

Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Jumlah Kasus Penyakit Menular Menggunakan Analisis Cluster Metode Ward

Anisa Janatil Arif, Meira Parma Dewi

Departemen Matematika,, Universitas Negeri Padang (UNP)

Article Info

Article history:

Received August 08, 2025

Revised October 07, 2025

Accepted October 13, 2025

Keywords:

Infectious Disease

Hierarchy

Dunn Index

Kata Kunci:

Penyakit Menular

Hirarki

Dunn Index

ABSTRACT

This study aims to cluster regions in West Sumatra Province using hierarchical cluster analysis with the Ward method, based on the number of cases of infectious diseases, including tuberculosis (TB), pneumonia, AIDS, and dengue fever (DF). The data used is secondary data obtained from the West Sumatra Provincial Statistics Agency (BPS) in 2024. The clustering process was carried out with a variation in the number of clusters between two and six groups. Based on the calculation results, the highest Dunn Index value of 0.6344 indicates that the best clustering is in three clusters. The areas included in the third cluster are identified as regions with a relatively high rate of infectious disease spread, thus requiring more attention in control efforts in West Sumatra Province.

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan daerah-daerah di Provinsi Sumatera Barat menggunakan analisis kluster hierarki dengan metode Ward, berdasarkan jumlah kasus penyakit menular yang meliputi Tuberkulosis (TBC), Pneumonia, AIDS, dan Demam Berdarah Dengue (DBD). Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Barat tahun 2024. Proses pengelompokan dilakukan dengan variasi jumlah kluster antara dua hingga enam kelompok. Berdasarkan hasil perhitungan, nilai Dunn Index tertinggi sebesar 0,6344 menunjukkan bahwa pengelompokan terbaik terdapat pada tiga kluster. Daerah-daerah yang termasuk dalam kluster ketiga diidentifikasi sebagai wilayah dengan tingkat penyebaran penyakit menular yang relatif tinggi, sehingga perlu mendapatkan perhatian lebih dalam upaya penanggulangan di Provinsi Sumatera Barat.

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



Penulis Korespondensi:

Anisa Janatil Arif

Departemen Matematika, Universitas Negeri Padang

Email: janatilanisa@gmail.com

1. PENDAHULUAN

Penyakit menular adalah jenis penyakit yang bisa berpindah dari satu orang ke orang lain, atau dari hewan ke manusia, melalui mikroorganisme seperti bakteri, virus, jamur, atau parasit yang menyebar lewat berbagai media[1].

Tuberkulosis (TBC) merupakan penyakit menular kronis yang disebabkan oleh bakteri *Mycobacterium tuberculosis*. Bakteri ini umumnya menyerang jaringan paru-paru dan menyebabkan TB paru, namun juga dapat menginfeksi organ lain di luar paru (TB ekstra paru) seperti pleura, kelenjar getah bening, tulang, dan organ tubuh lainnya [3]. Pneumonia merupakan penyakit infeksi pada paru-paru yang masih menjadi salah satu penyebab utama kesakitan dan kematian pada bayi serta anak-anak di seluruh dunia. Penyakit ini dapat disebabkan oleh virus, bakteri, jamur, atau kombinasi dari ketiganya, yang menimbulkan peradangan dan penumpukan cairan pada jaringan paru (parenkim paru). Gejala yang umum muncul pada penderita pneumonia antara lain batuk, napas cepat, demam (atau tanpa demam), batuk kering maupun berdahak, sesak napas, nyeri dada, dan rasa lelah berlebihan [4]. Penyakit Demam Berdarah Dengue (DBD), yang juga dikenal sebagai *Dengue Haemorrhagic Fever* (DHF), merupakan infeksi yang disebabkan oleh virus dengue dan menyebar melalui gigitan nyamuk *Aedes aegypti* maupun *Aedes albopictus*. Penyakit ini termasuk salah satu isu kesehatan masyarakat yang signifikan di tingkat global, terutama di wilayah dengan kondisi sosial ekonomi rendah [5]. AIDS (*Acquired Immune Deficiency Syndrome*) merupakan kumpulan gejala penyakit yang muncul akibat menurunnya sistem kekebalan tubuh. Penyakit ini disebabkan oleh virus HIV yang terdapat dalam cairan tubuh, terutama darah, cairan sperma, cairan vagina, dan air susu ibu. Karena lemahnya sistem kekebalan tubuh, infeksi yang biasanya ringan bagi orang sehat dapat menjadi berbahaya bahkan fatal bagi penderita AIDS [6].

