



Pemetaan Cuaca Berdasarkan Penyebaran Awan Menggunakan Analisis Korespondensi

Ni Komang Ayu Karisma Dewi, Ni Nyoman Swastika Sukma Suryani, Dewa Ayu Made Sri Wijayati Utami, Made Ayu Dwi Octavanny*

Matematika, Universitas Udayana, Jl. Kampus Unud, Jimbaran, Bali, Indonesia, 80361

*Penulis Korespondensi: octavanny@unud.ac.id

Abstract. *Weather and cloud cover are two fundamental atmospheric elements that are closely interconnected and play a vital role in understanding, classifying, and predicting atmospheric conditions. Accurate classification of weather conditions is essential not only for meteorological studies but also for practical applications in agriculture, aviation, transportation, and disaster management. Among the various statistical techniques available, correspondence analysis has proven to be an effective exploratory method for revealing associations among categorical variables by displaying their relationships in a two-dimensional graphical map. The primary objective of this study is to examine the relationship between different categories of weather conditions and levels of cloud cover using correspondence analysis. The data were sourced from the Kaggle platform, consisting of 13,200 observations with two key categorical variables: weather conditions (sunny, cloudy, rainy, and snowy) and cloud cover (clear, partly cloudy, cloudy, and overcast). This large dataset provides sufficient variability and reliability to capture the complex interdependence between the two variables. The findings reveal a statistically significant relationship between weather categories and cloud cover levels. The correspondence map visualization shows distinct patterns: clear weather is strongly associated with cloudless skies, while rainy and snowy weather conditions are predominantly linked to fully overcast skies. Cloudy and partly cloudy conditions occupy an intermediate position on the correspondence map, reflecting a transitional state between favorable and unfavorable weather. The first dimension of the map represents a spectrum from favorable weather conditions (sunny and clear) to adverse conditions (rainy and snowy), whereas the second dimension primarily reflects the intensity or density of cloud cover, ranging from thin to thick cloud layers. Overall, this study demonstrates the usefulness of correspondence analysis as an exploratory tool for meteorological data, offering valuable insights into the interrelation of weather and cloud cover categories.*

Keywords: *Atmospheric data visualization; Clouds; Correspondence; Weather mapping; Weather*

Abstrak. Cuaca dan tutupan awan merupakan dua elemen fundamental atmosfer yang saling berkaitan erat dan memainkan peran penting dalam memahami, mengklasifikasikan, dan memprediksi kondisi atmosfer. Klasifikasi kondisi cuaca yang akurat sangat penting tidak hanya untuk studi meteorologi tetapi juga untuk aplikasi praktis di bidang pertanian, penerbangan, transportasi, dan manajemen bencana. Di antara berbagai teknik statistik yang tersedia, analisis korespondensi telah terbukti menjadi metode eksploratif yang efektif untuk mengungkap hubungan antar variabel kategoris dengan menampilkan hubungannya dalam peta grafis dua dimensi. Tujuan utama studi ini adalah untuk mengkaji hubungan antara berbagai kategori kondisi cuaca dan tingkat tutupan awan menggunakan analisis korespondensi. Data bersumber dari platform Kaggle, yang terdiri dari 13.200 observasi dengan dua variabel kategoris utama: kondisi cuaca (cerah, berawan, hujan, dan bersalju) dan tutupan awan (cerah, sebagian berawan, berawan, dan mendung). Kumpulan data yang besar ini memberikan variabilitas dan reliabilitas yang memadai untuk menangkap interdependensi yang kompleks antara kedua variabel tersebut. Temuan ini mengungkapkan hubungan yang signifikan secara statistik antara kategori cuaca dan tingkat tutupan awan. Visualisasi peta korespondensi menunjukkan pola yang jelas: cuaca cerah sangat terkait dengan langit tanpa awan, sementara kondisi cuaca hujan dan bersalju sebagian besar terkait dengan langit yang sepenuhnya mendung. Kondisi berawan dan sebagian berawan menempati posisi menengah pada peta korespondensi, yang mencerminkan keadaan transisi antara cuaca yang menguntungkan dan yang tidak menguntungkan. Dimensi pertama peta mewakili spektrum dari kondisi cuaca yang menguntungkan (cerah dan cerah) hingga kondisi yang merugikan (hujan dan bersalju), sementara dimensi kedua terutama mencerminkan intensitas atau kepadatan tutupan awan, mulai dari lapisan awan tipis hingga tebal. Secara keseluruhan, studi ini menunjukkan kegunaan analisis korespondensi sebagai alat eksplorasi untuk data meteorologi, yang menawarkan wawasan berharga tentang interelasi kategori cuaca dan tutupan awan.

