



**UNIVERSITAS INDONESIA**

**ANALISIS FAKTOR-FAKTOR PENGARUH VARIASI KONSENTRASI  
LAPISAN OZON STRATOSFER DENGAN REGRESI**

**LAPORAN AKHIR KULIAH PEMODELAN MATEMATIS  
SEMESTER GANJIL 2022**

**Agustinus Bravy Tetuko Ompusunggu (2006521300)**

**Antonius Ranga Hapsoro Wicaksono (2006568790)**

**Carles Octavianus (2006568613)**

**Fenny Fadhilah Zakiyyatunnisa (2006568853)**

**Rafaella Garcinia Gayatri (2006572226)**

**Yohanes Bryan Sagala (2006568696)**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
PROGRAM STUDI MATEMATIKA**

**DEPOK**

**2023**

## KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa karena rahmat dan anugerah-Nya telah memberikan kesempatan pada penulis untuk menyelesaikan makalah ini. Atas rahmat dan anugerah-Nya lah penulis dapat menyelesaikan makalah yang berjudul ‘Analisis Faktor-Faktor Pengaruh Variasi Konsentrasi Lapisan Ozon Stratosfer dengan Regresi Linear’ secara tepat waktu.

Makalah ‘Analisis Faktor-Faktor Pengaruh Variasi Konsentrasi Lapisan Ozon Stratosfer dengan Regresi Linear’ disusun guna memenuhi tugas akhir mata kuliah pemodelan matematis dengan dosen pembimbing Bapak Dr. Hengki Tasman, S.Si., M.Si. di Universitas Indonesia. Selain itu, penulis juga berharap agar makalah ini dapat menambah wawasan bagi pembaca tentang analisis faktor-faktor pengaruh variasi konsentrasi lapisan ozon.

Penulis mengucapkan terima kasih sebesar-besarnya kepada Bapak Dr. Hengki Tasman, S.Si., M.Si. selaku dosen pembimbing penulis pada mata kuliah pemodelan matematis. Tugas yang telah diberikan ini dapat menambah pengetahuan dan wawasan terkait bidang yang ditekuni penulis. Penulis juga mengucapkan terima kasih pada seluruh dosen mata kuliah pemodelan matematis dan semua pihak yang telah membantu proses penyusunan makalah ini.

Penulis menyadari makalah ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun akan penulis terima demi kesempurnaan makalah ini.

Depok, 06 Januari 2023

Kelompok Ozon

## **ABSTRAK**

Nama : Kelompok Ozon  
Program Studi : Matematika  
Judul : Analisis Faktor-Faktor Pengaruh Variasi Konsentrasi  
Lapisan Ozon Stratosfer dengan Regresi Linear  
Pembimbing : Dr. Hengki Tasman, S.Si., M.Si.

Makalah ini membahas Analisis Faktor-Faktor Pengaruh Variasi Konsentrasi Lapisan Ozon Stratosfer dengan Regresi Linear Penelitian ini adalah penelitian kuantitatif dengan desain deskriptif. Hasil penelitian didapatkan bahwa variasi konsentrasi lapisan ozon dipengaruhi oleh beberapa faktor, dan faktor-faktor yang memengaruhinya juga bergantung pada lokasi penelitian dan juga rentang waktu penelitian.

Kata kunci:

Faktor, Lapisan Ozon, Regresi

## DAFTAR ISI

<b>KATA PENGANTAR</b>	<b>ii</b>
<b>ABSTRAK</b>	<b>ii</b>
<b>DAFTAR ISI</b>	<b>iv</b>
<b>DAFTAR GAMBAR</b>	<b>vi</b>
<b>DAFTAR TABEL</b>	<b>vii</b>
<b>1 Pendahuluan</b>	<b>1</b>
1.1 Latar Belakang . . . . .	1
1.2 Permasalahan . . . . .	1
1.2.1 Rumusan Masalah . . . . .	1
1.2.2 Batasan Permasalahan . . . . .	2
1.3 Tujuan . . . . .	2
1.4 Metodologi Penelitian . . . . .	2
<b>2 Pembahasan</b>	<b>3</b>
2.1 Landasan Teori . . . . .	3
2.1.1 Ketebalan Ozon . . . . .	3
2.1.2 Fenomena yang Berpotensi Mempengaruhi Ketebalan Ozon	3
2.1.3 Variabel Proxy . . . . .	5
2.2 Metode Statistik . . . . .	7
2.2.1 Regresi Linear . . . . .	7
2.2.2 Statistik Model Regresi . . . . .	7
2.2.3 Uji Statistik untuk Regresi . . . . .	9
2.2.4 Asumsi dalam Regresi Linear . . . . .	11
2.2.5 Seleksi Variabel dengan Stepwise Regression . . . . .	12
2.3 Data Penelitian . . . . .	14
2.3.1 Dataset . . . . .	14
2.3.2 Preprocessing . . . . .	15
2.3.3 Analisis Deskriptif . . . . .	15
2.4 Pemodelan . . . . .	16

2.4.1	Alur Pemodelan . . . . .	16
2.4.2	Seleksi Variabel dan Estimasi Parameter . . . . .	16
2.4.3	Pemeriksaan Asumsi Error . . . . .	20
2.4.4	Penilaian Kelayakan Model . . . . .	22
2.4.5	Analisis Koefisien Regresi . . . . .	22
2.5	Hasil Pemodelan . . . . .	23
2.5.1	Rangkuman Hasil Pemodelan . . . . .	23
2.5.2	Interpretasi . . . . .	23
<b>3</b>	<b>Penutup</b>	<b>27</b>
3.1	Kesimpulan . . . . .	27
3.2	Saran . . . . .	27
	<b>DAFTAR REFERENSI</b>	<b>28</b>

## DAFTAR GAMBAR

<b>Gambar 2.1</b>	Proses Pembentukan Ozon Secara Alami . . . . .	4
<b>Gambar 2.2</b>	Proses Perusakan Ozon di Lapisan Stratosfer . . . . .	5
<b>Gambar 2.3</b>	Tabel Analisis Deskriptif Untuk Samoa - April . . . . .	16
<b>Gambar 2.4</b>	Matriks Korelasi antar Variabel Untuk Samoa - April . . . .	17
<b>Gambar 2.5</b>	<i>Printout</i> Hasil <i>Stepwise Regression</i> di R . . . . .	18
<b>Gambar 2.6</b>	<i>Printout</i> Hasil <i>Fitting</i> model 1 di R . . . . .	19
<b>Gambar 2.7</b>	<i>Printout</i> Hasil <i>Fitting</i> model 2 di R . . . . .	19
<b>Gambar 2.8</b>	<i>Printout</i> Grafik <i>Residual vs Fitted</i> untuk Model 2 . . . . .	20
<b>Gambar 2.9</b>	<i>Printout</i> Grafik <i>Q-Q</i> untuk Model 2 . . . . .	21
<b>Gambar 2.10</b>	<i>Printout</i> Grafik <i>Scale-Location</i> untuk Model 2 . . . . .	22

