

Studi Karakteristik *Profile* Beban Kelistrikan Area Jaringan Tangerang Menggunakan Analisa *Multivariate*

Adella Wendy Putri¹, Adri Senen²

^{1,2}Fakultas Ketenagalistrikan dan Energi Terbarukan, Institut Teknologi PLN, Jakarta, Indonesia

Article Info

Article history:

Received Maret, 4, 2024

Revised Maret, 4, 2024

Accepted Maret, 4, 2024

Kata Kunci:

Beban listrik,
Principal Component
Analysis(PCA),
K-means,
Regresi Linear

Keywords:

Electrical load ,
Principal Component Analysis
(PCA),
K-means,
Linear Regression

ABSTRAK

Analisis multivariat pada profil beban listrik adalah pendekatan yang penting untuk memahami pola penggunaan energi secara holistik dan mendalam. Dengan mengintegrasikan berbagai faktor yang mempengaruhi konsumsi energi listrik. Studi ini bertujuan untuk melakukan profilisasi beban listrik menggunakan metode analisis komponen utama (Principal Component Analysis, PCA) dan algoritma K-Means, serta mengevaluasi pemodelan dengan Regresi Linier Berganda dan menganalisis korelasi antara variabel-variabel terkait. Data beban listrik dari suatu wilayah tertentu dikumpulkan dan diolah menggunakan teknik PCA untuk mengidentifikasi pola utama dalam distribusi beban. Selanjutnya, algoritma K-Means digunakan untuk mengelompokkan pola-pola tersebut menjadi kategori yang lebih terdefinisi. Model Regresi Linier Berganda kemudian diterapkan untuk mengevaluasi hubungan antara variabel-variabel beban listrik dengan faktor-faktor yang memengaruhinya. Analisis korelasi digunakan untuk mengeksplorasi keterkaitan antara variabel beban listrik dan faktor-faktor lingkungan atau ekonomi yang dapat memengaruhi konsumsi energi. Hasil dari penelitian ini variabel PDRB, Jumlah Rumah tangga, beban listrik pada sektor bisnis dan perumahan memiliki korelasi yang dapat dikategorikan kuat yang melebihi nilai koefisien korelasi diangka 0,6. Reduksi duabelas variabel menggunakan PCA yang mempengaruhi konsumsi energi listrik menjadi empat variabel utama yaitu PDRB, Jumlah Rumah tangga, beban listrik pada sektor bisnis dan perumahan. Pemodelan yang menggunakan metode PCA mempunyai hasil yang baik dengan nilai RMSE 1.09

ABSTRACT

Multivariate analysis of electricity load profiles is an important approach to understand energy use patterns in a holistic and in-depth manner. By integrating various factors that influence electrical energy consumption. This study aims to profile electrical loads using the principal component analysis (PCA) method and the K-Means algorithm, as well as evaluating modeling using Multiple Linear Regression and analyzing the correlation between related variables. Electrical load data from a particular area is collected and processed using PCA techniques to identify main patterns in load distribution. Next, the K-Means algorithm is used to group these patterns into more defined categories. The Multiple Linear Regression Model is then applied to evaluate the relationship between electricity load variables and the factors that influence them. Correlation analysis is used to explore the relationship between electricity load variables and environmental or economic factors that can influence energy consumption. The results of this research are that the variables GRDP, number of households, electricity load in the business and housing sectors have a correlation that can be categorized as strong, exceeding the correlation coefficient value of 0.6. Reducing twelve variables using PCA that influence electrical energy consumption into four main variables, GRDP, number

of households, electricity load in the business and housing sectors. Modeling using the PCA method has good results with an RMSE value of 1.09.

This is an open access article under the [CC BY](#) license.



Corresponding Author:

Adella Wendy Putri

Fakultas Ketenagalistrikan dan Energi Terbarukan, Institut Teknologi PLN,
Jakarta, Indonesia
Email: adellawendy25@icloud.com

1. PENDAHULUAN

Dalam pengembangan sistem kelistrikan dan manajemen energi, pemahaman yang mendalam tentang karakteristik beban atau profil beban suatu wilayah menjadi elemen kunci. Profil beban adalah gambaran tentang bagaimana konsumsi listrik berfluktuasi seiring waktu. Dimana hal tersebut memiliki implikasi signifikan terhadap perencanaan, pengoperasian, dan manajemen distribusi energi listrik. Profil beban yang akurat dan detail memainkan peran penting dalam penentuan kapasitas transmisi dan distribusi yang diperlukan, serta dalam perencanaan sumber energi terbarukan, pengendalian beban, dan manajemen krisis. Dengan pesatnya pertumbuhan populasi serta perkembangan teknologi, terdapat peningkatan signifikan dalam beban energi listrik yang mungkin mengarah pada tantangan dalam penyediaan energi yang efisien dan berkelanjutan.

