Analisis Potensi Pencemaran Air Sungai Di Lingkungan Hutan Tropis Lembap Kalimantan Timur Menggunakan Model Regresi Weibull

Analysis of River Water Pollution Potential in Tropical Rain Forest Environment of East Kalimantan Using the Weibull Regression Model

Yola Desty[1a), Suyitno[2], Ika Purnamasari[3]
[1,2,3]Program Studi Statistika, FMIPA Universitas Mulawarman, Indonesia
[1,2]Laboratorium Statistika Terapan, FMIPA Universitas Mulawarman, Indonesia
[3]Laboratorium Statistika Ekonomi dan Bisnis, FMIPA Universitas Mulawarman, Indonesia

[a)]*Corresponding author*: yoladesty12@gmail.com,

ABSTRACT
Weibull Regression Model (WR) is the Weibull distribution in which scale parameter is stated in the regression parameter. WR model derived from the interrelated functions of Weibull Distribution, consisting of Weibull survival regression model, Weibull cumulative distribution regression model, Weibull hazard regression model, and Weibull mean regression. The purpose of this study was to obtain the pollution potential information of river water in east Kalimantan and to obtain the factors that influence it through RW modeling on dissolved oxygen (DO) data in 2022. Research data is secondary data provided by life inveronment of East Kalimantan Province. The parameter estimation method is maximum likelihood estimation (MLE). The study concluded the pollution potential information of river water in east Kalimantan Timur based on modeling RW DO data consists of the chance the unpolluted river water is 0.6868, chance of polluted river water is 0.3132, the water pollution rate is 0.4349 locations/ ppm, and average river water DO is 5.6003 ppm. Factors that influence the pollution potential of river water is nitrite concentration, water temperature, and degree of water color.

Keywords: *DO, Weibull regression model, MLE, river water pollution potential in east Kalimantan*

1. Pendahuluan

Sungai memiliki peran penting dalam kehidupan masyarakat sejak dahulu hingga sekarang, salah satu peran potensi sungai di Kalimantan Timur adalah sebagai sumber air bersih dalam menunjang kebutuhan manusia untuk mencuci, memasak, mandi, dan sebagai sumber baku air minum (PDAM). Banyak akitivitas di sepanjang daerah aliran sungai (DAS) yaitu rumah penduduk, *supermarket*, hotel, jalur transportasi, pelabuhan, kegiatan pertambangan, industri kayu, perternakan dan perikanan (Inayah, Suyitno, & Siringoringo 2021).

Aktivitas-aktivitas disepanjang DAS tersebut berpotensi menghasilkan limbah domestik maupun non domestik. Semakin banyak limbah dalam aliran sungai maka akan berpotensi menimbulkan pencemaran dan mengakibatkan kesehatan masyarakat menjadi terancam. Maka dari itu, dibutuhkan upaya pencegahan pencemaran air, dan contoh bentuk usulan pencegahan pencemaran air sungai secara statistika ialah dengan memberi wawasan bagi manusia tentang hal-hal yang dapat menyebabkan air sungai di Kalimantan Timur terancam tercemar. Informasi atau faktor yang memengaruhi pencemaran air sungai dapat diperoleh melalui pemodelan statistika yaitu pemodelan regresi Weibull (RW) dalam data *Dissolved Oxygen* (DO).

Model RW adalah distribusi Weibull yang memuat kovariat atau peubah variabel bebas. Model RW dapat di peroleh dari fungsi yang saling berkaitan dalam distribusi Weibull, dengan menyatakan pengukuran skala dalam model regresi atau fungsi dari variabel bebas (Francis & Lawless, 1983). Berdasarkan fungsi yang saling berhubungan dalam distribusi Weibull, maka model-model RW terdiri dari model regresi *survival* Weibull, model regresi distribusi kumulatif Weibull, model regresi *hazard* Weibull, dan model regresi *mean* Weibull (Suyitno, 2017). Pada umumnya model RW diterapkan pada data waktu, dan pada penelitian ini penerapannya dikembangkan dalam data kontinu non negatif bukan data waktu, yaitu data DO. DO adalah total oksigen terlarut di air dalam bentuk milligram per liter (mg/l) maupun *part per million* (*ppm*). Semakin rendah total oksigen terlarut pada sebuah perairan, mengakibatkan semakin tingginya tingkat pencemaran air (Fardiaz, 1992). Berdasarkan tersebut, penelitian ini membahas model RW pada data DO air sungai di Kalimantan Timur tahun 2022 dengan tujuan memperoleh informasi potensi pencemaran air sungai dan memperoleh faktor yang berpengaruh terhadap potensi pencemaran air sungai di Kalimantan Timur.

2. Tinjauan Pustaka

2.1 Distribusi Weibull

Menurut Suyitno (2017) Distribusi Weibull ialah distribusi dari variabel acak kontinu non-negatif. Fungsi distribusi kumulatif dari variabel acak kontinu non negatif Y yang berdistribusi Weibull versi skala-bentuk ialah

p-ISSN 2085-7829
e-ISSN 2798-3455

$$F(y) = P(Y \leq y) = 1 - exp\left[-\left(\frac{y}{\lambda}\right)^{\gamma}\right]. \tag{1}$$

Fungsi kepadatan peluang distribusi Weibull versi skala-bentuk berdasarkan persamaan (1) adalah

$$f(y) = \frac{dF(y)}{dy} = \frac{\gamma}{\lambda}\left(\frac{y}{\lambda}\right)^{\gamma-1} exp\left[-\left(\frac{y}{\lambda}\right)^{\gamma}\right]. \tag{2}$$

Fungsi *survival* distribusi Weibull versi skala-bentuk berdasarkan persamaan (1) yaitu

$$S(y) = P(Y > y) = 1 - f(y) = 1 - \exp\left[-\left(\frac{y}{\lambda}\right)^{\gamma}\right]. \tag{3}$$

