



TUGAS AKHIR - SS141501

**REGRESI GAMMA DAN REGRESI LOGNORMAL
UNTUK MEMODELKAN BIOMASSA KLOORIFIL- α
DI SUNGAI DANUBE EROPA**

**KURNIATUL FAIZAH
NRP 062114 4000 0099**

**Dosen Pembimbing
Dr. Purnhadi, M.Sc.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS 141501

**REGRESI GAMMA DAN REGRESI LOGNORMAL
UNTUK MEMODELKAN BIOMASSA KLOOROFIL- α
DI SUNGAI DANUBE EROPA**

**KURNIATUL FAIZAH
NRP 062114 4000 0099**

**Dosen Pembimbing
Dr. Purhadi, M.Sc.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS 141501

**GAMMA REGRESSION AND LOGNORMAL
REGRESSION TO MODEL THE CHLOROPHYLL- α
BIOMASS IN THE DANUBE RIVER EUROPE**

**KURNIATUL FAIZAH
SN 062114 4000 0099**

**Supervisor
Dr. Purhadi, M.Sc.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

**REGRESI GAMMA DAN REGRESI LOGNORMAL
UNTUK MEMODELKAN BIOMASSA KLOOROFIL- α DI
SUNGAI DANUBE EROPA**

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Kurniatul Faizah

NRP. 062114 4000 0099

Disetujui oleh Pembimbing :
Dr. Purnhadi, M. Sc
NIP. 19620204 198701 1 001

()



Mengetahui,
Kepala Departemen

Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2018

Halaman Ini Sengaja Dikosongkan

REGRESI GAMMA DAN REGRESI LOGNORMAL UNTUK MEMODELKAN BIOMASSA KLOOROFIL- α DI SUNGAI DANUBE EROPA

Nama Mahasiswa : Kurniatul Faizah
NRP : 062114 4000 0099
Departemen : Statistika
Dosen Pembimbing : Dr. Puhadi, M.Sc.

Abstrak

Sungai Danube merupakan sungai terpanjang kedua di Eropa setelah Sungai Volga. Pada penelitian ini dilakukan pemodelan biomassa phytoplankton di Sungai Danube Eropa berdasarkan beberapa variabel. Variabel yang digunakan dalam pemodelan adalah kandungan nutrisi seperti konsentrasi nitrogen dan fosfor dalam air, derajat keasaman (pH), konduktivitas, dan suhu air sungai. Pada penelitian sebelumnya, biomassa klorofil- α digunakan sebagai ukuran biomassa phytoplankton, sehingga variabel respons pada penelitian ini adalah biomassa klorofil- α . Biomassa klorofil- α di Sungai Danube mengikuti distribusi gamma dan distribusi lognormal. Dengan demikian analisis regresi yang dapat digunakan untuk memodelkan biomassa klorofil- α di Sungai Danube adalah regresi gamma dan regresi lognormal. Berdasarkan nilai AICc, model regresi gamma memiliki performansi yang lebih baik daripada model regresi lognormal dalam memodelkan biomassa klorofil- α di Sungai Danube. Pengujian parameter menunjukkan bahwa variabel berpengaruh signifikan terhadap biomassa klorofil- α di Sungai Danube dalam model gamma adalah suhu dan derajat keasaman (pH). Hasil analisis menunjukkan bahwa semakin tinggi suhu dan derajat keasaman (pH) maka biomassa klorofil- α dari fitoplankton akan semakin besar.

Kata Kunci : biomassa klorofil- α , regresi gamma, regresi lognormal, Sungai Danube

Halaman Ini Sengaja Dikosongkan

GAMMA REGRESSION AND LOGNORMAL REGRESSION TO MODEL THE CHLOROPHYLL- α BIOMASS IN THE DANUBE RIVER EUROPE

Student Name : Kurniatul Faizah
Student Number : 062114 4000 0099
Departement : Statistics
Supervisor : Dr. Purhadi, M.Sc.

Abstract

Danube is the second longest river in Europe after the Volga. Danube River Basin extends into the territories of 19 countries in Europe. In this study, an analysis is conducted to determine factors affecting biomass of phytoplankton in Danube River. Some factors that might affect biomass of phytoplankton are nutrients concentration (e.g. nitrogen and phosphorus), acidity (pH), conductivity, and temperature of river water. In the previous studies, chlorophyll- α biomass was used as an indicator of phytoplankton biomass, hence chlorophyll- α biomass is used as response variable in this study. According to distribution test, phytoplankton klorofil- α biomass in Danube River Basin follows gamma distribution and lognormal distribution. Therefore, gamma regression and lognormal regression can be used to analyse factors affecting chlorophyll- α biomass in Danube River Basin, Europe. The AICc value of the models show that gamma regression model has better performance in this case than lognormal regression model. Parameter significance test of gamma regression model shows that temperature and acidity (pH) of river water significantly affect chlorophyll- α biomass in Danube River Basin. According to regression coefficient, the greater value of temperature and acidity, the value chlorophyll- α biomass will be larger.

Key Words : *chlorophyll- α biomass, Danube River, gamma regression, lognormal regression,*

This Page is Intentionally Left Blank

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah SWT karena berkat rahmat dan berkat-Nya penulis dapat menyelesaikan laporan Tugas Akhir dengan judul "**Regresi Gamma dan Regresi Lognormal untuk Memodelkan Biomassa Klorofil- α di Sungai Danube Eropa**".

Penyusunan dan penulisan laporan Tugas Akhir ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Orang tua penulis, Ibu Rohimah, yang telah memberikan dukungan moral sehingga penulis dapat menyelesaikan Tugas Akhir dengan baik.
2. Bapak Dr. Purhadi, M.Sc. selaku dosen pembimbing yang telah membimbing dan memberikan arahan serta masukan kepada penulis.
3. Bapak Dr. Sutikno dan Ibu Santi Wulan Purnami, S. Si, M. Si, Ph. D. selaku dosen penguji yang telah memberikan masukan untuk kesempurnaan tugas akhir ini.
4. Ibu Dra. Madu Ratna, M. Si selaku dosen wali selama masa perkuliahan di Departemen Statistika FMKSD ITS yang telah banyak memberikan saran dan arahan dalam proses belajar.
5. Para dosen Departemen Statistika FMKSD ITS yang telah memberikan ilmu dan motivasi selama masa perkuliahan.
6. Semua pihak yang telah membantu dalam penulisan laporan ini, yang tidak dapat penulis sebutkan satu per satu.

Demi perbaikan atas kekurangan pada penulisan laporan ini, saran dan kritik yang membangun akan penulis terima dengan senang hati. Semoga laporan Tugas Akhir ini dapat bermanfaat bagi semua pihak.

Surabaya, Juli 2018

Penulis

Halaman Ini Sengaja Dikosongkan

DAFTAR ISI

HALAMAN JUDUL	i
COVER PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	4
1.4 Manfaat	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Statistika Deskriptif.....	5
2.2 Pengujian Distribusi	6
2.3 Multikolinearitas	8
2.4 Regresi Gamma	9
2.5 Regresi Lognormal.....	19
2.6 Interpretasi Model	24
2.7 Kriteria Pemilihan Model Terbaik	25
2.8 <i>Danube River Basin</i>	26
BAB III METODOLOGI PENELITIAN	29
3.1 Sumber Data.....	29
3.2 Variabel Penelitian dan Struktur Data.....	29
3.3 Langkah-Langkah Analisis	31
BAB IV ANALISIS DAN PEMBAHASAN	33
4.1 Pencemaran Nutrisi di Sungai Danube.....	33
4.2 Pengujian Distribusi	38
4.3 Pemeriksaan Multikolinearitas	40

4.4 Analisis Regresi Gamma.....	41
4.5 Analisis Regresi Lognormal.....	43
4.6 Pemilihan Model Terbaik.....	44
4.7 Interpretasi Model Terbaik.....	45
BAB V KESIMPULAN DAN SARAN	49
5.1 Kesimpulan	49
5.2 Saran.....	50
DAFTAR PUSTAKA	51
LAMPIRAN.....	55

DAFTAR TABEL

Tabel 3.1	Variabel Penelitian	30
Tabel 3.2	Struktur Data.....	31
Tabel 4.1	Statistika Deskriptif Sungai Bagian <i>Upper</i>	34
Tabel 4.2	Statistika Deskriptif Sungai Bagian <i>Middle</i>	35
Tabel 4.3	Statistika Deskriptif Sungai Bagian <i>Lower</i>	36
Tabel 4.4	Hasil Uji Distribusi Variabel Respons.....	39
Tabel 4.5	Estimasi Parameter Distribusi.....	40
Tabel 4.6	VIF Setiap Variabel Prediktor	40
Tabel 4.7	Hasil Pengujian Serentak Model Regresi Gamma	42
Tabel 4.8	Hasil Estimasi dan Pengujian Paramater Model Regresi Gamma Secara Parsial.....	42
Tabel 4.9	Hasil Pengujian Serentak Model Regresi Lognormal	43
Tabel 4.10	Hasil Estimasi dan Pengujian Paramater Model Regresi Lognormal Secara Parsial.....	43
Tabel 4.11	Nilai AICc Setiap Model	44

Halaman Ini Sengaja Dikosongkan

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Tingkat Signifikan α dan p -value dalam Pengujian Hipotesis.....	7
Gambar 2.2 Kurva PDF Distribusi Gamma dengan Parameter Bentuk (α) dan Skala (β) yang Berbeda.....	10
Gambar 2.3 Kurva PDF Distribusi Lognormal dengan Parameter Lokasi (μ) dan Skala (σ) yang Berbeda.....	20
Gambar 2.4 Peta Daerah Aliran Sungai Danube	28
Gambar 3.1 Diagram Alir Analisis Data.....	32
Gambar 4.1 Rata-Rata Biomassa Klorofil- α , Nitrogen dan Fosfor.....	33
Gambar 4.2 <i>Fitting</i> Biomassa Klorofil- α terhadap Distribusi Normal, Lognormal, dan Gamma	39

Halaman Ini Sengaja Dikosongkan

DAFTAR LAMPIRAN

Lampiran 1.	Lokasi Pengambilan <i>Sampling</i> JDS3	55
Lampiran 2.	Data Penelitian	57
Lampiran 3.	Statistika Deskriptif.....	59
Lampiran 4.	Hasil Uji Distribusi dan Estimasi Parameter Distribusi.....	60
Lampiran 5.	Matriks Korelasi Antar Variabel.....	60
Lampiran 6.	Analisis Regresi Linear	61
Lampiran 7.	Analisis Regresi Gamma.....	62
Lampiran 8.	Analisis Regresi Lognormal.....	63
Lampiran 9.	Nilai Prediksi Biomassa Klorofil- α dengan Model Gamma	65
Lampiran 10.	Nilai Prediksi Biomassa Klorofil- α dengan Model Lognormal	67

Halaman Ini Sengaja Dikosongkan

BAB I PENDAHULUAN

1.1 Latar Belakang

Salah satu bioindikator untuk menilai kesehatan sungai adalah *phytoplankton*. *Phytoplankton* merupakan organisme yang hidup di perairan dengan melakukan fotosintesis. Keberadaan organisme tersebut di suatu wilayah perairan dipengaruhi oleh nutrisi seperti nitrogen (N) dan fosfor (P) yang terkandung dalam air. Hal tersebut dikarenakan *phytoplankton* membutuhkan nutrisi untuk menghasilkan energi melalui fotosintesis. Semakin tinggi konsentrasi nutrisi dalam perairan tersebut, maka akan semakin banyak *phytoplankton* yang tumbuh dalam perairan tersebut.

Kondisi dimana suatu perairan memiliki konsentrasi nutrisi yang tinggi disebut dengan *eutrophication*. Konsentrasi nutrisi yang terlalu tinggi dapat menyebabkan *algal bloom* yaitu kondisi dimana terjadi peningkatan populasi alga dan *phytoplankton* yang sangat cepat hingga permukaan air tertutup alga dan *phytoplankton* sehingga menyebabkan perubahan warna air akibat tingginya konsentrasi sel pigmen. Semakin banyak populasi alga dan *phytoplankton* maka ketika organisme tersebut mati akan dibutuhkan lebih banyak oksigen untuk proses dekomposisi. Hal tersebut menyebabkan penurunan konsentrasi oksigen terlarut dan air menjadi asam sehingga dapat menyebabkan ikan-ikan atau organisme lain yang hidup di perairan tersebut mati (Fondriest Environmental, Inc., 2014; Smith, *et al.*, 2015). Oleh karena itu *phytoplankton* sering digunakan sebagai indikator dari polusi nutrisi di suatu perairan seperti danau, estuari, sungai, pantai, dan laut. Pada penelitian-penelitian sebelumnya tentang *phytoplankton* di perairan

tertentu, klorofil- α digunakan sebagai indikator kualitas air karena klorofil- α merupakan indikator dari biomassa *phytoplankton* (Boyer, *et al.*, 2009; Magumba, *et al.*, 2013).

Faktor-faktor yang diduga mempengaruhi biomassa klorofil- α pada *phytoplankton* adalah kandungan nutrisi seperti konsentrasi nitrogen dan fosfor dalam air. Hasil penelitian oleh Magumba, *et al.*, (2013) menunjukkan bahwa berdasarkan hasil analisis regresi dengan metode *Ordinary Least Square* (OLS) fosfor total (TP) dan nitrogen total (TN) merupakan determinan atau faktor penentu untuk klorofil- α di danau. Dalam Lv, Wu dan Chen (2011), hasil dari analisis regresi mengindikasikan bahwa sebagian besar danau di wilayah perkotaan yang diteliti merupakan danau *hypertrophic* berdasarkan konsentrasi fosfor total (TP) dan nitrogen total (TN). Dalam penelitian tersebut disebutkan bahwa biomassa *phytoplankton* paling tinggi pada musim panas dan *limiting factor* untuk pertumbuhan *phytoplankton* sepanjang tahun adalah temperatur dan TP. Pada penelitian tersebut diketahui juga bahwa korelasi fosfor total (TP) dan klorofil- α lebih tinggi dibandingkan dengan korelasi antara nitrogen total (TN) dan klorofil- α , dan rasio TN:TP tidak selalu sesuai digunakan sebagai indeks untuk menentukan baik nitrogen atau fosfor membatasi biomassa *phytoplankton* di danau yang dangkal.

Penelitian oleh Dodds & Smith (2016) menunjukkan bahwa nutrisi (nitrogen dan fosfor) mempengaruhi *eutrophication* di sungai. Selain itu, variabel *physico-chemical* lain seperti derajat keasaman (pH), konduktivitas dan temperatur air juga diduga mempengaruhi biomassa klorofil- α . Dalam Rocha *et al.* (2009), konsentrasi klorofil- α di danau dataran banjir Sungai Paraná, Brazil dipengaruhi secara signifikan oleh pH, konduktivitas, total seston, dan nitrat.

Pada penelitian ini dilakukan analisis untuk mengetahui pengaruh *physico-chemical*, seperti kadar nitrogen; fosfor; temperatur; derajat keasaman; dan konduktivitas air sungai, terhadap biomassa klorofil- α di Sungai Danube, Eropa. Sungai tersebut merupakan sungai terpanjang kedua di Eropa setelah Sungai Volga. Daerah aliran Sungai Danube meliputi 19 negara di daratan Eropa. Sungai tersebut mengalir dari Jerman bagian barat dan berakhir di Laut Hitam (ICPDR, 2009).