## DAFTAR TABEL

<b>Tabel 2.1</b>	Rangkuman Hasil Pemodelan Pada Daerah Tropis . . . . .	24
<b>Tabel 2.2</b>	Rangkuman Pemodelan Daerah Subtropis Selatan . . . . .	24
<b>Tabel 2.3</b>	Rangkuman Pemodelan Daerah Subtropis Utara . . . . .	24
<b>Tabel 2.4</b>	Rangkuman Pemodelan Daerah Kutub Utara . . . . .	25
<b>Tabel 2.5</b>	Rangkuman Pemodelan Daerah Kutub Selatan . . . . .	25

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Lapisan ozon adalah wilayah stratosfer Bumi yang mengandung konsentrasi ozon (O<sub>3</sub>) tinggi dibandingkan di lapisan lain. Lapisan ozon utamanya ditemukan di bagian bawah stratosfer, dari sekitar 15 hingga 35 kilometer (9 hingga 22 mil) di atas Bumi. Ketebalan lapisan ozon bervariasi secara musiman dan geografis.

Lapisan ozon berperan penting sebagai penyerap radiasi ultraviolet Matahari. Lapisan ozon menyerap 97 hingga 99 persen sinar ultraviolet frekuensi menengah Matahari (dari panjang gelombang sekitar 200 nm hingga 315 nm). Jika tidak diserap, sinar ini berpotensi merusak kehidupan di permukaan Bumi.

*Ozone depletion* adalah fenomena penipisan (pemecahan) ozon pada lapisan ozon di bumi yang disebabkan oleh *ozone depleting substances*. *Ozone-depleting substances* (ODS) adalah gas halogen yang memuat klorin dan/atau bromin yang mempunyai potensi untuk memecah ozon di stratosfer. Beberapa contoh dari ODS antara lain klorofluorokarbon (CFC), hidroklorofluorokarbon (HCFC), metil klorida, metil bromida, dan halon. *Ozone-depleting substances* dapat disebabkan secara natural dan antropogenik.

Proses yang dapat memengaruhi variasi konsentrasi lapisan ozon antara lain, *Midlatitude Halogen Chemistry, Aerosol Effects, Quasi-Biennial Oscillation, Solar Cycle*.

### **1.2 Permasalahan**

Pada bagian ini akan dijelaskan mengenai rumusan masalah yang Penulis hadapi dan ingin diselesaikan serta asumsi dan batasan yang digunakan dalam menyelesaikannya.

#### **1.2.1 Rumusan Masalah**

1. Apa saja faktor-faktor yang memengaruhi variasi konsentrasi lapisan ozon?
2. Bagaimana model matematis untuk mengkuantifikasi pengaruh masing-masing faktor terhadap variasi konsentrasi lapisan ozon?



### **1.2.2 Batasan Permasalahan**

Pada penelitian ini penulis tidak menggunakan model untuk prediksi, melainkan penulis hanya memodelkan untuk interpretasi supaya dapat menentukan faktor-faktor yang memengaruhi variasi konsentrasi lapisan ozon.

### **1.3 Tujuan**

1. Mengetahui faktor-faktor yang memengaruhi variasi konsentrasi lapisan ozon
2. Membangun model matematis untuk mengkuantifikasi pengaruh masing-masing faktor terhadap variasi konsentrasi lapisan ozon

### **1.4 Metodologi Penelitian**

Metodologi penelitian yang penulis gunakan adalah metode kuantitatif dengan desain deskriptif.

## **BAB 2**

### **PEMBAHASAN**

#### **2.1 Landasan Teori**

##### **2.1.1 Ketebalan Ozon**

Ozon adalah molekul gas yang tersusun dari tiga atom oksigen yang secara alami terdapat di atmosfer bumi dan menyerap radiasi sinar ultraviolet pada panjang gelombang tertentu. Ozon terdapat di seluruh atmosfer, tetapi sebagian besar terdapat di lapisan stratosfer. Ozon inilah yang dikenal dengan istilah “Lapisan Ozon”.

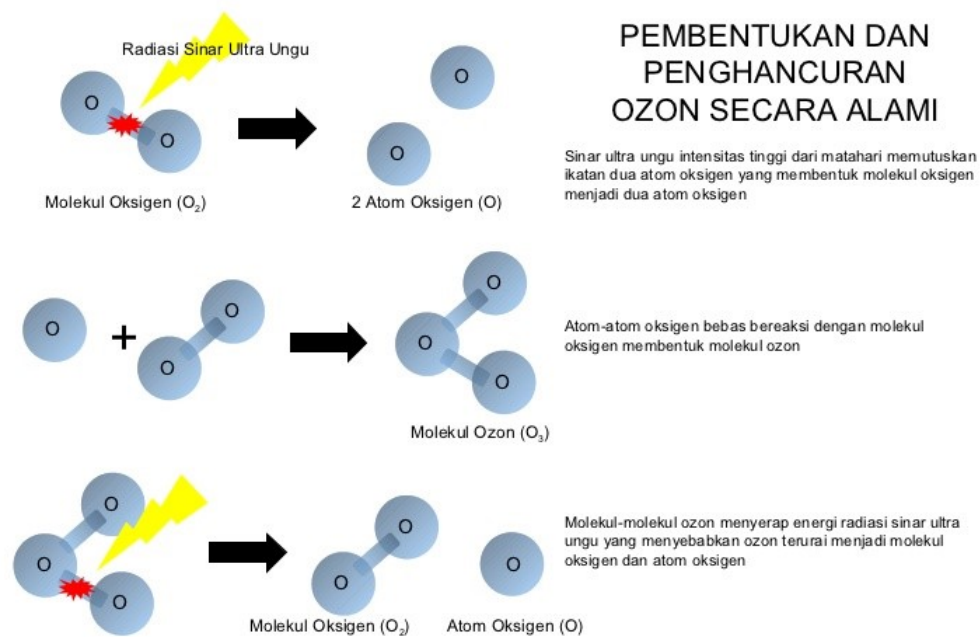
Secara alamiah ozon tersebar dalam stratosfer membentuk lapisan yang tebalnya kurang lebih 35 km. Konsentrasi ozon di lapisan stratosfer bervariasi menurut ketinggian. Lapisan ozon yang tipis ini apabila dibandingkan dengan tebalnya seluruh atmosfer bumi cukup efisien dalam menyaring semua sinar ultraviolet matahari yang berbahaya bagi makhluk hidup di bumi. Oleh karena itu, ozon penting sekali bagi kehidupan di muka bumi dari bahaya sengatan ultraviolet.

Ketebalan lapisan ozon di stratosfer dapat diukur dengan sebuah alat yang bernama Spektrofotometer Dobson dan satuan yang digunakan untuk mengukur ketebalan lapisan ozon adalah Dobson Unit (DU), dimana 1 DU adalah jumlah molekul ozon yang dibutuhkan untuk membentuk satu lapisan ozon setebal 0,01 mm pada suhu 0°C dan tekanan 1 atm.