Kata kunci: Awan; Cuaca; Korespondensi; Pemetaan cuaca; Visualisasi data atmosfer

1. LATAR BELAKANG

Analisis korespondensi merupakan salah satu metode analisis eksplorasi data peubah ganda yang bertujuan untuk menggambarkan secara simultan (bersamaan) hubungan antara baris dan kolom dari suatu tabel kontingensi dua arah (*crosstab*) (Husson et al., 2017). Secara umum, metode ini digunakan untuk menganalisis data kualitatif yang diklasifikasikan dalam beberapa kategori tertentu. Dalam analisis korespondensi, informasi utama yang diangkat adalah keterkaitan antara kategori baris dan kolom tabel, kemudian divisualisasikan dalam bentuk peta dua dimensi yang akan mempermudah interpretasi hubungan antar kategori (Greenacre, 2017).

Salah satu tujuan utama dari analisis korespondensi dalam penelitian ini adalah untuk mengevaluasi ada tidaknya ketergantungan atau hubungan antara kategori-kategori pada dua variabel kualitatif tertentu. Dengan menggunakan tabel kontingensi sebagai dasar analisis, keterkaitan antar kategori dapat diuji menggunakan statistik *chi-square*. Jika hasil pengujian menunjukkan bahwa kategori-kategori tersebut tidak saling bebas, maka dapat disimpulkan adanya hubungan signifikan antar kategori, yang selanjutnya divisualisasikan dalam peta korespondensi (Hair et al., 2010).

Dalam studi klimatologi dan meteorologi, pemahaman terhadap keterkaitan antara fenomena cuaca dan elemen atmosfer lainnya, seperti penyebaran awan, merupakan aspek penting dalam mendukung klasifikasi pola cuaca (Ahn et al., 2023). Cuaca merupakan suatu kondisi atmosfer yang bersifat dinamis dan terjadi di suatu tempat dalam periode yang relatif singkat (Rinaldi & Harmayani, 2025). Sementara itu, menurut Konferensi Iklim Dunia mendefinisikan cuaca termasuk keadaan atmosfer yang menyeluruh dengan meliputi variasi, perkembangan, dan hambatan terhadap suatu fenomena (Najah et al., 2025). Kondisi ini dipengaruhi oleh unsur-unsur seperti suhu, kelembapan, curah hujan, dan tekanan udara.

Proses terbentuknya awan melalui proses kondensasi uap air yang naik dan mendingin di atmosfer (Kholiviana et al., 2022). Beragam karakteristik awan yang meliputi jenis, ketinggian, dan tingkat tutupan sering kali dikaitkan dengan tipe-tipe cuaca tertentu, seperti cerah, berawan, hujan ringan, maupun hujan lebat. Keberadaan dan karakteristik awan tidak hanya menjadi indikator utama dalam memantau kondisi atmosfer, tetapi juga berperan dalam proses peramalan dan peringatan cuaca, serta memengaruhi keseimbangan energi, suhu, dan sistem cuaca di Bumi (Lin et al., 2023).

Umumnya, data klasifikasi cuaca terdiri atas kategori cerah, berawan, hujan atau bersalju, serta informasi tentang penyebaran awan. Namun demikian, hubungan antar kategori yang bersifat kualitatif tersebut tidak selalu efektif dianalisis dengan pendekatan

parametrik (Ardiansyahroni et al., 2023). Oleh karena itu, diperlukan pendekatan statistik eksploratori yang mampu menggambarkan keterkaitan antar kategori secara visual dan interpretatif, seperti analisis korespondensi.