Pada penelitian ini menggunakan 4 wilayah sampel yang berbeda yang terhubung dengan Area Jaringan Tangerang diantaranya Kota Tangerang, Kabupaten Tangerang, Kota Bogor, dan Jakarta Barat dengan 119 kelurahan berbeda. Dengan membandingkan karakteristik penggunaan listrik di antara kota-kota tersebut dengan dua belas variabel yang seragam. Faktor beban, faktor kapasitas dan faktor permintaan dari tahun 2013 hingga tahun 2021 mengalami ketidakstabilan kenaikan maupun penurunan persentase, sementara daya yang tersambung dari tahun 2013 hingga tahun 2021 mengalami peningkatan yang signifikan. Peningkatan daya yang tersambung dapat disebabkan oleh pertumbuhan ekonomi, peningkatan aktivitas industri, dan perkembangan infrastruktur di wilayah tersebut. Hal ini dapat mengakibatkan peningkatan beban listrik karena lebih banyak bisnis, rumah tangga, dan industri memerlukan akses listrik untuk beroperasi.

Pada penelitian ini, teknik yang dipakai dalam analisis multivariat ialah *Principal Component Analysis* (PCA), Analisa regresi linear berganda dan analisa korelasi serta melibatkan metode kmeans untuk pengelompokan variabel dominan. Principal Component Analysis (PCA) dan algoritma K-Means telah terbukti menjadi metode yang kuat dalam analisis data untuk mengidentifikasi pola-pola yang tersembunyi dan mengelompokkan data menjadi kategori yang berbeda. Dalam konteks profilisasi beban listrik, penggunaan PCA dan K-Means dapat membantu dalam mengurangi dimensi variabel dan mengungkap struktur pola konsumsi energi yang mendasari.

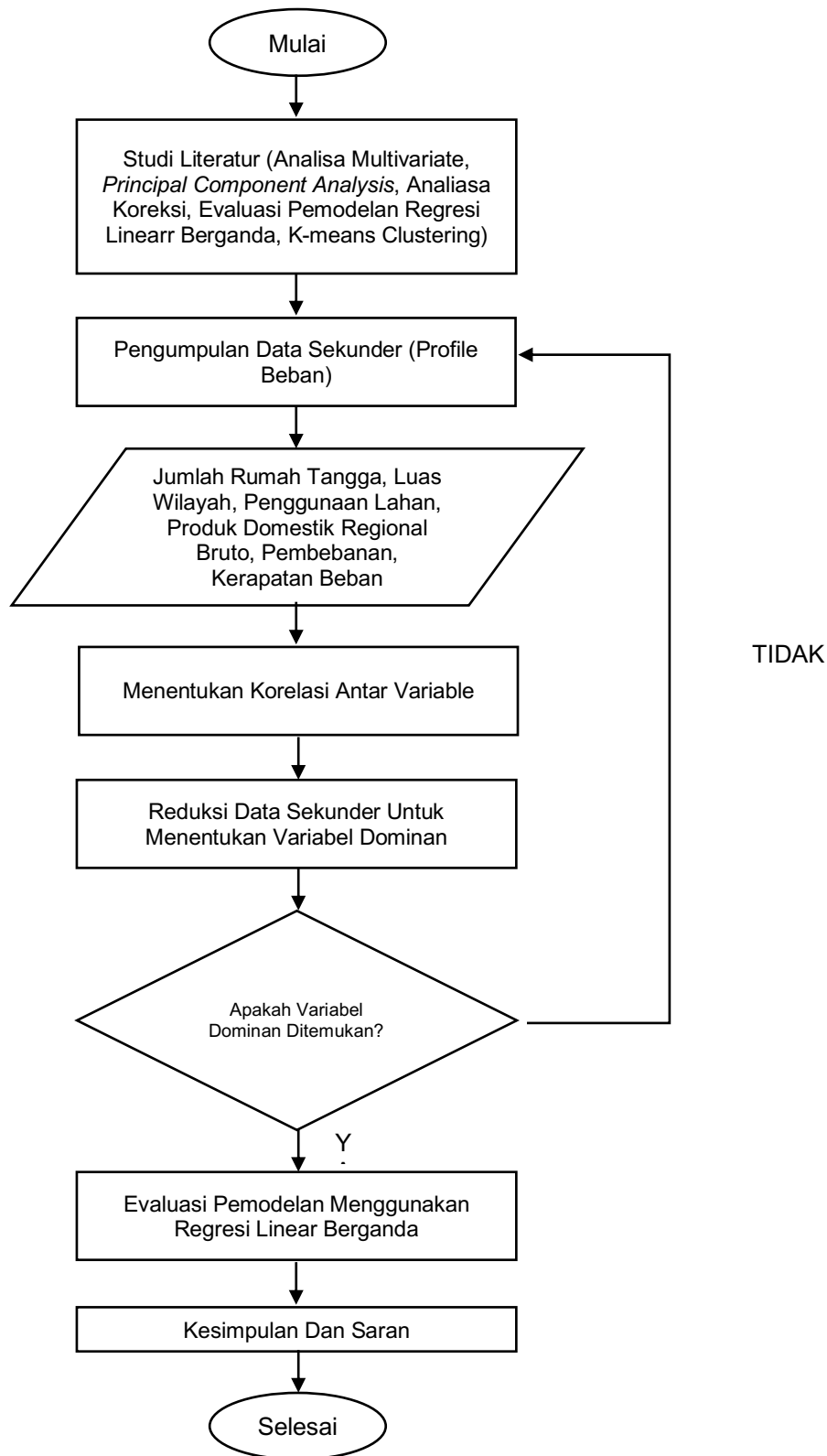
Oleh karena itu, penelitian yang menggabungkan PCA, K-Means, Regresi Linear Berganda, dan analisis korelasi dalam profilisasi beban listrik diharapkan dapat memberikan kontribusi yang signifikan dalam pemahaman tentang pola konsumsi energi listrik, pengembangan strategi pengelolaan energi yang lebih efisien, dan peningkatan kesadaran akan keberlanjutan dalam penggunaan energi

