

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI LEVEL POLUSI UDARA DENGAN METODE
REGRESI LOGISTIK BINER****Diah Kartikasari**Jurusan Matematika, FMIPA, Universitas Negeri Surabaya
e-mail : diahkartikasari16030214018@mhs.unesa.ac.id**Abstrak**

Perkembangan teknologi khususnya pada bidang transportasi membuat pencemaran udara semakin meningkat. Salah satu zat polutan berbahaya di udara adalah O_3 . Indeks Standar Pencemaran Udara (ISPU) dengan parameter O_3 digunakan untuk mengukur kualitas udara berdasarkan banyaknya kadar O_3 di udara. ISPU parameter O_3 dapat menunjukkan dua kemungkinan level kualitas udara yaitu berbahaya atau tidak sehingga data kualitas udara yang dilihat dari level ISPU parameter O_3 bersifat biner. Beberapa faktor meteorologi yang diduga berpengaruh antara lain kecepatan angin, temperatur, radiasi matahari, tekanan, arah angin, dan kelembaban relatif. Metode regresi logistik biner digunakan untuk menganalisis data polusi udara berdasarkan level ISPU parameter O_3 guna mengetahui faktor-faktor yang mempengaruhi level ISPU parameter O_3 diudara. Hasil penelitian yang didapatkan dapat diketahui bahwa faktor-faktor meteorologi berupa kecepatan angin, temperatur, radiasi matahari, tekanan, dan kelembaban relatif berpengaruh secara signifikan terhadap level ISPU parameter O_3 .

Kata kunci: regresi logistik biner, polusi udara, ISPU, O_3 **Abstract**

Technological developments especially in the field of transportation makes air pollution increase. One of the dangerous pollutants in the air is O_3 . Indeks Standar Pencemaran Udara (ISPU) with the O_3 parameter is used to measure air quality based on the amount of O_3 in the air. ISPU with O_3 parameter can indicate two possible levels of air quality that is dangerous or not so that air quality data seen from the ISPU level O_3 parameter is binary. Some meteorological factors that are thought to affect include wind speed, temperature, solar radiation, pressure, wind direction, and relative humidity. The binary logistic regression method is used to analyze air pollution data based on the ISPU level with O_3 parameter to find out the factors that affect the ISPU level with O_3 parameter in the air. The result obtained can be seen that meteorological factors such as wind speed, temperature, solar radiation, pressure, and relative humidity significantly affect the ISPU level of the O_3 parameter.

Keywords : binary logistics regression; air pollution; ISPU; O_3 **1. PENDAHULUAN**

Udara merupakan salah satu komponen penting dari alam yang sangat diperlukan oleh makhluk hidup. Seiring berkembangnya transportasi dan pembangunan industri pada era modern berpotensi mempengaruhi kualitas udara. Pencemaran udara yang dihasilkan oleh aktivitas masyarakat dan industri dapat mengakibatkan turunnya kualitas udara. Perubahan kualitas udara yang disebabkan masuknya zat pencemar berupa partikel ataupun gas dengan jumlah dan rentang waktu tertentu yang dapat membawa dampak buruk bagi makhluk hidup disebut dengan pencemaran udara (Keresztes & Rápó, 2017). Pencemaran udara dapat terjadi jika zat-zat berbahaya hasil aktifitas manusia yang dilepaskan ke udara jumlahnya melampaui kemampuan udara untuk menyerap zat-zat tersebut.

Terdapat beberapa jenis polutan penyebab pencemaran udara salah satunya yang berbahaya adalah polutan O_3 . Ozon atau yang lebih dikenal dengan sebutan O_3 berbentuk kabut asap yang terikat oleh oksigen. Paparan O_3 dalam waktu tertentu dapat menyebabkan iritasi tenggorokan, paru-paru, serta paparan jangka panjangnya dapat mengakibatkan pemburukan pada penyakit asma (Octavianti dkk., 2014).