Penelitian ini bertujuan untuk menerapkan analisis korespondensi guna mengeksplorasi dan memvisualisasikan hubungan antara tipe cuaca dan kategori penyebaran awan. Diharapkan, hasil penelitian ini dapat memberikan gambaran awal mengenai pola keterkaitan atmosferik yang dimanfaatkan dalam pengembangan sistem klasifikasi cuaca, maupun prediksi kondisi langit.

2. KAJIAN TEORITIS

A. Tabel Kontingensi Dua Arah

Tabel kontingensi dua arah merupakan representasi data dalam bentuk tabel dimana menyajikan frekuensi pengamatan dari dua variabel kategorik (Satria et al., 2023). Umumnya digunakan untuk menganalisis hubungan antara dua variabel kategorik dengan variabel dalam baris memiliki k kategori dan b kategori dalam variabel kolom. Tabel kontingensi digunakan uji statistik dengan menggunakan pendekatan non-parametrik (Asnawi et al., 2021). Tabel ini memiliki bentuk baris dan kolom yang masing-masing mewakili kategori dari variabel X dan Y .

Secara umum, jika terdapat dua variabel kategorik, yaitu variabel X dengan b kategori dan variabel Y dengan k kategori, maka data dapat disusun dalam sebuah tabel kontingensi $b \times k$ sebagai berikut:

Tabel 1 Tabel Kontingensi Dua Arah

Variabel	Y (k taraf)				Jumlah	
	1	2	...	k		
X (b taraf)	1	O_{11}	O_{12}	...	O_{1k}	n_{10}
	2	O_{21}	O_{22}	...	O_{2k}	n_{20}
	⋮	⋮	⋮	⋱	⋮	⋮
	b	O_{b1}	O_{b2}	...	O_{bk}	n_{b0}
Jumlah	n_{01}	n_{02}	...	n_{0k}	N	

B. Uji Independensi (Kebebasan)

Salah satu langkah penting sebelum melakukan analisis korespondensi adalah mengetahui apakah terdapat hubungan antara dua variabel kategorik yang akan dianalisis dengan uji independensi. Uji ini akan menggunakan uji statistik *chi-square* (χ^2), uji statistik

dengan pendekatan non-parametrik dengan menguji 2 variabel dengan skala nominal (Annisa et al., 2020).

Hipotesis

H_0 : Tidak ada hubungan antara variabel X dan Y (independen)

H_1 : Ada hubungan antara variabel X dan Y (dependen)

Statistik Uji

$$\chi^2 = \sum_{i=1}^b \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Menurut Agresti (2006), kriteria penolakan H_0 berada pada χ^2 hitung lebih besar dari $\chi^2_{(db, \alpha)}$ pada taraf signifikansi α dan $db = (i - 1)(j - 1)$ atau $p - value < \alpha$. Dengan demikian, dapat disimpulkan terdapat hubungan antara variabel X dan Y (dependen).

C. Analisis Korepondensi

Analisis korespondensi merupakan metode analisis multivariat yang digunakan untuk menganalisis dan memvisualisasikan hubungan antara dua variabel kategorik yang disajikan dalam bentuk tabel kontingensi dua arah (dua dimensi) dan dapat diperluas ke dalam bentuk multi arah (multi dimensi) (Abidin et al., 2021). Tujuan utama dari analisis korespondensi adalah menyederhanakan data yang kompleks dalam bentuk visual dua dimensi dengan tetap mempertahankan informasi di dalamnya. Metode ini berguna untuk mengevaluasi hubungan antara kategori baris dan kolom secara simultan. Sebagai metode non-parametrik, analisis ini tidak memerlukan pengujian asumsi klasik dalam proses analisisnya (Addini et al., 2024). Luaran dari analisis ini adalah terbentuknya dua dimensi utama, di mana setiap data akan dipetakan ke dalam koordinat titik, yang merepresentasikan informasi yang termuat dalam setiap dimensi (Afif & Fahmi, 2022).