2. METODE

Penelitian ini dimulai dengan mempelajari literatur mengenai analisis profil beban serta metode-metode yang digunakan dalam analisis karakteristik beban untuk mengetahui hasilnya. Selain itu juga mempelajari metode ekstraksi menggunakan *Principal Component Analysis* (PCA) baik dalam kasus *load profile* maupun pada topik lain yang berhubungan. Serta mempelajari.

Tahap selanjutnya yaitu pengumpulan data yang akan digunakan dalam penelitian. Data yang dimaksud berupa data sekunder penggunaan beban untuk studi kasus area Tangerang yang terdiri dari data jumlah konsumen rumah tangga, luas wilayah, luas lahan setiap sektor (perumahan, industri, bisnis, dan sosial), PDRB (Produk Domestik Regional Bruto), data penggunaan beban setiap sektor (perumahan, industri, bisnis dan sosial) serta data *load density* atau kepadatan bebannya.

Setelah data profil beban diperoleh, tahap selanjutnya adalah proses ekstraksi data menggunakan metode *Principal Component Analysis* (PCA) untuk mengetahui variabel-variabel apa saja yang paling dominan serta memiliki pengaruh lebih besar terhadap *load density* berdasarkan data profil beban yang diperoleh. selanjutnya mengevaluasi hasil pemodelan dari PCA menggunakan Analisa regresi linear berganda dan menggunakan Analisa korelasi untuk memvalidasi hasil dari variabel dominan yang di hasilkan oleh PCA.



Gambar 1. Diagram Alur Penelitian

2.1 Metode Analisis Data

Dalam penelitian, analisis multivariat pada data profil beban dilakukan menggunakan metode *Principal Component Analysis* (PCA). Pengolahan data dengan metode PCA dilakukan menggunakan perangkat lunak Matlab. PCA membantu mengidentifikasi komponen utama

yang menjelaskan variasi dalam data dan mengidentifikasi hubungan antara variabel-variabel tersebut.

Berdasarkan data profil beban studi kasus daerah Tangerang yang diperoleh, variabel-variabel berupa jumlah konsumen rumah tangga, luas wilayah, luas lahan setiap sektor (perumahan, industri, bisnis, dan sosial), PDRB (Produk Domestik Regional Bruto), data penggunaan beban setiap sektor (perumahan, industri, bisnis dan sosial) akan diidentifikasi untuk mengetahui variabel apa saja yang memiliki pengaruh dominan terhadap *load density*. Dalam proses pengolahan menggunakan metode PCA, akan dihitung nilai-nilai parameter PCA seperti nilai dari *eigen value* dan *eigen vector*. Persamaan-persamaan pada metode PCA dapat dirumuskan menggunakan Persamaan (2.7) sampai dengan Persamaan (2.11).

Langkah pertama untuk menentukan variabel dominan dari data yang sudah diperoleh adalah melakukan Analisa korelasi yang bertujuan untuk melihat seberapa kuat korelasi antar variabel terhadap *load density*. Rumus yang digunakan untuk mencari koefisien korelasi terdapat pada persamaan (2.14). Fungsi dan perintah yang digunakan untuk mencari nilai koefisien korelasi pada MATLAB adalah sebagai berikut :

```
R1 = corr(data input,load density);
```

Selanjutnya setelah diketahui nilai masing masing koefisien korelasi yang di hasilkan olehh masing masing variabel mengolah data dengan metode PCA di Matlab dapat menggunakan beberapa perintah dan fungsi yang tersedia, antara lain:

1. Persiapan data

Memastikan data dalam bentuk matriks di mana baris mewakili observasi (misalnya, sampel atau pengukuran) dan kolom mewakili variabel. Data perlu juga harus dinormalisasi jika variabel memiliki skala yang berbeda.

2. Menghitung rata-rata

Menghitung rata-rata dapat menggunakan fungsi “mean” dan kemudian gunakan `bsxfun` untuk mengurangi rata-rata dari setiap observasi seperti contoh berikut:

```
mu = mean(data);  
dataCentered = bsxfun(@minus, data, mu);
```

3. Menghitung matriks kovarian

Menghitung matriks kovarian dari data yang telah dinormalisasi menggunakan fungsi berikut:

```
covMatrix = cov(dataCentered);
```

4. Menghitung komponen utama

Terdapat fungsi ‘pca’ untuk menghitung komponen utama (principal components). Dengan fungsi ini, dapat menentukan berapa banyak komponen utama yang akan simpan. Contohnya, jika ingin menyimpan 2 komponen utama maka:

```
numComponentsToKeep = 2;  
[coeff, score, latent, tsquared, explained] = pca(dataCentered, 'NumComponents',  
numComponentsToKeep);
```

Keterangan:

- ‘coeff’ adalah matriks komponen utama (principal component matrix).
- ‘score’ adalah proyeksi data pada ruang komponen utama.
- ‘latent’ adalah varian yang dijelaskan oleh setiap komponen utama.