Tantangan dan permasalahan dalam pembangunan kesehatan kini semakin kompleks dan sulit diprediksi, terutama dengan munculnya berbagai penyakit menular. Oleh karena itu, Dinas Kesehatan Provinsi Sumatera Barat menetapkan upaya pencegahan dan pengendalian penyakit menular sebagai salah satu prioritas utama dalam meningkatkan derajat kesehatan masyarakat [7]. Untuk menentukan prioritas tersebut, terlebih dahulu dilakukan pengelompokan daerah kabupaten/kota agar dapat diketahui kelompok-kelompok wilayah yang memiliki kesamaan dalam angka kejadian penyakit menular. Berdasarkan permasalahan yang telah dijelaskan sebelumnya, penelitian ini melakukan pengelompokan terhadap 19 kabupaten/kota di Provinsi Sumatera Barat berdasarkan jumlah kasus penyakit menular, yaitu Tuberkulosis (TBC), Pneumonia, Demam Berdarah Dengue (DBD), dan AIDS, dengan menggunakan metode analisis kluster hierarki. Keempat jenis penyakit ini dipilih karena berdasarkan data yang tersedia, penyakit-penyakit tersebut menunjukkan tingkat kejadian yang relatif tinggi dan memerlukan perhatian khusus dalam upaya pencegahan serta pengendalian.

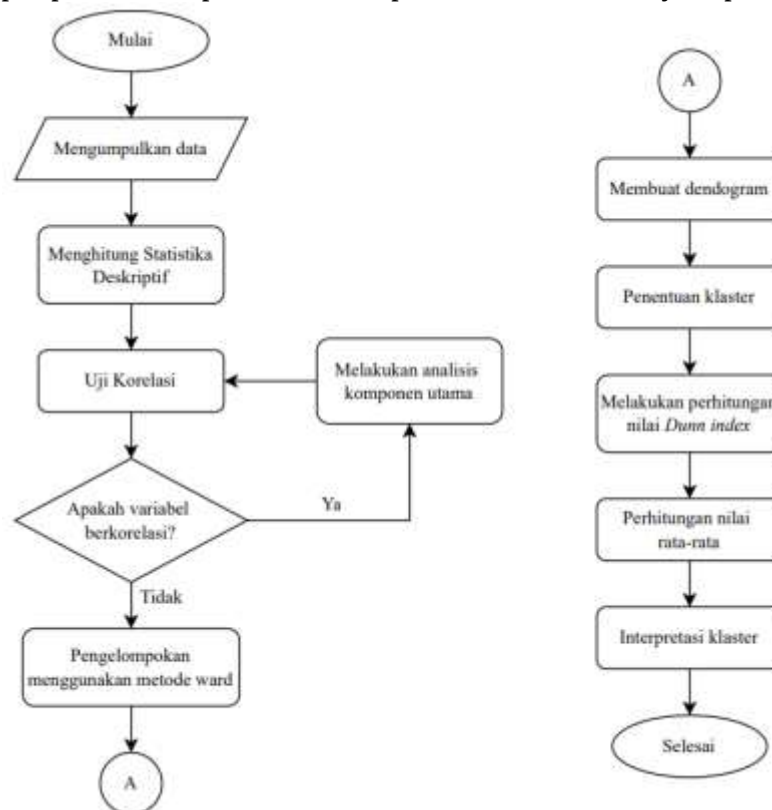
Data yang digunakan dalam penelitian ini merupakan data jumlah kasus penyakit menular di Provinsi Sumatera Barat, yang meliputi kasus Tuberkulosis (TBC), Pneumonia, AIDS, dan Demam Berdarah Dengue (DBD) pada tingkat kabupaten/kota. Data tersebut diperoleh dari publikasi *Sumatera Barat dalam Angka 2025* yang diterbitkan oleh publikasi BPS Sumatera Barat dalam angka 2025 [8] dan Dinas Kesehatan [9]. Dalam analisis data, metode yang digunakan untuk mengelompokkan individu ke dalam beberapa kelompok berdasarkan variabel-variabel tertentu dikenal sebagai analisis kluster. Pada penelitian ini, jenis analisis yang digunakan adalah analisis kluster hierarki, yang bertujuan untuk mengetahui struktur pengelompokan dari setiap objek yang diteliti. Hasil dari analisis ini disajikan dalam bentuk hierarki atau bertingkat, sehingga hubungan antar objek dapat terlihat secara berjenjang [10]. Metode yang digunakan peneliti adalah metode ward, Metode Ward bekerja dengan meminimalkan total varians di dalam setiap kluster. Pada setiap tahap pengelompokan, dua kluster yang digabung adalah kluster yang menghasilkan peningkatan paling kecil terhadap total varians, berdasarkan jarak terkecil antara pusat (centroid) kedua kluster tersebut. Metode Ward ini dipilih karena menurut [11] metode ini mampu menghasilkan pengelompokan yang lebih baik dibandingkan dengan metode kluster hierarki lainnya. Penentuan jumlah kluster terbaik dilakukan dengan menggunakan kriteria Dunn Index, yang dipilih karena proses perhitungannya relatif sederhana dan mudah diterapkan. Selain itu, menurut [12] *Dunn index* mampu mengidentifikasi hasil pengelompokan yang optimal pada tiga dataset dengan karakteristik yang berbeda, sementara metode lainnya tidak mampu menghasilkan pengelompokan yang tepat tersebut [13].