Analisis ini memanfaatkan profil baris dan profil kolom yang diekspresikan dalam ruang geometri dan memperhitungkan jarak antara profil tersebut menggunakan metrik jarak chi-square (*chi-square distance*) (Maulidia & Wulandari, 2022). Dengan menggunakan tabel kontingensi, analisis ini akan menghasilkan sejumlah sumbu utama, yang dimanfaatkan dalam visualisasi pola keterkaitan, mengelompokkan kategori serupa, serta mengurutkan objek (van Dam et al., 2021). Metode ini juga berkaitan dengan analisis faktor dan analisis komponen utama karena menggunakan singular value decomposition (SVD) sebagai teknik reduksi dimensi.

D. Matriks Data

Misalkan $N = [n_{ij}]$ adalah matriks kontingensi berukuran $I \times J$, yang berisi nilai frekuensi pengamatan (dengan $n_{ij} \geq 0$), dimana:

$i = 1, 2, \dots, I$ menyatakan katagori baris (variabel pertama)

$j = 1, 2, \dots, J$ menyatakan kategori kolom (variabel kedua)

Jumlah total dari seluruh elemen dalam matriks N adalah:

$$n_{..} = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$$

Matriks korespondensi $P = [p_{ij}]$ dibentuk dengan membagi setiap elemen dalam N dengan total observasi $n_{..}$:

$$p_{ij} = \frac{n_{ij}}{n_{..}} = \frac{\text{data baris ke } i \text{ kolom ke } j}{\text{nilai total}}$$

Selanjutnya, vektor jumlah baris dan jumlah kolom didefinisikan sebagai:

$r = P \cdot 1_J$: vektor jumlah baris

$c = P^T \cdot 1_I$: vektor jumlah kolom

Matriks diagonal dari vektor r dan c masing-masing disebut:

$D_r = \text{diag}(r)$ (ukuran $I \times I$)

$D_c = \text{diag}(c)$ (ukuran $J \times J$)

Profil baris dan kolom dari matriks P didefinisikan sebagai:

Matriks profil baris

$$R = D_r^{-1}P$$

Matriks profil kolom:

$$C = D_c^{-1}P^T$$

(Greenacre, 2017)

E. Penentuan Jarak Profil

Hubungan antara kategori baris atau kolom dapat diukur menggunakan jarak chi-square, yang memperhitungkan perbedaan proporsi relatif terhadap distribusi total. Persamaan penentuan jarak profil sebagai berikut (Lachenbruch et al., 1985):

a. Jarak antar baris i dan i' :

$$d(i, i')^2 = \sum_{j=1}^J \frac{1}{f_{\cdot j}} \left(\frac{f_{ij}}{f_{i\cdot}} - \frac{f_{i'j}}{f_{i'\cdot}} \right)^2$$

b. Jarak antar kolom j dan j' :

$$d(j, j')^2 = \sum_{i=1}^I \frac{1}{f_{i\cdot}} \left(\frac{f_{ij}}{f_{\cdot j}} - \frac{f_{ij'}}{f_{\cdot j'}} \right)^2$$

Keterangan:

f_{ij} : frekuensi relatif dari sel (i, j)

$f_{i\cdot}$: total frekuensi relatif dari baris ke- i

$f_{\cdot j}$: total frekuensi relatif kolom ke- j

F. Penguraian Nilai Singular (*Singular Value Decomposition*)

Untuk menyederhanakan struktur data dalam dimensi yang lebih rendah namun tetap mempertahankan informasi utama dari keragaman data (berdasarkan nilai *eigen* atau inersia terbesar), digunakan metode penguraian nilai singular. Teknik ini merupakan salah satu konsep penting dalam aljabar matriks dan berkaitan erat dengan *eigen decomposition*, yang melibatkan nilai *eigen* dan vektor *eigen* (Beh & Lombardo, 2014).