- 'explained' adalah persentase varian yang dijelaskan oleh masing-masing komponen utama.

5. Memvisualisasikan

Hasilnya dapat divisualisasikan dalam ruang komponen utama atau menggambarkan persentase varian yang dijelaskan oleh setiap komponen utama dengan menggunakan fungsi berikut:

```
gscatter(score(:,1), score(:,2));
```

Setelah ekstraksi data menggunakan PCA dan variabel dominan diketahui, langkah selanjutnya adalah mengevaluasi pemodelan yang dihasilkan oleh PCA menggunakan Analisa regresi linear berganda dengan menggunakan aplikasi MATLAB. Evaluasi pemodelan ini membandingkan nilai error yang dihasilkan antara variabel utama yang sudah diketahui dengan 12 variabel bebas. Pengolahan data menggunakan Analisa regresi linear berganda untuk evaluasi pemodelan menggunakan RMSE berdasarkan nilai yang dihasilkan oleh PCA pada MATLAB dapat menggunakan beberapa perintah dan fungsi, antara lain :

1. Persiapan Data

Data yang digunakan adalah nilai covarian matrix yang menghasilkan variabel dominan dan juga covarian matrix dari 11 variabel bebas sebagai input dan variabel Load Density sebagai output dengan menggunakan fungsi sebagai berikut :

```
X = [cov1 cov2 cov3 cov4 cov5 cov6 cov7 cov8 cov9 cov10 cov11];
```

Nilai x tersebut adalah nilai 11 variabel bebas, sedangkan untuk x variabel dominan hanya mengambil nilai covarian matrix yang sudah terekstaksi

```
Y = [cov12];
```

Dan nilai Y tersebut adalah nilai output aktual pada load density

2. Menghitung nilai prediksi

Menghitung korelasi variabel bebas terhadap variabel Y dapat menggunakan rumus yang ada pada persamaan (2.13) dengan perintah dan fungsi pada MATLAB sebagai berikut :

$$YP = 4.6464e-05 * X1 + 1.8025 * X2 + 1.7968 * X3 + 1.8217 * X4 + 1.7648 * X5 + 1.8008 * X6 + 0.031029 * X7 + 0.0012487 * X8 + 0.00042536 * X9 + 0.0031182 * X10 + 0.00065579 * X11 + 3.0839e-14$$

3. Evaluasi pemodelan menggunakan nilai RMSE

RMSE mengukur rata-rata dari selisih antara nilai aktual dan nilai prediksi yang dikuadratkan, kemudian diambil akar kuadratnya. Maka dapat menggunakan fungsi dan perintah sebagai berikut :

```
error = Y-YP;
```

```
squarederror = error.^2;
```

```
mse = mean(squarederror);
```

```
rmse = sqrt(mse);
```

Semakin kecil nilai RMSE yang dihasilkan maka semakin baik juga pemodelan yang dilakukan.

3. HASIL DAN PEMBAHASAN

Faktor beban, faktor kapasitas dan faktor permintaan dari tahun 2013 hingga tahun 2021 mengalami ketidakstabilan kenaikan maupun penurunan persentase sementara daya yang tersambung dari tahun 2013 hingga tahun 2021 mengalami peningkatan yang signifikan. Peningkatan daya yang tersambung dapat disebabkan oleh pertumbuhan ekonomi, peningkatan aktivitas industri, dan perkembangan infrastruktur di wilayah tersebut. Ini dapat mengakibatkan peningkatan beban listrik karena lebih banyak bisnis, rumah tangga, dan industri memerlukan akses listrik untuk beroperasi.

Berdasarkan data-data yang diperoleh melalui observasi di PT PLN(PERSERO) UID Banten, faktor beban, faktor kapasitas dan faktor permintaan dari tahun 2013 hingga tahun 2021 mengalami ketidakstabilan kenaikan maupun penurunan persentase sementara daya yang tersambung dari tahun 2013 hingga tahun 2021 mengalami peningkatan yang signifikan. Peningkatan daya yang tersambung dapat disebabkan oleh pertumbuhan ekonomi, peningkatan aktivitas industri, dan perkembangan infrastruktur di wilayah tersebut. Ini dapat mengakibatkan peningkatan beban listrik karena lebih banyak bisnis, rumah tangga, dan industri memerlukan akses listrik untuk beroperasi. Profiling beban listrik dapat menunjukkan peningkatan beban di berbagai jam sehari dan hari dalam seminggu, mencerminkan aktivitas penggunaan listrik yang lebih tinggi. Beban puncak harian di berbagai penghantar memiliki pola yang sama yaitu pada saat siang hari angka beban puncak lebih tinggi daripada siang hari, hal ini disebabkan oleh aktifitas rumah tangga yang lebih aktif pada siang hari terutama penggunaan alat elektronik, juga kegiatan industri dan komersial yang memiliki jam kerja di siang hari.

