

PENERAPAN ANALISIS KOMPONEN UTAMA FUNGSIONAL (AKUF) PADA DATA SPEKTROSKOPI (KASUS : DATA TECATOR DI SOFTWARE R)

Ana Risqa JL¹, Novian Riskiana Dewi², Arini Alhaq³, Siti Ulfa Nabila⁴

¹Pendidikan Matematika FTIK UIN Raden Intan Lampung [Email: anarisqa@radenintan.ac.id]

² Pendidikan Matematika FTIK UIN Raden Intan Lampung [Email: novianriskiana@radenintan.ac.id]

³ Pendidikan Matematika FTIK UIN Raden Intan Lampung [Email: arinialhaq@radenintan.ac.id]

⁴ Pendidikan Matematika FTIK UIN Raden Intan Lampung [Email: sitiulfanabila@radenintan.ac.id]

[§]Corresponding Author : Ana risqa JL

Received 22nd May 2024; Accepted 20th Jun 2024; Published 30th Jun 2024;

Abstrak

Analisis komponen utama (AKU) merupakan salah satu teknik analisis dalam peubah ganda yang digunakan untuk menyederhanakan suatu data, dengan cara mentransformasi data secara linier sehingga terbentuk sistem_koordinat baru dengan varians maksimum, mereduksi dimensi suatu data tanpa mengurangi karakteristik data tersebut secara signifikan dan juga sering digunakan untuk menyelesaikan masalah multikolinearitas antar peubah bebas dalam model regresi berganda. Namun AKU mengalami kesulitan dalam menganalisis data fungsional seperti data Spektroskopi karena jumlah dimensi yang terlalu besar. Untuk mengatasi kesulitan ini, AKUF menyediakan cara yang jauh lebih informatif memeriksa struktur sampel kovarian dari AKU, dan juga dapat melengkapi pemeriksaan langsung dari varians-kovarians struktur. Tujuan Penelitian ini adalah menerapkan dan membandingkan metode AKU dan AKUF pada data berbentuk Spektroskopi. Pada data tekator yang ada di software R ini setipe dengan data spektroskopi dapat dilihat bahwa penerapan AKUF lebih baik dibandingkan dengan AKU ini terlihat dari nilai RMSEP pada hasil AKUF lebih kecil dibanding AKU. Sehingga dapat disimpulkan bahwa untuk menganalisis data spektroskopi akan lebih baik menggunakan AKUF pada studi kasus untuk data Tekator yang ada di software.

Kata Kunci: AKU, AKUF, Spektroskopi

1. Pendahuluan

Kemometrik merupakan aplikasi matematika untuk memproses, mengevaluasi, dan menginterpretasi sejumlah besar data yang dihasilkan oleh suatu percobaan kimia. Teknik ini memungkinkan penggunaan model kalibrasi dalam penerapannya. Data Spektroskopi berupa

data fungsional yang memiliki pengamatan spektrum yang merupakan fungsi dari panjang gelombang. Dalam menganalisis data spektroskopi dapat digunakan metode AKU untuk membentuk model kalibrasinya. Data spektroskopi ini diperoleh dari FTIR (*Fourier*

Transform Infrared). Pengukuran dengan menggunakan *FTIR* akan menghasilkan pola spektrum yang tertentu sesuai dengan responnya, proses ini memerlukan biaya dan waktu yang tidak terlalu besar. Model Kalibrasi digunakan untuk pendugaan konsentrasi senyawa aktif yang ada pada tanaman obat. AKU dapat digunakan untuk menyederhanakan dimensi suatu data dengan mentransformasi data secara linear sehingga terbentuk sistem koordinat baru dengan keragaman yang maksimum, sehingga AKU dapat mengatasi masalah dimensi yang besar, multikolinear dan $p \gg n$. Namun AKU masih memiliki banyak kekurangan yaitu dengan adanya jumlah p yang besar proses perhitungan AKU biasanya akan mengalami kendala dalam matriks ragam peragamnya yang berukuran sangat besar, dan pola data spektroskopi yang tidak mulus.

Beberapa penelitian dalam penggunaan metode AKU dalam model kalibrasi yaitu, Khristiningrum (1997) yang menunjukkan metode AKU untuk menduga model kalibrasi lebih baik daripada metode bertatar, begitupula Herwindiati (1997) membandingkan metode Regresi Komponen Utama, Regresi Ridge dan Regresi PLS. Selain metode AKU ada beberapa metode lain yang dapat digunakan dalam model kalibrasi anantara lain regresi bayes, regresi ridge, regresi bertatar, regresi PLS, regresi linear berganda dengan wavelenght dan lain-lain.