Dari persamaan (2) dan (3) fungsi *hazard* distribusi Weibull versi skala-bentuk, yaitu

$$h(y) = \frac{\gamma}{\lambda}\left(\frac{y}{\lambda}\right)^{\gamma-1} = \gamma\lambda^{-\gamma}y^{\gamma-1}. \tag{4}$$

Momen ke-*r* distribusi Weibull versi skala-bentuk dengan FKP yang disajikan oleh persamaan (2) mampu dinyatakan dalam bentuk umum

$$E(Y^r) = \lambda^r \Gamma_r\left(\frac{r}{\gamma} + 1\right), \tag{5}$$

dengan Γ ialah fungsi gamma, dimana Γ_r diartikan oleh

$$\Gamma_r = \left(\frac{r}{\gamma} + 1\right). \tag{6}$$

Mean dari variabel acak Y didapatkan dari persamaan (5) dan (6), yakni

$$\mu_r = E(Y) = \lambda\Gamma\left(\frac{1}{\gamma} + 1\right). \tag{7}$$

2.2 Penaksiran Parameter Distribusi Weibull

Metode umum yang diterapkan dalam penaksiran parameter distribusi Weibull ialah *Maximum Likelihood Estimation* (MLE). MLE ialah penaksiran parameter yang memaksimumkan fungsi *likelihood*. Persamaan *likelihood* pada penaksiran distribusi Weibull adalah sistem persamaan nonlinier yang saling bergantungan, maka penyelesaian eksak guna memperoleh penaksiran ML tidak dapat dilaksanakan secara analitis, sehingga metode alternatif yang diterapkan guna menuntaskan persamaan *likelihood* ialah metode iteratif Newton-Raphson (Khairunnisa, Suyitno, & Mahmuda, 2023).

2.3 Pengujian Distribusi Data

Adapun satu metode uji distribusi data ialah pengujian Kolmogorov-Smirnov (K-S). Contohnya, akan mengetahui apakah sebuah data berdistribusi Weibull dengan fungsi distribusi kumulatif $\hat{F}(y)$, maka hipotesis uji distribusi sebagai berikut

$H_0: F(y) = \hat{F}(y)$

(Data pengamatan berdistribusi Weibull dengan fungsi distribusi kumulatif $\hat{F}(y)$)

$H_1: F(y) \neq \hat{F}(y)$

(Data pengamatan tidak berdistribusi Weibull dengan fungsi kumulatif $\hat{F}(y)$)

Statistik uji ialah

$$D = \max_{1 \leq i \leq n} |\hat{F}(y_i) - F_0(y_i)|, \tag{8}$$

dimana $F_0(y_i)$ adalah fungsi distribusi empiris yang didefinisikan oleh

$$F_0(y_i) = \frac{\text{banyaknya data respon } (Y) \text{ yang } \leq y_i}{n}, \tag{9}$$

$\hat{F}(y)$ adalah taksiran fungsi distribusi teoritis berdasarkan data pengamatan. H_0 ditolak pada taraf signifikan α jika $D > D_{(\alpha,n)}$ (Fajriati, Suyitno, & Wasono, 2022), dimana nilai $D_{(\alpha,n)}$ diperoleh dari tabel K-S (Susetyo, 2010).

2.4 Pendeteksian Multikolinearitas

Multikolinearitas merupakan kondisi dimana terdapat hubungan linier yang kuat antra variabel bebas di dalam model regresi. Metode pendeteksian multkolinearitas adalah *Variance Inflation Factor* (VIF) disertai nilai VIF diperoleh berdasarkan rumus antara lain.

$$VIF = \frac{1}{1 - R_k^2} \tag{10}$$

variabel bebas lain (Chairina, Suyitno, & Siringoringo, 2020).

2.5 Model Regresi Weibull

Model RW ialah pengembangan dari distribusi Weibull, yakni distribusi Weibull yang dipengaruhi langsung oleh variabel bebas. Model RW dapat diperoleh dari distribusi Weibull dengan parameter skala (λ), pada bentuk fungsi parameter regresi. Parameter skala (λ) bernilai riil positif, sehingga diungkapkan pada bentuk fungsi parameter regresi antara lain (Suyitno, 2017)**.**

$$\lambda(\mathbf{x}) = \exp(\boldsymbol{\beta}^T\mathbf{x}) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3) + \cdots + \beta_p X_p), \tag{11}$$

dengan, $\boldsymbol{\beta}^T = [\beta_0, \beta_1, \dots \beta_k]$ ialah vektor parameter regresi yang berdimensi $p + 1$ sementara $\mathbf{x} = \left[X_0, X_1, \dots X_p\right]^T$ ialah vektor variabel bebas dengan $X_0 = 1$. Subsitusi persamaan (11) ke persamaaan (3) didapatkan model regresi *survival* Weibull, yakni

$$S(y, \mathbf{x}) = \exp[-y^\gamma \exp(-y\boldsymbol{\beta}^T\mathbf{x})]. \tag{12}$$

Berdasarkan hubungan pada persamaan (2) dan memperhatikan persamaan (11) diperoleh model regresi distribusi kumulatif Weibull yaitu

$$F(y, \mathbf{x}) = 1 - \exp[-y^\gamma \exp(-y\boldsymbol{\beta}^T\mathbf{x})]. \tag{13}$$

dan persamaan (11) disubsitusikan ke persamaan (4) didapatkan model regresi *hazard* Weibull, yaitu

$$h(y, \mathbf{x}) = \gamma y^{\gamma-1} \exp[-\gamma\boldsymbol{\beta}^T\mathbf{x}]. \tag{14}$$

Berdasarkan persamaaan (7) dan persamaan (11) didapatkan model RW untuk *mean* peubah acak *Y*, yaitu

$$\mu_Y(\mathbf{x}) = \Gamma\left(\frac{1}{\gamma} + 1\right) \exp[\boldsymbol{\beta}^T\mathbf{x}]. \tag{15}$$