Data yang digunakan dalam penelitian ini diambil dari 64 lokasi pengamatan yang terletak di berbagai negara Eropa. Berdasarkan hasil pengujian distribusi, data biomassa klorofil- α di Sungai Danube mengikuti distribusi lognormal dan gamma. Dengan demikian, analisis regresi dengan menggunakan fungsi distribusi dari kedua distribusi tersebut dapat digunakan untuk menganalisis pengaruh faktor-faktor di atas terhadap biomassa klorofil- α di Sungai Danube, Eropa. Pada penelitian ini, hasil analisis dengan menggunakan dua jenis analisis regresi tersebut dibandingkan untuk mendapatkan model yang terbaik di antara keduanya.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka permasalahan yang diangkat dalam penelitian ini adalah pengaruh *physico-chemical* air sungai, yang meliputi konsentrasi nitrogen; konsentrasi fosfor; temperatur; derajat keasaman; dan konduktivitas terhadap biomassa klorofil- α dari *phytoplankton* di Sungai Danube, Eropa berdasarkan model regresi lognormal dan gamma. Hasil analisis dari kedua jenis regresi tersebut akan dibandingkan untuk mengetahui model yang terbaik diantara keduanya.

1.3 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut.

1. Memperoleh model terbaik dari dua model yang diusulkan, yaitu model regresi gamma dan model regresi lognormal, untuk memodelkan faktor-faktor yang mempengaruhi biomassa klorofil- α di Sungai Danube, Eropa.
2. Mengetahui bagaimana pengaruh kandungan nutrisi (nitrogen dan fosfor), temperatur, derajat keasaman, dan konduktivitas air sungai terhadap biomassa klorofil- α di Sungai Danube berdasarkan model terbaik yang diperoleh.

1.4 Manfaat

Pada penelitian ini, pemodelan dengan regresi lognormal dan regresi gamma digunakan untuk menganalisis pengaruh *physico-chemical* air sungai terhadap biomassa klorofil- α di Sungai Danube, Eropa. Penelitian ini diharapkan dapat memberikan wawasan bagi pembaca mengenai pemodelan dengan regresi lognormal dan regresi gamma.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah data yang digunakan merupakan data *cross-sectional* yang diperoleh dari *Joint Danube Survey 3* (JDS3). Survei tersebut dilakukan enam tahun sekali dan pertama kali dilaksanakan pada tahun 2001. Survei terakhir adalah JDS3 dilaksanakan pada tahun 2013 dan berlangsung dari Agustus 2013 hingga September 2013. Data yang digunakan merupakan data hasil *sampling* di bagian tengah badan sungai.

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistik deskriptif adalah metode statistika yang digunakan untuk menggambarkan atau mendeskripsikan data yang telah dikumpulkan menjadi sebuah informasi yang berguna. Statistika deskriptif hanya memberikan informasi mengenai data yang ada dan tidak menarik kesimpulan (Walpole, *et al.* 2010). Informasi yang didapatkan tersebut selanjutnya akan diuji secara statistik sehingga dapat ditarik kesimpulan.

Terdapat dua macam ukuran dalam statistika deskriptif, yaitu ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan adalah nilai tunggal yang mewakili keseluruhan data dan menunjukkan karakteristik data secara umum. Ukuran pemusatan menunjukkan pusat dari nilai data, meliputi *mean* median, modus dan sejenisnya. Selain itu juga terdapat ukuran lokasi yang meliputi kuartil, desil, kuantil, dan persentil. Sedangkan ukuran penyebaran adalah suatu nilai atau ukuran yang menunjukkan besarnya simpangan data dari pusatnya. Ukuran penyebaran juga dapat menunjukkan homogenitas data. Semakin besar nilai suatu ukuran penyebaran data maka semakin rendah homogenitas data (artinya data semakin tidak homogen).

1. *Mean*

Mean adalah jumlah nilai pada data dibagi dengan banyaknya data tersebut. Ukuran pemusatan ini sensitif terhadap keberadaan *outlier* (Montgomery & Runger, 2014). Ada pun perumusan untuk *mean* adalah sebagai berikut.

$$\bar{X} = \frac{\sum x_i}{n} \quad (2.1)$$

Dengan:

x_i = data ke- i

n = banyak data

2. Standar Deviasi

Standar deviasi adalah suatu nilai yang menunjukkan besarnya simpangan rata-rata seluruh nilai yang ada dalam kelompok data dengan nilai pusatnya dengan cara menghilangkan kemungkinan nilai nol dengan jalan dikuadratkan. Semakin kecil nilai simpangannya berarti variasi nilai data semakin kecil. Jika ukuran ini bernilai 0, maka nilai semua datanya adalah sama. Standar deviasi ditentukan dengan rumusan berikut.

$$S = \sqrt{\frac{1}{n} \sum (X_i - \bar{X})^2} \quad (2.2)$$

2.2 Pengujian Distribusi

Suatu variabel random dapat dikatakan mengikuti suatu distribusi tertentu setelah dilakukan pengujian statistik. Uji Kolmogorov-Smirnov dan Anderson-Darling dapat digunakan untuk menentukan distribusi tertentu dari suatu data. Pada penelitian ini, pengujian distribusi dilakukan menggunakan uji Anderson-Darling. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut.

$H_0 : F_{(y)} = F_{(y)}^0$ (Distribusi suatu variabel random sesuai dengan distribusi tertentu)

$H_1 : F_{(y)} \neq F_{(y)}^0$ (Distribusi suatu variabel random tidak sesuai dengan distribusi tertentu)

Sementara statistik uji yang digunakan menurut Stephens (1974) adalah sebagai berikut.

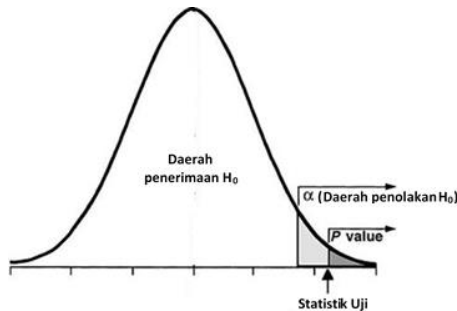
$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) (\ln(F_{(y_i)}^0) + \ln(1 - F_{(y_{n+1-i})}^0)) \quad (2.3)$$

dengan:

$F_{(y_i)}^0$: fungsi distribusi kumulatif dari distribusi tertentu

n : ukuran data

Penolakan terhadap hipotesis awal (H_0) terjadi ketika $A^2_{hitung} > A^2_{\alpha}$, dengan α merupakan tingkat signifikan (Stephens, 1974).



Gambar 2.1 Ilustrasi Tingkat Signifikan α dan p -value dalam Pengujian Hipotesis

Penolakan terhadap hipotesis awal (H_0) juga terjadi apabila p -value lebih kecil daripada tingkat signifikan α . Tingkat signifikan α merupakan probabilitas hipotesis awal (H_0) ditolak ketika hipotesis tersebut benar, sedangkan p -value merupakan probabilitas untuk memperoleh statistik uji dengan asumsi bahwa hipotesis awal benar. Secara matematis, p -value dalam pengujian hipotesis distribusi ini dapat dirumuskan

sebagai $P(A^2_{hitung} \leq A^2_\alpha)$. Ilustrasi tingkat signifikan α dan p -value dalam pengujian hipotesis satu arah dapat dilihat pada Gambar 2.1.

2.3 Multikolinearitas

Salah satu asumsi yang harus dipenuhi dalam regresi adalah tidak adanya korelasi antar variabel prediktor atau tidak ada multikolinearitas. Adanya korelasi antar variabel prediktor dalam model regresi dapat menyebabkan nilai estimasi parameter model regresi yang dihasilkan memiliki error yang besar (Gujarati & Porter, 2008). Oleh karena itu, perlu dilakukan pengecekan adanya multikolinearitas sebelum dilakukan pemodelan.

Salah satu cara untuk mendeteksi adanya multikolinearitas adalah dengan melihat nilai *Variance Inflation Factors* (VIF). Nilai VIF untuk setiap variabel prediktor ke- j dapat dihitung dengan menggunakan persamaan di bawah ini.

$$VIF_j = \frac{1}{1 - R_j^2}; \quad j = 1, 2, \dots, k \quad (2.4)$$

dimana R_j^2 merupakan koefisien determinasi antara variabel prediktor ke- j (X_j) dengan variabel prediktor lainnya. Tidak ada batasan atau *threshold* VIF yang secara formal mengindikasikan adanya kasus multikolinearitas. Menurut Kleinbaum, Kupper dan Muller dalam Gujarati & Porter (2008), apabila nilai VIF_j lebih besar dari 10 maka terdapat indikasi bahwa terdapat kolinearitas pada variabel prediktor ke- j .

Penanganan terhadap pelanggaran asumsi multikolinearitas dapat dilakukan dengan berbagai cara. Salah satunya adalah dengan mengeliminasi variabel penyebab multikolinearitas atau menggantinya dengan variabel lain. Solusi lain adalah

dengan menggunakan regresi *ridge*, mentransformasi data dan menggunakan *Principal Component Analysis* (PCA).

2.4 Regresi Gamma

Beberapa variabel random bernilai nonnegatif dan memiliki distribusi data yang tidak simetrik (*skewed*). Variabel random yang demikian tidak memenuhi distribusi normal. Salah satu distribusi variabel random kontinu yang menghasilkan kurva yang memiliki *skewness* adalah distribusi gamma. Distribusi gamma merupakan fungsi distribusi kontinu yang mempunyai dua parameter bernilai positif, yaitu α dan θ (Wackerly, *et al.*, 2008). Distribusi gamma terjadi pada data yang bernilai positif seperti data waktu antar kerusakan mesin, data meteorologi seperti curah hujan harian, data klaim asuransi, data biologi, dan data lainnya yang bernilai positif.

Fungsi densitas dari distribusi gamma dengan parameter α dan θ dalam Wackerly *et al.* (2008) adalah sebagai berikut.

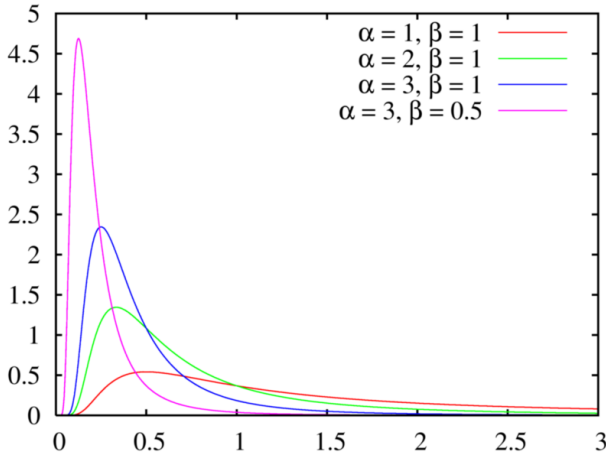
$$f(y; \alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} y^{\alpha-1} e^{-y/\theta}; \quad \alpha, \theta > 0; \quad 0 < y < \infty \quad (2.5)$$

dengan $\Gamma(\alpha)$ merupakan fungsi gamma sebagai berikut.

$$\Gamma(\alpha) = (\alpha - 1)! = (\alpha - 1)(\alpha - 2) \dots (2)(1) \quad (2.6)$$

Pada fungsi densitas distribusi gamma, parameter α dikenal sebagai parameter bentuk. Parameter tersebut mempengaruhi bentuk kurva fungsi densitas distribusi gamma; nilai α yang berbeda akan menghasilkan bentuk kurva yang berbeda. Sementara parameter θ merupakan parameter skala karena perkalian antara variabel random yang berdistribusi gamma dengan konstanta positif menghasilkan variabel random yang juga berdistribusi gamma dengan nilai α yang sama

namun dengan θ yang berbeda (Wackerly, et al., 2008). Kurva fungsi densitas distribusi gamma dengan parameter bentuk α dan parameter skala β yang berbeda-beda ditunjukkan oleh Gambar 2.2.



Gambar 2.2 Kurva PDF Distribusi Gamma dengan Parameter Bentuk (α) dan Skala (β) yang Berbeda

Nilai harapan μ dan varians dari distribusi gamma dengan parameter α dan θ adalah sebagai berikut (Casella & Berger, 2002).

$$\mu = E(Y) = \alpha\theta \quad (2.7)$$

$$\text{Var}(Y) = \alpha\theta^2 \quad (2.8)$$

Karena α dan θ bernilai positif, maka nilai μ juga harus bernilai positif. Sedangkan nilai varians bergantung pada nilai μ .

2.4.1 Model Regresi Gamma

Regresi gamma merupakan bentuk regresi yang digunakan untuk memodelkan hubungan antara satu atau beberapa variabel prediktor (X_1, X_2, \dots, X_k) dengan satu variabel

respon (Y) yang berdistribusi gamma. Dengan demikian, variabel respon merupakan variabel random kontinu yang bernilai positif dan koefisien variasi yang konstan. Model yang digunakan pada penelitian ini adalah model log linear sehingga didapatkan nilai harapan (μ) yang bernilai positif. Jika variabel yang mempengaruhi variabel respons Y adalah X_1, X_2, \dots, X_k , maka nilai μ untuk Y pada observasi ke- i menurut model log linear adalah sebagai berikut (Dobson & Barnett, 2008).

$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}) \quad (2.9)$$

dengan $\mu_i = E(Y_i)$ dan $g(\mu_i)$ merupakan fungsi transformasi.

Jika $\mathbf{x}_i = [1 \ x_{i1} \ x_{i2} \ \dots \ x_{ik}]^T$ merupakan vektor variabel prediktor untuk observasi ke- i dan vektor untuk koefisien regresi adalah $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k]^T$, maka μ_i dapat diekspresikan sebagai berikut.

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \quad (2.10)$$

Persamaan (2.7) dan persamaan (2.10) menghasilkan persamaan dibawah ini.

$$\alpha \theta_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

$$\theta_i = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\alpha} \quad (2.11)$$

Dengan demikian, diperoleh fungsi gamma dengan mensubstitusikan persamaan (2.11) ke dalam persamaan (2.5) sebagai berikut.

$$f(y_i) = \frac{1}{\Gamma(\alpha) \left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\alpha} \right)^\alpha} y_i^{\alpha-1} \exp \left(\frac{-y_i}{\left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\alpha} \right)} \right) \quad (2.12)$$

2.4.2 Estimasi Parameter Model Regresi Gamma

Pada persamaan (2.12), yang menjadi parameter model regresi gamma adalah α dan $\boldsymbol{\beta}$. Estimasi parameter tersebut dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Setelah didapatkan fungsi untuk estimasi masing-masing parameter, dilakukan optimasi numerik untuk mendapatkan hasil estimasi yang konvergen. Metode optimasi numerik yang digunakan adalah Broyden-Fletcher-Goldfarb-Shanno (BFGS).

a. Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE) merupakan metode yang sering digunakan dalam estimasi parameter model. MLE merupakan suatu prosedur untuk mendapatkan nilai dari parameter suatu model yang memaksimalkan fungsi *likelihood* (Cox & Hinkley, 1974). Metode MLE memungkinkan diperolehnya estimasi yang konsisten dan varians yang lebih kecil sehingga tidak terpengaruh oleh *sampling error*, *robust* terhadap pelanggaran asumsi, serta mudah dipahami. Namun, metode ini membutuhkan komputasi sehingga didapatkan estimasi parameter yang konvergen.

