chemolab.2005.04.007.
- Serneels S, Filzmoser P, Croux C, Van Espen PJ. 2005. Robust continuum regression. *Chemom. Intell. Lab. Syst.* 76(2):197–204.doi:10.1016/j.chemolab.2004.11.002.
- Setiawan. 2007. Pendekatan Regresi Kontinum dalam Model Kalibrasi. :111.
- Setiawan, S., & Notodiputro, K. A. 2009. Continuum regression with discrete Wavelet transformation preprocessing. *Jurnal ILMU DASAR*, 8(2):103-109.
- Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. V. 2016. A review of machine learning techniques using decision tree and support vector machine. In 2016 international conference on computing communication control and automation (ICCUBEA) (pp. 1-7).
- Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF, Kraft P, Chen R, Kallberg HJ, Kurreeman FAS, et. al. 2012. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat. Genet.* 44(5):483–489.doi:10.1038/ng.2232.
- Stone M, Brooks RJ. 1990. Continuum Regression: Cross-Validated Sequentially Constructed Prediction Embracing Ordinary Least Squares, Partial Least Squares and Principal Components Regression. :34.
- Suma'inna S, Gumilar G. 2013. Implementasi Transformasi Wavelet Daubechies pada Kompresi Citra Digital. *CAUCHY*. 2(4):211.doi:10.18860/ca.v2i4.3117.
- Sutarno. 2010. Analisis Perbandingan Transformasi Wavelet pada Pengenalan Citra Wajah. *J. GENERIC*. 5(2):15–21.
- Wang J-L, Chiou J-M, Müller H-G. 2016. Functional Data Analysis. *Annu. Rev. Stat. Its Appl.* 3(1):257–295.doi:10.1146/annurev-statistics-041715-033624.
- Yao, F., Müller, H. G., & Wang, J. L. 2005. Functional data analysis for sparse longitudinal data. *Journal of the American statistical association*, 100(470): 577-590.

- Yazici, M. 2016. The weighted least squares ratio (WLSR) method to M-estimators. In 2016 SAI Computing Conference (SAI) (pp. 434-445). IEEE.
- Zanon M, Sparacino G, Facchinetti A, Talary M, Mueller M, Caduff A, Cobelli C. 2013. Non-Invasive Continuous Glucose Monitoring with Multi-Sensor Systems: A Monte Carlo-Based Methodology for Assessing Calibration Robustness. *Sensors*. 13(6):7279–7295.doi:10.3390/s130607279.
- Zhou Z. 2019. Functional continuum regression. *J. Multivar. Anal.* 173:328–346.