Dari persamaan (5), (12) dan persamaan (14) didapatkan FKP yang langsung dipengaruhi oleh variabel bebas (parameter regresi) yakni

$$f(y, \mathbf{x}) = \gamma y^{\gamma-1} \exp[-\gamma \exp(\boldsymbol{\beta}^T\mathbf{x})] \exp[-\gamma^\gamma \exp(-\gamma\boldsymbol{\beta}^T\mathbf{x})]. \tag{16}$$

2.6 Penaksiran Parameter Regresi Weibull

Penaksiran parameter RW memakai sebuah metode MLE. Misalkan diberikan data sampel bersukuran *n*, yaitu $(y_i, \mathbf{x}_i, \delta_i)$ dengan asumsi $y_i = 1, 2, \dots, n$ saling independent berdistribusi indentik, $y_i \sim W(\gamma, \lambda(\mathbf{x}_i)$ dengan $\lambda(\mathbf{x}_i)$ diberikan oleh persamaan (11). δ_i adalah status klasifikasi induvidu ke- *i*, yang didefinisikan

$$\delta_i = \begin{cases} 1, \text{jika } y_i < y^* \\ 0, \text{jika } y_i \geq y^* \end{cases}$$

dengan y^* adalah suatu bilangan rill positif yang diketahui. Diketahui $P(Y = y_i | \delta_i = 1) = f(y_i, \mathbf{x}_i)$ dan $P(Y = y_i | \delta_i = 0) = S(y_i, \mathbf{x}_i)$ dengan $S(y_i, \mathbf{x}_i)$ dan $f(y_i, \mathbf{x}_i)$ berturut-turut diberikan oleh persamaan (12), (14) dan (16). Fungsi *likelihood* didefinisikan

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left(f(y_i)\right)^{\delta_i} \left(S(y_i)\right)^{1-\delta_i} \tag{17}$$

dengan $\boldsymbol{\theta} = [\gamma \quad \boldsymbol{\beta}^T]$, dimana $\boldsymbol{\beta}^T = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_p]$. Berdasarkan persamaan (4), fungsi *likelihood* (17) disederhanakan menjadi

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \left(h(y_i)\right)^{\delta_i} \left(S(y_i)\right) \\ &= \prod_{i=1}^{n} \left(\gamma y^{\gamma-1} \exp(-\gamma\boldsymbol{\beta}^T\mathbf{x}_i)\right)^{\delta_i} \exp\left(-y_i^\gamma \exp(-\gamma\boldsymbol{\beta}^T\mathbf{x}_i)\right), \end{aligned} \tag{18}$$

dengan $\mathbf{x}_i = [x_{i0} \quad x_{i1} \quad x_{ip}]^T; x_{i0} = 1.$ Implementasi logaritma natural pada dua ruas persamaan (18), didapatkan fungsi *log-likelihood* antara lain:

$$\ell(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \delta_i \left([\ln\gamma + (\gamma - 1)\ln y_i - \gamma\boldsymbol{\beta}^T\mathbf{x}_i] - y_i^\gamma \exp(-\boldsymbol{\beta}^T\mathbf{x}_i)\right). \tag{19}$$

Penaksir ML model RW didapatkan dengan menuntaskan persamaan *likelihood* antara lain:

$$\frac{\partial\ell(\boldsymbol{\theta})}{\ell(\boldsymbol{\theta})} = \mathbf{0}, \tag{20}$$

dimana $\mathbf{0}$ ialah vektor 0 yang berdimensi $p + 2$. Ruas kanan persamaan (20) ialah vektor gradien yang diungkapkan dalam bentuk umum antara lain:

$$\mathbf{g} = \frac{\partial\ell(\boldsymbol{\theta})}{\ell(\boldsymbol{\theta})} = \left[\frac{\partial\ell(\boldsymbol{\theta})}{\partial\gamma} \quad \frac{\partial\ell(\boldsymbol{\theta})}{\partial\beta_p}\right]^T. \tag{21}$$

Persamaan *likelihood* (20) meliputi dari persamaan nonlinier dan saling bergantungan, sehingga penyelesaian guna memperoleh penaksir eksak ML tidak dapat ditentukan secara analisis. Metode alternatif yang diterapkan yaitu metode iteratif Newton-Raphson. Metode iteratif Newton-Raphson membutuhkan perhitungan vektor gradien dan matriks Hessian. Vektor gradien disajikan oleh persamaan (21) dan matriks Hessian ialah matriks dari turunan orde kedua dari fungsi *likelihood* persamaan (20) terhadap seluruh kombinasi komponen vektor $\boldsymbol{\theta}$.Bentuk umum matriks Hessian ialah (Suyitno, Nohe, Purnamasari, Siringoringo, Goejantoro, & Rahmah, 2022).

$$\mathbf{H}(\boldsymbol{\theta}) = \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^T}\right]_{(p+2)\times(p+2)} \tag{22}$$

Algoritma iterasi Newton-Raphson disajikan oleh

$$\hat{\boldsymbol{\theta}}^{(q+1)} = \hat{\boldsymbol{\theta}}^{(q)} - \left[\mathbf{H}\left(\hat{\boldsymbol{\theta}}^{(q)}\right)\right] \mathbf{g}\left(\hat{\boldsymbol{\theta}}^{(q)}\right); q = 0,1,2,\ldots \tag{23}$$

Dari matriks Hessian dalam persamaan (22), dapat ditemukan matriks informasi Fisher yakni

$$\left[\mathbf{I}(\hat{\boldsymbol{\theta}})\right] = -E\left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^T}\right) = -E(\mathbf{H}(\hat{\boldsymbol{\theta}})). \tag{24}$$