Fungsi *likelihood* dari model regresi gamma pada persamaan (2.12) menurut Putri (2016) adalah sebagai berikut.

$$\begin{aligned}
L(\alpha, \boldsymbol{\beta}) &= \prod_{i=1}^n f(y_i) \\
&= \prod_{i=1}^n \left[\frac{1}{\Gamma(\alpha) \left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\alpha} \right)^\alpha} y_i^{\alpha-1} \exp\left(\frac{-\alpha y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right) \right] \\
&= \left[\left(\frac{1}{\Gamma(\alpha)} \right)^n \left(\frac{\exp(\sum_{i=1}^n \mathbf{x}_i^T \boldsymbol{\beta})}{\alpha^n} \right)^{-\alpha} \left(\prod_{i=1}^n (y_i^{\alpha-1}) \right) A \right] \quad (2.13)
\end{aligned}$$

dimana:

$$A = \exp\left(\sum_{i=1}^n \frac{-\alpha y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right)$$

Sedangkan fungsi *ln likelihood* dari persamaan (2.13) adalah sebagai berikut.

$$\begin{aligned}
\ln L(\alpha, \boldsymbol{\beta}) &= \ln \left[\left(\frac{1}{\Gamma(\alpha)} \right)^n \left(\frac{\exp(\sum_{i=1}^n \mathbf{x}_i^T \boldsymbol{\beta})}{\alpha^n} \right)^{-\alpha} \left(\prod_{i=1}^n (y_i^{\alpha-1}) \right) A \right] \\
&= -n \ln \Gamma(\alpha) - \alpha \sum_{i=1}^n \mathbf{x}_i^T \boldsymbol{\beta} + \alpha n \ln \alpha + (\alpha - 1) \sum_{i=1}^n \ln y_i + B \quad (2.14)
\end{aligned}$$

dimana:

$$B = \sum_{i=1}^n \frac{-\alpha y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}$$

Turunan parsial pertama dari persamaan (2.14) terhadap masing-masing parameter, yaitu α dan $\boldsymbol{\beta}$, yang kemudian

disamadengankan nol digunakan untuk estimasi parameter regresi gamma. Turunan parsial pertama fungsi *ln likelihood* terhadap parameter α dan disamadengankan nol dalam Putri (2016) adalah sebagai berikut.

$$\frac{\partial \ln L(\alpha, \boldsymbol{\beta})}{\partial \alpha} = -n \frac{1}{\Gamma(\alpha)} \Gamma(\alpha)' - \sum_{i=1}^n \mathbf{x}_i^T \boldsymbol{\beta} + n \ln(\alpha) + n + C = 0 \quad (2.15)$$

dimana:

$$C = \sum_{i=1}^n \frac{-y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}$$

Sedangkan turunan parsial pertama fungsi *ln likelihood* terhadap parameter $\boldsymbol{\beta}^T$ dan disamadengankan nol dalam Putri (2016) ditunjukkan oleh persamaan di bawah ini.

$$\frac{\partial \ln L(\alpha, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} = -\alpha \sum_{i=1}^n \mathbf{x}_i + \alpha \sum_{i=1}^n \frac{\mathbf{x}_i y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} = 0 \quad (2.16)$$

Persamaan (2.15) dan (2.16) tidak dapat diselesaikan secara analitik sehingga diperlukan metode optimasi numerik agar didapatkan estimasi parameter yang konvergen. Salah satu jenis optimasi numerik yang dapat digunakan pada kedua persamaan tersebut adalah algoritma *Fisher Scoring*. Uraian mengenai algoritma tersebut dijelaskan di bawah ini.

b. Algoritma *Fisher Scoring*

Metode *Fisher Scoring* merupakan salah satu bentuk metode iterasi Newton yang digunakan untuk menyelesaikan persamaan *maximum likelihood* secara numerik. Estimasi parameter dengan metode *Fisher Scoring* membutuhkan vektor *score* dan matriks informasi Fisher. Vektor *score* merupakan

vektor dimana elemennya merupakan turunan pertama fungsi *Inlikelihood* terhadap masing-masing parameter. Sedangkan matriks informasi Fisher adalah matriks yang elemennya merupakan nilai ekspektasi dari turunan kedua fungsi *Inlikelihood* terhadap masing-masing parameter. Penggunaan matriks Hessian dalam metode *Newton-Rhapson* memiliki kelemahan, yaitu hasil iterasinya tidak selalu konvergen. Oleh karena itu, pada metode *Fisher Scoring* matriks Hessian dimodifikasi dengan matriks informasi Fisher

Vektor *score* untuk estimasi parameter model gamma adalah sebagai berikut.

Algoritma metode *Fisher Scoring* untuk estimasi parameter regresi gamma di bawah populasi dijelaskan sebagai berikut.

1. Inisialisasi $\boldsymbol{\gamma}_0$ dan menentukan toleransi error (ε)

Vektor $\boldsymbol{\gamma}_0$ merupakan nilai awal parameter regresi gamma yang memuat parameter $\boldsymbol{\beta}_{(0)}$ dan $\alpha_{(0)}$. Dengan demikian, maka $\boldsymbol{\gamma}_0 = [\beta_{0(0)} \ \beta_{1(0)} \ \dots \ \beta_{k(0)} \ \alpha_{(0)}]^T$. Nilai $\boldsymbol{\beta}_{(0)}^T$ diperoleh dari hasil estimasi parameter menggunakan metode *Ordinary Least Square* (OLS). Estimasi $\boldsymbol{\beta}_{(0)}^T$ dengan OLS dapat dinyatakan dengan $\widehat{\boldsymbol{\beta}}_{(0)} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$. Sedangkan $\alpha_{(0)}$ merupakan sembarang nilai positif sebagaimana pada persamaan (2.3) tentang fungsi densitas distribusi gamma.

2. Menghitung nilai $\mathbf{V}(\boldsymbol{\gamma}_m)$ dan $\mathbf{I}(\boldsymbol{\gamma}_m)$

$$\mathbf{V}(\boldsymbol{\gamma}_m) = \begin{bmatrix} \frac{\partial \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \boldsymbol{\beta}^T} \\ \frac{\partial \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \alpha} \end{bmatrix}$$

$$I(\boldsymbol{y}_m) =$$

$$\begin{bmatrix} E \left(\frac{\partial^2 \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \boldsymbol{\beta}^T \partial \boldsymbol{\beta}} \right) & E \left(\frac{\partial^2 \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \boldsymbol{\beta}^T \partial \alpha} \right) \\ E \left(\frac{\partial^2 \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \alpha \partial \boldsymbol{\beta}^T} \right) & E \left(\frac{\partial^2 \ln L(\alpha, \boldsymbol{\beta} | y_1, y_2, \dots, y_n)}{\partial \alpha^2} \right) \end{bmatrix}$$

3. Menghitung $\boldsymbol{y}_{m+1} = \boldsymbol{y}_m + I^{-1}(\boldsymbol{y}_m) \boldsymbol{V}(\boldsymbol{y}_m)$, dengan $m = 0, 1, 2, \dots$
4. Menghitung $\|\boldsymbol{y}_{m+1} - \boldsymbol{y}_m\|$, jika $\|\boldsymbol{y}_{m+1} - \boldsymbol{y}_m\| < \varepsilon$ maka iterasi dihentikan
5. Mengulangi langkah 3 dan seterusnya dengan $m = m + 1$ hingga didapatkan $\|\boldsymbol{y}_{m+1} - \boldsymbol{y}_m\| < \varepsilon$.

2.4.3 Pengujian Signifikansi Parameter Model Regresi Gamma

Pengujian parameter dilakukan untuk mengetahui apakah variabel prediktor memberikan pengaruh terhadap variabel respons, baik secara serentak atau parsial. Statistik uji yang digunakan dalam pengujian secara serentak dilakukan menggunakan *Maximum Likelihood Ratio Test* (MLRT). Sedangkan pengujian secara parsial dilakukan menggunakan uji-Z.

a. Pengujian Serentak

Tujuan dari pengujian parameter secara serentak adalah untuk mengetahui apakah variabel prediktor dalam suatu model secara bersama-sama (simultan) berpengaruh signifikan terhadap variabel respons. Rumusan hipotesis dalam pengujian ini adalah sebagai berikut.

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

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

Statistik uji yang digunakan dalam pengujian serentak adalah MLRT, dimana pada metode ini dihitung *odd ratio* dari fungsi *likelihood* di bawah hipotesis *null* ($L(\hat{\omega})$) terhadap fungsi *likelihood* di bawah populasi ($L(\hat{\Omega})$). Himpunan parameter di bawah populasi dinyatakan dengan $\Omega = \{\beta_0, \beta_1, \beta_2, \dots, \beta_k, \alpha\} = \{\boldsymbol{\beta}, \alpha\}$ dan himpunan parameter di bawah hipotesis *null* dinyatakan dengan $\omega = \{\beta_0, \alpha_0\}$. Dengan demikian, fungsi $L(\Omega)$ adalah nilai maksimum *likelihood* untuk model lengkap. Sedangkan fungsi $L(\omega)$ merupakan nilai maksimum *likelihood* untuk model tanpa menyertakan variabel prediktor.

Odd ratio dapat diekspresikan dengan persamaan berikut.

$$\Lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

$$\ln \Lambda = \ln L(\hat{\omega}) - \ln L(\hat{\Omega}) \quad (2.17)$$

Fungsi *ln likelihood* di bawah populasi $L(\Omega)$ dalam Putri (2016) adalah

$$\begin{aligned} \ln L(\Omega) &= \ln \prod_{i=1}^n f(y_i | \Omega) \\ &= -n \ln \Gamma(\alpha) - \alpha \sum_{i=1}^n \mathbf{x}_i^T \boldsymbol{\beta} + \alpha n \ln \alpha + (\alpha - 1) \sum_{i=1}^n \ln y_i + B \end{aligned} \quad (2.18)$$

$$L(\hat{\Omega}) = \max_{\Omega} L(\Omega)$$

dimana:

$$B = \sum_{i=1}^n \frac{-\alpha y_i}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}$$

Sedangkan fungsi *ln likelihood* di bawah hipotesis *null* $L(\omega)$ dinyatakan dengan

$$\begin{aligned} \ln L(\omega) &= \ln \prod_{i=1}^n f(y_i | \omega) \\ &= -n \ln \Gamma(\alpha_0) - \alpha_0 \beta_0 + \alpha_0 n \ln \alpha_0 + (\alpha_0 - 1) \sum_{i=1}^n \ln y_i + D \end{aligned} \quad (2.19)$$

$$L(\hat{\omega}) = \max_{\omega} L(\omega)$$

dimana:

$$D = \sum_{i=1}^n \frac{-\alpha_0 y_i}{\exp(\beta_0)}$$

Nilai $L(\hat{\Omega})$ dan $L(\hat{\omega})$ dapat dihitung setelah didapatkan nilai parameter dari hasil estimasi menggunakan MLE dan optimasi *Fisher Scoring*.

Setelah estimasi parameter dari model lengkap dan model tanpa variabel prediktor didapatkan, maka statistik uji dapat dihitung dengan persamaan di bawah ini (Putri, 2016).

$$G^2 = -2 \ln \Lambda = -2 \ln L(\hat{\omega}) + 2 \ln L(\hat{\Omega}) \quad (2.20)$$

Nilai G^2 mengikuti distribusi χ^2 dengan derajat bebas k , dimana k merupakan banyak variabel prediktor. Hipotesis *null* (H_0) ditolak pada tingkat signifikan α apabila $G^2 > \chi^2_{\alpha, k}$.

b. Pengujian Parsial

Pengujian secara parsial dilakukan untuk mengetahui adanya pengaruh suatu variabel prediktor terhadap variabel respons. Pengujian ini dilakukan apabila pada pengujian parameter secara serentak diputuskan untuk menolak Hipotesis *null* (H_0). Keputusan tersebut menunjukkan bahwa paling tidak terdapat satu variabel prediktor yang signifikan terhadap variabel respons. Dengan demikian, pengujian parsial dilakukan pada masing-masing parameter dari setiap variabel. Pengujian secara parsial juga dilakukan untuk mengetahui apakah *intercept* dalam model (β_0) signifikan atau tidak terhadap variabel respons. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut.

$$H_0 : \beta_j = 0$$

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

Sedangkan statistik uji yang digunakan adalah nilai Z sebagai berikut.

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.21)$$

dimana $SE(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)}$ dan $Var(\hat{\beta}_j)$ merupakan elemen diagonal ke- $j+1$ dari matriks $-\mathbf{H}^{-1}(\hat{\mathbf{Y}})$. Penolakan terhadap hipotesis awal (H_0) terjadi jika $|Z_{hitung}| > Z_{\alpha/2}$.

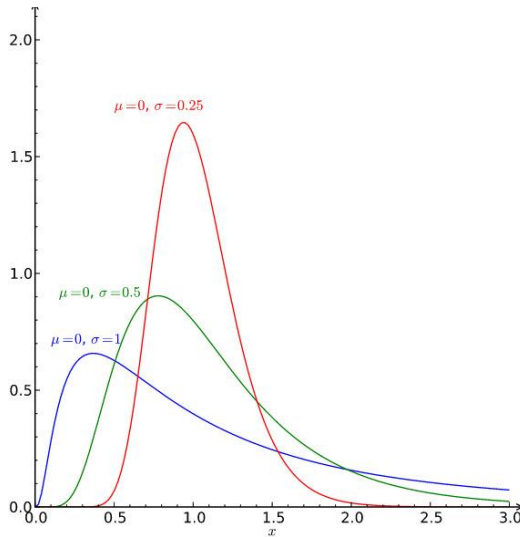
2.5 Regresi Lognormal

Disribusi lognormal merupakan bentuk distribusi normal dengan variabel randon yang ditransformasi ke bentuk *log*. Fungsi densitas dari distribusi lognormal adalah sebagai berikut.

$$f(y; \mu, \sigma,) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln y - \mu)^2}{2\sigma^2}\right);$$

$$y, \sigma > 0, -\infty < \mu < \infty \quad (2.22)$$

Pada persamaan di atas, parameter μ merupakan parameter lokasi, sedangkan parameter σ merupakan parameter skala. Kurva fungsi densitas distribusi lognormal dengan nilai parameter μ dan σ yang berbeda-beda ditunjukkan oleh Gambar 2.3 di bawah ini.



Gambar 2.3 Kurva PDF Distribusi Lognormal dengan Parameter Lokasi (μ) dan Skala (σ) yang Berbeda

Nilai harapan dan varians dari distribusi lognormal dengan parameter μ dan σ adalah sebagai berikut (Casella & Berger, 2002).

$$E(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad (2.23)$$

$$\begin{aligned} \text{Var}(Y) &= [\exp(\sigma^2) - 1] \exp(2\mu + \sigma^2) \\ &= \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2) \quad (2.24) \end{aligned}$$

2.5.1 Model Regresi Lognormal

Jika nilai μ dapat diestimasi menggunakan beberapa variabel prediktor (X_1, X_2, \dots, X_k) sehingga $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$ maka didapatkan fungsi densitas sebagai berikut

$$f(y_i; \boldsymbol{\beta}, \sigma) = \frac{1}{y_i \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2\sigma^2}\right) \quad (2.25)$$

Pada model regresi lognormal, nilai Y dapat diestimasi menggunakan persamaan (2.23).