Maka dari itu Variabel variabel yang dibutuhkan untuk profiling beban listrik tahun 2021 di wilayah yang terhubung dengan Area Jaringan Tangerang diantaranya, Jumlah Rumah tangga, luas wilayah, penggunaan lahan di sektor perumahan, industri, bisnis dan sosial, Produk Domestik Regional Bruto, Beban yang terpakai di sektor perumahan, industri, bisnis dan sosial serta kerapatan beban di wilayah yang terhubung dengan AJ Tangerang.

Data yang telah di peroleh melalui observasi di PT PLN (PERSERO) Unit Induk Distribusi Tangerang dan Badan Pusat Statistik Tangerang terdiri 4(empat) Kota madya yaitu Jakarta Barat, Kota Tangerang, Kabupaten Tangerang dan Bogor, 119 kelurahan serta 11 Variabel yang mempengaruhi kerapatan beban diantaranya, jumlah rumah tangga, luas wilayah, penggunaan lahan (land use) di sektor perumahan, industri, bisnis dan sosial, PDRB, Pembebanan di sektor perumahan, industri, bisnis dan sosial. Serta kerapatan beban (Load Density) sebagai output.

Dari variabel variabel yang mempengaruhi kerapatan beban akan di reduksi menjadi variabel yang paling dominan mempengaruhi kerapatan beban tersebut menggunakan metode Principal Component Analysis.

3.1. Pengolahan Data

Tahap pertama dalam penelitian ini adalah pengolahan data. Sejumlah 119 data sampling tiap kelurahan area Jakarta Barat, Tangerang, Kabupaten Tangerang dan Bogor digunakan dengan variabel proses adalah 12 data. Variabel proses yang dimaksud diantaranya adalah data jumlah rumah tangga, luas wilayah, penggunaan lahan dalam hektare (Ha) untuk sektor perumahan, industri, bisnis dan sosial, produk domestik regional bruto (PDRB), beban dalam

kilo Watt (kW) untuk sektor perumahan, industri, bisnis dan sosial serta terakhir adalah *load density* (kW/Ha).

3.2 Analisis Korelasi

Analisis korelasi pada penelitian ini digunakan untuk mengetahui variabel input yang paling berpengaruh dalam menentukan *load density*. Koefisien korelasi didasarkan pada pengamatan tiap variabel input terhadap variabel output, persamaan korelasi yang digunakan pada penelitian ini terdapat pada persamaan.

Dengan r adalah koefisien korelasi, x adalah variabel input, dan y adalah variabel output. Berikut adalah tabel korelasi tiap variabel meliputi data jumlah rumah tangga, luas wilayah, penggunaan lahan dalam hekta are (Ha) untuk sektor perumahan, industri, bisnis dan sosial, produk domestik regional bruto (PDRB), beban dalam kilo Watt (kW) untuk sektor perumahan, industri, bisnis dan sosial terhadap variabel output berupa *load density* (kW/Ha). Berikut merupakan perintah dan fungsi yang digunakan pada aplikasi MATLAB untuk mendapatkan koefisien korelasi :

```
% analisis korelasi
R1 = corr(d1,d12);
R2 = corr(d2,d12);
R3 = corr(d3,d12);
R4 = corr(d4,d12);
R5 = corr(d5,d12);
R6 = corr(d6,d12);
R7 = corr(d7,d12);
R8= corr(d8,d12);
R9 = corr(d9,d12);
R10 = corr(d10,d12);
R11 = corr(d11,d12);
```

Parameter luas wilayah, land use sektor industri dan land use sektor bisnis memiliki koefisien korelasi definit negatif dengan nilai relatif kecil yaitu dikisaran -0.2. Dapat diartikan hubungan ketiga variabel input tersebut berbanding terbalik dengan variabel *load density* dan berpengaruh kecil terhadap perubahan *load density*. Untuk parameter jumlah rumah tangga, land use sektor perumahan, land use sektor sosial, PDRB, Beban sektor perumahan, industri, bisnis dan sosial memiliki nilai koefisien korelasi definit positif atau perubahan kenaikan nilai input berbanding lurus dengan parameter output. Parameter input yang dominan terhadap hasil *load density* adalah produk domestik regional bruto (PDRB) dengan nilai $r = 0.8292$, beban sektor perumahan dengan $r = 0.8203$, Jumlah Rumah Tangga dengan nilai $r = 0,7413$ dan beban sektor bisnis dengan $r = 0.6471$. Untuk variabel dengan pengaruh terkecil adalah land use sektor sosial dengan $r = 0.0839$.