Diketahui bahwa $\hat{\boldsymbol{\theta}}$ adalah penaksir dari $\boldsymbol{\theta}$, sehingga pada persamaan (24) menjadi

$$\left[\mathbf{I}(\hat{\boldsymbol{\theta}})\right] = -[\mathbf{H}(\hat{\boldsymbol{\theta}})], \tag{25}$$

dan

$$\text{cov}(\hat{\boldsymbol{\theta}}) = [\mathbf{I}(\hat{\boldsymbol{\theta}})]^{-1}. \tag{26}$$

2.7 Pengujian Hipotesis Parameter Regresi Weibull Terbaik Secara Serentak dan Parsial

Pengujian hipotesis parameter terdiri dari uji signifikansi parameter regresi secara serentak serta secara parsial. Uji parameter secara serentak diterapkan guna menganalisis signifikasi parameter regresi terhadap variabel respon secara keseluruhan. Hipotesis parameter regresi secara serentak ialah

$H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_p = 0$

(Secara serentak variabel bebas tidak berpengaruh pada model RW)

H_1: minimal ada satu, $\beta_k \neq 0, k = 1,2,\ldots p$

(Minimal terdapat satu variabel bebas berpengaruh pada model RW)

Nilai maksimum fungsi *log-likelihood* di bawah model populasi (model RW lengkap) ialah

$$\ell(\hat{\boldsymbol{\theta}}) = \ln(L(\hat{\boldsymbol{\theta}})) = \sum_{i=1}^{n}[\delta_i(\ln\hat{\gamma} + (\hat{\gamma}-1)\ln y_i - \gamma\hat{\boldsymbol{\beta}}^T\mathbf{x}_i)] - y_i^{\hat{\gamma}} \exp\,[\hat{\gamma}\hat{\boldsymbol{\beta}}^T\mathbf{x}_i]], \tag{27}$$

dan nilai maksimum fungsi *log-likelihood* model RW dibawah hipotesis nol ialah

$$\ell(\hat{\boldsymbol{\theta}}_0) = \ln(L(\hat{\boldsymbol{\theta}}_0)) = \sum_{i=1}^{n}[\delta_i(\ln\hat{\gamma}_0 + (\hat{\gamma}_0-1)\ln y_i - \hat{\gamma}_0\hat{\beta}_{00})] - y_i^{\hat{\gamma}} \exp[-\hat{\gamma}_0\hat{\beta}_{00}]] \tag{28}$$

Statistik uji yang diterapkan ialah statistik pengujian G, yaitu

$$G = 2(\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\mathbf{0}})), \tag{29}$$

Daerah kritis uji ini adalah H_0 ditolak dalam taraf signifikansi α apabila nilai $G \geq \chi^2_{(\alpha,p)}$ maupun jika $P_{value} \leq \alpha$, dimana

$$P_{value} = P(G_v > G) = 1 - F(G), \tag{30}$$

dengan G_v adalah peubah acak berdistribusi $\chi^2_{(\alpha,p)}$ dan F adalah fungsi distribusi kumulatif (Suyitno, 2017).

Pengujian Parameter regresi secara parsial guna menganalisis variabel bebas tertentu secara individual berpengaruh terhadap model RW. Hipotesis pengujian parameter regresi secara parsial β_k tertentu $k = 1,2,\ldots p$ adalah

$H_0: \beta_k = 0$

(Variabel bebas X_k tidak berpengaruh terhadap model RW)

$H_1: \beta_k \neq 0$

(Variabel bebas X_k berpengaruh terhadap model RW)

Statistik uji ialah

$$W_0 = \frac{\hat{\beta}_k}{\sqrt{\text{var}(\hat{\beta}_k)}} \sim N(0,1) \tag{31}$$

dimana $\text{var}(\hat{\beta}_k)$ ialah elemen diagonal utama ke $(k+1)$ dari matriks $\left[\mathbf{H}(\hat{\boldsymbol{\theta}})\right]^{-1}$ yang disajikan oleh persamaan (26). Daerah kritis uji hipotesis nol ialah H_0 ditolak dengan taraf signifikansi α apabila nilai $|W_0| > Z_{1-\alpha/2}$

p-ISSN 2085-7829
e-ISSN 2798-3455

atau jika $P_{value} \leq \alpha$, dimana

$$p_{value} = P(|W| > W_0) = 1 - 2P(W > |W_0|) \tag{32}$$

dengan W adalah peubah acak berdistribusi $N(0,1)$, (Suyitno, 2017).

2.8 Interpretasi Model Regresi Weibull

Interpretasi Model RW berlandaskan dari nilai rasio model regresi *survival* Weibull, model regresi *Hazard* Weibull, model regresi *mean* Weibull dan model regresi distribusi kumulatif Weibull berdasarkan variabel yang berpengaruh. Nilai rasio regresi *survival* Weibull mengacu kenaikan nilai variabel bebas X_k kontinu satu-satuan adalah (Khairunnisa, Suyitno, & Mahmuda, 2023).

$$\begin{aligned} R_s(X_k) &= \frac{\hat{S}(y, \mathbf{x}|X_k + 1)}{\hat{S}(y, \mathbf{x})} \\ &= \frac{\exp\left[-\hat{\gamma}\exp\left(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k(X_k + 1) + \cdots + \hat{\beta}_p X_p\right)\right]}{\exp\left[-\hat{\gamma}\exp\left(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k + \cdots + \hat{\beta}_p X_p\right.\right.} \end{aligned} \tag{33}$$

Nilai rasio regresi distribusi kumulatif Weibull berdasarkan kenaikan nilai variabel bebas X_k kontinu satu-satuan adalah

$$R_F(X_k) = \frac{\hat{F}(y, \mathbf{x}|X_k + 1)}{\hat{F}(y, \mathbf{x})} = \frac{\hat{S}(y, \mathbf{x}|X_k + 1)}{1 - \hat{S}(y, \mathbf{x})} \tag{34}$$