2.5.2 Estimasi Parameter Model Regresi Lognormal

Estimasi parameter model regresi lognormal dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Fungsi *likelihood* yang dimaksimumkan untuk mendapatkan estimasi parameter adalah sebagai berikut.

$$\begin{aligned} L(\boldsymbol{\beta}, \sigma) &= \prod_{i=1}^n f(y_i) \\ &= \prod_{i=1}^n \left[\frac{1}{y_i \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2\sigma^2}\right) \right] \\ &= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(\sum_{i=1}^n -\frac{(\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2\sigma^2}\right) \frac{1}{\prod_{i=1}^n y_i} \quad (2.26) \end{aligned}$$

Sedangkan fungsi *ln likelihood* dari persamaan (2.26) adalah sebagai berikut.

$$\begin{aligned}
\ln L(\boldsymbol{\beta}, \sigma) &= \ln \prod_{i=1}^n f(y_i) \\
&= \ln \left[\frac{1}{\sigma^n (2\pi)^{n/2}} \exp \left(-\sum_{i=1}^n \frac{(\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2\sigma^2} \right) \frac{1}{\prod_{i=1}^n y_i} \right] \\
&= -n \ln \sigma - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 - \sum_{i=1}^n \ln(y_i)
\end{aligned} \tag{2.27}$$

Estimasi parameter dilakukan dengan menggunakan turunan parsial pertama dari persamaan (2.27) terhadap masing-masing parameter, yaitu σ dan $\boldsymbol{\beta}$, yang kemudian disamadengkan nol. Turunan parsial pertama fungsi *ln likelihood* terhadap parameter σ dan disamadengkan adalah sebagai berikut.

$$\begin{aligned}
\frac{\partial \ln L(\sigma, \boldsymbol{\beta})}{\partial \sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 = 0 \\
\sigma^2 &= \frac{1}{n} \sum_{i=1}^n (\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2
\end{aligned} \tag{2.28}$$

Sedangkan turunan parsial pertama fungsi *ln likelihood* terhadap parameter $\boldsymbol{\beta}^T$ dan disamadengkan nol ditunjukkan oleh persamaan di bawah ini.

$$\begin{aligned}
\frac{\partial \ln L(\alpha, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \mathbf{x}_i (\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta}) = 0 \\
\boldsymbol{\beta} &= (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \ln \mathbf{Y}
\end{aligned} \tag{2.29}$$

2.5.3 Pengujian Signifikansi Paramater Regresi

Lognormal

Pengujian signifikansi parameter dilakukan secara serentak dan parsial. Hipotesis yang diuji dalam pengujian ini adalah sebagai berikut.

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

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

Pengujian signifikansi parameter model regresi lognormal secara serentak dilakukan menggunakan *Maximum Likelihood Ratio Test* (MLRT), dimana pada metode ini dihitung *odd ratio* dari fungsi *likelihood* di bawah hipotesis *null* ($L(\hat{\omega})$) terhadap fungsi *likelihood* di bawah populasi ($L(\hat{\Omega})$). Himpunan parameter di bawah populasi dinyatakan dengan $\Omega = \{\beta_0, \beta_1, \beta_2, \dots, \beta_k, \sigma\} = \{\boldsymbol{\beta}, \sigma\}$ dan himpunan parameter di bawah hipotesis *null* dinyatakan dengan $\omega = \{\beta_0, \sigma_0\}$. Fungsi *ln likelihood* di bawah populasi $L(\Omega)$ adalah

$$\begin{aligned} \ln L(\Omega) &= \ln \prod_{i=1}^n f(y_i | \Omega) \\ &= -n \ln \sigma - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (\ln y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 - \sum_{i=1}^n \ln(y_i) \end{aligned} \quad (2.30)$$

$$L(\hat{\Omega}) = \max_{\Omega} L(\Omega)$$

Sedangkan fungsi *ln likelihood* di bawah hipotesis *null* $L(\omega)$ dinyatakan dengan

$$\begin{aligned}
\ln L(\omega) &= \ln \prod_{i=1}^n f(y_i|\omega) \\
&= -n \ln \sigma_0 - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma_0^2} \sum_{i=1}^n (\ln y_i - \beta_0)^2 - \sum_{i=1}^n \ln(y_i)
\end{aligned} \tag{2.31}$$

$$L(\hat{\omega}) = \max_{\omega} L(\omega)$$

Setelah estimasi parameter dari model lengkap dan model tanpa variabel prediktor didapatkan, maka statistik uji dapat dihitung dengan persamaan (2.20).

Sedangkan hipotesis yang diuji dalam pengujian signifikansi parameter secara parsial adalah sebagai berikut.

$$H_0 : \beta_j = 0$$

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

Sementara statistik uji yang digunakan adalah nilai Z seperti yang dirumuskan dalam persamaan (2.21).

2.6 Interpretasi Model

Secara umum, model yang digunakan dalam penelitian ini adalah *generalized linear model* dengan fungsi link ($g(\mu_i)$) yang digunakan adalah *log*. Jika variabel yang mempengaruhi variabel respons Y adalah X_1, X_2, \dots, X_k , maka nilai μ untuk Y pada observasi ke- i adalah $\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$ atau $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$ (Dobson & Barnett, 2008). Pada model tersebut, untuk variabel dengan skala rasio atau interval setiap penambahan nilai variabel X_l sebesar satu satuan akan menyebabkan variabel respons menjadi lebih besar atau lebih kecil sebesar $\exp(\beta_1)$ satu satuan. Jika koefisien β_1 bernilai positif maka variabel respons akan lebih besar dan jika

koefisien β_1 bernilai negatif maka nilai variabel respons akan lebih kecil. Hal yang sama berlaku untuk $\beta_2, \beta_3, \dots, \beta_k$.

Jika variabel prediktor berskala kategorik maka sebelum dilakukan regresi perlu dibentuk variabel *dummy*. Pada pembentukan variabel *dummy* tersebut, salah satu kategori atau kelas dijadikan sebagai basis sehingga terdapat *dummy* sebanyak $r-1$ dimana r merupakan banyaknya kategori atau kelas. Dengan begitu maka akan terbentuk koefisien regresi sebanyak $r-1$ untuk satu variabel kategorik tersebut. Misal terdapat suatu variabel yang terdiri dari kategori 1, 2, dan 3 dengan kategori 3 dijadikan sebagai basis, maka diperoleh dua variabel *dummy* dengan koefisien regresi β_{d1} dan β_{d2} . Jika variabel tersebut kategori 1 maka variabel respons akan lebih besar atau lebih kecil sebanyak $\exp(\beta_{d1})$ kali dibandingkan dengan kategori 3.

2.7 Kriteria Pemilihan Model Terbaik

Pemilihan model terbaik dapat dilakukan menggunakan nilai AIC (*Akaike Information Criterion*). AIC pertama kali diperkenalkan oleh Hirotugu Akaike melewati *proceeding* berjudul “*Information Theory and an Extension of the Maximum Likelihood Principles*” yang dipublikasikan secara formal pada tahun 1974. Sejak saat itu, AIC merupakan kriteria pemilihan model yang banyak digunakan. Nilai AIC dirumuskan dalam persamaan berikut.

$$AIC = 2(k) - 2 \ln L(\hat{\beta}) \quad (2.32)$$

dimana k merupakan banyak parameter dalam model dan $L(\hat{\beta})$ merupakan fungsi *maximum likelihood* dari model. Pada penggunaan AIC sebagai kriteria pemilihan model terbaik, model dugaan terbaik merupakan model yang memiliki nilai AIC minimum.

Jika ukuran sampel kecil, terdapat kemungkinan bahwa AIC akan menunjukkan model dengan jumlah parameter yang lebih banyak sebagai model yang terbaik (Claeskens & Hjort, 2008). Oleh karena itu dikembangkan AICc untuk mengatasi AIC yang *overfitting* dimana AICc tersebut merupakan AIC yang disesuaikan untuk sampel berukuran kecil. Jika n adalah besar sampel dan k adalah banyak parameter dalam model maka nilai AICc adalah sebagai berikut.

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1} \quad (2.33)$$

2.8 *Danube River Basin*

Sungai Danube merupakan sungai terpanjang kedua di Eropa setelah Sungai Volga. Sungai tersebut mempunyai total area sebesar 801.463 km². Daerah aliran Sungai Danube meliputi 19 negara di daratan Eropa. Sungai Danube mengalir dari Black Forest di Jerman dan berakhir di Laut Hitam. Peta daerah aliran Sungai Danube ditunjukkan oleh Gambar 2.4. Negara utama di dalam daerah aliran Sungai Danube yaitu: Jerman, Austria, Republik Ceko, Slovakia, Hungaria, Slovenia, Kroasia, Bosnia dan Herzegovina, Serbia, Montenegro, Romania, Bulgaria, Moldova, dan Ukraina. Terdapat lebih dari 2.000 km² area Sungai Danube yang terdapat di masing-masing negara tersebut (ICPDR, 2009).

Daerah aliran Sungai Danube dapat dibedakan menjadi tiga *sub-region*, yaitu *upper*, *middle* dan *lower basin*. *Upper Basin* membentang dari Jerman bagian barat ke Bratislava di Slovakia. Bagian *Upper* memiliki arus yang berkisar antara 8 hingga 9 km/jam. Anak sungai pada bagian tersebut membawa aliran dari Pegunungan Alpens bagian utara dan *Central European Highlands* bagian selatan. Kedalaman sungai pada

bagian ini adalah antara 1 hingga 8 meter. Kemudian, dari Bratislava sampai Iron Gate Gorge di perbatasan antara Serbia dan Romania merupakan *Middle Basin*. Bagian tersebut merupakan bagian yang terbesar dari daerah aliran Sungai Danube. Sedangkan *Lower Basin* terdiri dari dataran rendah, dataran tinggi dan pegunungan di Romania dan Bulgaria. Sebelum mencapai Laut Hitam, Sungai Danube terbagi menjadi tiga cabang utama yang membentuk Danube Delta, yang meliputi area seluas 6.750 km² (ICPDR, 2009).



Gambar 2.4 Peta Daerah Aliran Sungai Danube
Sumber: icpdr.org

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data hasil ekspedisi penelitian sungai Danube di Eropa pada tahun 2013 yang dikenal sebagai *Joint Danube Survey 3* (JDS3). JDS3 dikoordinasikan oleh *International Commission for the Protection of the Danube River* (ICPDR), dimana ekspedisi tersebut merupakan kerjasama antara 14 negara yang menjadi daerah aliran Sungai Danube (*Danube River Basin*) dan *European Commission*. Ekspedisi tersebut berlangsung selama enam minggu, dari Agustus hingga September 2013. Pengambilan data dilakukan di sungai utama dan juga di anak sungai. Data yang digunakan dalam penelitian ini merupakan data yang diambil dari stasiun-stasiun pengamatan yang ditunjukkan dalam Lampiran 2. Data JDS3 dapat diakses di *website Danube River Basin Water Quality Database* dengan alamat <http://www.icpdr.org/wq-db>.

3.2 Variabel Penelitian dan Struktur Data

Pada penelitian ini, regresi dilakukan dengan menggunakan satu variabel respons dan 5 variabel prediktor. Variabel respons (Y) dalam penelitian ini adalah biomassa klorofil- α pada *phytoplankton* di Sungai Danube, Eropa. Sedangkan variabel prediktor yang digunakan untuk memodelkan variabel respons tersebut adalah *physico-chemical* air sungai yang meliputi nitrogen total (TN), fosfor total (TP), temperatur, derajat keasaman (pH), dan konduktivitas. Uraian mengenai masing-masing variabel dalam penelitian ini ditunjukkan oleh Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Kode	Nama Variabel	Satuan	Skala
Y	Biomassa klorofil- α	$\mu\text{g/L}$	Rasio
X_1	Nitrogen total (TN)	mg/L	Rasio
X_2	Fosfor total (TP)	mg/L	Rasio
X_3	Temperatur	$^{\circ}\text{C}$	Interval
X_4	Derajat keasaman (pH)	-	Interval
X_5	Konduktivitas	$\mu\text{S/cm}$	Rasio

Sedangkan definisi operasional dari masing-masing variabel tersebut adalah sebagai berikut.

- Biomassa klorofil- α** : konsentrasi klorofil- α dalam satu liter air yang digunakan untuk mengukur biomassa fitoplankton dimana ukuran ini bukan merupakan hasil reaksi kimia variabel prediktor.
- Nitrogen total (TN)** : jumlahan unsur nitrogen (N) dalam nitrat (NO_3), nitrit (NO_2), ammonia (NH_3), dan nitrogen dalam senyawa organik per liter air.
- Fosfor total (TP)** : jumlahan unsur fosfor (P) dalam berbagai bentuk senyawa yang mengandung fosfor per satu liter air.
- Temperatur** : suhu air sungai yang diukur di setiap lokasi pengamatan
- Derajat keasaman (pH)** : ukuran yang digunakan untuk menyatakan derajat keasaman, semakin kecil nilai pH maka air sungai semakin asam, jika $\text{pH} = 7$ maka air sungai bersifat netral, dan jika di atas 7 maka air sungai bersifat basa.
- Konduktivitas** : ukuran kemampuan air sungai dalam menghantarkan arus listrik.

Berdasarkan variabel pada Tabel 3.2, struktur data dalam penelitian ini adalah sebagai berikut.

Tabel 3.2 Struktur Data

Lokasi ke- i	X_1	X_2	X_3	X_4	X_5	Y
1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	y_1
2	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	y_2
3	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	y_3
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
n	x_{n1}	x_{n2}	x_{n3}	x_{n4}	x_{n5}	y_n

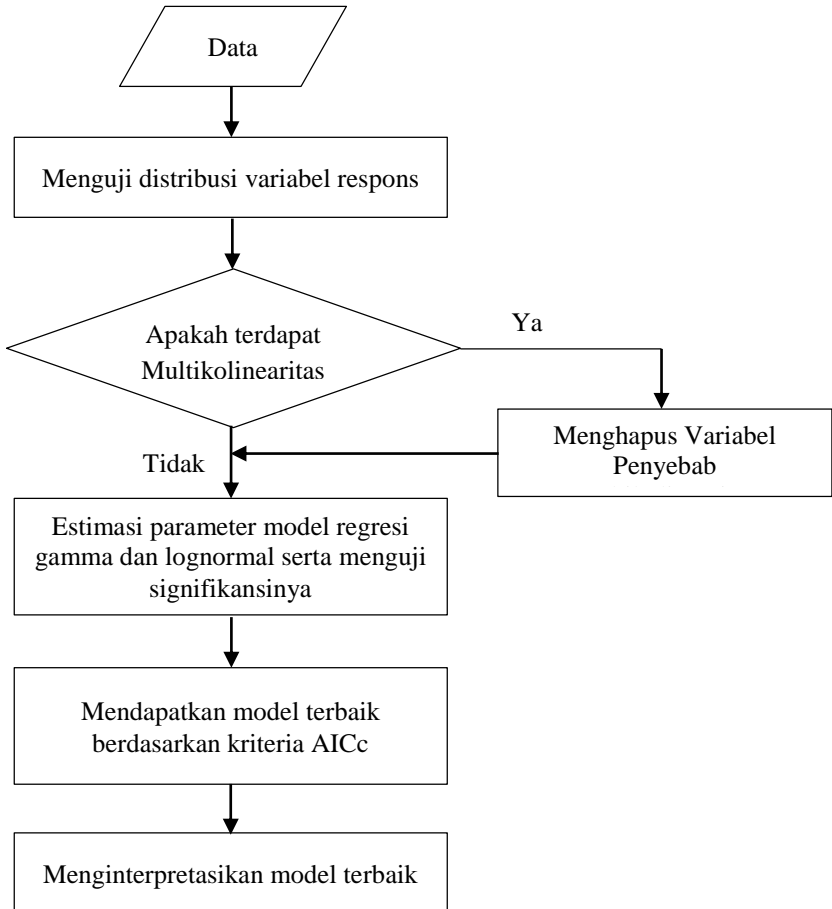
3.3 Langkah-Langkah Analisis

Tahapan analisis dalam penelitian ini setelah didapatkan data seperti pada adalah sebagai berikut.