Kualitas udara berdasarkan kadar O_3 diukur setiap hari melalui stasiun pemantauan kualitas udara ambien (SUF) dan dapat dilihat berdasarkan Indeks Standar Kualitas Udara (ISPU) dengan parameter O_3 . Level ISPU parameter O_3 bernilai 1 apabila konsentrasi O_3 diudara berada pada level berbahaya untuk makhluk hidup atau nilai ISPU lebih dari 50 dan bernilai 0 untuk sebaliknya. Sehingga data kualitas udara berdasarkan level ISPU parameter O_3 merupakan data biner. Data biner mengikuti distribusi

Binomial (Luo dkk., 2018). Faktor meteorologi seperti kecepatan angin, temperatur, radiasi matahari, tekanan, arah angin, dan kelembaban relatif akan digunakan sebagai variabel bebas yang dapat mempengaruhi tingkat polusi udara yang dilihat dari level ISPU parameter O_3 (Puc & Bosiaccka, 2011).

Metode regresi logistik biner adalah salah satu bentuk khusus dari regresi yang tergolong dalam regresi non-linier yang digunakan khusus saat variabel respon terdiri dari 2 kategori (sukses atau gagal) atau lebih (Tranmer & Elliot, 2011). Faktor-faktor meteorologi yang mempengaruhi level ISPU parameter O_3 akan dianalisis menggunakan metode regresi logistik biner agar dapat diketahui apakah terdapat perbedaan signifikan yang mempengaruhi penelitian dalam memprediksi hubungan antara faktor meteorologi yang merupakan variabel bebas (x) dan level ISPU parameter O_3 yang merupakan variabel respon (y).

2. KAJIAN TEORI

Polusi Udara

Polusi atau pencemaran udara merupakan suatu proses turunnya mutu udara ambien sampai tingkat level tertentu yang disebabkan oleh masuknya komponen lain berupa gas ataupun energi ke dalam udara ambien yang dipicu oleh aktivitas manusia (Pujiastuti dkk, 2013). Kualitas udara dapat dilihat dari Indeks Standar Pencemaran Udara (ISPU) yang diukur setiap hari di stasiun pemantauan kualitas udara ambien (SUF). Nilai ISPU merupakan suatu angka yang tidak memiliki satuan dan tidak berdimensi dengan rentang 0 sampai dengan 500 yang ditetapkan berdasarkan hasil pengubahan kadar pencemaran udara yang telah diukur.

O_3 atau ozon merupakan salah satu jenis polutan berbahaya di udara yang bersifat oksidator kuat, sehingga pencemaran yang terjadi karena O_3 dapat berdampak buruk bagi keberlangsungan makhluk hidup yang ditinjau dari aspek kesehatan (Bernardini dkk., 2019). Keberadaan O_3 di udara dapat dideteksi melalui aroma khas yang dapat ditemukan di udara sebagai hasil pembakaran dari mesin-mesin bertenaga listrik. Pembentukan O_3 dinilai sangat cepat karena dapat terbentuk hanya dengan kurun waktu kurang dari atau sama dengan satu jam (Amann dkk., 2008). Kontak dengan O_3 pada konsentrasi lebih dari 9,0 ppm selama beberapa waktu mampu menimbulkan permasalahan kesehatan seperti pusing berat, kehilangan koordinasi bagi orang-orang yang sensitif, peningkatan resiko penyakit pernapasan terutama asma hingga menyebabkan *edema pulmonari* (Ghorani-Azam dkk., 2016).

Pencemaran udara berdasarkan level ISPU parameter O_3 dikategorikan menjadi 2 level yaitu level berbahaya dan level tidak berbahaya. Level ISPU parameter O_3 dikatakan berbahaya untuk makhluk hidup khususnya tumbuhan yang

sensitif apabila nilai ISPU dengan parameter O_3 lebih dari 50 (Budiyono, 2010). Sedangkan nilai ISPU parameter O_3 dikatakan tidak berbahaya bagi kesehatan manusia, nilai estetika, dan makhluk hidup lainnya apabila berada pada nilai kurang dari atau sama dengan 50. Karena terdapat dua kemungkinan level ISPU parameter O_3 maka data polusi udara dilihat dari level ISPU parameter O_3 merupakan data biner.