AKUF memiliki cara kerja yang sama dengan AKU sehingga dapat mengatasi masalah dimensi data yang besar, multikolinear serta $p \gg n$ namun tidak menghilangkan informasi yang ada dalam data. Dalam hal ini datanya berupa fungsi

atau kontinu, sehingga pada proses pereduksian dimensi dilakukan dengan tidak menghilangkan pola data dan data spektrum menjadi lebih mulus. Metode AKUF mengkonversi matriks data non fungsional menjadi matriks data fungsional yang merupakan kombinasi linier dari basis fungsi antara lain basis polinomial, basis B-spline dan basis Fourier. Dwi Nabila Lestari (2014) menerapkan AKUF dan Regresi Komponen Utama Fungsional (RKUF) untuk memprediksi curah hujan dengan menggunakan data GCM dan Shang (2011) melakukan kajian penggunaan metode RKUF. Aguilera et al (2013) melakukan penelitian menggunakan metode RKUF pada data spektroskopi. Pada penelitian ini membahas penggunaan AKUF untuk pemodelan kalibrasi pada data spektroskopi.

2. Landasan Teori

Analisis Komponen Utama Fungsional

AKUF banyak digunakan untuk menganalisis data fungsional dengan cara kerja yang sama dengan *AKU*. Ide utama AKUF ialah mengganti vector menjadi fungsi dan matriks ragam peragam dengan matriks ragam peragam fungsional. Jika terdapat peubah acak X_1, X_2, \dots, X_p yang akan direduksi. Unit sample yang digunakan sebanyak n , sehingga kombinasi linier peubah fungsional yang diperoleh ialah $\mathbf{x}(\mathbf{s}) = [x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_n(\mathbf{s})]$. Metode *AKUF* juga merupakan kombinasi linier dari $\mathbf{x}(\mathbf{s}), \mathbf{s} \in [x_1, x_p]$. Elemen matriks ragam peragam untuk data fungsional $\mathbf{v}(\mathbf{s}_j, \mathbf{s}_k)$ dapat dituliskan sebagai berikut :

Persamaan 2.1

$$v(s_j, s_k) = \frac{1}{n} \sum_{i=1}^n x_i(s_j)x_i(s_k)$$

Untuk $j = 1, \dots, p; k = 1, \dots, p$ dan $j \neq k$.

Pada data fungsional rumus umum mencari nilai akar cirri dan vector cirri dapat dituliskan sebagai berikut :

Persamaan 2.2

$$\int v(s_j, s_i)\xi(s_k)ds_k = \rho\xi(s_k)$$

Dengan ξ ialah vector ciri fungsional, ρ ialah akar ciri fungsional dan v ialah matriks ragam peragam. Nilai vector ciri fungsional memenuhi $\int_{x_1(s)}^{x_p(s)} \xi^2(s)ds = 1$ untuk memaksimalkan ragam komponen utama dan menjamin keunikan nilai vector ciri fungsional. Nilai skor komponen utama fungsional (kf) dapat dinotasikan sebagai berikut :