Nilai rasio regresi *hazard* Weibull berdasarkan kenaikan nilai variabel bebas X_k kontinu satu-satuan adalah

$$R_h(X_k) = \frac{\hat{h}(y, \mathbf{x}|X_k + 1)}{\hat{h}(y, \mathbf{x})} = \exp\left(-\hat{\gamma}\hat{\beta}_k\right) \tag{35}$$

Nilai rasio regresi *mean* Weibull berdasarkan kenaikan nilai variabel bebas X_k kontinu satu-satuan adalah

$$R_\mu(X_k) = \frac{\hat{\mu}(\mathbf{x}|X_k + 1)}{\mathbf{x}(x)} = \exp\left(\hat{\beta}_k\right) \tag{36}$$

2.9 Ukuran Kebaikan Model

Salah satu ukuran kebaikan model regresi adalah *Bayesian Information Criterion* (BIC). Semakin kecil nilai BIC mengakibatkan semakin baiknya model RW. Nilai BIC dapat diperoleh dari persamaan berikut.

$$BIC = -2\ell(\hat{\boldsymbol{\theta}}) + k\ell(n) \tag{37}$$

dimana $\ell(\hat{\boldsymbol{\theta}})$ ialah nilai maksimum fungsi *log-likelihood*, k didefinisikan sebagai beragam parameter di masing-masing model dan n adalah banyaknya data penelitian (Suryadiningrat & Sihombing, 2022).

3. Metodologi Penelitian

Data kajian ini ialah data skunder yang didapatkan dari Dinas Lingkungan Hidup (DLH). Data variabel penelitian mencakup dari data variabel terikat dan variabel bebas yang diamati dalam aliran sungai di Kalimantan Timur. Data penelitian ini mencakup dari data variabel terikat (Y) ialah *dissolved Oxygen* (DO) air sungai di Kalimantan Timur (golongan air kelas 1), data variabel bebas mencakup dari konsentrasi nitrit (X_1), konsentrasi amonia (X_2), *Total Dissolved Solid* (X_3), suhu air (X_4), konsentrasi ferrum (X_5), konsentrasi fosfat (X_6), dan derajat warna air (X_7). Populasi riset ini ialah sungai di Kalimantan Timur. Sampel penelitian ini ialah titik-titik lokasi pengamatan sungai di Kalimantan Timur yang telah ditetapkan oleh DLH Provinsi Kalimantan Timur tahun 2022. Teknik sampling dalam kajian ialah *purposive sampling*, yakni teknik pemilihan sampel berlandaskan pertimbangan tertentu. Adapun prosedur analisis data sebagai berikut:

1. Menganalisis statistik deskriptif pada data penelitian.
2. Penaksiran parameter FKP distribusi Weibull versi skala-bentuk.
3. Pengujian distribusi Weibull pada data variabel terikat dengan menggunakan uji K-S.
4. Pendeteksian multikolonearitas antar variabel bebas menggunakan nilai VIF.
5. Penaksiran parameter model RW menerapkan metode MLE.
6. Pengujian hipotesis parameter model RW secara serentak serta parsial.
7. Pemilihan model RW terbaik berdasarkan BIC.
8. Interpretasi model RW dari nilai rasio *survival*, nilai rasio distribusi kumulatif, nilai rasio *hazard* dan nilai rasio *mean*.

4. Hasil dan Pembahasan

4.1 Analisis Statistik Deskriptif

Deskripsi data penelitian pada bentuk statistik deskriptif yang mencakup dari rata-rata, simpangan baku, nilai minimum dan maksimum. Deskripsi data bertujuan untuk mengetahui informasi awal dalam analisis

p-ISSN 2085-7829
e-ISSN 2798-3455

model RW menggunakan *software octave* 8.2.0. Statistik deskriptif data kajian akan mengacu di Tabel 1.

Tabel 1. Hasil Statistik Data Penelitian

Variabel	Rata-rata	Simpangan baku	Minimum	Maksimum
Y	5,3636	1,2255	3,6	7
X_1	0,0296	0,0230	0,0059	0,09
X_2	0,3445	0,9299	0,01	5
X_3	59,7560	47,0110	12	184
X_4	27,8790	1,4739	25	32
X_5	0,6093	0,7585	0,04	4
X_6	0,0309	0,0272	0,025	0,18
X_7	86,7880	46,6390	21	170

Dari statistik deskriptif di Tabel 4.1 rata-rata DO air sungai di Kalimantan Timur adalah 5,3636 mg/l, nilai simpangan baku DO adalah 1,2255 mg/l. Rata-rata konsentrasi nitrit air sungai di Kalimantan Timur adalah 0,0296 mg/l. Nilai simpangan baku konsentrasi nitrit adalah 0,0230 mg/l. Rata-rata konsentrasi amonia air sungai di Kalimantan Timur adalah 0,3445 mg/l. Nilai simpangan baku konsentrasi amonia adalah 0,9299 mg/l. Rata-rata TDS adalah 59,7560 mg/l. Nilai simpangan baku TDS adalah 47,0110 mg/l. Rata-rata suhu air sungai di Kalimantan Timur adalah 27,8790 ºC. Nilai simpangan baku suhu air adalah 1,4739 ºC. Rata-rata konsentrasi ferrum air sungai di Kalimantan Timur adalah 0,6093 mg/l. Nilai simpangan baku konsentrasi ferrum adalah 0,7585 mg/l. Rata-rata konsentrasi fosfat air sungai di Kalimantan Timur adalah 0,0309 mg/l. Nilai simpangan baku konsentrasi fosfat adalah 0,0272 mg/l. Rata-rata derajat warna air sungai di Kalimantan Timur adalah 86,7880 Pt-Co Unit. Nilai simpangan baku derajat warna air adalah 46,6390 Pt-Co Unit.

4.2 Penaksiran Parameter Distribusi Weibull

Penaksiran parameter distribusi Weibull dilakukan pada data DO menerapkan metode MLE yang dituntaskan memakai metode iteratif Newton-Raphson. Hasil penaksiran parameter distribusi Weibull dapat di lihat di Tabel 2.