Menurut Greenacre (2017), penguraian nilai singular diterapkan pada matriks A berukuran $I \times J$ dengan rank K . Matriks ini diuraikan menggunakan bentuk:

$$A = U D_{\lambda} V^T$$

Atau dalam bentuk jumlah dari komponen-komponen singular:

$$A = \sum_{k=1}^K \lambda_k u_k v_k^T$$

Dengan:

U : matriks vektor-vektor eigen dari AA^T

D_{λ} : matriks diagonal yang berisi nilai-nilai singular λ , yaitu akar dari nilai eigen yang tidak nol

V : matriks vektor-vektor eigen dari $A^T A$

K : rank dari matriks A

Sifat-sifat dari dekomposisi ini mencangkup bahwa vektor-vektor kolom U dan V bersifat ortonormal, serta nilai-nilai $\lambda_1, \lambda_2, \dots, \lambda_K$ memenuhi urutan $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$

Berdasarkan dekomposisi ini, dapat dibentuk dua matriks koordinat sebagai berikut:

$F = UD_\lambda$: menyatakan koordinat baris

$G = VD_\lambda$: menyatakan koordinat kolom

G. Penguraian Nilai Singular Umum

Secara umum, bentuk penguraian nilai singular dari matriks korespondensi P dapat dituliskan sebagai:

$$P - rc^T = AD\mu B^T$$

dengan syarat sebagai berikut:

$$A^T D_r A = I, \quad B^T D_c B = I, \quad \mu_1 \geq \mu_2 \geq \dots \geq \mu_K > 0$$

dan definisi dari masing-masing komponen adalah:

A : matriks yang kolom-kolomnya terdiri dari vektor-vektor eigen dari matriks $(P - rc^T)^T(P - rc^T)$,

B : matriks yang kolom-kolomnya terdiri atas vektor-vektor eigen dari matriks $(P - rc^t)(P - rc^T)^T$

D_μ : matriks diagonal yang memuat nilai-nilai singular μ , yang merupakan akar dari nilai-nilai *eigen* tak nol dari hasil dekomposisi.

H. Dekomposisi Inersia

Inersia dalam analisis korespondensi merepresentasikan besarnya kontribusi dari suatu kategori terhadap total penyebaran (dispersi) data dalam ruang *chi-square*. Nilai inersia dari baris ke- I menunjukkan seberapa besar baris tersebut berperan dalam pembentukan struktur data secara keseluruhan, sedangkan inersia total didefinisikan sebagai jumlah tertimbang dari kuadrat jarak antara titik-titik (kategori) terhadap titik pusat (*centroid*) dengan mempertimbangkan massa (frekuensi marginal) dan metrik jarak yang digunakan. Persamaan inersia total baris dan kolom dapat dituliskan sebagai berikut (Greenacre, 2017):

Inersia total untuk baris:

$$in(I) = \sum_{i=1}^I r_i (r_i - c)^T D_c^{-1} (r_i - c)$$

Inersia total untuk kolom:

$$in(J) = \sum_{j=1}^J c_j (c_j - r)^T D_r^{-1} (c_j - r)$$

Setiap sumbu utama (komponen) yang diperoleh dari dekomposisi nilai singular menyumbang bagian tertentu dari inersia total. Nilai inersia pada dimensi ke- k dinyatakan dengan $\lambda_k = \mu_k^2$, yang merupakan kuadrat dari nilai singular ke- k , μ_k . Nilai λ_k ini disebut sebagai inersia utama ke- k . Persamaan inersia utama untuk baris dan kolom dapat dituliskan sebagai berikut (Greenacre, 2017):

Inersia utama untuk baris:

$$F^T D_r F = D_\mu^2 = D_\lambda$$

Inersia utama untuk kolom:

$$G^T D_c G = D_\mu^2 = D_\lambda$$

3. METODE PENELITIAN

Data yang digunakan berupa dataset klasifikasi tipe cuaca yang unggah oleh Narayan (2024) dan tersedia pada laman *Kaggle* (kaggle.com). Dataset ini terdiri dari 13.200 baris dan menggunakan variabel tertera pada Tabel 2.