##### **2.1.2 Fenomena yang Berpotensi Mempengaruhi Ketebalan Ozon**

Ozon terbentuk secara alami melalui siklus Chapman, dimana reaksi pemecahan molekul Oksigen ( $O_2$ ) oleh sinar UV menjadi dua atom oksigen yang kemudian bereaksi dengan molekul oksigen lain menjadi molekul  $O_3$ . Pembentukan molekul ozon paling banyak terbentuk di daerah tropis karena intensitas sinar UV paling optimum di daerah tersebut.

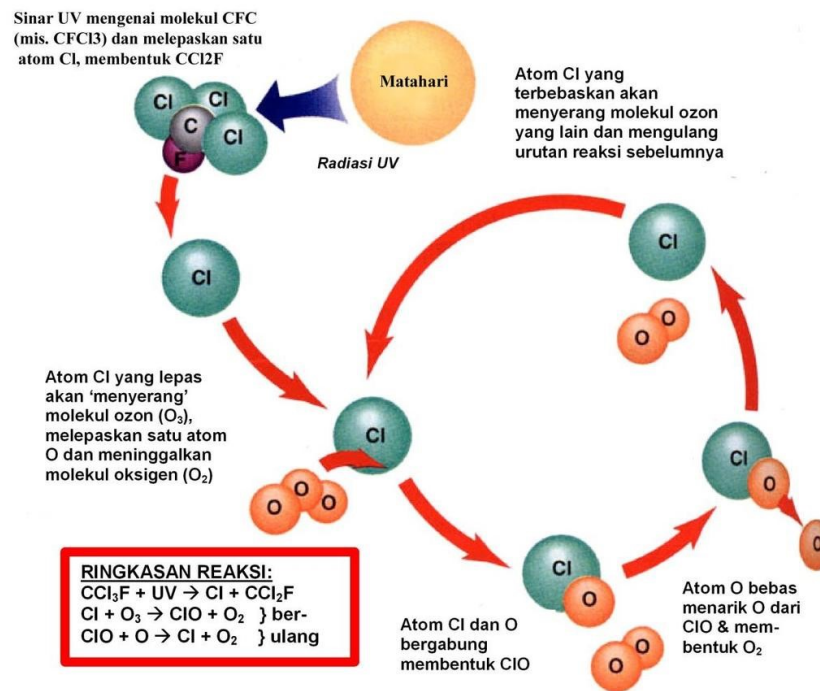
Secara alami, ozon bereaksi dengan berbagai molekul yang mengandung nitrogen, hidrogen dan klorin. Jumlah molekul-molekul tersebut sangatlah kecil sehingga



**Gambar 2.1:** Proses Pembentukan Ozon Secara Alami

tidak mengganggu kelimpahan ozon di stratosfer. Kelimpahan ozon akan terganggu oleh *Ozone Depleting Substances* (ODS) yang merupakan senyawa kimia yang terdiri dari unsur karbon, hidrogen, klorin dan/atau bromin. Senyawa tersebut tidak mudah terurai pada lapisan atmosfer bawah (troposfer) dan bersifat sangat stabil. Sebagai contoh, ODS banyak digunakan dalam peralatan pendingin buatan manusia (refrigeran), contohnya senyawa CFC (*chlorofluorocarbon*) yang mengandung klorin. Klorin yang terlepas dari CFC akan menguraikan ikatan  $O_3$ , sehingga kerapatan lapisan ozon akan berkurang jika proses tersebut berlanjut. Dari fenomena ini, penipisan lapisan ozon terjadi.

Fenomena adanya pengaruh ODS di atmosfer dapat memengaruhi ketebalan lapisan ozon menipis, padahal ozon pada lapisan stratosfer berfungsi sebagai filter sekaligus pelindung bumi dari pengaruh berbahaya radiasi matahari. Radiasi ultraviolet (UV) yang berasal dari matahari berbahaya bagi kehidupan di bumi. Meningkatnya jumlah radiasi UV (UV-B) dapat merusak rantai makanan yang ada di laut. Di samping itu, terdapat hubungan yang kuat antara meningkatnya UV dengan meningkatnya kasus-kasus penyakit kanker kulit dan katarak mata pada manusia. Pada dasarnya atmosfer bertindak sebagai perisai terhadap radiasi matahari melalui penyebaran atau penyerapan oleh molekul- molekul gas yang



**Gambar 2.2:** Proses Perusakan Ozon di Lapisan Stratosfer

ada di dalam atmosfer bumi. Dalam hal ini, ozonlah yang paling efektif menyerap radiasi UV. Secara alami molekul-molekul ozon terbentuk dan rusak di atmosfer bumi. Secara alami pula penipisan lapisan ozon terjadi di atas kutub selatan setiap musim semi.

### 2.1.3 Variabel Proxy

#### 1. EESC (*Equivalent Effective Stratospheric Chlorine*)

Gas sumber halogen yang berkontribusi terhadap penipisan ozon sebagian besar adalah bahan kimia yang mengandung klorin dan bromin yang memiliki masa hidup yang sangat lama di atmosfer. EESC adalah parameter yang mudah digunakan untuk memperkirakan jumlah penipisan ozon stratosfer karena senyawa tersebut. Semakin tinggi nilai EESC, semakin banyak klorin tersedia untuk penghancuran ozon. Satuan EESC adalah *parts per billion by volume* (ppbv, volume klorin setara dengan volume udara).

EESC merupakan proxy yang digunakan pada *Midlatitude Halogen Chemistry*, yaitu reaksi kimia zat halogen yang terjadi pada *midlatitudes* (daerah yang berada di antara 35 hingga 55 derajat lintang selatan maupun utara. Pada lintang ini, terjadi reaksi halogen karena adanya zat halogen bebas akibat reaksi *photochemical* dan reaksi heterogen yang disebabkan oleh *volcanic*

*sulfate aerosols*. Zat halogen yang terdapat pada lintang ini adalah Cl dan Br, di mana kedua senyawa ini akan mengganggu pembentukan zat-zat yang membentuk lapisan ozon.

## 2. *Solar Flux*

*Solar flux* merupakan *proxy variable* yang digunakan untuk merepresentasikan proses *solar cycle*. *Solar cycle* adalah siklus perubahan besarnya radiasi matahari dengan besar periode sekitar 11 tahun. Matahari kita adalah bola gas panas bermuatan listrik yang sangat besar. Gas bermuatan ini bergerak, menghasilkan medan magnet yang kuat. Medan magnet Matahari melewati sebuah siklus, yang disebut siklus matahari.