Tabel 2. Taksiran Parameter Distribusi Weibull

Parameter	Taksiran
Skala ($\boldsymbol{\lambda}$)	4,6694
Bentuk ($\boldsymbol{\gamma}$)	6,7000

Dari hasil penaksir parameter distribusi Weibull di Tabel 2 didapatkan taksiran fungsi *survival* berdasarkan persamaan (12) ialah

$$\hat{S}(y) = \exp\left[-\left(\frac{y}{4,6694}\right)^{6,7000}\right] \tag{38}$$

dan taksiran fungsi distribusi kumulatif berdasarkan persamaan (13) adalah

$$\hat{F}(y) = 1 - \exp\left[-\left(\frac{y}{4,6694}\right)^{6,7000}\right] \tag{39}$$

4.3 Pengujian Distribusi Data DO

Uji distribusi data DO menerapkan pendekatan Kolmogorov-Smirnov. Rumusan hipotesis pengujian distribusi adalah

$H_0: F(y) = \hat{F}(y)$

(Data DO berdistribusi Weibull dengan fungsi distribusi $\hat{F}(y)$ disajikan pada persamaan (39))

$H_1: F(y) \neq \hat{F}(y)$

(Data DO tidak berdistribusi Weibull dengan fungsi distribusi $\hat{F}(y)$ disajikan pada persamaan (39))

Hasil perhitungan pengujian distribusi data dapat dilihat pada Tabel 3.

Tabel 3. Uji Distribusi Weibull Data DO

Statistik Uji (D)	$D_{(0,05;18)}$
0,2331	0,3090

Dari hasil perhitungan statistik pengujian yang dipaparkan di Tabel 4.3 dinyatakan gagal menolak H_0 dalam taraf signifikansi 0,05, perihal tersebut didukung oleh nilai statistik uji D = 0,2331 < $D_{(0,05;18)}$ = 0,3090. Simpulan dari pengujian hipotesis ini adalah bahwa data DO berdistribusi Weibull dengan fungsi distribusi pada persamaan (39).

4.4 Pendeteksian Multikolinearitas

Pendeteksian multikolinearitas menerapkan nilai VIF. Nilai VIF dihitung berdasarkan persamaan (8) dan nilai VIF > 10 mengidentifikasikan terdapat multikolinearitas antara variabel bebas. Hasil perhitungan nilai

VIF menggunakan *software Octave*, berturut-turut sebesar 2,5838; 2,2856; 1,6667; 1,5898; 2,8428; 1,0443; dan 1,7697; berdasarkan hasil perhitungan nilai VIF setiap variabel >10, sehingga kesimpulannya bahwa tidak mengalami multikolinearitas, maka pemodelan RW pada penelitian ini dapat melibatkan 7 variabel bebas yaitu, konsentrasi nitrit, konsentrasi amonia, TDS, suhu air, konsentrasi ferrum, konsentrasi fosfat, dan derajat warna air, berdasarkan pada Tabel 1.

4.5 Pemilihan Model Regresi Weibull Terbaik

Kriteria pemilihan model terbaik yaitu, model RW layak, semua variabel bebas berpengaruh, variabel bebas yang berpengaruh paling banyak, dan memberikan nilai BIC terkecil. Model RW yang terbentuk dari semua kombinasi 7 variabel bebas pada Tabel 1 adalah sebanyak $2^7 - 1 = 127$ model. Berdasarkan seleksi 127 model RW, maka didapatkan model terbaik ialah model RW dengan 3 variabel bebas, yakni konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7). Model RW terbaik ini memberikan nilai BIC terkecil sebesar 84,7505, semua variabel berpengaruh, dan model layak.

4.6 Penaksiran Parameter Regresi Weibull Terbaik

Penaksiran parameter model RW terbaik menerapkan metode MLE yang dituntaskan dengan metode Iteratif Newton-Raphson. Hasil penaksiran parameter dapat dilihat di Tabel 6.

Tabel 6. Penaksiran Parameter Model RW Terbaik

Parameter	Taksiran
$\boldsymbol{\gamma}$	5,9291
$\boldsymbol{\beta_0}$	4,5925
$\boldsymbol{\beta_1}$	-4,9939
$\boldsymbol{\beta_4}$	-0,0880
$\boldsymbol{\beta_7}$	-0,0022

Berdasarkan Tabel 6, diperoleh model RW terbaik sebagai berikut.
Model regresi *survival* Weibull berdasarkan persamaan (12), yaitu

$$\hat{S}(y,\mathbf{x}) = \exp\left[-y^{5,9291}\exp\left[-27,2317 + 29,6093X_1 + 0,5217X_4 + 0,0130X_7\right]\right]. \quad (40)$$

Model regresi distribusi kumulatif Weibull berdasarkan persamaan (13), yaitu

$$\hat{F}(y,\mathbf{x}) = 1 - \exp\left[y^{5,9291}\exp\left[-27,2317 + 29,6093X_1 + 0,5217X_4 + 0,0130X_7\right]\right]. \quad (41)$$

Model regresi *hazard* Weibull berdasarkan persamaan (14), yaitu

$$\hat{h}(y,\mathbf{x}) = 5,9291y^{4,9291}\exp\left[-27,2317 + 29,6093X_1 + 0,5217X_4 + 0,0130X_7\right]. \quad (42)$$

Model regresi *mean* Weibull berdasarkan persamaan (15), yaitu

$$\hat{\mu}(\mathbf{x}) = 0,9271\exp\left[4,5925 - 4,9939X_1 - 0,0880X_4 - 0,0022X_7\right]. \quad (43)$$

Model regresi *survival* Weibull yang disajikan pada persamaan (40) adalah model peluang air sungai di Kalimantan Timur tidak tercemar, model regresi distribusi kumulatif Weibull diberikan pada persamaan (41) adalah peluang air sungai di Kalimantan Timur tercemar, model regresi *hazard* Weibull diberikan pada persamaan (42) adalah laju pencemaran air sungai di Kalimantan Timur dan model regresi *mean* Weibull diberikan pada persamaan (43) adalah model rata-rata DO air sungai di Kalimantan Timur.