2. METODE

Penelitian ini termasuk dalam jenis penelitian terapan yang diawali dengan kajian teori, kemudian dilanjutkan dengan proses pengumpulan data. Data yang digunakan merupakan data sekunder yang bersumber dari publikasi Badan Pusat Statistik (BPS) Provinsi Sumatera Barat tahun 2025. Data tersebut memuat informasi mengenai jumlah kasus penyakit menular, yaitu Tuberkulosis (TBC), Pneumonia, AIDS, serta Demam Berdarah Dengue (DBD) di seluruh wilayah Provinsi Sumatera Barat pada tahun 2024. Dalam penelitian ini, proses analisis data dilakukan melalui sejumlah tahapan, yang dapat dijelaskan sebagai berikut.

1. Mengumpulkan data yang akan digunakan dalam analisis.
2. Menghitung statistika deskriptif untuk setiap variabel penelitian.
3. Memeriksa nilai koefisien korelasi antar variabel. Jika nilai korelasi antar variabel cukup tinggi, maka dilakukan analisis komponen utama (PCA).
4. Membentuk cluster dengan menggunakan metode *cluster hierarki* dengan pendekatan ward.
5. Menghitung nilai *Dunn index* untuk menentukan jumlah *cluster* terbaik.
6. Menghitung nilai rata-rata untuk setiap variabel jumlah kasus penyakit tuberkulosis, jumlah kasus penyakit pneumonia, jumlah kasus penyakit AIDS, dan jumlah kasus penyakit DBD pada *cluster* terbaik yang terbentuk.
7. Melakukan interpretasi terhadap hasil pengelompokan yang dihasilkan.

Struktur atau tahapan pelaksanaan penelitian ini dapat dilihat secara lebih jelas pada Gambar 1.



Gambar 1. Alur Penelitian

3. HASIL DAN PEMBAHASAN

3.1 Analisis Deskriptif

Analisis deskriptif dalam penelitian ini bertujuan untuk memberikan gambaran mengenai data yang digunakan dalam pengelompokan kasus penyakit menular di Provinsi Sumatera Barat tahun 2024. Data yang dianalisis merupakan data sekunder yang diperoleh dari publikasi Badan Pusat Statistik