Persamaan 2.3

$$kf_i(s) = \int_{x_1(s)}^{x_n(s)} \xi_i(s)x_i(s) ds. \quad \text{untuk}$$

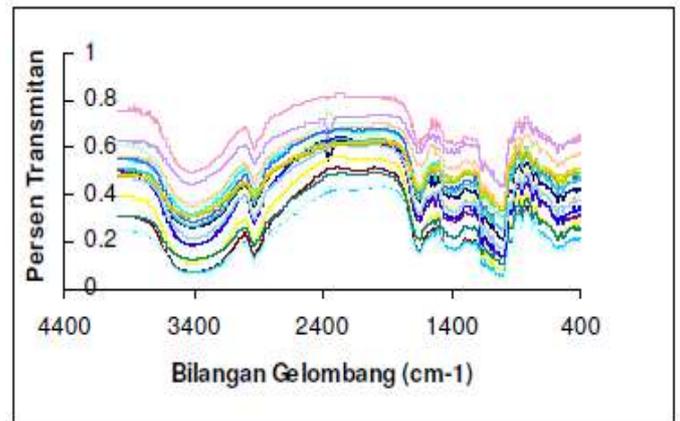
$1 \leq i \leq n$

Seluruh KU bersifat orthogonal yang artinya KU yang dihasilkan sudah tidak saling berkorelasi. AKU dan AKUF memiliki perbedaan yang paling mendasar yaitu tentang jumlah maksimum dari pasangan nilai akar cirri dan fungsi akar cirri. bagian ini tidak harus berbunyi Landasan Teori. Pada bagian ini diterangkan beberapa konsep dasar dalam bidang terkait judul artikel, yang benar-benar digunakan dalam bagian Hasil dan Pembahasan.

3. Hasil Dan Pembahasan

3.1 Eksplorasi data

Pada Data tekator dilakukan pengulangan sebanyak 1000 kali. Gambar 1 menyajikan spektrum hasil keluaran FTIR simulasi.



Gambar 1. Plot spektrum kluaran FTIR

Secara umum terlihat bahwa keseluruhan spektrum memiliki pola yang sama. Kemudian data dilakukan dua tahap pertama data awal di analisis menggunakan RKU dan kemudian data awal yang telah ditarnsformasi menjadi data fungsional menggunakan transformasi polinomial kemudian dilakukan analisis menggunakan RKUF.

3.2 Regresi Komponen Utama (RKU)

Metode RKU menggunakan data hasil simulasi data tekator. Hasil analisis dengan menggunakan RKU pada Tabel 1, menunjukkan bahwa rata-rata komponen utama pertama (Z_1) mampu menjelaskan sebanyak 90.09% dari keragaman yang ada dan rata-rata komponen kedua (Z_2) mampu menjelaskan 5.21% dari keragaman yang ada dan rata-rata komponen ketiga (Z_3) mampu menjelaskan 3,93% dari keragaman.

Tabel 3.1 Keragaman Komponen model RKU

Rata-rata Komponen			
	Z ₁	Z ₂	Z ₃
Proporsi Keragaman	0.9009	0.0521	0.0393
Proporsi Kumulatif	0.9009	0.9533	0.9926

Rata-rata Komponen Z₁ dan Z₂ mampu menjelaskan keragaman pada data sebesar 95.33 % sehingga peubah yang digunakan untuk analisis pada metode RKU adalah Z₁ dan Z₂. Sedangkan Penduga Parameter dan statistik uji pada satu kali pengulangan terdapat pada Tabel 3.2.

Tabel 3.2 Penduga parameter dan Statistik ujinya.

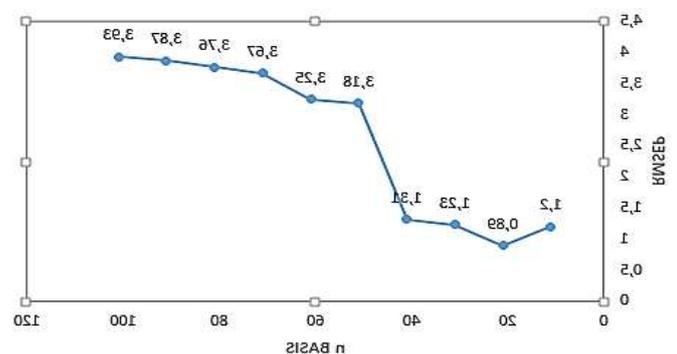
Parameter	Penduga	Galat Baku	t-hitung	Nilai-p
Intersep	1.12804	0,14234	7.87503	7.58e-06
Z ₁	-0.01900	0.02438	-0.77978	0.45278
Z ₂	0.02689	0.11140	0.24156	0.81432

Tabel 3.2 menunjukkan penduga parameter dan statistik ujinya dan tidak ada satupun komponen yang berpengaruh terhadap hasil pengukuran serbuk kurkumin. Ini juga dijelaskan dengan hasil nilai kebaikan model yang hanya 7.2 % yang berarti bahwa peubah Z₁ dan Z₂ hanya mampu menerangkan total keragaman hasil pengukuran sebesar 7.2%. Model akhir metode RKU dapat ditulis sebagai berikut :

$$Y = 1.12804 - 0.019 Z_1 + 0.02689 Z_2 \quad (iii)$$

3.3 Regresi Komponen Utama Fungsional (RKUF)

Metode RKUF menggunakan peubah prediktor data yang telah di transformasi menjadi data fungsional. Penentuan basis dalam transformasi dapat dengan menggunakan nilai *RMSEP*. Pemilihan basis terbaik yakni basis yang memiliki nilai *RMSEP* paling kecil yang digunakan dalam penelitian. Gambar 2 menunjukkan bahwa basis yang baik digunakan dalam metode AKUF adalah basis 21 dengan nilai *RMSEP* terkecil yakni 0,89.