## 3. QBO (*Quasi-biennial Oscillation*)

*Quasi-biennial Oscillation* (QBO) adalah variasi iklim secara berkala yang diidentifikasi berdasarkan pola angin atau massa udara yang berhembus pada lapisan stratosfer di atas ekuator yang ditandai oleh fenomena berputarnya angin tropikal di stratosfer bawah dari arah timur ke arah barat yang terjadi dalam waktu dua tahun sekali. Perputaran ini akan mempengaruhi transpor ozon. Bila di stratosfer berhembus angin baratan maka ozon pada lintang menengah dan kutub akan berkurang sekitar 6% – 8%. Sebaliknya, bila angin timuran berhembus akan meningkatkan konsentrasi ozon. Massa udara tersebut menjalar mengelilingi bumi sepanjang ekuator dan setiap 14 bulan atau lebih, arahnya akan berbalik menuju arah yang berlawanan. Dengan demikian, total dalam 1 siklus pada arah timuran dan baratan kemudian kembali ke timuran, QBO membutuhkan waktu sekitar 28 bulan. Dua proxy digunakan untuk menjelaskan kemungkinan pergeseran fasa sementara dari sinyal QBO, yaitu QBO-30 dan QBO-50.

## 4. *Aerosol Backscatter Ratio*

*Aerosol Backscatter Ratio* (ABR) merupakan *proxy variable* yang digunakan untuk merepresentasikan jumlah zat aerosol yang berada dalam atmosfer. Untuk pengukurannya sendiri dilakukan dengan menggunakan LiDAR, yaitu sensor cahaya yang memancarkan sinar ke daerah tertentu dan mengolah sinar yang terpantulkan kembali untuk memperoleh informasi tentang hal tersebut. Fenomena terpantulnya kembali cahaya ke sensor disebut *backscattering*. *Backscattering* dipengaruhi oleh sifat zat yang sedang diteliti, antara lain ukuran partikel dan ketebalan. Oleh karena itu, atmosfer yang tidak mengandung aerosol akan menghasilkan backscattering yang berbeda dengan atmosfer yang mengandung aerosol. Rasio dari kemampuan *backscattering* atmos-

fer yang mengandung aerosol dengan kemampuan *backscattering* atmosfer biasa (tidak mengandung aerosol) disebut *aerosol backscatter ratio*.

### 5. Column Ozone

TCO (*Total Column Ozone*) yaitu *proxy variable* untuk pengukuran jumlah total ozon atmosfer dalam kolom tertentu. Biasanya, TCO diukur dalam Dobson Units (DU), dimana 1 DU adalah jumlah molekul ozon yang dibutuhkan untuk membentuk satu lapisan ozon setebal 0,01 mm pada suhu 0°C dan tekanan 1 atm.

## 2.2 Metode Statistik

### 2.2.1 Regresi Linear

Regresi Linear merupakan metode pembentukan model prediksi. Tujuan dari model ini adalah untuk mengekspresikan relasi terhadap variabel dependen  $y \in R^n$  terhadap satu atau lebih variabel independen  $X_i \in R^n$ . Model yang diperoleh dapat digunakan untuk memprediksi nilai  $y$  berdasarkan variabel-variabel independen yang dimiliki. Contoh model regresi linear:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \varepsilon, \varepsilon \sim N(0, \sigma^2)$$

Dengan variabel  $a_i$  merepresentasikan parameter atau koefisien untuk setiap variabel  $X_i$

### 2.2.2 Statistik Model Regresi

#### (a) Korelasi Pearson ( $r$ )

Tipe korelasi yang menjelaskan hubungan antar dua variabel, misal  $x$  dan  $y$ . Koefisien Pearson memiliki *range* nilai -1 sampai dengan 1.

Koefisien Pearson didapat dengan:

$$r = \frac{\sum (X_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (X_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Dengan:

- $r$  = Koefisien Pearson
- $X_i$  = nilai ke- $i$  variabel  $x$

- $\bar{x}$  = Nilai rata - rata variabel x
- $y_i$  = nilai ke-i variabel y
- $\bar{y}$  = Nilai rata - rata variabel y

Interpretasi nilai  $r$ :

- Bila korelasi  $r$  berada pada *range*  $-1 \leq r < 0$ , maka kedua variabel memiliki korelasi negatif.
- Bila korelasi  $r = 0$ , maka kedua variabel tidak berkorelasi sama sekali.
- Bila korelasi  $r$  berada pada *range*  $0 < r \leq 1$ , maka kedua variabel memiliki korelasi positif.

(b) Mallow's  $C_p$

Metrik yang digunakan untuk mengevaluasi suatu model regresi. Mallow's  $C_p$  digunakan untuk memiliki model regresi dengan nilai *fit* terbaik diantara model - model regresi lainnya yang telah dibuat.

Mallow's  $C_p$  didapat dengan:

$$C_p = \frac{RSS_p}{s^2} - n + 2(p - 1)$$

Dengan:

- $C_p$  = Metric Mallow's  $C_p$ .
- $RSS_p$  = *Residual Sum of Squares* pada model dengan jumlah variabel prediktor  $p$ .
- $s^2$  = *Residual Mean Square* pada model (Diestimasikan dengan *Mean Square Error*).
- $n$  = jumlah sampel data
- $p$  = jumlah variabel prediktor

(c) Akaike Information Criterion (AIC)

Skor angka yang dapat digunakan untuk menentukan model *machine learning* mana yang terbaik dalam suatu *dataset* jika *dataset* tersebut tidak mudah untuk dites. AIC bagus digunakan untuk model runtun waktu (*time series*) karena sebagian besar dari data runtun waktu memiliki data yang paling bagus pada data terkini. Semakin rendah skor dari AIC, model tersebut semakin bagus.

Skor AIC didapat dengan:

$$AIC = -2\frac{l}{n} + 2\frac{k}{n}$$

$$l = -\frac{n}{2}(1 + \ln(2\pi) + \ln(\frac{1}{n} \sum (y_i - \hat{y}_i)^2))$$

Dengan:

- AIC = Skor AIC
- $l$  = fungsi log *likelihood*
- $n$  = jumlah sampel pada *dataset*
- $y_i$  = Data ke- $i$  pada *dataset*
- $\hat{y}_i$  = Data  $y_i$  yang diprediksi

(d) *Multiple and Adjusted ( $R^2$  dan  $R_a^2$ )*

*Multiple R Squared ( $R^2$ )* adalah teknik pengukuran model untuk melihat *fitting* model regresi linear ganda yang didapat. *Adjusted  $R^2$  ( $R_a^2$ )* memiliki tujuan yang sama seperti *Multiple  $R^2$* , namun  $R_a^2$  memakai jumlah sampel  $n$  pada model dan jumlah parameter  $k$  pada model.