3.3 Reduksi Variable menggunakan Principal Component Analysis dan Kmeans Clustering

3.3.1 Reduksi Data dengan PCA

Setelah nilai simpangan masing – masing data diketahui, maka ekstraksi nilai eigen value (*score*). Dimana ke 12 variabel di reduksi menggunakan Principal Component Analysis dan di kelompokkan menjadi 4 variabel dominan berdasarkan nilai koefisien korelasi yang lebih dari 0,6 dengan menggunakan Kmeans Clustering.

```
[idx c sumd]= kmeans(dataset,4);
[coeff,score,latent,tsquared,expalined,mu]= pca(dataset);
pc1 = score(:,1);
pc2 = score(:,2);
figure;
gscatter(pc1,pc2,idx);
```

Berdasarkan script tersebut diketahui nilai principal component 1 dan nilai principal component 2 yang sudah diperintahkan melaui syntax `pca(dataset)` berdasarkan fungsi “score”. skor PCA adalah hasil dari proyeksi data ke dalam ruang fitur baru yang terbentuk oleh komponen utama. Lebih spesifik, skor PCA adalah hasil dari perkalian matriks antara matriks data yang telah distandarisasi (di mana setiap variabel telah dikurangi mean dan dibagi oleh standar deviasi) dengan matriks dari eigenvectors (komponen utama) yang dipilih dari analisis PCA. Dimana setiap baris mewakili skor PCA untuk satu observasi dalam ruang fitur baru yang terdefinisi oleh komponen utama

3.3.2 Pengelompokan Reduksi Data menggunakan Kmeans Clustering

Setelah ke 12 Variabel tersebut tereduksi oleh PCA dan dari nilai latent kita sudah dapat mengetahui variabel mana saja yang mempengaruhi variabel laiannya, selanjutnya kmeans dengan nilai `c(coeff)` nya mengelompokan menjadi 4 variabel utama yang mempengaruhi beban.

Tabel 1. Pengelompokan Reduksi 12 Variabel menjadi 4 Variabel Utama

Pengelompokan Reduksi 12 Variabel menjadi 4 Variabel Utama												
	RT	LW	LUPR	LUID	LUBS	LUSO	PDRB	BPR	BID	BBS	BSO	LD
PDRB	3,74	3,23	1,29	6,62	6,38	6,37	3,26	6,00	3,75	2,19	3,04	4,64
RT	2,26	6,17	4,00	7,89	8,91	4,85	2,28	3,43	7,76	6,34	9,86	3,51
B_BS	1,32	5,56	3,37	6,54	6,69	8,59	2,04	5,42	6,43	4,76	1,74	3,37
B_PR	2,24	3,41	1,44	1,88	1,25	1,38	1,01	5,33	9,46	2,39	7,68	4,30

Dari table tersebut dapat diketahui bahwa nilai koefisien variabel PDRB terhadap *Load Density* memiliki nilai yang paling tinggi yaitu sebesar 4,64 , diikuti dengan variabel beban perumahan dengan nilai koefisien 4,30 , variabel rumah tangga dengan koefisien 3,51 dan beban bisnis dengan koefisien 3,37. Artinya semakin tinggi nilai koefisien nya maka semakin tinggi variabilitas tersebut mempengaruhi kerapatan bebannya.

3.4 Pemodelan Regresi Linear Berganda

Berikut adalah script pemodelan regresi berganda yang digunakan pada penelitian ini:

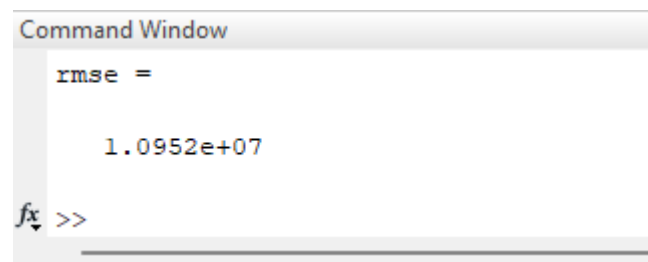
```
%pemodelan regresi linear berganda
X = [cov1 cov7 cov8 cov10]
X1 = cov1;
X2 = cov7;
X3 = cov8;
X4 = cov10;
Y = [cov12];
YP = 4.6464e-05*X1 - 1.8025*X2 + 1.7968*X3 + 1.8217*X4;
error = Y-YP;
squareerror = error.^2;
```

```
mse = mean(squarederror);rmse = sqrt(mse)
```

Hasil pemodelan regresi linear berganda berdasarkan persamaan (2.13) regresi sebagai berikut:

$$YP = 4.6464e-05 \cdot X_1 + 1.8025 \cdot X_2 + 1.7968 \cdot X_3 + 1.8217 \cdot X_4;$$

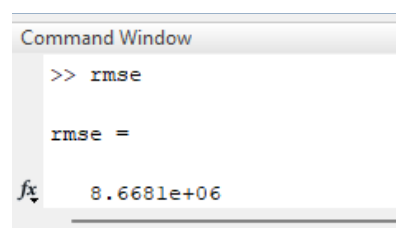
Hasil nilai output prediksi *load density* digunakan untuk mendapatkan akurasi performa pemodelan regresi linear berganda yang digunakan pada penelitian ini. Akurasi performa model dievaluasi dengan menggunakan parameter RMSE (*Root Mean Squared Error*) atau nilai akar rata – rata kesalahan data prediksi terhadap data asli. Parameter RMSE yang dihasilkan dari pemodelan regresi pada penelitian ini adalah 1.09.