4.7 Pengujian Parameter Model Regresi Weibull Terbaik Secara Serentak dan Parsial

Pengujian parameter secara serentak memiliki tujuan menganalisis apakah variabel bebas secara serentak berpengaruh pada model RW. Hipotesis uji parameter secara serentak adalah

$H_0: \beta_1 = \beta_4 = \beta_7 = 0$
(Secara serentak variabel bebas tidak berpengaruh pada model RW)
H_1: minimal ada satu, $\beta_k \neq 0, k = 1,4,7$
(Minimal ada satu variabel bebas berpengaruh pada model RW)

Hasil uji hipotesis parameter RW secara bersama dipaparkan di Tabel 7.

Tabel 7. Uji hipotesis parameter RW secara serentak

Statistik Uji (G)	$D_{(0,05;18)}$	P_{value}
27,2808	7,8047	$5,1409 \times 10^{-6}$

Dari hasil perhitungan uji parameter RW secara serentak, maka diputuskan menolak H_0 dalam taraf signifikansi $\alpha = 0,05$, perihal tersebut didorong oleh nilai statistik pengujian $G = 27,2808 > \chi^2_{(0,05;3)} = 7,8047$ dan $P_{value} = 5,1409 \times 10^{-6} < \alpha = 0,05$. Kesimpulan uji hipotesis ini adalah konsentrasi nitrit, suhu air, dan derajat warna air secara serentak berpengaruh pada model RW terbaik atau dapat disimpulkan model RW terbaik dengan variabel bebas konsentrasi nitrit, suhu air, dan derajat warna air adalah layak (*fit*).

Uji parameter secara parsial memiliki tujuan menganalisis pengaruh setiap variabel bebas pada model

p-ISSN 2085-7829
e-ISSN 2798-3455

RW. Hipotesis uji parameter secara parsial untuk parameter dengan k = 0,1,4,7 adalah

$H_0: \beta_k = 0$

(Variabel bebas X_k tidak berpengaruh terhadap model RW)

$H_1: \beta_k \neq 0$

(Variabel bebas X_k berpengaruh terhadap model RW)

Statistik uji ialah statistik disajikan oleh persamaan (32) dengan $W_0 \sim N(0,1)$. Hasil uji hipotesis parameter RW secara parsial disajikan di Tabel 7.

Tabel 7. Hasil Uji Hipotesis Parameter RW Terbaik Secara Parsial

Variabel	Penaksir	SE	W_0	P_{value}	Keputusan
−	4,5926	0,7005	6,5558	$5{,}5322 \times 10^{-11}$	Menolak H_0
X_1	-4,9939	1,8720	2,6676	0,0076	Menolak H_0
X_4	-0,0880	0,0239	3,6777	$2{,}3530 \times 10^{-4}$	Menolak H_0
X_7	-0,0022	0,0009	2,4754	0,0133	Menolak H_0

Berdasarkan nilai statistik uji yang diperoleh pada Tabel 7, bahwa keputusan uji hipotesis semua parameter RW secara parsial menolak H_0 pada taraf signifikansi 5% diambil kesimpulan bahwa variabel-variabel konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7), masing-masing secara individual berpengaruh pada model RW. Perihal tersebut didukung oleh nilai statistik uji W_0 ketiga variabel tersebut masing-masing lebih dari 1,96 dan *p-value* ketiga variabel tersebut seluruhnya kurang dari 0,05.

Penerapan model-model RW yang diberikan pada oleh persamaan (40), (41), (42), dan (43) pada data sampel dapat diperoleh ukuran potensi pencemaran air sungai di Kalimantan Timur yang mengacu informasi di Tabel 8.

Tabel 8. Informasi Potensi Pencemaran Air Sungai di Kalimantan Timur

$\hat{S}(y, \mathbf{x})$	$\hat{F}(y, \mathbf{x})$	$\hat{h}(y, \mathbf{x})$	$\hat{\mu}(\mathbf{x})$
0,6868	0,3132	0,4349	5,6003

Berdasarkan Tabel 8 diperoleh informasi potensi pencemaran air sungai di Kalimantan Timur, yaitu peluang air sungai di Kalimantan Timur tidak tercemar sebesar 0,6868, peluang air sungai di Kalimantan Timur tercemar sebesar 0,3132, laju pencemaran air sungai di Kalimantan Timur sebesar 0,4349 lokasi/*ppm* DO atau setara dengan 4 lokasi pengamatan tercemar setiap 10 *ppm* DO, dan rata-rata DO air Sungai di Kalimantan Timur sebesar 5,6003 mg/l.

4.8 Interpretasi Model Regresi Weibull

Interpretasi model RW dilaksanakan berlandaskan nilai rasio pada variabel konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7). Nilai rasio regresi *survival* Weibull, Regresi distribusi kumulatif Weibull, regresi *hazard* Weibull, dan regresi *mean* Weibull berturut-turut diberikan pada persamaan (40), (41), (42) dan (43). Hasil perhitungan rasio regresi *survival*, rasio regresi distribusi kumultif, rasio regresi *hazard* dan rasio regresi *mean* keterangan dipaparkan dalam Tabel 9.