Gambar 2 Penentuan basis dalam metode AKUF

Hasil analisis menggunakan metode RKUF dengan transformasi pada Tabel 5, menunjukkan bahwa rata-rata komponen utama fungsional basis pertama (*kfb₁*) mampu menjelaskan sebanyak 92.08% dari keragaman yang ada dan rata-rata komponen utama fungsional basis kedua (*kfb₂*) mampu menjelaskan 5.55% dari keragaman yang ada dan rata-rata komponen utama fungsional basis ketiga (*kfb₃*) mampu menjelaskan 1.37 % dari keragaman.

Tabel 3.3 proporsi rata-rata keragaman setiap komponen

	Komponen		
	kfb_1	kfb_2	kfb_3
Proporsi rata-rata Keragaman	0.9208	0.0555	0.0137
Proporsi Komulatif	0.9208	0.9755	0.9894

Tabel 3.5 memberikan informasi bahwa metode AKUF dapat memepertahankan keragaman data lebih baik dibandingkan dengan AKU, dilihat dari proporsi rata-rata keragaman kfb_1 yakni 92.08 % yang sudah cukup tinggi di atas 80%. Namun untuk menjaga kesamaan dalam membandingkan kedua metode tersebut, maka skor komponen utama fungsional yang di gunakan pada RKUF sebanyak dua yaitu kfb_1 dan kfb_2 sesuai dengan jumlah komponen utama yang digunakan pada metode RKU dalam membentuk model regresi.

Tabel 3.4 Penduga Parameter dan Statistik ujinya

Parameter	Penduga	Galat Baku	t- hitung	Nilai-p
Intersep	1.10583	0,086	12.831	5.12e-05
kf_1	0.05172	0.015	1.299	0.0342
Kf_2	0.06142	0.069	0.634	0.005

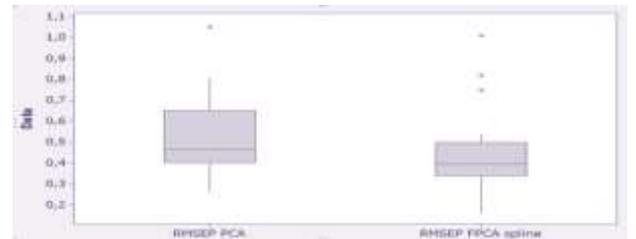
Tabel 3.4 merupakan tabel penduga parameter untuk metode RKUF dengan menggunakan dua skor komponen utama fungsional dengan proporsi rata-rata keragaman sebesar 96.94% dan model RKUF menghasilkan nilai kebaikan model sebesar 79.0% yang berarti bahwa peubah kfb_1 dan kfb_2 mampu menjelaskan total keragaman hasil pengukuran serbuk

kurkumin sebesar 79.0%. Ini lebih baik dibandingkan dengan menggunakan metode RKU. Model akhir metode RKUF sebagai berikut :

$$y = 1.105833 + 0.0517261 kf_1 + 0.0614291 kf_2$$
(v)

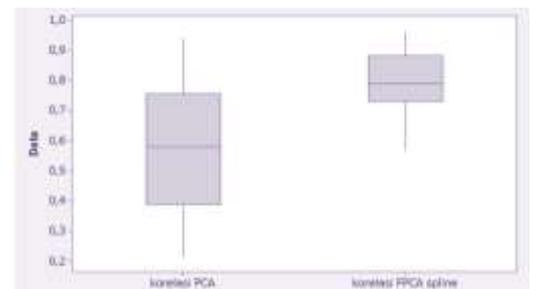
3.4 Validasi

Validasi merupakan salah satu tahap yang paling penting untuk dilakukan karena validasi mencerminkan keakuratan hasil prediksi model yang dihasilkan. Pada gambar 3 terlihat bahwa nilai *RMSEP* pada RKU lebih menyebar dan lebih cenderung besar nilainya, lalu pada RKUF nilai *RMSEP* nya lebih mengumpul dan cenderung lebih kecil nilainya dibandingkan pada RKU.