Gambar 2. Nilai RMSE yang di hasilkan menggunakan 4 variabel komponen utama pada tampilan workspace MATLAB

Perbandingan pemodelan regresi berganda dengan menggunakan 12 variabel untuk profile beban

```
X = [cov1 cov2 cov3 cov4 cov5 cov6 cov7 cov8 cov9 cov10 cov11];  
X1 = cov1;  
X2 = cov2;  
X3 = cov3;  
X4 = cov4;  
X5 = cov5;  
X6 = cov6;  
X7 = cov7;  
X8 = cov8;  
X9 = cov9;  
X10 = cov10;  
X11 = cov11;  
Y = [cov12];  
YP = 4.6464e-05*X1 + 1.8025*X2 + 1.7968*X3 + 1.8217*X4 + 1.7648*X5 + 1.8008*X6 +  
0.031029*X7 + 0.0012487*X8 + 0.00042536*X9 + 0.0031182*X10 + 0.00065579*X11 + 3.0839e-14;  
error = Y-YP;  
squarederror = error.^2;  
mse = mean(squarederror);  
rmse = sqrt(mse);
```



Gambar 3. Parameter RMSE yang dihasilkan oleh 12 variabel

Tabel 2. Perbandingan Evaluasi Pemodelan Regresi Linear Berganda

Jenis Variabel	Persamaan	Hasil	Keterangan
Variabel PCA	$YP = 4.6464e-05 \cdot X1 + 1.8025 \cdot X2 + 1.7968 \cdot X3 + 1.8217 \cdot X4;$	1.09	Baik
Variabel Bebas	$YP = 4.6464e-0.5 \cdot X1 + 1.8025 \cdot X2 + 1.7968 \cdot X3 + 1.8217 \cdot X4 + 1.7648 \cdot X5 + 1.8008 \cdot X6 + 0.031029 \cdot X7 + 0.0012487 \cdot X8 + 0.00042536 \cdot X9 + 0.0031182 \cdot X10 + 0.00065579 \cdot X11 + 3.0839e-14;$	8,66	Kurang Baik

Parameter RMSE yang dihasilkan oleh 12 variabel adalah 8.66 sedangkan parameter RMSE yang dihasilkan oleh 4 variabel komponen utama adalah 1.09. Nilai RMSE yang baik adalah Ketika mendekati 0, maka dapat disimpulkan bahwa dengan pemodelan menggunakan Principal Component Analysis pemodelan yang baik.

4. KESIMPULAN

Penerapan Analisa multivariate pada profile beban suatu wilayah menggunakan analisa korelasi untuk mengetahui koefisien korelasi yang paling kuat antar variabel, *Principal Component Analysis* digunakan untuk mereduksi beberapa variabel menjadi 4 variabel utama, dan Analisa regresi digunakan untuk evaluasi pemodelan.

Analisis korelasi digunakan untuk mengetahui pengaruh variabel input terhadap *load density*. Variabel input yang paling berpengaruh terhadap penentuan *load density* adalah produk domestik regional bruto (PDRB) dengan nilai $r = 0.829$ dan disusul beban sektor perumahan dengan $r = 0.820$ dan variabel dengan pengaruh terkecil adalah land use sektor sosial dengan $r = 0.083$. Parameter luas wilayah, land use sektor industri dan land use sektor bisnis memiliki distribusi perubahan yang berbanding terbalik dengan *load density*.

Untuk menentukan variabel dominan menggunakan nilai PCA maka harus didapatkan nilai koefisien utama pada variabel nya, semakin besar nilai koefisien utama pada variabel tersebut maka makin tinggi juga pengaruh variabilitasnya. Untuk memvisualisasikan nya dapat menggunakan metode kmeans clustering dan mengelompokannya menjadi 4 variabel berdasarkan hasil koefisien korelasi yang dihasilkan.

Akurasi dalam pemodelan menggunakan PCA dikategorikan dengan evaluasi model yang baik dikarenakan hasil RMSE yang mendekati 0 yaitu dengan nilai 1.09 sedangkan untuk yang tidak menggunakan PCA dikategorikan dengan evaluasi yang kurang baik dengan nilai RMSE 8,66.

Dari hasil profiling yang didapatkan maka dapat dilakukan beberapa *treatment* lanjutan yang dapat dilakukan diantaranya, segmentasi pengguna, penyusunan tarif, penyusunan kebijakan serta monitoring dan evaluasi

Hasil dari Reduksi PCA dapat membantu dalam pembuatan master plan jaringan distribusi, informasi tentang PDRB dan beban listrik rumah tangga, bisnis, dan perumahan dari analisis PCA dapat digunakan dalam perencanaan kapasitas jaringan distribusi. Data ini dapat membantu dalam menentukan kapasitas yang diperlukan untuk memenuhi kebutuhan listrik yang berkembang di setiap segmen wilayah dan konsumen.