Model Regresi Logistik Biner

Regresi logistik merupakan suatu model yang mendeskripsikan hubungan antara beberapa variabel bebas dengan variabel respon yang bersifat dikotomis (dua kategori) atau polikotomis (lebih dari dua kategori) (Agresti, 2007). Jika variabel respon terdiri dari dua kemungkinan yaitu sukses atau gagal maka dapat digunakan bentuk regresi logistik biner. Pada regresi logistik biner variabel respon y mengikuti distribusi Binomial (Ruspriyanti & Sofro, 2018). Misalkan y_t adalah nilai dari variabel respon y ke t merupakan ukuran pengamatan maka fungsi probabilitas y adalah:

$$f(y_t) = \pi(x_t)^{y_t} (1 - \pi(x_t))^{1-y_t} \quad (1)$$

Dimana nilai $y_t = 0; 1$, $\pi(x_t)$ menyatakan probabilitas berhasil/sukses, $1 - \pi(x_t)$ menyatakan probabilitas tidak berhasil/gagal, dan $t = 1, 2, \dots, n$ menyatakan indeks observasi dengan n merupakan banyak pengamatan. Model regresi logistik biner yang digunakan adalah (Hosmer & Lemeshow, 2000):

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} \quad (2)$$

dimana k adalah banyaknya variabel bebas. Model regresi logistik pada persamaan (2) ditransformasikan menggunakan logit didapatkan :

$$g(x) = \log \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (3)$$

Estimasi parameter β dengan $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ dilakukan menggunakan metode *Maximum Likelihood Estimation* (MLE) yang akan dibahas berikutnya.

Maximum Likelihood Estimation (MLE)

Metode *Maximum Likelihood Estimation* (MLE) merupakan metode estimasi parameter yang dapat digunakan untuk mengestimasi parameter-parameter yang terdapat dalam model regresi logistik biner. Metode ini mengestimasi parameter β dimana $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ dengan cara memaksimumkan fungsi likelihood (Liu, 2016). Adapun fungsi likelihood $L(\beta)$ sebagai berikut:

$$L(\beta) = \prod_{t=1}^n f(x_t; \beta)$$

$$L(\beta) = \prod_{t=1}^n \pi(x_t)^{y_t} (1 - \pi(x_t))^{1-y_t} \quad (4)$$

Variabel respon (y) diasumsikan mengikuti distribusi Binomial dengan probabilitas sukses dilambangkan dengan

ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI LEVEL POLUSI UDARA DENGAN METODE REGRESI LOGISTIK BINER

$\pi(x_t)$ dan probabilitas gagal dilambangkan dengan $1 - \pi(x_t)$.

Untuk mempermudah perhitungan secara matematis, digunakan fungsi *log-likelihood* sebagai berikut.

$$l(\beta) = \ln L(\beta) = \sum_{t=1}^n \ln f(x_t; \beta) \quad (5)$$

Sehingga persamaan (4) dapat ditulis kembali sebagai.

$$l(\beta) = \sum_{j=0}^k \left[\sum_{t=1}^n y_t x_{tj} \right] \beta_j - \sum_{t=1}^n \ln \left[1 + \exp \left(\sum_{j=0}^k \beta_j x_{tj} \right) \right]$$

Nilai parameter β diperoleh dengan memaksimalkan fungsi *likelihood* dengan cara mencari turunan parsial pertama terhadap setiap parameter sebagai berikut.

$$l \frac{\partial l(\beta)}{\partial \beta_1} = 0; \quad l \frac{\partial l(\beta)}{\partial \beta_2} = 0; \quad l \frac{\partial l(\beta)}{\partial \beta_k} = 0 \quad (6)$$

Hasil penurunan pertama persamaan (16) terkadang tidak mendapatkan hasil yang eksplisit sehingga digunakan metode numerik Newton Raphson untuk mengoptimalkan persamaan (6).