1. Menguji distribusi variabel respons, yaitu biomassa klorofil- α pada *phytoplankton* menggunakan statistik uji pada persamaan (2.3).
2. Mendeteksi adanya multikolinearitas dalam variabel prediktor menggunakan kriteria VIF. Jika terdapat multikolinearitas, maka perlu dilakukan penanganan terlebih dahulu sebelum dilakukan regresi.
3. Melakukan estimasi parameter dengan memaksimumkan fungsi pada persamaan (2.14) untuk model regresi gamma dan memaksimumkan fungsi pada persamaan (2.27) untuk model regresi lognormal.
4. Menguji signifikansi parameter secara serentak dengan menggunakan statistik uji pada persamaan (2.20) dan pengujian secara serentak parsial dengan statistik uji pada persamaan (2.21) untuk setiap model regresi.
5. Menghitung nilai AICc model regresi gamma dan lognormal, sehingga dapat dibandingkan model mana yang lebih baik diantara kedua model tersebut.

6. Menginterpretasikan model terbaik yang didapatkan.

Berdasarkan uraian diatas, langkah-langkah analisis pada peneitian ini dapat dijelaskan oleh diagram alir di bawah ini.



Gambar 3.1 Diagram Alir Analisis Data

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Pencemaran Nutrisi di Sungai Danube

Sungai Danube merupakan sungai besar yang melewati 19 negara di daratan Eropa. Sungai ini terbagi ke dalam 3 *sub-region*, yaitu *upper*, *middle* dan *lower region* sebagaimana yang dijelaskan dalam subbab 2.6. Gambar 4.1 menunjukkan rata-rata biomassa klorofil- α , total nitrogen, total fosfor serta rasio nitrogen terhadap fosfor pada masing-masing *sub-region* dan Sungai Danube secara keseluruhan.



Gambar 4.1 Rata-Rata Biomassa Klorofil- α , Nitrogen dan Fosfor

Gambar 4.1 memperlihatkan bahwa rata-rata biomassa klorofil- α di bagian *upper* lebih tinggi dari bagian lainnya dan Sungai Danube secara keseluruhan. Begitu juga dengan rata-rata total nitrogen dan rasio antara total nitrogen dengan fosfor (N/P). Namun, sungai bagian *lower*, yaitu sungai yang mengalir setelah *Iron Gate* hingga ke Laut Hitam, mempunyai rata-rata total fosfor yang lebih tinggi dibandingkan *sub-region* lain.

Tabel 4.1 Statistika Deskriptif Sungai Bagian *Upper*

Variabel	Mean	Stdev	Min	Max
Biomassa Klorofil- α	8,40	9,28	1,10	31,00
Nitrogen	2,12	0,55	1,50	3,00
Fosfor	0,07	0,02	0,05	0,10
Temperatur	20,60	1,07	18,00	22,00
pH	8,18	0,19	7,90	8,50
Konduktivitas	426,70	81,40	320,00	580,00

Sungai bagian *upper* mempunyai rata-rata biomassa klorofil- α sebesar 8,40 $\mu\text{g/L}$. Biomassa klorofil- α paling tinggi terdapat di stasiun pengamatan JDS12 atau di Austria, yaitu Sungai Morava pada kilometer ke-0,08. Pada stasiun tersebut, biomassa klorofil- α yang terukur adalah sebesar 31 $\mu\text{g/L}$. Sementara kandungan total nitrogen dan total fosfor yang terukur di lokasi tersebut adalah sebesar 1,50 mg/L dan 0,06 mg/L secara berurutan. Nilai tersebut tidak lebih tinggi dari rata-rata total nitrogen dan fosfor di bagian *upper*. Selain JDS12, nilai biomassa klorofil- α yang terukur lebih tinggi dari rata-rata juga dijumpai di stasiun JDS9, JDS10 dan JDS11 yang terletak di Klostemeuburg, hulu Morava, dan Wildungsmayer. Pada ketiga stasiun tersebut, total nitrogen dan fosfor yang terukur juga tidak lebih tinggi dari nilai rata-rata sungai secara keseluruhan. Total nitrogen paling tinggi di bagian *upper* terukur paling tinggi di stasiun JDS4 atau di Deggendorf, Jerman. Pada stasiun tersebut, total nitrogen yang terukur adalah sebesar 3 mg/L dan biomassa klorofil- α yang terukur adalah sebesar 1,1 $\mu\text{g/L}$.

Pada bagian *upper*, suhu air sungai berkisar antara 18°C hingga 22°C dengan rata-rata 20,60 \pm 1,07 °C. Suhu paling tinggi terukur di stasiun JDS9 dan paling rendah di JDS6 (Jochenstein, Jerman). Sementara derajat keasaman yang terukur di bagian *upper* rata-rata adalah sebesar 8,18 dengan

nilai terendah yang terukur adalah 7,9 dan yang tertinggi adalah 8,5. Dibandingkan dengan *region* yang lain, air sungai di bagian *upper* lebih basa. Nilai konduktivitas bagian *upper* paling tinggi adalah sebesar 580 $\mu\text{S/cm}$ dan paling rendah adalah sebesar 320 $\mu\text{S/cm}$. Sedangkan rata-rata konduktivitas di bagian *upper* adalah sebesar $426,7 \pm 81,4 \mu\text{S/cm}$.

Tabel 4.2 Statistika Deskriptif Sungai Bagian *Middle*

Variabel	Mean	Stdev	Min	Max
Biomassa Klorofil- α	8,22	4,66	1,50	20,00
Nitrogen	1,56	0,33	0,69	2,00
Fosfor	0,08	0,02	0,05	0,11
Temperatur	21,48	0,96	20,00	24,00
pH	8,11	0,18	7,60	8,30
Konduktivitas	391,90	32,52	295,00	476,00

Pada bagian *middle region*, rata-rata biomassa klorofil- α yang terukur adalah sebesar 8,22 $\mu\text{g/L}$. Biomassa klorofil- α terukur paling tinggi di stasiun JDS18 yang berlokasi di Vah kilometer ke-8. Nilai yang terukur adalah sebesar 20 $\mu\text{g/L}$. Pada stasiun tersebut, total nitrogen dan fosfor yang terukur adalah sebesar 1,50 mg/L dan 0,11 mg/L . Nilai biomassa klorofil- α yang terukur lebih tinggi dari rata-rata di bagian *middle* juga dijumpai di Rackeve-Soroksar Danube pada kilometer ke-59, hulu Sungai Drava, di bawah jembatan Ilok-Backa Palanka, hulu dan hilir Sungai Novi-Sad, serta di hulu Sungai Tisa. Total fosfor yang terukur di lokasi pengamatan tersebut sama atau lebih besar dari rata-rata yang terukur di *middle region*. Selain di Rackeve-Soroksar Danube rkm-59, total nitrogen yang terukur pada lokasi-lokasi tersebut lebih tinggi dari nilai rata-rata di *middle region*.

Total nitrogen yang terukur di *middle region* rata-rata adalah sebesar $1,56 \pm 0,33 \text{ mg/L}$. Nilai paling tinggi terukur di

stasiun JDS16 yang terletak di Hungaria dengan nilai sebesar 2 mg/L dan paling rendah terukur di Sungai Tisa kilometer pertama (JDS35) dan Sungai Sava kilometer ke-7 (JDS37). Sedangkan total fosfor yang terukur di *middle region* rata-rata adalah sebesar $0,08 \pm 0,02$ mg/L dengan nilai terendah adalah sebesar 0,05 mg/L dan nilai tertinggi adalah sebesar 0,11 mg/L. Total fosfor paling rendah terukur di Sungai Sava kilometer ke-7 (JDS37), hulu Pancevo (JDS38), hilir Pancevo (JDS39), dan hulu Velika Morava (JDS40). Sedangkan total fosfor paling tinggi terukur di Vah kilometer ke-0,8 (JDS18) dan hilir Drava (JDS30).

Rata-rata suhu air sungai di bagian *middle* adalah sebesar $21,48 \pm 0,96^\circ\text{C}$. Suhu air sungai di bagian ini lebih hangat dibandingkan dengan *region* lain. Suhu terendah yang terukur di bagian *middle* adalah sebesar 20°C dan paling tinggi adalah 24°C . Sementara berdasarkan derajat keasaman, air sungai di *region* ini lebih asam dibandingkan dengan *region* lain. Nilai pH yang terukur di *middle region* paling rendah adalah 7,6 dan paling tinggi adalah 8,3. Konduktivitas air sungai di *middle region* juga lebih rendah dari *region* lain. Rata-rata konduktivitas air sungai yang terukur di *middle region* adalah sebesar $391,9 \pm 32,52$ $\mu\text{S/cm}$.

Tabel 4.3 Statistika Deskriptif Sungai Bagian *Lower*

Variabel	Mean	Stdev	Min	Max
Biomassa Klorofil- α	3,26	1,59	0,33	5,70
Nitrogen	1,68	1,29	1,10	5,80
Fosfor	0,12	0,14	0,03	0,55
Temperatur	20,22	1,68	17,00	23,00
pH	8,11	0,15	7,80	8,40
Konduktivitas	435,57	111,32	375,00	822,00

Biomassa klorofil- α yang terukur di *lower region* merupakan yang terendah diantara ketiga *region*. Rata-rata biomassa klorofil- α yang terukur pada bagian ini adalah sebesar 3,26 $\mu\text{g/L}$ dengan nilai terendah yang terukur adalah sebesar 0,33 $\mu\text{g/L}$ dan yang tertinggi adalah sebesar 5,7 $\mu\text{g/L}$. Nilai yang tertinggi tersebut bahkan lebih kecil dari rata-rata yang terukur di bagian *upper* dan *middle*. Biomassa tertinggi terukur di stasiun JDS60 di Silistra, Romania. Pada stasiun tersebut, total nitrogen dan fosfor yang terukur tidak lebih besar dari nilai rata-rata yang terukur di bagian ini, yaitu sebesar 1,20 mg/L dan 0,06 mg/L.

Total nitrogen yang terukur di *lower region* rata-rata adalah sebesar $1,68 \pm 1,29$ mg/L dengan nilai terendah adalah sebesar 1,1 mg/L dan nilai tertinggi adalah 5,8 mg/L. Nilai paling tinggi terukur di stasiun JDS58 (Arges, Romania) dan paling rendah terukur di hilir Olt (JDS52), hilir Jantra (JDS55), hilir Ruse (JDS57) dan di cabang Vilkova-Chilia yang terletak di delta (JDS66). Jika tanpa mengikutsertakan pengukuran di stasiun JDS56 dan JDS58, dimana di stasiun tersebut total nitrogen yang terukur sangat berbeda dengan stasiun-stasiun lain, total nitrogen di *lower region* akan memiliki keragaman yang kecil.

Sedangkan total fosfor yang terukur di *lower region* rata-rata adalah sebesar $0,12 \pm 0,14$ mg/L dengan nilai terendah adalah sebesar 0,03 mg/L dan nilai tertinggi adalah sebesar 0,55 mg/L. Total fosfor paling rendah terukur di hilir Timok (JDS47). Sedangkan total fosfor paling tinggi terukur di stasiun JDS58. Selain di stasiun JDS58, total fosfor yang sangat tinggi juga terukur di stasiun JDS54, JDS56, JDS58, dan JDS64. Berdasarkan analisis tersebut, total nitrogen dan total fosfor

paling tinggi terukur di stasiun JDS58. Biomassa klorofil- α yang terukur di JDS58 adalah sebesar 3,3 $\mu\text{g/L}$.

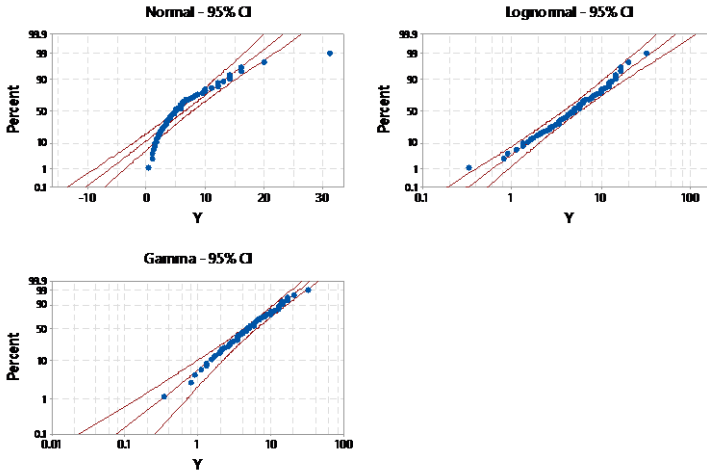
Rata-rata suhu air sungai di *lower region* adalah sebesar $20,22 \pm 1,68^\circ\text{C}$. Suhu air sungai di bagian ini cenderung lebih dingin dibandingkan dengan *region* lain. Suhu terendah yang terukur di bagian *middle* adalah sebesar 17°C dan paling tinggi adalah 23°C . Sementara nilai pH yang terukur di *lower region* rata-rata adalah 8,11 dengan nilai paling rendah adalah 7,8 dan paling tinggi adalah 8,4. Konduktivitas air sungai di *lower region* cenderung lebih tinggi dari *region* lain. Rata-rata konduktivitas air sungai yang terukur di *region* ini adalah sebesar 435,57 $\mu\text{S/cm}$ dengan nilai paling rendah adalah sebesar 375 $\mu\text{S/cm}$ dan yang paling tinggi adalah 822 $\mu\text{S/cm}$.

Salah satu bioindikator pencemaran nutrisi adalah *phytoplankton*. Pada penelitian ini, keberadaan *phytoplankton* di Sungai Danube diukur menggunakan biomassa klorofil- α . Secara teori, semakin tinggi kandungan nutrisi seperti nitrogen dan fosfor dalam suatu perairan maka kelimpahan *phytoplankton* di perairan tersebut juga semakin banyak, yang berarti biomassa klorofil- α dari *phytoplankton* di perairan tersebut juga semakin besar. Namun, data yang digunakan dalam penelitian ini tidak menunjukkan hal yang demikian secara konsisten.

4.2 Pengujian Distribusi

Pada penelitian ini, analisis regresi digunakan untuk menganalisis pengaruh dari beberapa variabel tertentu terhadap biomassa klorofil- α di Sungai Danube. Sebelum dilakukan analisis, distribusi dari variabel respons diuji terlebih dahulu untuk menentukan fungsi yang sesuai dalam estimasi parameter. Gambar 4.2 menunjukkan *fitting* data biomassa

klorofil- α terhadap distribusi normal, distribusi lognormal, dan distribusi gamma.



Gambar 4.2 *Fitting* Biomassa Klorofil- α terhadap Distribusi Normal, Lognormal, dan Gamma

Gambar 4.2 menunjukkan bahwa hasil *fitting* distribusi normal menghasilkan beberapa *dot* (warna biru di bagian kanan dan kiri) yang berada di luar batas interval (ditunjukkan oleh garis berwarna merah). Sementara *fitting* distribusi lognormal dan gamma menghasilkan *dot* yang tersebar di dalam batas interval. Hasil pengujian distribusi ditunjukkan oleh Tabel 4.4 di bawah ini.