$R^2$  dan  $R_a^2$  didapat dengan:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{(y_i - \bar{y})^2}, 0 \leq R^2 \leq 1$$

$$R_a^2 = 1 - \left[ \frac{n-1}{n-(k+1)} \right] (1 - R^2), R_a^2 \leq R^2$$

Dengan:

- $R^2$  = Angka *Multiple R Squared*
- $R_a^2$  = Angka *Adjusted R Squared*
- $n$  = Jumlah sampel pada model
- $k$  = Jumlah variabel pada model

### 2.2.3 Uji Statistik untuk Regresi

(a) Uji t

Uji-t digunakan untuk melihat linearitas variabel dependen suatu data dengan variabel - variabel prediktornya. Tes ini dimulai dengan membuat hipotesis nul dan hipotesis alternatif dimana (Contohnya pada kasus ini adalah uji-t untuk regresi):

$H_0 : \beta = 0$  (Gradien sama dengan nol)

$H_a : \beta \neq 0$  (Gradien tidak sama dengan nol)



Lalu dihitung nilai t untuk setiap variabel dengan:

$$t_i = \frac{\beta_i}{SE(\beta_i)}$$

Dengan:

- $\beta_i$  = Koefisien variabel prediktor
- $SE(\beta_i)$  = Error standar untuk koefisien prediktor

Bila nilai-p yang didapat dari tabel nilai-p lebih kecil dibandingkan derajat signifikan  $\alpha$ , hipotesis nol ditolak, sehingga variabel yang diuji signifikan terhadap variabel dependennya.

(b) Uji F

Uji statistik yang menguji tingkat signifikansi dari sejumlah variabel dan melihat apakah setiap variabel itu *textitjointly* significant atau tidak. Terdapat 3 cara untuk melakukan Uji-F:

i. *Left-Tailed Test*

$$H_0 : \sigma_1^2 = \sigma_0^2$$

$$H_1 : \sigma_1^2 < \sigma_0^2$$

Kriteria = Jika variansi variabel pertama lebih kecil dibandingkan yang kedua, maka hipotesis nol-nya dapat ditolak

ii. *Right-Tailed Test*

$$H_0 : \sigma_1^2 = \sigma_0^2$$

$$H_1 : \sigma_1^2 > \sigma_0^2$$

Kriteria = Jika variansi variabel pertama lebih besar dibandingkan yang kedua, maka hipotesis nol-nya dapat ditolak

iii. *Two-Tailed Test*

$$H_0 : \sigma_1^2 = \sigma_0^2$$

$$H_1 : \sigma_1^2 \neq \sigma_0^2$$

Kriteria = Jika variansi variabel pertama tidak sama dengan yang kedua, maka hipotesis nol-nya dapat ditolak

Menghitung nilai F:

$$F = \frac{\frac{SS_{y,y} - SSE}{y}}{\frac{SSE}{n - (k+1)}}$$

Dengan:

- F = Nilai F
- $SS_{y,y}$  = *Sum of Squares* dari y

- $SSE = \text{Sum of Square Errors}$  dari model
- $n = \text{jumlah sampel pada dataset}$
- $k = \text{jumlah variabel pada model}$

Hipotesis nol ditolak apabila  $F > F_\alpha$ , dimana  $k$  adalah derajat kebebasan untuk pembilang dan  $n - (k + 1)$  derajat kebebasan untuk penyebut, atau  $\alpha < \text{nilai-}p$ , dimana  $\text{nilai-}p = P(F > F_C)$ ,  $F_C$  adalah nilai yang dihitung pada uji statistik

(c) Nested F

Uji-F bersifat *nested* dilakukan apabila kedua model yang ingin diuji bersifat *nested*, yaitu ketika model pertama memiliki semua variabel yang dimiliki model kedua, dan setidaknya satu variabel yang beda.

Langkah - langkah Uji-F *Nested*:

- Ambil hipotesis nol dan hipotesis alternatif, dimana:

$$H_0 : \beta_{g+1} = \beta_{g+2} = \dots = \beta_k$$

$$H_a : \text{Minimal satu dari parameter } \beta \neq 0$$

- Hitung nilai F dengan:

$$F = \frac{\frac{SSE_R - SSE_C}{k - g}}{\frac{SSE_C}{n - (k + 1)}} = \frac{\frac{(SSE_R - SSE_C)}{\text{Jumlah koefisien } \beta \text{ yang diuji}}}{MSE_C}$$

Dengan:

- $SSE_R = \text{Sum of Squared Error}$  untuk model yang tereduksi
  - $SSE_C = \text{Sum of Squared Error}$  untuk model yang komplit
  - $MSE_C = \text{Mean Squared Error}$  untuk model yang komplit
- Hipotesis nol ditolak apabila  $F > F_\alpha$ , dimana  $v_1 = k - g = \text{derajat kebebasan untuk pembilang}$  dan  $v_2 = n - (k + 1) = \text{derajat kebebasan untuk penyebut}$ , atau  $\alpha < \text{nilai-}p$ , dimana  $\text{nilai-}p = P(F > F_C)$ ,  $F_C$  adalah nilai yang dihitung pada uji statistik

## 2.2.4 Asumsi dalam Regresi Linear

Dalam regresi linear, terdapat beberapa asumsi yang digunakan untuk membangun model. Pada regresi linear biasa, error dari model diasumsikan memenuhi sifat-sifat berikut.

- (a) Variabel-variabel penjelas  $X_i$  saling independen.
- (b) Mean error adalah 0.
- (c) Error berdistribusi normal.
- (d) Homoskedastisitas error (variansi dari error adalah konstan).
- (e) Error saling independen satu dengan yang lain.

Dengan asumsi-asumsi diatas, dapat dibuktikan bahwa kedua hal berikut benar.

- (a) Estimator *sum of squared residuals*  $s^2 = \sum_{i=1}^n (\hat{y}_i - y_i)^2$  merupakan estimator yang *unbiased* untuk variansi  $\sigma^2$ .
- (b) Distribusi taksiran  $\beta_i$ , yaitu  $\hat{\beta}_i$ , nilai yang kita peroleh saat melakukan *fitting* data sampel, berdistribusi normal. Ini memperbolehkan kita untuk menggunakan uji-t pada koefisien-koefisien yang kita peroleh dalam memeriksa signifikansi statistik dari suatu variabel.
- (c) Kita dapat menggunakan model regresi dan tidak perlu menggunakan model runtun waktu.

### 2.2.5 Seleksi Variabel dengan Stepwise Regression

*Stepwise Regression* adalah suatu metode untuk menkonstruksi model regresi dengan menyeleksi variabel - variabel independen pada suatu *dataset* secara iteratif. Setiap variabel akan diuji tingkat signifikannya dalam model yang akan dibuat. *Stepwise Regression* memiliki tiga metode:

- **Forward Selection** dimulai dengan model yang tidak memiliki variabel, dan akan ditambahkan variabel - variabel yang memiliki tingkat signifikan statistik yang baik. Proses ini diulang hingga hasilnya optimal
- **Backward Elimination** Dimulai dengan semua variabel independen pada suatu *dataset*, lalu akan diuji tingkat signifikansi suatu variabel dengan menghapusnya dari model.
- **Bidirectional Elimination** adalah gabungan dari kedua metode diatas.