Uji Signifikansi Parameter

Uji signifikansi parameter dilakukan pada seluruh koefisien parameter variabel bebas terhadap variabel respon. Uji ini dilakukan secara serentak dan secara parsial (Agresti, 2007).

Uji serentak dilakukan untuk menguji pengaruh koefisien β secara keseluruhan dalam model, dengan hipotesis sebagai berikut :

$$H_0: \beta_1 = \beta_2 = \dots = \beta_j = 0$$

$$H_1: \text{minimal terdapat satu } \beta_j \neq 0 \text{ dengan } j = 1, 2, \dots, k$$

Statistik uji serentak yang digunakan untuk model adalah (Engle, 1984):

$$G^2 = -2 \ln \left[\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right] \quad (7)$$

Pengujian serentak menggunakan uji-G dengan rumus pada persamaan (7) mengikuti distribusi Chi-Square χ^2 dengan derajat kebebasan k , dimana H_0 akan ditolak jika nilai $G > \chi^2_{(k;\alpha)}$ atau $p\text{-value} < \alpha$, yang memiliki makna bahwa variabel bebas secara serentak memberikan pengaruh signifikan pada variabel respon.

Pengujian secara parsial dilakukan untuk menentukan efek dari parameter variabel bebas secara individu dengan membandingkan standar error. Hipotesis uji parsial sebagai berikut:

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, j = 1, 2, \dots, k$$

Statistik uji parsial yang digunakan adalah

$$W = \left[\frac{\hat{\beta}_j}{SE \hat{\beta}_j} \right]^2 \quad (8)$$

Dimana $\hat{\beta}_j$ adalah penaksir β_j dan $SE \hat{\beta}_j$ adalah standar error, W adalah uji Wald mengikuti sebaran distribusi Normal. H_0 ditolak jika $|W| > Z_{(1;\alpha)/2}$, atau $p\text{-value} < \alpha$, sehingga memberi arti bahwa variabel bebas berpengaruh pada variabel respon.

3. DATA

Dalam penelitian ini digunakan data sekunder tentang polusi udara Kota Surabaya yang ditinjau dari level ISPU parameter O_3 yang diperoleh dari Dinas Lingkungan Hidup Jawa Timur. Data terdiri dari 7054 observasi yang merupakan data harian per 30 menit dari bulan September 2019 hingga Januari 2020. Data polusi udara akan dianalisis menggunakan metode regresi logistik biner.

Variabel respon (y) pada penelitian ini adalah level ISPU parameter O_3 , bernilai 1 untuk level ISPU berbahaya dan 0 untuk sebaliknya. Variabel bebas pada penelitian ini adalah kecepatan angin (x_1) dengan satuan knot, temperatur udara (x_2) dengan satuan Celcius, radiasi matahari (x_3) dengan satuan cal/cm^2 , tekanan udara (x_4) dengan satuan milibar, arah angin (x_5) dengan satuan derajat, kelembaban relatif (x_6) dengan satuan persen.

4. PEMBAHASAN

Analisis data polusi udara yang dilihat berdasarkan level ISPU parameter O_3 diawali dengan pembentukan model regresi logistik biner. Adapun spesifikasi model yang terdapat pada persamaan (2) dengan menggunakan metode MLE untuk mengestimasi parameter β . Hasil estimasi yang diperoleh disajikan sebagai berikut.