Tabel 4.4 Hasil Uji Distribusi Variabel Respons

Distribusi	AD	<i>P-value</i>	Keputusan
Normal	2,641	< 0,005	Tolak H_0
Lognormal	0,235	0,784	Gagal tolak H_0
Gamma	0,217	> 0,250	Gagal tolak H_0

Tabel 4.4 menunjukkan bahwa variabel biomassa klorofil- α mengikuti distribusi lognormal dan distribusi gamma.

Hal tersebut ditunjukkan oleh p -value dari hasil pengujian masing-masing distribusi yang lebih tinggi dari tingkat signifikan 0,05. Tingginya p -value tersebut menyebabkan hipotesis *null* pengujian distribusi gagal ditolak pada tingkat signifikan 5%. Hasil estimasi parameter untuk masing-masing distribusi ditunjukkan oleh Tabel 4.5.

Tabel 4.5 Estimasi Parameter Distribusi

Distribusi	Lokasi	Bentuk	Skala
Lognormal	1,53573	-	0,86990
Gamma	-	1,66105	3,89053

4.3 Pemeriksaan Multikolinearitas

Salah satu asumsi yang harus dipenuhi dalam analisis regresi adalah tidak adanya korelasi antar variabel prediktor atau disebut multikolinearitas. Pada penelitian ini pemeriksaan terhadap multikolinearitas dilakukan dengan mengguakan nilai *Variance Inflation Factors* (VIF). Nilai VIF untuk masing-masing variabel prediktor dalam penelitian ini ditunjukkan oleh Tabel 4.6.

Tabel 4.6 VIF Setiap Variabel Prediktor

Variabel	VIF
TN	2,903
TP	3,004
Temperatur	1,456
pH	1,185
Konduktivitas	1,794

Menurut Kleinbaum, Kupper dan Muller dalam Gujarati & Porter (2008), apabila nilai VIF_j lebih besar dari 10 maka terdapat indikasi bahwa terdapat kolinearitas pada variabel prediktor ke- j . Tabel 4.2 menunjukkan bahwa nilai VIF untuk masing-masing prediktor dalam penelitian ini lebih kecil dari 10. Dengan demikian, diketahui bahwa tidak terdapat

multikolinearitas dalam data sehingga analisis regresi dapat dilakukan untuk mengetahui faktor-faktor yang mempengaruhi biomassa klorofil- α .

4.4 Analisis Regresi Gamma

Hasil pengujian distribusi menunjukkan bahwa biomassa klorofil- α berdistribusi gamma. Pemodelan regresi gamma dilakukan untuk mengetahui adanya pengaruh variabel-variabel prediktor terhadap biomassa klorofil- α . Estimasi parameter model dilakukan dengan menggunakan *Maximum Likelihood Estimation* (MLE) dan optimasi numerik dengan metode *Fisher Scoring*. Pada metode MLE, estimasi dilakukan dengan memaksimalkan fungsi *Inlikelihood* dari fungsi densitas distribusi gamma yang ditunjukkan oleh persamaan (2.14). Sementara optimasi numerik diperlukan dalam estimasi parameter karena turunan kedua dari persamaan (2.14) terhadap masing-masing parameter bukan merupakan persamaan yang *closed-form*.

Setelah didapatkan model, dilakukan pengujian signifikansi parameter secara serentak untuk mengetahui apakah variabel-variabel yang diduga mempengaruhi biomassa klorofil- α berpengaruh secara signifikan. Pengujian signifikansi parameter dilakukan dalam dua tahap. Tahap pertama adalah pengujian serentak untuk mengetahui apakah variabel-variabel prediktor mempengaruhi biomassa klorofil- α secara serentak. Pengujian secara serentak dilakukan dengan menggunakan statistik uji *Maximum Likelihood Ratio Test* (MLRT) seperti pada persamaan (2.20). Apabila hasil pengujian serentak menunjukkan bahwa setidaknya terdapat satu variabel prediktor yang secara signifikan berpengaruh terhadap variabel respons maka pengujian dilanjutkan ke tahap kedua untuk mengetahui

variabel-variabel tertentu yang berpengaruh signifikan terhadap biomassa klorofil- α . Hasil pengujian signifikansi parameter secara serentak ditunjukkan oleh Tabel 4.7.

Tabel 4.7 Hasil Pengujian Serentak Model Regresi Gamma

G^2	Df	$\chi^2_{0,05;5}$	Keputusan
26,0633	5	11,070	Tolak H_0

Tabel 4.7 menunjukkan nilai G^2 lebih besar dari nilai tabel χ^2 pada tingkat signifikan 0,05 dan derajat bebas 5 sehingga hipotesis awal yang menyatakan bahwa parameter regresi sama dengan nol atau tidak signifikan dapat ditolak. Dengan demikian, dapat disimpulkan bahwa secara serentak variabel-variabel prediktor berpengaruh signifikan terhadap biomassa klorofil- α .

Tabel 4.8 Hasil Estimasi dan Pengujian Paramater Model Regresi Gamma Secara Parsial

Parameter	Koefisien	SE	Z	<i>p-value</i>	Keputusan
Konstan	-25,4199	5,0561	5,0276	$1,29 \cdot 10^{-6}$	Tolak H_0
TN	-0,1373	0,1650	0,8321	0,2822	Gagal tolak H_0
TP	0,9487	1,7025	0,5572	0,3416	Gagal tolak H_0
Temperatur	0,2962	0,0711	4,1670	$6,7710 \cdot 10^{-5}$	Tolak H_0
pH	2,5567	0,5184	4,9316	$2,09 \cdot 10^{-6}$	Tolak H_0
Konduktivitas	0,0009	0,0014	0,6515	0,3227	Gagal Tolak H_0

Karena secara serentak variabel-variabel prediktor berpengaruh signifikan terhadap biomassa klorofil- α , yang artinya terdapat minimal satu prediktor yang signifikan, maka pengujian dilanjutkan ke tahap pengujian secara parsial. Hasil estimasi parameter dan pengujian secara parsial dapat dilihat pada Tabel 4.8. Pada tabel tersebut dapat dilihat bahwa variabel yang secara signifikan berpengaruh terhadap biomassa klorofil- α adalah suhu air sungai dan derajat keasaman (pH). Sementara variabel yang berkaitan dengan pencemaran nutrisi, seperti

konsentrasi nitrogen total dan fosfor total dalam air sungai tidak berpengaruh signifikan terhadap biomassa klorofil- α dari fitoplankton di Sungai Danube.

4.5 Analisis Regresi Lognormal

Parameter model regresi diestimasi menggunakan metode *Maximum Likelihood Estimation* (MLE). Setelah didapatkan estimasi parameter, dilakukan pengujian parameter secara serentak menggunakan *Maximum Likelihood Ratio Test* (MLRT). Hasil pengujian secara serentak ditunjukkan oleh Tabel 4.9.

Tabel 4.9 Hasil Pengujian Serentak Model Regresi Lognormal

G^2	Df	$\chi^2_{0,05;5}$	Keputusan
23,843	5	11,070	Tolak H_0

Nilai hitung statistik uji G^2 lebih besar dari nilai χ^2 pada tingkat signifikan 0,05 dan derajat bebas 5. Hal tersebut mengakibatkan terjadinya penolakan terhadap hipotesis awal pada tingkat signifikan 5%. Dengan demikian, disimpulkan bahwa terdapat minimal satu variabel prediktor dalam model regresi lognormal yang mempengaruhi biomassa klorofil- α secara signifikan.

Tabel 4.10 Hasil Estimasi dan Pengujian Paramater Model Regresi Lognormal Secara Parsial

Parameter	Koefisien	SE	Z	<i>p-value</i>	Keputusan
Konstan	-23,8504	4,8754	4,8920	$2,54 \cdot 10^{-6}$	Tolak H_0
TN	-0,1005	0,1905	0,5276	0,3471	Gagal tolak H_0
TP	-3,3379	2,4343	1,3712	0,1558	Gagal tolak H_0
Temperatur	0,2236	0,0675	3,3124	0,0017	Tolak H_0
Ph	2,4055	0,5308	4,5321	$1,38 \cdot 10^{-5}$	Tolak H_0
Konduktivitas	0,0045	0,0012	3,8903	0,0002	Tolak H_0

Setelah diketahui bahwa variabel-variabel prediktor secara serentak berpengaruh signifikan terhadap biomassa klorofil- α , maka selanjutnya dilakukan pengujian parameter secara parsial untuk mengetahui variabel tertentu yang berpengaruh signifikan terhadap respons. Hasil pengujian parameter regresi secara parsial ditunjukkan oleh Tabel 4.10. Berdasarkan hasil pengujian signifikansi parameter secara parsial, variabel yang secara signifikan berpengaruh terhadap biomassa klorofil- α dari fitoplankton di Sungai Danube pada tingkat signifikan 5% adalah variabel suhu air sungai, pH dan konduktivitas air sungai.

4.6 Pemilihan Model Terbaik

Setelah dilakukan pemodelan dengan menggunakan regresi gamma dan lognormal, dilakukan perhitungan nilai AICc untuk mengetahui model yang terbaik diantara kedua model yang terbentuk. Semakin kecil nilai AICc maka model yang dihasilkan semakin baik. Hasil perhitungan AICc untuk masing-masing model yang terbentuk ditunjukkan oleh Tabel 4.11.

Tabel 4.11 Nilai AICc Setiap Model

Model	AICc
Gamma	348,227
Lognormal	389,145

Tabel 4.11 menunjukkan nilai AICc yang terkecil dimiliki oleh model regresi gamma. Hal tersebut menunjukkan bahwa model regresi gamma lebih baik dari model regresi lognormal untuk kasus biomassa klorofil- α di Sungai Danube.

4.7 Interpretasi Model Terbaik

Berdasarkan nilai AICc, diketahui bahwa model gamma menghasilkan model dengan performansi yang lebih baik dari model regresi lognormal dalam memodelkan faktor-faktor yang mempengaruhi biomassa klorofil- α di Sungai Danube Eropa. Menurut hasil pengujian parameter pada model regresi gamma seperti yang ditunjukkan oleh Tabel 4.8, variabel yang secara signifikan mempengaruhi biomassa klorofil- α dari fitoplankton di Sungai Danube Eropa adalah suhu air sungai dan derajat keasaman (pH) air sungai.

Pada penelitian ini, model yang digunakan adalah *generalized linear model* dengan fungsi *link* ($g(\mu_i)$) yang digunakan adalah *log*. Dengan demikian, model yang terbentuk dalam penelitian ini adalah $y_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$. Pada model tersebut, penambahan nilai variabel X_i , dimana variabel tersebut berskala rasio atau interval, sebesar satu satuan akan menyebabkan variabel respons menjadi lebih besar atau lebih kecil sebesar $\exp(\beta_i)$ satu satuan.

Tabel 4.8 menunjukkan nilai koefisien regresi untuk variabel temperatur atau suhu air sungai adalah sebesar 0,2962. Berdasarkan koefisien tersebut, maka penambahan satu satuan suhu air sungai akan memperbesar biomassa klorofil- α sebesar 1,3447 satu satuan. Hal ini berarti bahwa semakin hangat air sungai maka biomassa klorofil- α akan semakin tinggi. Hal tersebut dikarenakan *phytoplankton* membutuhkan suhu yang hangat untuk berfotosintesis. *Algal bloom*, suatu kondisi dimana organisme menyerupai tumbuhan seperti alga dan *phytoplankton* berkembangbiak dengan sangat cepat, cenderung terjadi pada saat musim panas dimana intensitas cahaya lebih besar dan suhu di sekitar lebih hangat dibandingkan musim-musim lain.

Pada saat terjadi *algal bloom* sungai akan menjadi keruh karena tertutup alga dan *phytoplankton*. Kandungan oksigen terlarut dalam air juga menjadi berkurang karena digunakan untuk dekomposisi alga dan *phytoplankton* yang mati. Hal ini menyebabkan air menjadi semakin asam (pH rendah). Koefisien regresi untuk variabel pH air sungai menunjukkan bahwa semakin tinggi pH, yang berarti air menjadi semakin basa, maka biomassa klorofil- α akan semakin tinggi. Berdasarkan koefisien regresi variabel pH, penambahan derajat keasaman sebesar satu satuan akan memperbesar biomassa klorofil- α sebesar 12,8937 satu satuan.

Nilai derajat keasaman atau pH merupakan nilai konsentrasi ion hidrogen (H^+) sehingga semakin banyak ion hidrogen yang dilepaskan ke air maka air akan semakin asam, dimana hal tersebut menyebabkan perairan kurang produktif. Semakin tinggi konsentrasi ion H^+ dalam air (H_2O) maka ion OH^- akan semakin rendah. Pada kondisi asam tersebut konsentrasi oksigen terlarut (DO) dalam air akan berkurang. Organisme yang menyerupai tumbuhan seperti fitoplankton dan alga selain melakukan fotosintesis juga melakukan respirasi. Pada siang hari, organisme tersebut menyerap CO_2 untuk melakukan fotosintesis dan pada malam hari menyerap oksigen untuk respirasi. Menurunnya pH (ion H^+ meningkat) dan oksigen dalam air dapat menyebabkan kelimpahan biota perairan dan biomasnya menurun.

Sementara variabel yang berkaitan dengan pencemaran nutrisi seperti total nitrogen dan total fosfor dalam air sungai tidak berpengaruh secara signifikan terhadap biomassa klorofil- α . Hasil analisis regresi dengan model lognormal juga menunjukkan bahwa kedua variabel tersebut tidak berpengaruh signifikan terhadap biomassa klorofil- α di Sungai Danube. Hal

ini sesuai dengan analisis pada subbab 4.1 yang menunjukkan bahwa semakin tinggi kandungan nitrogen dan fosfor dalam air sungai belum tentu menyebabkan biomassa klorofil- α menjadi semakin tinggi.

Menurut penelitian sebelumnya, semakin tinggi kandungan nutrisi dalam air maka *phytoplankton* yang tumbuh dan berkembangbiak dalam perairan tersebut akan semakin banyak sehingga biomassa klorofil- α *phytoplankton* dalam perairan akan semakin besar. Namun, hal ini tidak dapat ditunjukkan oleh hasil pemodelan yang menggunakan data JDS3. Kemungkinan hal ini disebabkan karena *phytoplankton* yang terjaring pada saat dilakukan pengukuran berasal dari tempat lain yang terbawa arus hingga ke lokasi pengukuran. *Phytoplankton* merupakan organisme yang melayang di dalam wilayah perairan, sehingga penelitian mengenai *phytoplankton* di sungai menghasilkan analisis yang kurang baik. Penelitian seperti ini lebih cocok dilakukan di danau, dimana airnya tidak mengalir. Penelitian tentang klorofil dapat dilakukan di sungai namun bioindikator yang digunakan adalah *phytobenthos*, karena organisme tersebut memiliki akar sehingga tidak terpengaruh arus sungai.

Halaman Ini Sengaja Dikosongkan

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pengujian distribusi, diketahui bahwa biomassa klorofil- α di Sungai Danube, Eropa dapat mengikuti distribusi gamma dan distribusi lognormal. Dengan demikian fungsi yang dapat digunakan untuk mengestimasi biomassa klorofil- α dengan menggunakan beberapa variabel prediktor adalah fungsi gamma dan fungsi lognormal. Variabel prediktor yang diduga mempengaruhi biomassa klorofil- α adalah kandungan total nitrogen dan total fosfor dalam air sungai, suhu air sungai, pH air sungai dan konduktivitas air sungai.