Langkah - langkah dari algoritma *Stepwise Regression*:

- (a) Akan dibuat semua model satu variabel yang mungkin dalam bentuk:

$$E(y) = \beta_0 + \beta_1 X_i$$

Dimana  $X_i$  adalah variabel independen ke- $i$ ,  $i = 1, 2, \dots, n$ . Untuk setiap model, akan diuji hipotesis nolnya dengan hipotesis alternatifnya dengan uji- $t$  atau uji- $F$ :

$$H_0 = \beta_1 = 0$$

$$H_a = \beta_1 \neq 0$$

Variabel independen yang menghasilkan nilai- $t$  terbesar adalah prediktor satu variabel terbaik untuk  $y$ , sehingga variabel tersebut dinamakan  $X_1$ .

- (b) Algoritma ini sekarang akan mulai mencari variabel - variabel  $(k - 1)$  sisanya untuk mendapatkan model dua variabel terbaik dengan:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_i$$

Proses ini dilakukan dengan melakukan *fitting* semua model dua variabel yang didapat yang berisi variabel  $X_1$  dengan setiap pilihan  $(k - 1)$  untuk menjadi variabel  $X_2$ . nilai- $t$  untuk uji  $H_0 : \beta_2 = 0$  dihitung untuk setiap model  $(k - 1)$ , dan variabel dengan nilai- $t$  tertinggi akan menjadi variabel  $X_2$

Sebelum algoritma melanjutkan prosesnya ke langkah ketiga, algoritma ini akan mengecek nilai- $t$  dari  $\hat{\beta}_1$  setelah model telah ditambahkan  $\hat{\beta}_2 X_2$ . Jika nilai- $t$  menjadi insignifikan pada suatu nilai  $\alpha$  (misal nilai  $\alpha = 0.05$ , variabel  $X_1$  dihapus dari model dan algoritmanya akan mencari variabel independen yang memiliki parameter  $\beta$  yang memiliki nilai- $t$  yang paling signifikan dengan adanya  $\hat{\beta}_2 X_2$

Nilai- $t$  untuk  $X_1$  dapat berubah pada langkah kedua adalah karena arti dari koefisien  $\hat{\beta}_1$  berubah. Pada langkah kedua, aproksimasinya dilakukan pada ruang tiga dimensi (karena adanya  $y$ ,  $X_1$ , dan  $X_2$ ), sehingga *fitting* terbaik untuk  $\hat{\beta}_1$  akan memiliki nilai yang berbeda dibandingkan nilai *fitting* pada langkah pertama. Karena nilai  $\hat{\beta}_1$  dapat berubah, akan dicek kembali nilai- $t$  pada setiap langkah.

- (c) Algoritma akan mencari variabel independen ketiga untuk dimasukkan

kedalam model, sehingga dicari model terbaik dengan bentuk:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Algoritma akan membuat dan melakukan *fitting* untuk semua model  $(k - 2)$  dengan menggunakan  $X_1$ ,  $X_2$ , dan setiap sisa  $(k - 2)$  variabel  $X_i$ , sebagai variabel  $X_3$  yang memungkinkan. Seperti langkah - langkah sebelumnya, variabel independen dengan nilai- $t$  terbaik akan dimasukkan ke dalam model sebagai  $X_3$  dan algoritma akan mengecek nilai- $t$  untuk koefisien dari  $X_1$  dan  $X_2$ , dan mengubah variabel - variabel yang nilai- $t$ -nya menjadi tidak signifikan.

Hasil akhir dari algoritma ini adalah model yang berisi variabel - variabel yang memiliki nilai- $t$  yang signifikan dengan nilai- $\alpha$  yang telah ditentukan.

## 2.3 Data Penelitian

Bagian ini akan menjelaskan data - data yang akan digunakan pada topik makalah pemodelan ini.

### 2.3.1 Dataset

Dataset yang akan digunakan pada topik ini terdiri dari *proxy* yang merupakan representasi dari masing - masing faktor yang telah dijelaskan sebelumnya:

- EESC (*Equivalent Effective Stratospheric Chlorine*) yang merupakan *proxy* dari *Midlatitude Halogen Chemistry*
- *Solar Flux* yang merupakan *proxy* dari besarnya daya dari cahaya matahari yang mengenai Bumi.
- QBO (*Quasi - biennial Oscillation*) sebagai *proxy* dari pola hembusan angin di sekitar khatullistiwa dan untuk topik ini akan diambil 2 QBO dengan tekanan angin yang berbeda yaitu, QBO-30 dan QBO-50.
- *Aerosol back scatter* sebagai *proxy* untuk efek aerosol.
- *Column Ozone* sebagai *proxy* dari ketebalan ozon yang dihitung dengan satuan Dobson Unit (DU) dan data ini akan dipilih sebagai variabel target.

## Daerah

Untuk daerah, topik ini memilih 5 daerah yaitu;

- Samoa untuk di daerah tropis.
- New Zealand sebagai perwakilan dari daerah subtropis bagian selatan.
- Boulder sebagai perwakilan dari daerah subtropis bagian utara.
- Kutub Utara (*North Pole*).
- Kutub Selatan (*South Pole*).

Dipilihnya daerah ini, diakibatkan karena terbatasnya lokasi penelitian mengenai ketebalan ozone dan akses yang ada untuk mendapatkan data. Rentang waktu data set yang akan digunakan dimulai dari bulan Januari 1992 hingga bulan Desember 2011.

### 2.3.2 Preprocessing

Data preprocessing adalah proses persiapan data yang akan dimodelkan. Data preprocessing dimulai dengan pengumpulan data. Lalu, dilanjutkan dengan *data cleaning*. Pada tahapan *data cleaning*, kami mencari ada atau tidaknya kekurangan dalam data, seperti *missing value handling*, memeriksa sumber observasi data, dan sebagainya. Setelah *data cleaning*, mengelompokkan data berdasarkan bulan. Tujuan dari proses ini adalah untuk mencegah adanya efek *seasonality*, karena data kita merupakan data runtun waktu. Proses ini dilakukan dengan membuat 3 dataset berdasarkan daerah yaitu, SP (Kutub selatan), NP (Kutub Utara), serta Subtropis dan Tropis. Lalu dipisahkan menjadi bulannya dimana tiap dataset ini berisi 5 faktor, dan diurutkan berdasarkan tahun .

### 2.3.3 Analisis Deskriptif

Analisis Deskriptif adalah analisis yang dilakukan untuk menilai karakteristik dari sebuah data. Jenis data yang kita miliki adalah data numerik dimana data ini terdiri dari angka - angka dan angka tersebut memiliki tipe float atau bilangan berkoma. Berikut adalah tabel yang menunjukkan analisis deskriptif dari data set Samoa untuk bulan April. Pada Baris Mode dapat terlihat bahwa EESC, Solar Flux, dan QBO 30 tidak memiliki nilai yang terulang sehingga pada tabel diberikan garis.