Tabel 1. Hasil estimasi parameter regresi logistik biner

Koefisien	Parameter	Estimasi	SE	z-value	p-value	Exp (Coef.)
(Intercept)	$\hat{\beta}_0$	-131.8	19.93	-6.61	0.00	0.00
(x_1)	$\hat{\beta}_1$	-0.048	0.013	-3.58	0.00	0.95
(x_2)	$\hat{\beta}_2$	0.363	0.025	14.26	0.00	1.43
(x_3)	$\hat{\beta}_3$	-0.007	0.000	-12.76	0.00	0.99
(x_4)	$\hat{\beta}_4$	0.115	0.019	5.99	0.00	1.12
(x_5)	$\hat{\beta}_5$	-0.0006	0.000	-1.79	0.072	0.99
(x_6)	$\hat{\beta}_6$	0.005	0.004	10.52	0.00	1.00

*)signifikansi pada $\alpha = 5\%$

Besarnya pengaruh masing-masing variabel bebas yang signifikan dalam penelitian dapat diketahui melalui interpretasi koefisien (*odd ratio*). Interpretasi koefisien (*odd ratio*) yang didapat pada Tabel 1 menyatakan bahwa temperatur udara yang tinggi memiliki resiko 1.43 kali lebih besar menjadikan level ISPU parameter O_3 berada pada level berbahaya dibandingkan dengan temperatur udara dalam keadaan normal yaitu kurang dari 27°C. Tekanan udara yang tinggi memberikan resiko kemungkinan 1.12 kali lebih besar menjadikan level ISPU parameter O_3 berada pada level berbahaya dibandingkan dengan dengan tekanan udara dalam keadaan normal yaitu kurang dari 1008.6 milibar. Adapun koefisien *odd ratio* variabel lain yang nilainya sama dengan 1 mengindikasikan bahwa tidak ada kemungkinan perubahan level ISPU

parameter O_3 seiring meningkatnya nilai variabel-variabel tersebut.

Hasil pengujian secara serentak dengan nilai uji statistik G sebesar 310.53 dimana nilai tersebut lebih besar dari nilai $\chi^2_{(6;0.05)}$ serta didapatkan p -value yang lebih kecil dari nilai signifikansi yang ditentukan yaitu $\alpha = 5\%$. Berdasarkan nilai $G > \chi^2_{(6;0.05)}$ dan p -value $< \alpha$ maka dapat disimpulkan bahwa menolak H_0 . Hal ini memberi makna bahwa secara serentak faktor-faktor meteorologi berupa kecepatan angin, temperatur, radiasi matahari, tekanan, arah angin, dan kelembaban relatif dalam penelitian mempengaruhi variabel respon berupa level ISPU parameter O_3 secara signifikan.

Berdasarkan hasil pengujian secara parsial model regresi logistik yang terdapat pada Tabel 1 dengan level signifikansi sebesar 5% menunjukkan bahwa terdapat 5 faktor meteorologi dan konstanta yang berpengaruh terhadap level ISPU parameter O_3 secara signifikan antara lain variabel kecepatan angin (x_1) dengan nilai uji statistik W sebesar -3.58, temperatur udara (x_2) dengan nilai W sebesar 14.26, radiasi matahari (x_3) dengan nilai W sebesar -12.76, tekanan udara (x_4) dengan nilai W sebesar 5.99, dan kelembaban relatif (x_6) dengan nilai W sebesar 10.52. Dari nilai-nilai tersebut diketahui bahwa $|W|$ lebih besar dari nilai $Z_{(1;\alpha)/2}$ dengan $\alpha = 5\%$. Dalam hal ini, juga didapatkan nilai p -value dari masing-masing x_1, x_2, x_3, x_4 , dan x_6 variabel sebesar 0.00 yang kurang dari nilai $\alpha = 0.05$. Karena $|W| > Z_{(1;\alpha)/2}$, atau p -value $< \alpha$ dapat disimpulkan bahwa menolak H_0 . Hal ini memberi makna bahwa kecepatan angin, temperatur udara, radiasi matahari, tekanan udara, dan kelembaban relatif mempengaruhi level ISPU parameter O_3 secara parsial.