Berdasarkan nilai AICc, model yang didapatkan dari hasil estimasi dengan fungsi gamma menghasilkan AICc sebesar 348,227. Sedangkan model lognormal menghasilkan AIC sebesar 389,145. Karena AICc model regresi gamma lebih kecil dibandingkan dengan model regresi lognormal, maka model regresi gamma lebih baik daripada model regresi lognormal untuk memodelkan biomassa klorofil- α di Sungai Danube, Eropa. Hasil analisis regresi gamma menunjukkan bahwa variabel prediktor yang secara signifikan berpengaruh terhadap biomassa klorofil- α adalah derajat keasaman dan suhu air sungai. Berdasarkan koefisien regresi, penambahan satu satuan suhu air sungai akan memperbesar biomassa klorofil- α sebesar 1,3447 satu satuan. Sedangkan penambahan satu satuan derajat keasaman akan memperbesar biomassa klorofil- α sebesar 12,8937.

Variabel yang berkaitan dengan pencemaran nutrisi seperti total nitrogen dan total fosfor dalam air sungai tidak berpengaruh secara signifikan terhadap biomassa klorofil- α

baik dalam model regresi gamma maupun model regresi lognormal. Hasil analisis secara deskriptif juga menunjukkan hal yang sama: bahwa tingginya kandungan nitrogen dan fosfor dalam air di Sungai Danube belum tentu menyebabkan biomassa klorofil- α dari fitoplankton semakin besar.

5.2 Saran

Hasil analisis secara deskriptif untuk masing-masing *region* sungai menunjukkan bahwa daerah *upper*, *middle* dan *lower basin* mempunyai karakteristik yang berbeda sehingga untuk memodelkan biomassa klorofil- α di Sungai Danube diperlukan pertimbangan mengenai adanya heterogenitas secara spasial. Sungai Danube mengalir dari Jerman bagian barat hingga dan bermuara di Laut Hitam. Sungai tersebut telah mengalami banyak perubahan di berbagai tempat sebagai akibat aktivitas manusia, seperti pembangunan bendungan untuk pembangkit energi. Oleh karena itu jika dilakukan analisis dengan memperhatikan kondisi spasial maka peneliti perlu memahami kondisi sungai secara keseluruhan.

Selain itu, penelitian tentang klorofil- α dari *phytoplankton* lebih cocok dilakukan di perairan yang tidak berarus seperti di danau. Hal ini dikarenakan *phytoplankton* bukan merupakan organisme yang mempunyai akar sehingga dapat menetap di satu tempat, namun organisme ini melayang di wilayah perairan dan mengikuti arus. Penelitian tentang klorofil dapat dilakukan di sungai namun bioindikator yang digunakan adalah *phytobenthos*, karena organisme tersebut memiliki akar sehingga tidak terpengaruh arus sungai.

DAFTAR PUSTAKA

- Boyer, J. N., Kelble, C. R., Ortner, P. B. & Rudnick, D. T., 2009. Phytoplankton bloom status: Chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. *Ecological Indicators*, Volume 9S, pp. s56-s67.
- Casella, G. & Berger, R. L., 2002. *Statistical Inference*. 2nd ed. California: Duxbury Press.
- Claeskens, G. & Hjort, N. L., 2008. *Model Selection and Model Averaging*. Cambridge: Cambridge University Press.
- Cox, D. R. & Hinkley, D. V., 1974. *Theoretical Statistics*. London: Chapman & Hall.
- Dobson, A. J. & Barnett, A. G., 2008. *An Introduction to Generalized Linear Models*. 3rd ed. Boca Raton: Chapman & Hall.
- Dodds, W. K. & Smith, V. H., 2016. Nitrogen, phosphorus, and eutrophication in streams. *Inland Waters*, 6(2), pp. 155-164.
- Fondriest Environmental, Inc, 2014. *Algae, Phytoplankton and Chlorophyll*. [Online]. Tersedia di: <http://fondriest.com/environmentalmeasurements/parameters/water-quality/algae-phytoplankton-and-chlorophyll> [Diakses pada 15 Februari 2018].
- Gujarati, D. N. & Porter, D. C., 2008. Chapter 10 Multicollinearity: What Happens If the Regressors Are Correlated. In: *Basic Econometrics*. 5th ed. New York: McGraw-Hill.

- ICPDR, 2009. *River Basin*. [Online]. Tersedia di: <http://icpdr.org/danube-basin/river-basin> [Diakses pada 28 Januari 2018].
- Lv, J., Wu, H. & Chen, M., 2011. Effect of nitrogen and phosphorus on phytoplankton composition and biomass in 15 subtropical, urban shallow lakes in Wuhan, China. *Limnologica*, Volume 41, pp. 48-56.
- Magumba, D., Maruyama, a., Takagaki, M., Kato, A. & Kikuchi, M., 2013. Relationship between Chlorophyll-a, Phosphorus and Nitrogen as Fundamental Controlling Phytoplankton Biomass in Lakes. *Environmental Control in Biology*, 51(4), pp. 179-185.
- Montgomery, D. C. & Runger, G. C., 2014. *Applied Statistics & Probability for Engineers*. 6th ed. New Jersey: Wiley.
- Putri, D. E., 2016. *Pemodelan Geographically Weighted Gamma Regression (GWGR), Studi Kasus: Pencemaran Sungai di Kota Surabaya Tahun 2013*, Surabaya: s.n.
- Putri, D. E., Purnadi & Prastyo, D. D., 2017. *Parameter estimation and hypothesis testing on geographically weighted gamma regression*. Bali, Indonesia, IOP Publishing.
- Rocha, R., Thomaz, S., Carvalho, P. & Gomes, L., 2009. Modeling chlorophyll-a and dissolved oxygen concentration in tropical floodplain lakes (Parana River, Brazil). *Braz. J. Biol.*, 69(2), pp. 491-500.
- Santoso, F. P., 2015. *Model Geographically Weighted Univariate Weibull Regression, Studi Kasus: Indikator Pencemaran Sungai di Surabaya*, Surabaya: s.n.

- Smith, L., Porter, K., Hiscock, K., Porter, M.J. & Benson, D., eds., 2015. *Cathment and River Basin Management: Integrating Science and Governance*. New York: Routledge.
- Stephens, M. A., 1974. EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, September, 69(347), pp. 730-737.
- Wackerly, D. D., Mendenhall, W. & Scheaffer, R. L., 2008. *Mathematical Statistics with Applications*. 7th ed. California: Thomson Brooks/Cole.
- Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K., 2010. *Probability & Statistics for Engineers & Scientist*. 9th ed. Boston: Prentice Hall.

Halaman Ini Sengaja Dikosongkan

LAMPIRAN

Lampiran 1. Lokasi Pengambilan *Sampling* JDS3

<i>Station Code</i>	<i>Location</i>	<i>Country</i>
JDS3	Geisling power plant	Germany
JDS4	Deggendorf	Germany
JDS5	Mühlau	Germany
JDS6	Jochenstein	Germany
JDS7	Upstream dam Abwinden-Asten	Austria
JDS8	Oberloiben	Austria
JDS9	Klostemeuburg	Austria
JDS10	Wildungsmauer	Austria
JDS11	Upstream Morava (Hainburg)	Austria
JDS12	Morava (rkm 0.08)	Austria
JDS13	Bratislava	Slovak Republic
JDS14	Gabcikovo resevoir	Slovak Republic
JDS15	Medvedov/Medve	Slovak Republic
JDS16	Moson Danube Arm	Hungary
JDS17	Klizska Nema	Slovak Republic
JDS18	Vah (rkm 0.8)	Slovak Republic
JDS19	Iza/Szony	Slovak Republic
JDS20	Szob	Hungary
JDS21	Budapest upstream	Hungary
JDS22	Budapest downstream	Hungary
JDS23	Rackeve-Soroksar Danube Arm	Hungary
JDS24	Dunafoldvar	Hungary
JDS25	Paks	Hungary
JDS26	Baja	Hungary
JDS27	Hercegszanto	Hungary
JDS28	Upstream Drava	Croatia
JDS29	Drava (rkm 1.4)	Croatia
JDS30	Downstream Drava (Erduť/Bogojevo)	Croatia
JDS31	Ilok/Backa Palanka	Croatia
JDS32	Upstream Novi-Sad	Republic of Serbia
JDS33	Downstream Novi-Sad	Republic of Serbia

Lampiran 1. Lokasi Pengambilan *Sampling* JDS3 (Lajutan)

<i>Station Code</i>	<i>Location</i>	<i>Country</i>
JDS34	Upstream Tisa (Stari Slankamen)	Republic of Serbia
JDS35	Tisa (rkm 1.0)	Republic of Serbia
JDS36	Downstream Tisa/Upstream Sava	Republic of Serbia
JDS37	Sava (rkm 7.0)	Republic of Serbia
JDS38	Upstream Pancevo/Downstream Sava	Republic of Serbia
JDS39	Downstream Pancevo	Republic of Serbia
JDS40	Upstream Velika Morava	Republic of Serbia
JDS41	Velika Morava	Republic of Serbia
JDS42	Downstream Velika Morava	Republic of Serbia
JDS43	Banatska Palanka/Bazias	Republic of Serbia
JDS44	Irongate reservoir (Golubac/Koronin)	Republic of Serbia
JDS45	Irongate reservoir (Tekija/Orsova)	Republic of Serbia
JDS46	Vrbica/Simijan	Republic of Serbia
JDS47	Upstream Timok (Rudujevac/Gruia)	Republic of Serbia
JDS49	Pristol/Novo Selo Harbour	Romania
JDS50	Downstream Kozloduy	Bulgaria
JDS51	Iskar (rkm 0.3)	Bulgaria
JDS52	Downstream Olt	Romania
JDS54	Jantra (rkm 1.0)	Bulgaria
JDS55	Downstream Jantra	Romania
JDS56	Russenski Lom	Bulgaria
JDS57	Downstream Ruse/Giurgiu	Bulgaria
JDS58	Arges	Romania
JDS59	Downstream Arges, Oltenita	Romania
JDS60	Chiciu/Silistra	Romania
JDS61	Giurgeni	Romania
JDS62	Braila	Romania
JDS63	Siret (rkm 1.0)	Romania
JDS64	Prut (rkm 1.0)	Romania
JDS65	Reni	Romania
JDS66	Vilkova - Chilia arm/Kilia arm	Romania
JDS67	Sulina - Sulina arm	Romania
JDS68	Sf.Gheorghe - Sf.Gheorghe arm	Romania

Lampiran 2. Data Penelitian

<i>Station</i>	<i>Y</i>	<i>NT</i>	<i>PT</i>	<i>T(C)</i>	<i>pH</i>	<i>Conductivity</i>
JDS3	3,40	2,90	0,07	21	8,00	487
JDS4	1,10	3,00	0,06	21	7,90	492
JDS5	1,30	2,80	0,08	21	8,00	497
JDS6	2,00	1,80	0,10	18	8,10	377
JDS7	2,20	1,80	0,09	20	8,20	371
JDS8	6,00	1,80	0,06	21	8,20	320
JDS9	12,00	1,90	0,05	22	8,50	385
JDS10	11,00	1,80	0,07	21	8,40	374
JDS11	14,00	1,90	0,07	20	8,30	384
JDS12	31,00	1,50	0,06	21	8,20	580
JDS13	5,90	1,90	0,07	21	8,30	378
JDS14	9,70	1,80	0,08	20	8,30	386
JDS15	8,20	1,80	0,07	20	8,30	389
JDS16	6,10	2,00	0,10	20	8,10	456
JDS17	4,50	1,80	0,08	21	8,30	388
JDS18	20,00	1,50	0,11	22	8,30	476
JDS19	7,90	1,60	0,09	21	8,30	387
JDS20	8,60	1,70	0,08	21	8,30	395
JDS21	9,50	1,60	0,08	21	8,30	391
JDS22	6,60	1,80	0,08	21	8,20	394
JDS23	12,00	0,75	0,10	23	7,90	358
JDS24	1,60	1,70	0,09	21	8,00	395
JDS25	4,00	1,70	0,07	21	8,10	396
JDS26	9,80	1,70	0,08	22	8,10	396
JDS27	7,50	1,70	0,07	22	8,10	397
JDS28	14,00	1,70	0,10	22	8,20	398
JDS29	5,60	1,20	0,08	21	8,20	295
JDS30	16,00	1,80	0,11	22	8,20	375
JDS31	16,00	1,80	0,09	22	8,30	376
JDS32	14,00	1,60	0,08	22	8,20	377
JDS33	13,00	1,70	0,09	21	8,10	377
JDS34	12,00	1,60	0,08	21	8,10	377

Lampiran 2. Data Penelitian (Lanjutan)

<i>Station</i>	Y	NT	PT	T(C)	pH	<i>Conductivity</i>
JDS35	4,50	0,69	0,10	24	7,80	472
JDS36	9,90	1,60	0,08	21	8,10	378
JDS37	1,50	0,69	0,05	24	8,10	400
JDS38	7,00	1,40	0,05	21	8,00	384
JDS39	3,30	1,40	0,05	21	8,00	381
JDS40	6,30	1,50	0,05	21	7,90	383
JDS41	5,00	1,80	0,09	22	7,60	428
JDS42	2,80	1,40	0,06	22	7,90	383
JDS43	1,90	1,50	0,06	22	7,80	383
JDS44	1,30	1,50	0,06	21	7,90	385
JDS45	5,60	1,30	0,05	22	8,00	386
JDS46	0,33	1,30	0,05	22	7,90	388
JDS47	0,79	1,20	0,03	23	7,90	386
JDS49	0,87	1,30	0,04	22	7,90	388
JDS50	3,30	1,20	0,05	22	8,20	380
JDS51	4,70	1,50	0,14	21	8,00	391
JDS52	3,70	1,10	0,05	22	8,30	375
JDS54	1,70	1,40	0,26	19	8,10	418
JDS55	4,90	1,10	0,06	21	8,10	395
JDS56	2,10	5,70	0,46	18	8,40	822
JDS57	3,90	1,10	0,04	21	8,10	381
JDS58	3,30	5,80	0,55	21	7,80	489
JDS59	3,30	1,20	0,05	20	8,20	383
JDS60	5,70	1,20	0,06	20	8,20	386
JDS61	5,60	1,20	0,06	20	8,20	393
JDS62	4,00	1,20	0,06	20	8,10	391
JDS63	2,50	1,70	0,13	17	8,30	678
JDS64	2,60	1,50	0,22	17	8,20	587
JDS65	4,20	1,20	0,07	18	8,20	402
JDS66	4,90	1,10	0,07	20	8,20	399
JDS67	2,60	1,40	0,09	19	8,20	408
JDS68	3,00	1,40	0,05	19	8,20	407

Lampiran 3. Statistika Deskriptif

Descriptive Statistics: Y, NT (X1), PT (X2), Temperatur (X3), pH (X4), Konduktivitas (X5)