(BPS) Provinsi Sumatera Barat. Data tersebut mencakup jumlah kasus penyakit menular di 19 wilayah administratif, yang terdiri dari 12 kabupaten dan 7 kota. Kabupaten yang termasuk dalam penelitian ini yaitu Kepulauan Mentawai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Padang Pariaman, Agam, Lima Puluh Kota, Pasaman, Solok Selatan, Dharmasraya, dan Pasaman Barat. Sementara itu, wilayah kota meliputi Kota Padang, Kota Solok, Kota Sawahlunto, Kota Padang Panjang, Kota Bukittinggi, Kota Payakumbuh, dan Kota Pariaman. Data yang digunakan dalam penelitian ini terdiri atas empat variabel, yaitu banyaknya kasus penyakit Tuberkulosis (X1), banyaknya kasus penyakit Pneumonia (X2), banyaknya kasus penyakit AIDS (X3), dan banyaknya kasus penyakit Demam Berdarah Dengue (DBD) (X4). Statistik deskriptif dari masing-masing variabel yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Statistika Deskriptif Setiap Variabel Penelitian

Variabel	Minimum	Maximum	Total	Rata-rata	Simpangan Baku
X ₁	654	25,418	104635	5507,1053	185,34949
X ₂	29	2298	8546	449,7895	525,62044
X ₃	46	542	2749	144,6842	111,57262
X ₄	2	2643	4308	224,4737	599,58304

3.2 Korelasi Antar Variabel

Nilai korelasi antar variabel:

$$R = \begin{bmatrix} 1 & 0,800 & 0,827 & 0,816 \\ 0,800 & 1 & 0,835 & 0,840 \\ 0,827 & 0,835 & 1 & 0,830 \\ 0,816 & 0,840 & 0,830 & 1 \end{bmatrix}$$

Berdasarkan matriks korelasi (R) dan merujuk pada tabel 2, diketahui bahwa beberapa variabel memiliki korelasi yang cukup kuat hingga sangat kuat. Seluruh variabel menunjukkan nilai korelasi dengan nilai absolut lebih besar dari 0,70, yang menandakan adanya hubungan yang tinggi antarvariabel. Oleh karena itu, sebelum dilakukan proses pengelompokan (klasterisasi) terhadap objek penelitian, perlu dilakukan analisis komponen utama (Principal Component Analysis) untuk memperoleh variabel-variabel baru yang tidak saling berkorelasi.

3.3 Analisis Komponen Utama

Analisis komponen utama dilakukan untuk memperoleh komponen utama beserta nilai skornya pada setiap variabel, sehingga antarvariabel tidak lagi saling berkorelasi. Nilai eigen serta proporsi keragaman kumulatif dari masing-masing komponen utama dapat dilihat pada Tabel 2.

Tabel 2. Nilai eigen setiap komponen

Komponen Utama	λ_i	Proporsi λ_i (%)	Kumulatif (%)
Y ₁	3,474	86,848	86,848
Y ₂	0,205	5,124	91,972
Y ₃	0,168	4,206	96,178
Y ₄	0,153	3,822	100,000
Total	4	100	

Nilai koefisien komponen utama setiap variabel dapat dilihat pada Tabel 3.

Tabel 3. Nilai Koefisien Komponen Utama

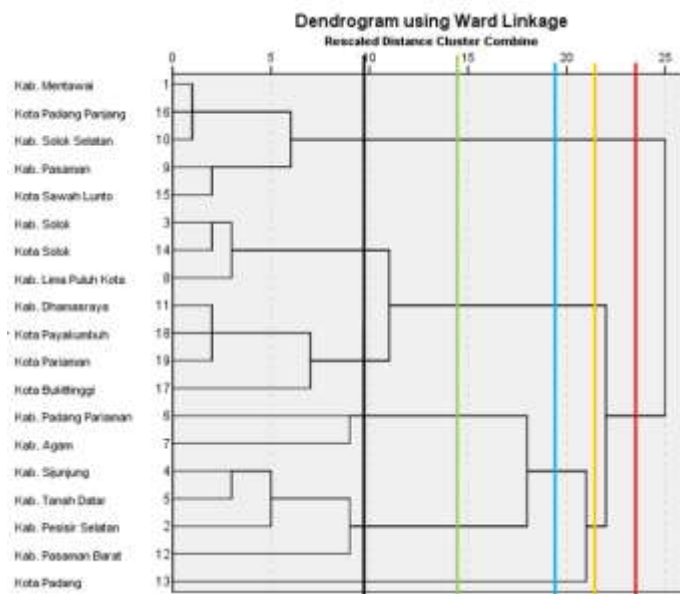
Variabel	Komponen Utama			
	Y ₁	Y ₂	Y ₃	Y ₄
Tuberkulosis (Z ₁)	0,923	0,355	0,075	0,128
Pneumonia (Z ₂)	0,932	-0,252	-0,056	0,254
AIDS (Z ₃)	0,937	0,024	-0,291	-0,192
DBD (Z ₄)	0,935	-0,123	0,237	-0,188