Tabel 2. Variabel Penelitian

Variabel	Skala	Kategori
Cuaca	Nominal	1. Berawan (<i>cloudy</i>)
		2. Hujan (<i>rainy</i>)
		3. Bersalju (<i>snowy</i>)
		4. Cerah (<i>sunny</i>)
Penyebaran awan	Nominal	1. Tidak ada awan (<i>clear</i>)
		2. Sebagian berawan (<i>partly cloudy</i>)
		3. Berawan (<i>cloudy</i>)
		4. Sepenuhnya tertutup awan (<i>overcast</i>)

Langkah-langkah penerapan analisis korespondensi pada penelitian ini sebagai berikut:

1. Membentuk tabel kontingensi.
2. Menguji hubungan kebebasan antara variabel status pekerjaan dengan pendapatan per tahun dan status pekerjaan dengan pengeluaran per kunjungan dengan uji *chi-square*.
3. Membentuk matriks korespondensi dan matriks diagonal.
4. Membentuk profil baris dan kolom.
5. Menentukan nilai inersia proporsi baris dan kolom berdasarkan profil baris dan kolom.
6. Membentuk plot korespondensi dan interpretasi.

4. HASIL DAN PEMBAHASAN

Pengujian hubungan kebebasan antara 2 variabel ini dengan menggunakan uji Pearson chi-square. Diperoleh nilai χ^2_{hitung} sebesar 9407,359 dan $p - value$ sebesar 0,000. Dengan menggunakan taraf signifikansi (α) sebesar 5%, dapat diketahui bahwa uji Pearson *chi-square* $p - value < 0,05$. Oleh karena itu, keputusan akan menolak H_0 dan dapat disimpulkan bahwa terdapat hubungan antara cuaca dengan penyebaran awan.

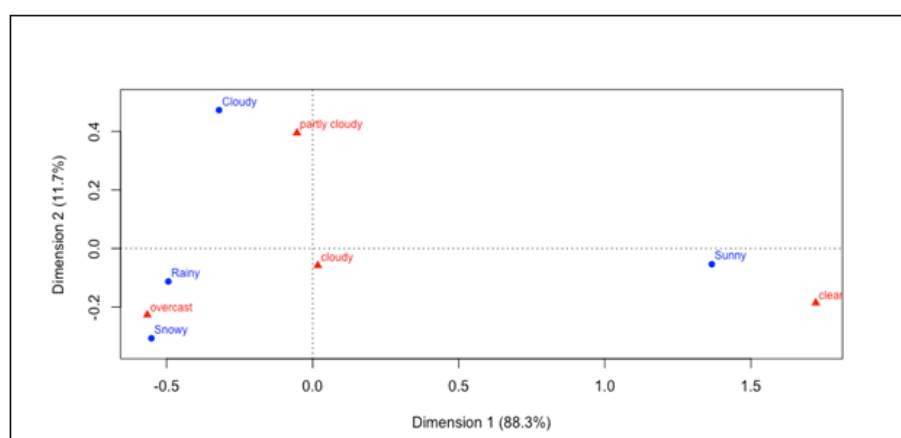
Tabel 3 Profil Baris terhadap Penyebaran Awan dengan Cuaca

	Berawan	Hujan	Bersalju	Cerah
Tidak ada awan	0	0	0	1
Sebagian berawan	0,224	0,255	0,260	0,260
Berawan	0,214	0,360	0,409	0,017
Sepenuhnya tertutup awan	0,417	0,220	0,154	0,209
Massa	0,25	0,25	0,25	0,25

Tabel 4 Profil Kolom terhadap Cuaca dengan Penyebaran Awan

	Berawan	Hujan	Bersalju	Cerah	Massa
Tidak ada awan	0	0	0	0,648	0,162
Sebagian berawan	0,028	0,032	0,032	0,032	0,345
Berawan	0,395	0,665	0,754	0,031	0,031
Sepenuhnya tertutup awan	0,577	0,304	0,213	0,288	0,461

Nilai massa menunjukkan nilai proporsi setiap kategori terhadap keseluruhan data dan menyatakan bobot dari setiap titik, baik pada baris maupun kolom. Diketahui dari Tabel 3 bahwa nilai massa keempat kategori cuaca memiliki nilai yang sama, sebesar 0,35. Sedangkan pada Tabel 4, nilai massa pada profil kolom terbesar untuk kategori penyebaran awan adalah sepenuhnya tertutup awan, yang selanjutnya diikuti oleh sebagian berawan, tidak ada awan, dan berawan.