Statistics

Variable	Sub-region	Mean	TrMean	StDev	Min	Max
Y	LW	3.256	3.279	1.594	0.33	5.70
	MD	8.216	7.985	4.658	1.50	20.00
	UP	8.40	6.49	9.28	1.10	31.00
NT (X1)	LW	1.678	1.510	1.294	1.10	5.80
	MD	1.5623	1.5981	0.3289	0.69	2.00
	UP	2.120	2.087	0.551	1.50	3.00
PT (X2)	LW	0.1174	0.1010	0.1356	0.03	0.55
	MD	0.07968	0.07963	0.01703	0.05	0.11
	UP	0.07100	0.07000	0.01524	0.05	0.10
Temperatur (X3)	LW	20.217	20.238	1.678	17.00	23.00
	MD	21.484	21.407	0.962	20.00	24.00
	UP	20.600	20.750	1.075	18.00	22.00
pH (X4)	LW	8.1130	8.1143	0.1546	7.80	8.40
	MD	8.1097	8.1259	0.1814	7.60	8.30
	UP	8.1800	8.1750	0.1874	7.90	8.50
Konduktivitas (X5)	LW	435.6	420.0	111.3	375.0	822.0
	MD	391.90	390.67	32.52	295.00	476.00
	UP	426.7	420.9	81.4	320.0	580.0

Lampiran 4. Hasil Uji Distribusi dan Estimasi Parameter Distribusi

Goodness of Fit Test

Distribution	AD	P
Normal	2.641	<0.005
Lognormal	0.235	0.784
Gamma	0.217	>0.250

ML Estimates of Distribution Parameters

Distribution	Location	Shape	Scale	Threshold
Normal*	6.46234		5.42166	
Lognormal*	1.53573		0.86990	
Gamma		1.66105	3.89053	

* Scale: Adjusted ML estimate

Lampiran 5. Matriks Korelasi Antar Variabel

Correlation: Y, NT, PT, Temperatur, pH, Konduktivitas Correlations

	Y	NT	PT	Temperatur	pH
NT	-0.077 0.543				
PT	-0.097 0.444	0.795 0.000			
Temperatur	0.206 0.102	-0.242 0.054	-0.315 0.011		
pH	0.368 0.003	0.017 0.894	-0.014 0.913	-0.367 0.003	
Konduktivitas	-0.022 0.865	0.575 0.000	0.582 0.000	-0.431 0.000	0.092 0.469

Cell Contents

Pearson correlation

P-Value

Lampiran 6. Analisis Regresi Linear**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.535 ^a	.286	.225	4.77364

a. Predictors: (Constant), Conductivity, pH, Temperature, NT, PT

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	530.166	5	106.033	4.653	.001 ^b
	Residual	1321.683	58	22.788		
	Total	1851.849	63			

a. Dependent Variable: CHLA.BIOMASS

b. Predictors: (Constant), Conductivity, pH, Temperature, NT, PT

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	-167.755	37.067		-4.526	.000		
NT	-.629	1.210	-.098	-.520	.605	.344	2.903
PT	2.444	12.481	.038	.196	.845	.333	3.004
Temperature	1.771	.521	.455	3.399	.001	.687	1.456
pH	16.436	3.801	.522	4.324	.000	.844	1.185
Conductivity	.011	.010	.161	1.083	.283	.558	1.794

a. Dependent Variable: CHLA.BIOMASS

Lampiran 7. Analisis Regresi Gamma**Goodness of Fit^a**

	Value	df	Value/df
Deviance	29.021	58	.500
Scaled Deviance	68.447	58	
Pearson Chi-Square	26.374	58	.455
Scaled Pearson Chi-Square	62.204	58	
Log Likelihood ^b	-166.113		
Akaike's Information Criterion (AIC)	346.227		
Finite Sample Corrected AIC (AICC)	348.227		
Bayesian Information Criterion (BIC)	361.339		
Consistent AIC (CAIC)	368.339		

Dependent Variable: CHLA.BIOMASS

Model: (Intercept), NT, PT, Temperature, pH, Conductivity

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Omnibus Test^a

Likelihood Ratio Chi- Square	df	Sig.
26.063	5	.000

Dependent Variable:

CHLA.BIOMASS

Model: (Intercept), NT, PT,
Temperature, pH, Conductivity

a. Compares the fitted model
against the intercept-only model.

Lampiran 7. Analisis Regresi Gamma (Lanjutan)**Parameter Estimates**

Parameter	B	Std. Error	Hypothesis Test			Exp(B)
			Wald Chi-Square	df	Sig.	
(Intercept)	-25.4199	5.0561	25.2767	1	.000	9.13E-12
NT	-0.1373	0.1650	0.6924	1	.405	0.8717
PT	0.9487	1.7025	0.3105	1	.577	2.5823
Temperature	0.2962	0.0711	17.3638	1	.000	1.3447
pH	2.5567	0.5184	24.3208	1	.000	12.8937
Conductivity	0.0009	0.0014	0.4244	1	.515	1.0009
(Scale)	0.4240 ^a	0.0703				

Dependent Variable: CHLA.BIOMASS

Model: (Intercept), NT, PT, Temperature, pH, Conductivity

a. Maximum likelihood estimate.

Lampiran 8. Analisis Regresi Lognormal**Goodness of Fit^a**

	Value	df	Value/df
Deviance	1275.889	58	21.998
Scaled Deviance	64.000	58	
Pearson Chi-Square	1275.889	58	21.998
Scaled Pearson Chi-Square	64.000	58	
Log Likelihood ^b	-186.573		
Akaike's Information Criterion (AIC)	387.145		
Finite Sample Corrected AIC (AICC)	389.145		
Bayesian Information Criterion (BIC)	402.257		
Consistent AIC (CAIC)	409.257		

Dependent Variable: CHLA.BIOMASS

Model: (Intercept), NT, PT, Temperature, pH, Conductivity

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Lampiran 8. Analisis Regresi Lognormal (Lanjutan)**Omnibus Test^a**

Likelihood Ratio Chi-Square	df	Sig.
23.843	5	.000

Dependent Variable:

CHLA.BIOMASS

Model: (Intercept), NT, PT,
Temperature, pH, Conductivity

a. Compares the fitted model
against the intercept-only model.

Parameter Estimates

Parameter	B	Std. Error	Hypothesis Test			Exp(B)
			Wald Chi-Square	df	Sig.	
(Intercept)	-23.8504	4.8754	23.93	1.00	0.00	4.3845E-11
NT	-0.1005	0.1905	0.28	1.00	0.60	0.9044
PT	-3.3379	2.4343	1.88	1.00	0.17	0.0355
Temperature	0.2236	0.0675	10.97	1.00	0.00	1.2506
pH	2.4055	0.5308	20.54	1.00	0.00	11.0838
Conductivity (Scale)	0.0045	0.0012	15.13	1.00	0.00	1.0045
	19.94 ^a	3.52				

Dependent Variable: CHLA.BIOMASS

Model: (Intercept), NT, PT, Temperature, pH, Conductivity

a. Maximum likelihood estimate.

Lampiran 9. Nilai Prediksi Biomassa Klorofil- α dengan Model Gamma

Model:

$$\hat{Y} = \exp(-25,412 - 0,137X_1 + 0,949X_2 + 0,296X_3 + 2,557X_4 + 0,0009X_5)$$

Stasiun	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS3	3,40	3,9105	-0,5105	0,2606
JDS4	1,10	2,9722	-1,8722	3,5051
JDS5	1,30	4,0388	-2,7388	7,5012
JDS6	2,00	2,2490	-0,2490	0,0620
JDS7	2,20	5,1734	-2,9734	8,8413
JDS8	6,00	6,4554	-0,4554	0,2074
JDS9	12,00	19,3733	-7,3733	54,3657
JDS10	11,00	11,4134	-0,4134	0,1709
JDS11	14,00	6,5424	7,4576	55,6152
JDS12	31,00	8,5184	22,4816	505,4237
JDS13	5,90	8,7496	-2,8496	8,1205
JDS14	9,70	6,7083	2,9917	8,9503
JDS15	8,20	6,6631	1,5369	2,3621
JDS16	6,10	4,2507	1,8493	3,4200
JDS17	4,50	9,0369	-4,5369	20,5832
JDS18	20,00	14,1122	5,8878	34,6665
JDS19	7,90	9,3685	-1,4685	2,1566
JDS20	8,60	9,2203	-0,6203	0,3847
JDS21	9,50	9,3138	0,1862	0,0347
JDS22	6,60	7,0363	-0,4363	0,1904
JDS23	12,00	6,7318	5,2682	27,7544
JDS24	1,60	4,3226	-2,7226	7,4128
JDS25	4,00	5,4820	-1,4820	2,1964
JDS26	9,80	7,4418	2,3582	5,5611
JDS27	7,50	7,3782	0,1218	0,0148
JDS28	14,00	9,8117	4,1883	17,5417
JDS29	5,60	6,9837	-1,3837	1,9145

Lampiran 9. Nilai Prediksi Biomassa Klorofil- α dengan Model Gamma (Lanjutan)

Stasiun	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS30	16,00	9,5682	6,4318	41,3681
JDS31	16,00	12,1345	3,8655	14,9420
JDS32	14,00	9,5760	4,4240	19,5713
JDS33	13,00	5,4915	7,5085	56,3782
JDS34	12,00	5,5148	6,4852	42,0575
JDS35	4,50	7,8388	-3,3388	11,1473
JDS36	9,90	5,5198	4,3802	19,1859
JDS37	1,50	15,0787	-13,5787	184,3811
JDS38	7,00	4,2936	2,7064	7,3246
JDS39	3,30	4,2819	-0,9819	0,9642
JDS40	6,30	3,2766	3,0234	9,1408
JDS41	5,00	2,1245	2,8755	8,2682
JDS42	2,80	4,5095	-1,7095	2,9223
JDS43	1,90	3,4445	-1,5445	2,3854
JDS44	1,30	3,3139	-2,0139	4,0556
JDS45	5,60	5,8640	-0,2640	0,0697
JDS46	0,33	4,5493	-4,2193	17,8023
JDS47	0,79	6,0743	-5,2843	27,9235
JDS49	0,87	4,5063	-3,6363	13,2228
JDS50	3,30	9,8597	-6,5597	43,0299
JDS51	4,70	4,6419	0,0581	0,0034
JDS52	3,70	12,8497	-9,1497	83,7177
JDS54	1,70	3,8597	-2,1597	4,6644
JDS55	4,90	5,8913	-0,9913	0,9826
JDS56	2,10	5,9752	-3,8752	15,0168
JDS57	3,90	5,7075	-1,8075	3,2671
JDS58	3,30	2,4874	0,8126	0,6602
JDS59	3,30	5,4678	-2,1678	4,6994
JDS60	5,70	5,5350	0,1650	0,0272
JDS61	5,60	5,5703	0,0297	0,0009

Lampiran 9. Nilai Prediksi Biomassa Klorofil- α dengan Model Gamma (Lanjutan)

Stasiun	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS62	4,00	4,3058	-0,3058	0,0935
JDS63	2,50	3,8238	-1,3238	1,7523
JDS64	2,60	3,0519	-0,4519	0,2042
JDS65	4,20	3,1355	1,0645	1,1331
JDS66	4,90	5,7323	-0,8323	0,6927
JDS67	2,60	4,2035	-1,6035	2,5712
JDS68	3,00	4,0433	-1,0433	1,0885

SSE 1425,9340

MSE 22,2802

Lampiran 10. Nilai Prediksi Biomassa Klorofil- α dengan Model Lognormal

Model:

$$\hat{Y} = \exp(-23,850 - 0,1005X_1 - 3,338X_2 + 0,224X_3 + 2,405X_4 + 0,004X_5)$$

Stasiun	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS3	3,40	5,7602	-2,3602	5,5705
JDS4	1,10	4,7408	-3,6408	13,2553
JDS5	1,30	5,8858	-4,5858	21,0297
JDS6	2,00	2,3098	-0,3098	0,0960
JDS7	2,20	4,6244	-2,4244	5,8779
JDS8	6,00	5,0839	0,9161	0,8392
JDS9	12,00	17,9312	-5,9312	35,1795
JDS10	11,00	10,1380	0,8620	0,7431
JDS11	14,00	6,5995	7,4005	54,7681
JDS12	31,00	16,8399	14,1601	200,5090
JDS13	5,90	8,0337	-2,1337	4,5528
JDS14	9,70	6,5054	3,1946	10,2052

Lampiran 10. Nilai Prediksi Biomassa Klorofil- α dengan Model Lognormal (Lanjutan)

Stasiun	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS15	8,20	6,8175	1,3825	1,9114
JDS16	6,10	5,0481	1,0519	1,1064
JDS17	4,50	8,2090	-3,7090	13,7563
JDS18	20,00	14,2108	5,7892	33,5150
JDS19	7,90	8,0644	-0,1644	0,0270
JDS20	8,60	8,5566	0,0434	0,0019
JDS21	9,50	8,4892	1,0108	1,0217
JDS22	6,60	6,6301	-0,0301	0,0009
JDS23	12,00	4,4562	7,5438	56,9092
JDS24	1,60	4,0216	-2,4216	5,8642
JDS25	4,00	5,4930	-1,4930	2,2292
JDS26	9,80	6,6440	3,1560	9,9606
JDS27	7,50	6,9004	0,5996	0,3595
JDS28	14,00	7,9763	6,0237	36,2846
JDS29	5,60	4,5148	1,0852	1,1776
JDS30	16,00	6,8879	9,1121	83,0300
JDS31	16,00	9,4081	6,5919	43,4537
JDS32	14,00	7,8380	6,1620	37,9700
JDS33	13,00	4,7181	8,2819	68,5900
JDS34	12,00	4,9275	7,0725	50,0202
JDS35	4,50	7,3543	-2,8543	8,1472
JDS36	9,90	4,9497	4,9503	24,5057
JDS37	1,50	12,9425	-11,4425	130,9312
JDS38	7,00	4,5085	2,4915	6,2078
JDS39	3,30	4,4481	-1,1481	1,3182
JDS40	6,30	3,4934	2,8066	7,8772
JDS41	5,00	2,2061	2,7939	7,8058
JDS42	2,80	4,2680	-1,4680	2,1550
JDS43	1,90	3,3219	-1,4219	2,0219
JDS44	1,30	3,4092	-2,1092	4,4486

Lampiran 10. Nilai Prediksi Biomassa Klorofil- α dengan Model Lognormal (Lanjutan)

Stasiun	Y	Model Lognormal		
		\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
JDS45	5,60	5,7465	-0,1465	0,0215
JDS46	0,33	4,5586	-4,2286	17,8813
JDS47	0,79	6,1010	-5,3110	28,2069
JDS49	0,87	4,7134	-3,8434	14,7714
JDS50	3,30	9,1412	-5,8412	34,1198
JDS51	4,70	3,4107	1,2893	1,6622
JDS52	3,70	11,4838	-7,7838	60,5878
JDS54	1,70	2,1191	-0,4191	0,1757
JDS55	4,90	6,0054	-1,1054	1,2220
JDS56	2,10	7,1240	-5,0240	25,2410
JDS57	3,90	6,0289	-2,1289	4,5321
JDS58	3,30	0,5408	2,7592	7,6134
JDS59	3,30	5,9242	-2,6242	6,8867
JDS60	5,70	5,8075	-0,1075	0,0116
JDS61	5,60	5,9929	-0,3929	0,1544
JDS62	4,00	4,6695	-0,6695	0,4482
JDS63	2,50	10,5506	-8,0506	64,8128
JDS64	2,60	4,1649	-1,5649	2,4489
JDS65	4,20	3,8590	0,3410	0,1163
JDS66	4,90	6,0146	-1,1146	1,2424
JDS67	2,60	4,5453	-1,9453	3,7842
JDS68	3,00	5,1713	-2,1713	4,7144

SSE 1275,8892

MSE 19,9358

Halaman Ini Sengaja Dikosongkan