Tabel 3 digunakan untuk membentuk skor komponen utama setiap pengamatan melalui persamaan:

$$\begin{aligned} Y_1 &= 0,923Z_1 + 0,932Z_2 + 0,937Z_3 + 0,935Z_4 \\ Y_2 &= 0,355Z_1 - 0,252Z_2 + 0,024Z_3 - 0,123Z_4 \\ Y_3 &= 0,075Z_1 - 0,056Z_2 - 0,291Z_3 + 0,273Z_4 \\ Y_4 &= 0,128Z_1 + 0,254Z_2 - 0,192Z_3 - 0,188Z_4 \end{aligned}$$

3.4 Hasil Analisis Cluster

Analisis klasterisasi kabupaten/kota pada penelitian ini dilakukan menggunakan metode Ward dengan bantuan perangkat lunak SPSS untuk memperoleh hasil pengelompokan. Hasil pengelompokan tersebut dapat dilihat pada Gambar 2.



Gambar 2. Dendrogram Pengklasteran Daerah

Penentuan jumlah klaster terbaik dilakukan dengan menggunakan nilai Dunn Index. Pada Gambar 2 ditampilkan potongan dendrogram yang diberi beberapa garis penanda, di mana garis merah menunjukkan pemotongan untuk dua klaster, garis kuning untuk tiga klaster, garis biru untuk empat klaster, garis hijau untuk lima klaster, dan garis hitam untuk enam klaster. Hasil pengelompokan dengan enam klaster tersebut disajikan pada Tabel 4.

Tabel 4. Hasil pengelompokan menjadi 6 klaster

Klaster	Banyak Anggota	Nama Anggota
1	5	Kab.Mentawai, Kota Padang Panjang, Kab.Solok Selatan, Kab.Pasaman, Kota sawah Lunto.
2	3	Kab.Solok, Kota Solok, Kab.Lima Puluh Kota.
3	4	Kab.Dhamasraya, Kota Payakumbuh, Kota Pariaman, Kota Bukittinggi.
4	2	Kab.Padang Pariaman, Kab.Agam.
5	4	Kab.Sijunjung, Kab.Tanah Datar, Kab.Pesisir Selatan, Kab.Pasaman Barat.
6	1	Kota Padang

3.5 Penentuan Cluster Terbaik

Pada tabel 5 disajikan nilai *Dunn Index* dari hasil pembentukan kluster yang telah dilakukan sebelumnya.

Tabel 5. Nilai Dun Index	
Banyak Kluster	<i>Dunn Index</i>
2	0,2436
3	0,6344
4	0,1712
5	0,3170
6	0,2782

Berdasarkan Tabel 5, dapat dilihat bahwa nilai *Dunn Index* terbesar diperoleh pada pengelompokan menjadi 3 kluster dengan nilai sebesar 0,6344. Dengan demikian, jumlah kluster terbaik yang digunakan dalam penelitian ini adalah sebanyak 3 kluster.

3.6 Variabel Pembeda

Tabel 6 menyajikan nilai rata-rata dan simpangan baku dari masing-masing variabel pada kluster terbaik yang telah terbentuk.

Tabel 6. Rata-rata dan Simpangan Baku			
Variabel	Cluster	Rata-rata	Simpangan Baku
Tuberkulosis	1	1847	1061,093
	2	4018,142	1840,400
	3	9610,43	7818,895
Pneumonia	1	110,800	90,358
	2	241,428	109,837
	3	900,29	654,324
AIDS	1	15,800	14,686
	2	159,142	194,288
	3	438,86	972,482
DBD	1	141	36,523
	2	77	27,166
	3	215	1061,093