Gambar 1. Keterkaitan Cuaca dengan Penyebaran Awan

Berdasarkan Gambar 1 kategori cuaca cerah cenderung memiliki kondisi tidak ada awan di langit. Kategori cuaca berawan memiliki kecenderungan kondisi langit yang sebagian berawan atau berawan. Kategori cuaca bersalju dan cuaca hujan memiliki penyebaran awan yang cenderung sepenuhnya tertutup awan dan kondisi langit berawan cenderung terjadi pada cuaca hujan.

Pada dimensi 1 yang menjelaskan sebesar 88,3% variasi, yang merepresentasikan perbedaan utama antara jenis cuaca berdasarkan kondisi penyebaran awan yang menyertainya. Pada dimensi ini, kategori cuaca cerah dan tidak ada awan menempati sisi kanan dengan nilai positif yang tinggi, menunjukkan asosiasi yang kuat antara cuaca cerah dengan kondisi langit yang tidak berawan. Sebaliknya, kategori seperti hujan, bersalju, dan kondisi langit sepenuhnya tertutup awan berada di sisi kiri (nilai negatif), yang mencerminkan kondisi cuaca yang cenderung buruk dan diasosiasikan dengan langit yang berawan atau tertutup awan tebal. Dimensi 1 pada kasus ini dapat menggambarkan sebagai spektrum dari cuaca buruk ke cuaca cerah, yang berkaitan erat dengan penyebaran awan. Kategori yang berada dekat dengan titik pusat tidak terlalu kuat diasosiasikan dengan kategori manapun di sisi lain dari variabel. Kategori kondisi sebagian berawan dan kondisi langit berawan tidak secara kuat dikaitkan hanya dengan salah satu jenis cuaca, melainkan cenderung muncul di berbagai jenis cuaca.

Dimensi 2, yang menjelaskan 11,7% variasi, menggambarkan perbedaan dalam hubungan antara cuaca dan penyebaran awan. Selanjutnya, kategori cuaca berawan dan kondisi langit yang sebagian berawan berada di bagian atas sumbu ini. Sementara itu, kategori kondisi langit sepenuhnya tertutup awan, cuaca bersalju, dan cuaca saat hujan berada di bagian bawah kiri, sedangkan cuaca cerah dan tidak ada awan berada di bagian bawah kanan dimensi ini, mengindikasikan asosiasi kondisi cuaca dengan penyebaran awan yang tinggi. Dengan demikian, Dimensi 2 dapat dipahami sebagai indikator tingkat intensitas atau ketebalan awan yang menyertai berbagai jenis cuaca.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa terdapat pola hubungan yang jelas antara kategori cuaca dan penyebaran awan yang dilakukan menggunakan analisis korespondensi. Cuaca cerah cenderung terjadi saat langit tidak berawan, sedangkan cuaca hujan dan bersalju biasanya disertai kondisi langit yang sepenuhnya tertutup awan. Cuaca berawan menunjukkan hubungan dengan kondisi langit

yang berawan sebagian. Dari analisis dimensi pertama (Dimensi 1), yang menjelaskan 88,3% variasi data, dapat dipahami bahwa perbedaan utama antara jenis cuaca terletak pada spektrum dari cuaca buruk (hujan dan salju) ke cuaca baik (cerah), yang ditentukan oleh penyebaran awan. Sementara itu, Dimensi 2 menjelaskan 11,7% variasi menggambarkan intensitas atau ketebalan awan yang menyertai cuaca. Pada dimensi ini, perbedaan antar cuaca tidak hanya terkait jenis awan, tetapi juga seberapa tebal atau intens penyebaran awan tersebut.

Penelitian selanjutnya dapat mempertimbangkan penambahan variabel kualitatif lain, seperti wilayah geografis, waktu kejadian, atau musim untuk dianalisis secara simultan melalui *multiple correspondance analysis* (MCA). Selain itu, penggabungan data observasi secara langsung dan hasil dari analisis korespondensi berpotensi membentuk model klasifikasi cuaca berdasarkan penyebaran awan.