	Aerosol Backscatter Ratio	EESC	solar_flux	QBO_30	QBO_50	Ozone_Column_Samoa
Mean	0.01527222222	1831.611111	944.8333333	-43	24.22222222	252.0555556
Median	0.0035	1844	793.5	73	58	255
Mode	0.0021	-	-	-	58	237
Max	0.1386	1936	1720	151	134	271
Min	0.002	1687	629	-298	-201	237
Range	0.1366	249	1091	449	335	34
Standard Dev.	0.03334267369	85.63131566	352.3209182	181.3183745	108.8390909	10.64105381
Sample Size	18	18	18	18	18	18

**Gambar 2.3:** Tabel Analisis Deskriptif Untuk Samoa - April

## 2.4 Pemodelan

Bagian ini akan menjelaskan pemodelan yang telah kami lakukan untuk analisis faktor-faktor pengaruh variasi konsentrasi lapisan ozon.

### 2.4.1 Alur Pemodelan

Secara garis besar, alur dari proses pemodelan yang telah dilakukan adalah sebagai berikut.

- (a) Seleksi Variabel dan Estimasi Parameter
- (b) Pemeriksaan Asumsi Error
- (c) Penilaian Kelayakan Model

Tahapan-tahapan di atas akan dijelaskan lebih lanjut dan disertakan dengan hasilnya di bagian-bagian selanjutnya.

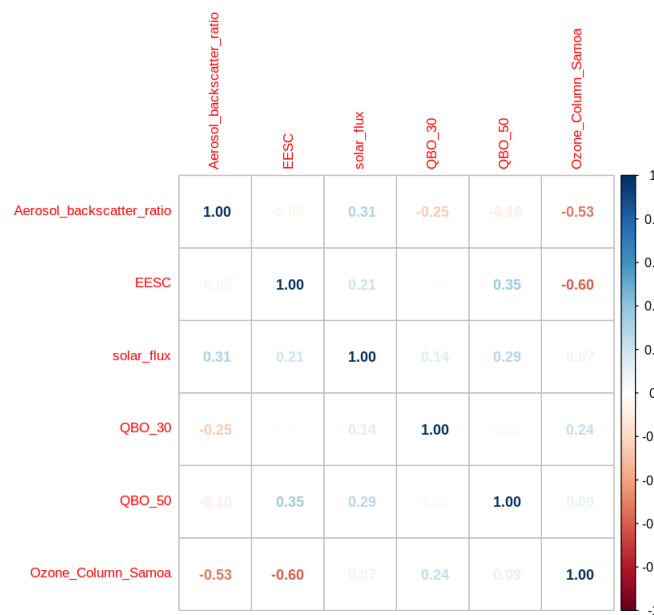
Untuk memberikan gambaran akan alur pemodelan yang dilakukan, kami akan membahas alur untuk pemodelan daerah Samoa pada Bulan April. Proses pemodelan untuk bulan dan daerah lain adalah sama, sehingga kami hanya akan merangkum hasilnya di bagian 2.5.

### 2.4.2 Seleksi Variabel dan Estimasi Parameter

Pada tahapan ini, kami mencari variabel-variabel yang layak untuk digunakan pada model. Hal pertama yang kami periksa adalah multikolinearitas. Multikolinearitas adalah penyelidikan ada atau tidaknya indikasi hubungan linear antara variabel-variabel penjelas. Nilai multikolinearitas yang tinggi mengindikasikan adanya beberapa variabel yang mengandung informasi

yang sama. Hal ini dapat mengurangi kemampuan interpretasi model. Sebab, ini artinya variabel penjelasnya tidak independen satu dengan yang lain.

Uji multikolinearitas dapat dilakukan dengan berbagai cara, antara lain dengan matriks korelasi dan *Variance Inflation Factor* (VIF). Karena jumlah variabel penjelas tidak terlalu banyak, kita akan menggunakan matriks korelasi. Dari matriks korelasi diatas, terlihat bahwa tidak ada variabel



**Gambar 2.4:** Matriks Korelasi antar Variabel Untuk Samoa - April

penjelas yang memiliki korelasi Pearson kuat. Artinya, tidak ada indikasi multikolinearitas dari data variabel penjelas kita. Jadi, kita dapat menggunakan semua variabel yang ada.

Setelah memeriksa multikolinearitas, seleksi variabel dilanjutkan dengan *stepwise regression*. Model yang akan kita ajukan adalah model dengan interaksi orde 2. Dengan kata lain, model terluas yang mungkin adalah sebagai berikut.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_1 X_2 + \beta_7 X_1 X_3 + \beta_8 X_1 X_4 + \beta_9 X_1 X_5 + \beta_{10} X_2 X_3 + \beta_{11} X_2 X_4 + \beta_{12} X_2 X_5 + \beta_{13} X_3 X_4 + \beta_{14} X_3 X_5 + \beta_{15} X_4 X_5 + \epsilon, \epsilon \sim N(0, \sigma^2)$$

Berikut adalah representasi variabel yang digunakan pada persamaan di atas.

(a)  $X_1$  : Aerosol\_backscatter\_ratio



- (b)  $X_2$  : EESC
- (c)  $X_3$  : solar\_flux
- (d)  $X_4$  : QBO\_30
- (e)  $X_5$  : QBO\_50
- (f)  $Y$  : Ozone\_Column\_Samoa

Model tersebut adalah model yang lumayan besar dengan 16 derajat kebebasan. *Stepwise regression* akan menyeleksi variabel-variabel yang paling berpengaruh.

Setelah menjalankan *stepwise regression*, berikut adalah hasil yang kami dapatkan.

Step	Variable Entered	R-Square	Adj. R-Square	C(p)	AIC	RMSE
1	EESC	0.3609	0.3209	238.6169	133.1250	8.7688
2	Aerosol_backscatter_ratio:EESC	0.6790	0.6362	114.8708	122.7284	6.4181
3	solar_flux	0.8452	0.8121	51.1764	111.5990	4.6132
4	QBO_50	0.8751	0.8367	41.3555	109.7342	4.3000
5	QBO_30	0.8753	0.8234	43.2780	111.7059	4.4721
6	Aerosol_backscatter_ratio	0.9254	0.8847	25.4782	104.4570	3.6126

**Gambar 2.5:** Printout Hasil *Stepwise Regression* di R

Pada tabel tersebut, kolom *step* menunjukkan nomor iterasi. Variabel-variabel yang dimasukkan ke dalam model pada setiap step dituliskan pada kolom *entered*. Kolom-kolom lainnya merupakan statistik-statistik untuk model-model yang dibangun pada setiap *step*, yaitu  $R^2$ , *adjusted*  $R^2$ , Mallows's  $C_p$ , Akaike Information Criterion (AIC), serta RMSE. Makna dari masing-masing statistik tersebut sesuai dengan yang telah dijelaskan pada bagian 2.2.2. Kita menggunakan hasil *stepwise regression* untuk membangun model pertama kita. Dalam bentuk persamaan, model pertama kita dapat dituliskan dengan persamaan berikut.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_1 X_2 + \epsilon, \epsilon \sim N(0, \sigma^2)$$