4. KESIMPULAN

Berdasarkan hasil pengelompokan yang telah dijelaskan pada bagian hasil dan pembahasan, diperoleh tiga kelompok daerah. Kelompok pertama terdiri atas lima kabupaten/kota, yaitu Kabupaten Kepulauan Mentawai, Kota Padang Panjang, Kabupaten Solok Selatan, Kabupaten Pasaman, dan Kota Sawahlunto. Kelompok kedua mencakup tujuh daerah, yaitu Kabupaten Solok, Kota Solok, Kabupaten Lima Puluh Kota, Kabupaten Dharmasraya, Kota Payakumbuh, Kota Pariaman, dan Kota Bukittinggi. Sementara itu, kelompok ketiga juga terdiri dari tujuh kabupaten/kota, yaitu Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten Sijunjung, Kabupaten Tanah Datar, Kabupaten Pesisir Selatan, Kabupaten Pasaman Barat, dan Kota Padang.

Hasil uji validitas menggunakan *Dunn Index* menunjukkan bahwa jumlah kluster terbaik adalah tiga, karena memberikan nilai *Dunn Index* tertinggi. Daerah-daerah yang tergabung dalam kluster ketiga, yaitu Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten Sijunjung, Kabupaten Tanah

Datar, Kabupaten Pesisir Selatan, Kabupaten Pasaman Barat, dan Kota Padang, memiliki rata-rata jumlah kasus penyakit Tuberkulosis (TBC), Pneumonia, AIDS, dan Demam Berdarah Dengue (DBD) yang lebih tinggi dibandingkan dengan daerah pada klaster pertama dan kedua. Berdasarkan hasil tersebut, daerah-daerah dalam klaster ketiga dapat dikategorikan sebagai wilayah dengan tingkat penyebaran penyakit menular yang tinggi di Provinsi Sumatera Barat, sehingga perlu mendapatkan perhatian lebih dalam upaya penanggulangan penyakit.

REFERENSI

- [1] World Health Organization, "Pengertian Penyakit Menular".
- [2] A. Darmawan, "Epidemiologi Penyakit Menular Dan Penyakit Tidak Menular," *Jmj*, vol. 4, no. 2, pp. 195–202, 2016,.
- [3] D. K. Kemenkes, "Mengenal Tuberculosis," no. 16, 2024.
- [4] Kemenkes DitjenP2, "Pneumonia Menjadi Ancaman Kesakitan dan Kematian di Dunia".
- [5] R. Anliyanita, C. Anwar, and N. A. Fajar, "Effect of physical environment and community behavior on dengue hemorrhagic fever (DHF): A literature review," *Community Res. Epidemiol.*, vol. 3, no. 2, pp. 74–76, 2023.
- [6] A. S. Aresta and W. Jumaiyah, "Pengetahuan Dan Dukungan Keluarga Dengan Kepatuhan Dalam Menjalankan Pengobatan Antiretroviral (ARV) Pada Pasien HIV/AIDS," *Indones. J. Nurs. Pract.*, vol. 2, no. 1, pp. 51–61, 2019.
- [7] Kementerian Kesehatan Republik Indonesia, "Tuberkulosis," 2024, [Online].
- [8] P. B. S. O. BPS, Barat Sumatera Provinsi/Statistic-BPS, "BADAN PUSAT STATISTIK PROVINSI SUMATERA BARAT DALAM ANGKA 2025," vol. 55, p. 852, 2020.
- [9] Kementrian Kesehatan, "Satu Sehat KemenKes," 2024.
- [10] S. M. Purnamasari, "Analisis Kelompok (Cluster Analysis)," *Program*, no. 18209007, pp. 1–6, 2011, [Online].
- [11] M. Paramadina, S. Sudarmin, and M. K. Aidid, "Perbandingan Analisis Cluster Metode Average Linkage dan Metode Ward (Kasus: IPM Provinsi Sulawesi Selatan)," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 1, no. 2, p. 22, 2019, doi: 10.35580/variansiunm9357.
- [12] F. Kovács, C. Legány, and A. Babos, "Cluster Validity Measurement Techniques," *Proc. 6th Int. Symp. Hungarian Res. Comput. Intell.*, pp. 1–11, 2005.
- [13] N. Ilc, "Modified Dunn's cluster validity index based on graph theory," *Prz. Elektrotechniczny (Electrical Rev.)*, vol. 88, no. 2, pp. 126–131, 2012.