DAFTAR REFERENSI

- Abidin, J., Satyahadewi, N., & Martha, S. (2021). Analisis korespondensi untuk mengetahui keterkaitan alasan mahasiswa memilih jurusan di FMIPA Universitas Tanjungpura. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, 10(1), 109–116. <https://doi.org/10.26418/bbimst.v10i1.44686>
- Addini, V., Permana, D., Amalita, N., & Salma, A. (2024). Pemetaan indikator pertumbuhan ekonomi di Provinsi Sumatera Barat menggunakan analisis korespondensi berganda. *UNP Journal of Statistics and Data Science*, 2(3), 310–317. <https://doi.org/10.24036/ujsds/vol2-iss3/190>
- Afif, A., & Fahmi, M. A. (2022). Penyebaran penyakit tidak menular berdasarkan wilayah puskesmas di Kota Kediri menggunakan analisis korespondensi. *JMPM: Jurnal Matematika dan Pendidikan Matematika*, 6(1). <https://doi.org/10.26594/jmpm.v6i1.2076>
- Agresti, A. (2006). *An introduction to categorical data analysis* (2nd ed.). Wiley. <https://doi.org/10.1002/0470114754>
- Ahn, H., Lee, S., Ko, H., Kim, M., Han, S. W., & Seok, J. (2023). Searching similar weather maps using convolutional autoencoder and satellite images. *ICT Express*, 9(1). <https://doi.org/10.1016/j.ict.2022.03.013>
- Ardiansyahroni, A., Tjalla, A., & Mahdiah, M. (2023). Data kategorik dalam penelitian: Review bibliometrik. *Jurnal Ilmiah Mandala Education*, 9(1). <https://doi.org/10.58258/jime.v9i1.4814>
- Asnawi, M. H., Rahmah, P. F., Nugraha, A., & Purwandari, T. (2021). Pemetaan kabupaten/kota di Jawa Barat berdasarkan jenis usaha pertanian menggunakan analisis korespondensi. *Seminar Nasional Statistika X*, 44–51. <https://doi.org/10.1234/pns.v10i.73>

- Beh, E. J., & Lombardo, R. (2014). Correspondence analysis: Theory, practice and new strategies. Wiley. <https://doi.org/10.1002/9781118762875>
- Greenacre, M. (2017). Correspondence analysis in practice (3rd ed.). CRC Press. <https://doi.org/10.1201/9781315369983>
- Husson, F., Lê, S., & Pagès, J. (2017). Exploratory multivariate analysis by example using R. CRC Press. <https://doi.org/10.1201/b21874>
- Kholiviana, P. A., Ruhiat, Y., & Saefullah, A. (2022). Analisis vertical wind shear pada pertumbuhan awan cumulonimbus di wilayah Kabupaten Tangerang. *Newton-Maxwell Journal of Physics*, 3(1). <https://doi.org/10.33369/nmj.v3i1.21080>
- Lin, F., Zhang, Y., & Wang, J. (2023). Recent advances in intra-hour solar forecasting: A review of ground-based sky image methods. *International Journal of Forecasting*, 39. <https://doi.org/10.1016/j.ijforecast.2021.11.002>
- Maulidia, N., & Wulandari, S. P. (2022). Analisis cluster dan korespondensi terhadap indikator pertumbuhan penduduk Kota Surabaya tahun 2020. *Jurnal Sains dan Seni ITS*, 11(1). <https://doi.org/10.12962/j23373520.v11i1.62843>
- Satria, T. A. I., Imroah, N., & Huda, N. M. (2023). Penerapan model log linier dalam menganalisis tabel kontingensi dua arah. *Buletin Ilmiah Math. Stat. dan Terapannya (BIMASTER)*, 12(6), 545–552. <https://doi.org/10.26418/bbimst.v12i6.72242>
- van Dam, A., Dekker, M., Morales-Castilla, I., Rodríguez, M., Wichmann, D., & Baudena, M. (2021). Correspondence analysis, spectral clustering and graph embedding: Applications to ecology and economic complexity. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-87971-9>