Model regresi logistik biner yang diperoleh sebagai berikut:

$$\pi(x) = \frac{\exp g(x)}{1 + \exp g(x)}$$

dengan model logit

$$g(x) = (-131.8) - (0.048)x_1 + (0.363)x_2 - (0.007)x_3 + (0.115)x_4 - (0.0006)x_5 + (0.005)x_6$$

5. PENUTUP

Simpulan

Hasil yang telah diperoleh menyatakan bahwa variabel bebas berupa faktor meteorologi yang terdiri dari kecepatan angin, temperatur udara, radiasi matahari, tekanan udara serta kelembaban relatif mempengaruhi tingkat polusi udara yang dilihat dari level ISPU parameter O_3 secara signifikan dengan level signifikansi sebesar 5%.

Saran

Untuk penelitian selanjutnya hasil pemodelan data polusi udara dapat dikembangkan sebagai bekal untuk memprediksi level ISPU parameter O_3 periode selanjutnya serta dapat menggunakan variabel bebas berupa faktor meteorologi lain seperti curah hujan.

DAFTAR PUSTAKA

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2nd Editio). Wiley-Inter-science A John Wiley & Sons, Inc.
- Amann, M., Derwent, D., Forsberg, B., Hanninen, O., & Hurley, F. (2008). *Health Risk of Ozone From Long-range Transboundary Air Pollution*.
- Bernardini, F., Attademo, L., Trezzi, R., Gobbicchi, C., & Balducci, P. M. (2019). Air Pollutants and Daily Number of Admissions to Psychiatric Emergency Services: Evidence for Detrimental Mental Health Effects of Ozone. *Epidemiology and psychiatric sciences*, 29.
- Budiyono, A. (2010). Index Kualitas Udara. *Berita Dirgantara*, 3(1), 1–14.
- Engle, R. F. (1984). *Wald, Likelihood Ratio, and Lagrange Multiplier Test in Econometrics. II*.
- Ghorani-Azam, A., Riahi-Zanjani, B., & Balali-Mood, M. (2016). Effects Of Air Pollution On Human Health and Practical Measures for Prevention in Iran. *Journal of Research in Medical Sciences*, 21(5).
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (Second Edi). John Wiley & Sons, Inc.
- Keresztes, R., & Rápó, E. (2017). Statistical analysis of air pollution with specific regard to factor analysis in the Ciuc basin, Romania. *Studia Universitatis Babes-Bolyai Chemia*, 62(3), 283–292.
- Liu, X. (2016). *A Multi-Indexed Logistic Model for Time Series*.
- Luo, F., Wang, Q., Sun, F. M., Xu, D., & Sun, C. H. (2018). Farmers' Willingness to Participate in the Management of Small-Scale Irrigation in China from a Social Capital Perspective. *Irrigation and Drainage*, 67(4), 594–604.
- Octavianti, T., Tjahyani, S., & Utami, B. (2014). *Faktor Risiko Kejadian Gangguan Pernapasan Akibat Ozon (O3) Udara Ambien Di Kecamatan Jagakarsa Tahun 2014*.
- Puc, M., & Bosiacka, B. (2011). Effects of meteorological factors and air pollution on urban pollen concentrations. *Polish Journal of Environmental Studies*, 20(3), 611–618.
- Pujiastuti, P., Soemirat, J., & Dirgawati, M. (2013). *Karakteristik Anorganik PM10 Di Udara Ambien Terhadap Mortalitas Dan Morbiditas Pada*

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI LEVEL POLUSI UDARA DENGAN METODE
REGRESI LOGISTIK BINER**

Kawasan Industri di Kota Bandung. 1(1), 1–11.

Ruspriyanty, D., & Sofro, A. (2018). *Analysis of Hypertension Disease using Logistic and Probit Regression Analysis of Hypertension Disease using Logistic and Probit Regression.*

Tranmer, M., & Elliot, M. (2011). Binary Logistic Regression. *Best Practices in Quantitative Methods*, 358–384.