Tabel 9. Rasio Model RW

Variabel Berpengaruh	Rasio Model-model RW			
	$R_s(X_k)$	$R_F(X_k)$	$R_h(X_k)$	$R_\mu(X_k)$
X_1	0,0010	3,1907	4,3952	0,0068
X_4	0,7731	1,4976	1,6850	0,9158
X_7	0,9951	1,0109	1,0132	0,9978

Interpretasi model RW pada data DO air Sungai di Kalimantan Timur sebagai berikut. Berdasarkan Tabel 9 bahwa nilai rasio model regresi *survival* pada kolom $R_s(X_k)$ berturut-turut adalah 0,0010, 0,7731, dan 0,9951. Interpretasinya adalah setiap kenaikan satu satuan konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7) masing-masing akan menurunkan peluang air sungai di Kalimantan Timur tidak tecemar menjadi 0,0010 kali; 0,7731 kali; dan 0,9951 kali.

Nilai rasio model regresi distribusi kumulatif pada kolom $R_F(X_k)$ berturut-turut adalah 3,1907, 1,4975, dan 1,0109. Interpretasinya adalah setiap kenaikan satu satuan konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7) masing-masing dapat meningkatkan peluang air sungai di Kalimantan Timur tercemar sebanyak 3,1907 kali; 1,4975 kali; dan 1,0109 kali.

p-ISSN 2085-7829
e-ISSN 2798-3455

Nilai rasio model regresi *hazard* pada kolom $R_h(X_k)$ berturut-turut adalah 4,3952, 1,6851, dan 1,0132. Interpretasinya adalah setiap kenaikan satu satuan konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7) masing-masing dapat meningkatkan peluang air sungai di Kalimantan Timur tercemar sebanyak 4,3952 kali; 1,6851 kali; dan 1,0132 kali.

Nilai rasio model regresi *mean* pada kolom $R_\mu(X_k)$ berturut-turut adalah 0,0068, 0,9158, dan 0,9978. Interpretasinya adalah setiap kenaikan satu satuan konsentrasi nitrit (X_1), suhu air (X_4), dan derajat warna air (X_7) masing-masing dapat menurunkan rata-rata DO air sungai di Kalimantan Timur sebanyak 0,0068 kali; 0,9158 kali; dan 0,9978 kali.

5. Kesimpulan

Berdasarkan hasil pemodelan RW pada DO air sungai di Kalimantan Timur Tahun 2022 diambil simpulan antara lain:

1. Model regresi *survival* Weibull ialaj model peluang air sungai di Kalimantan Timur tidak tercemar, model regresi distribusi kumulatif Weibull adalah peluang air sungai di Kalimantan Timur tercemar, model regresi *hazard* Weibull adalah laju pencemaran air sungai di Kalimantan Timur dan model regresi *mean* Weibull adalah model rata-rata DO air sungai di Kalimantan Timur.
2. Informasi potensi pencemaran air sungai di Kalimantan Timur yang diperoleh berdasarkan pemodelan RW pada tahun 2022 yaitu, peluang air sungai di Kalimantan Timur tidak tercemar sebesar 0,6868, peluang air sungai di Kalimantan Timur tercemar sebesar 0,3132, laju pencemaran air sungai di Kalimantan Timur sebesar 0,4349 lokasi/ 1 *ppm* DO, dan rata-rata DO air sungai di Kalimantan Timur sebesar 5,6003 *ppm*.
   Faktor yang berpengaruh terhadap potensi pencemaran air sungai di Kalimantan Timur berlandaskan pemodelan RW pada data DO tahun 2022 adalah konsentrasi nitrit, suhu air, dan derajat warna air.

6. Daftar Pustaka

Chairina, P., Suyitno, S., & Siringoringo, M. (2021). Model-Model Regresi Weibull Univariat pada Indikator Pencemaran Air Dissolved Oxygen di Daerah Aliran Sungai Lingkungan Hutan Hujan Tropis Kalimantan Timur. *Jurnal Eksponensial*, *11*(1), 19. https://doi.org/10.30872/eksponensial.v11i1.641

Fajriati, N. A., Suyitno, & Wasono. (2022). Model Regresi Hazard Rate Weibull Kesembuhan Pasien Rawat Inap Demam Berdarah Dengue (DBD) Di RSUD Panglima Sebaya Tanah. *Jurnal Eksponensial*, *13*(1), 35–43.

Fardiaz, S. (1992). *Polusi Air dan Udara*. Anggota IKAPI.

Francis, B., & Lawless, J. F. (1983). Statistical Models and Methods for Lifetime Data. *Biometrics*, *39*(3), 820. https://doi.org/10.2307/2531129

Inayah, U. R., Suyitno, & Siringoringo, M. (2021). Upaya Pencegahan Pencemaran Air Sungai Mahakam Melalui Pemodelan Geographically Weighted Logistic Regression pada Data BOD. *Jurnal Eksponensia*, *12*(1), 17–26.

Khairunnisa, S. F., Suyitno, S., & Mahmuda, S. (2023). Weibull Regression Model On Hospitalization Time Data Of COVID-19 Patients At Abdul Wahab Sjahranie Hospital Samarinda. *Jurnal Matematika, Statistika dan Komputasi*, *19*(2), 286–303.

Suryadiningrat, & Sihombing, R. P. (2022). Komparasi Pemodelan Logit, Probit dan Clog-Log Pada Regresi Beta (Studi Kasus: Pengaruh IPG dan IDG Terhadap IKG di Indonesia Tahun 2020). *Nusantara Journal of Behavioral and Social Science*, *1*(4), 101–104. https://doi.org/10.47679/njbss.202215

Susetyo, B. (2010). *Penaksiran Parameter dan Pengujian Hipotesis Model Regresi Weibull Univariat*. PT. Refika Aditama.

Suyitno. (2017). Penaksiran Parameter dan Pengujian Hipotesis Model Regresi Weibull Univariat. *Jurnal Eksponensial*, *8*(2), 179–183.

Suyitno, Nohe, D. A., Purnamasari, I., Siringoringo, M., Goejantoro, R., & Rahmah, M. N. (2022). Pemodelan Regresi Weibull pada Potensi Pencemaran Sungai Mahakam. *Anggota IKAPI, Jawa Barat*.