Gambar 3. Diagram kotak garis untuk nilai *RMSEP* pada RKU dan RKUF.

Pada gambar 4 terlihat nilai korelasi dengan menggunakan metode RKU lebih kecil dibandingkan nilai korelasi pada metode RKUF.



Gambar 4. Diagram kotak garis untuk nilai korelasi pada RKU dan RKUF.

3.5 Konsistensi Prediksi

Pada Tabel 3.4 dibawah ini merupakan hasil analisis konsistensi prediksi untuk kedua

metode yaitu metode RKU dan metode RKUF. Konsistensi model dilihat dari nilai simpangan baku *RMSEP* dan simpangan baku korelasi antara data awal dengan prediksinya pada 150 kali pengulangan.

Tabel 3.7 Konsistensi Prediksi

Metode	Simpangan Baku	
	<i>RMSEP</i>	Korelasi
RKU	0.159196	0.266652
RKUF	0.090051	0.103161

Metode RKU lebih baik untuk memprediksi data yang berjumlah sedikit namun jika untuk memprediksi data yang berjumlah besar kurang baik. Metode RKUF lebih konsisten dalam memprediksi data dalam jumlah besar hal ini berdasarkan dari nilai simpangan baku *RMSEP* (0.090051) dan simpangan baku r (0.103161).

4. Kesimpulan Dan Saran

Secara umum metode RKUF lebih baik dan konsisten dibandingkan dengan metode RKU dalam memprediksi data spektroskopi. Pada aplikasinya metode RKUF cenderung lebih rumit dibandingkan metode RKU karena tidak di ketahui batas tertentu dalam menentukan jumlah awalan knot yang akan digunakan untuk menentukan n basis yang digunakan.

Daftar Pustaka

- [1] Aguilera AM, Escabias M, Valderrama MJ, Morillo MCA. 2013. Functional Analysis of Chemometric Data. *Journal of Statistics* (3): 334-343.
- [2] Alfeeli B. 2005. Miniature Gas Sensing Device Based On Near-Infrared Spectroscopy. Blacksburg: Virginia Polytechnic Institute and State University.
- [3] Bellman RE. 1961. Adaptive Control Process : a Guided Tour, Princeton University Press, New Jersey : Princeton.
- [4] Croux C, Ruiz – Gazen A. 2005. High breakdown estimators for principal component : The projection – pursuit approach revisited. *Journal of Multivariate Analysis* 95: 206 – 226.
- [5] Ferraty F. & Vieu, P. 2006. Nonparametric Functional Data Analysis : Theory and Practise. New York : Springer.
- [6] Ingrassia S, Costanzo GD. 2005. Functional Principal Componen Analysis of Financial Timeseries, Springer – Verlag, Inc., Berlin.
- [7] Johnson RA, Wichern DW. 1998. Applied Multivariate Statistical Analysis . United States of America : Prentice Hall International. Inc.
- [8] Jolliffe IT. 2002 . Principal Componen Analysis Second Edition. Ed ke-2. New York (US) : Springer-Verlag.
- [9] Lestari DN. 2014. Pemodelan statistical downscaling dengan analisis komponen utama fungsional untuk prediksi curah hujan [tesis]. Bogor (ID): Sekolah Pascasarjana, Institut Pertanian Bogor.
- [10] Nur MA, Adijuwana H. 1989. Teknik Spektroskopi Dalam Analisis Biologi. Bogor: Pusat antar Universitas Ilmu Hayat, Institut Pertanian Bogor.
- [11] Ramsay OJ, Silverman WB . 2005. Functional Data Analysis. Ed ke-2. New York (US) : Springer.
- [12] Risqa JL Ana. 2015. Pemodelan Regresi Komponen Utama Fungsional Pada Data Spektroskopi [Tesis]. Bogor (ID): Sekolah Pascasarjana, Institut Pertanian Bogor.

- [13] Shang HL. 2011. A Survey of functional principal component analysis. Departement of econometrics and business Statistics Monash University, (working papers, 06/11).

- [14] Tran NM. 2008. An Introduction to Theoretical Properties of Functional Principal Component Analysis [tesis]. Melbourne (AU): University of Melbourne