REFERENSI

- [1] Senen, A., Widyastuti, C., Handayani, O., & Putera, P. (2021). Development of Micro-Spatial Electricity Load Forecasting Methodology Using Multivariate Analysis for

- Dynamic Area in Tangerang, Indonesia. *Pertanika Journal of Science & Technology*, 29(4).
- [2] Anggaini, D., Senen, A., & Putri, T. W. O. (2022). Analisis Multivariat untuk Penentuan Pola Kebutuhan Beban Listrik secara Spasial Menggunakan Metode Geographically Weighted Regression. *PROSIDING-SNEKTI*, 3.
- [3] Adityawan, R., & Umar, S. T. (2020). *Analisis Faktor Beban dan Kebutuhan Daya Listrik di Wilayah PLN APJ Madiun dengan Objek Pelanggan Bisnis*. Doctoral Dissertation. Universitas Muhammadiyah Surakarta.
- [4] Damara, D. B. O., & Yasa, I. N. M. (2019). Analisis Faktor-Faktor Yang Mempengaruhi Permintaan Energi Listrik di Provinsi Bali. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 8(1), 211-238.
- [5] Damara, D. B. O., & Yasa, I. N. M. (2019). Analisis faktor-faktor yang mempengaruhi permintaan energi listrik di Provinsi Bali. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 8(1), 211-238.
- [6] Fernandes, K. C., Sardinha, R., Rebelo, S., & Singh, R. (2019, April). Electric load analysis and forecasting using artificial neural networks. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 1274-1278.
- [7] Granato, D. (2018). Use Of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) for Multivariate Association Between Bioactive Compounds and Functional Properties in Foods: A Critical Perspective. *Trends in Food Science & Technology*, 72, 83-90.
- [8] Hayqal, H. H. Q., Soesanto, O., & Sukmawaty, Y. (2022). K-Means Clustering dan Principal Component Analysis (PCA) Dalam Radial Basis Function Neural Network (RBFNN) Untuk Klasifikasi Data Multivariat. *Journal of Mathematics: Theory and Applications*, 1-7.
- [9] Johannesen, N. J., Kolhe, M. L., & Goodwin, M. (2021). Comparing recurrent neural networks using principal component analysis for electrical load predictions. In *2021 6th International Conference on Smart and Sustainable Technologies (SpliTech)* (pp. 1-6). IEEE.
- [10] Kurniawati, U. F., Handayani, K. D. M. E., Nurlaela, S., Idajati, H., Firmansyah, F., Pratomoadojo, N. A., & Septriadi, R. S. (2020). Pengolahan Data Berbasis Sistem Informasi Geografis (Sig) Untuk Kebutuhan Penyusunan Profil di Kecamatan Sukolilo. *Sewagati*, 4(3), 190-196.
- [11] Leny, E. M., & Haryudo, S. I. (2019). Sistem Current Limiter Dan Monitoring Arus Serta Tegangan Menggunakan Sms Untuk Proteksi Pada Penggunaan Beban Rumah Tangga. *Jurnal Teknik Elektro*, 8(1).
- [12] Nisa, S. K., Sawitri, D. R., & Suprijono, H. *Identifikasi Karakteristik Sinyal Arus Motor Induksi Tiga Fase Akibat Ketidakseimbangan Tegangan* (Unbalance Voltage).
- [13] Rais, M., Goejantoro, R., & Prangga, S. (2021). Optimalisasi K-Means Cluster dengan Principal Component Analysis pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Tingkat Pengangguran Terbuka. *EKSPONENSIAL*, 12(2), 129-136.
- [14] Rizki, R. (2022). Pengaplikasian matlab pada perhitungan matriks. *Papanda Journal of Mathematics and Science Research*, 1(2), 81-93.
- [15] Wustqa, D. U., Listyani, E., Subekti, R., Kusumawati, R., Susanti, M., & Kismiantini, K. (2018). Analisis Data Multivariat dengan Program R. *Jurnal Pengabdian Masyarakat MIPA Dan Pendidikan MIPA*, 2(2), 83-86.
- [16] Yuniarti, E., Saputra, A. A., Malulu, A., & Santoso, B. (2020). Profil Beban Penyulang Pandu Pada Gardu Induk Talang Kelapa 150 kV/20 kV. SNPPM-2 (Seminar Nasional Penelitian dan Pengabdian kepada Masyarakat) Tahun 2020.

- [17] Faruque, M.O., & Mohd Radzi, M.A. (2019). Load Profiling and Analysis for Residential, Commercial, and Industrial Sector: A Review. *Energies*, 12(8), 1444.
- [18] Jain, P., Tyagi, B., Kumar, A., & Tyagi, S. (2019). Load Profile and Its Classification: A Review. *Renewable and Sustainable Energy Reviews*, 112, 354-366.
- [19] Sari, F., & Firmansyah, A. (2022). Analisis Multivariate untuk Mendeteksi Pola Konsumsi Listrik Rumah Tangga di Kota Bandung. *Jurnal Teknik Elektro dan Informatika*, 9(2), 89-101.
- [20] Nugroho, B., & Raharja, S. (2022). K-Means Clustering untuk Identifikasi Pola Pemakaian Listrik Rumah Tangga di Perkotaan Indonesia. *Jurnal Teknologi Informasi dan Komunikasi*, 9(2), 112-125