Model diatas akan kita sebut sebagai **model 1**. Kita dapat menjalankan algoritma regresi *Ordinary Least Squares* untuk mencari taksiran dari koefisien-koefisien regresi  $\beta_i$ . Berikut adalah hasilnya.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.852e+02  2.390e+01  16.113 5.35e-09 ***
EESC         -7.465e-02  1.367e-02  -5.461 0.000197 ***
solar_flux   7.814e-03  3.216e-03   2.430 0.033424 *
QBO_50       2.484e-02  9.288e-03   2.674 0.021624 *
QBO_30       4.157e-03  5.329e-03   0.780 0.451785
Aerosol_backscatter_ratio 6.787e+03  2.497e+03   2.718 0.019997 *
EESC:Aerosol_backscatter_ratio -3.875e+00  1.382e+00  -2.804 0.017165 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.613 on 11 degrees of freedom
Multiple R-squared:  0.9254,    Adjusted R-squared:  0.8847
F-statistic: 22.75 on 6 and 11 DF,  p-value: 1.348e-05

```

**Gambar 2.6:** Printout Hasil *Fitting* model 1 di R

Pada gambar di atas, nilai taksiran untuk koefisien-koefisien regresi diberikan pada kolom *Estimate*.

Kolom  $Pr(> |t|)$  adalah nilai probabilitas yang digunakan untuk uji-t masing-masing variabel. Perhatikan bahwa  $X_4$  tidak memenuhi uji-t. Oleh karena itu, kita akan mencoba menghilangkan variabel tersebut dan melakukan *fitting* dengan model berikut.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_5 + \beta_5 X_1 X_2 + \epsilon, \epsilon \sim N(0, \sigma^2)$$

Model ini akan kita sebut sebagai Model 2. Hasil dari proses *fitting* dengan R adalah sebagai berikut.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.875e+02  2.332e+01  16.617 1.20e-09 ***
EESC         -7.642e-02  1.326e-02  -5.765 8.96e-05 ***
solar_flux   8.651e-03  2.982e-03   2.902  0.0133 *
QBO_50       2.364e-02  9.010e-03   2.624  0.0222 *
Aerosol_backscatter_ratio 6.347e+03  2.392e+03   2.653  0.0211 *
EESC:Aerosol_backscatter_ratio -3.635e+00  1.325e+00  -2.743  0.0178 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.553 on 12 degrees of freedom
Multiple R-squared:  0.9213,    Adjusted R-squared:  0.8885
F-statistic: 28.09 on 5 and 12 DF,  p-value: 3.139e-06

```

**Gambar 2.7:** Printout Hasil *Fitting* model 2 di R

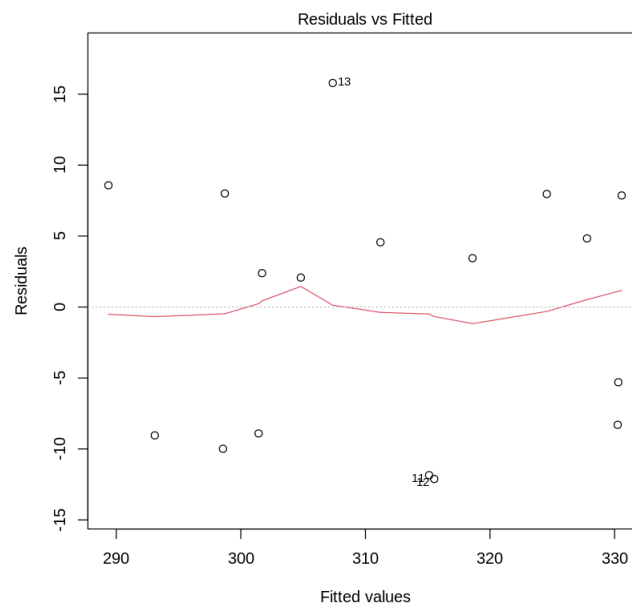
### 2.4.3 Pemeriksaan Asumsi Error

Setelah kita telah mendapatkan model, yaitu model 2, kita memeriksa apakah model kita memenuhi asumsi-asumsi error dalam regresi biasa. Pemeriksaan asumsi penting untuk memastikan bahwa model kita merupakan model yang baik, sehingga kita dapat menginterpretasikan model dengan tepat. Asumsi-asumsi error yang digunakan dalam regresi biasa adalah sebagai berikut.

- (a) Mean dari error adalah 0.
- (b) Error berdistribusi normal
- (c) Homoskedastisitas error
- (d) Independensi error

Untuk memeriksa keempat kriteria berikut, kita akan menggambar berbagai grafik yang mengandung nilai-nilai residual dari model kita.

Pertama, kita akan melihat grafik dari *residual vs fitted*. Sumbu  $x$  dari grafik ini adalah nilai-nilai data dari variabel dependen, yaitu  $y_i$ . Sedangkan sumbu  $y$  dari grafik ini adalah residual untuk setiap  $y_i$ , yaitu  $\hat{y}_i - y_i$ .

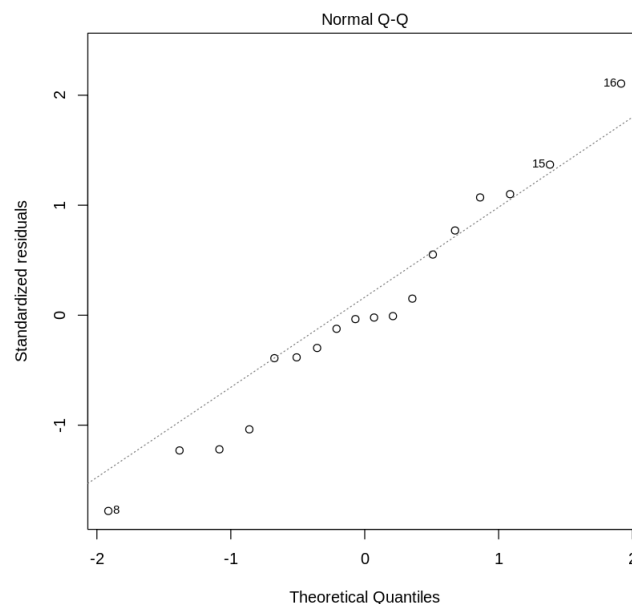


**Gambar 2.8:** Printout Grafik *Residual vs Fitted* untuk Model 2

Dari grafik diatas, terlihat bahwa mean dari residual adalah 0. Ini mengindikasikan bahwa asumsi pertama kita terpenuhi, yaitu mean dari error

adalah 0.

Untuk memeriksa bahwa distribusi dari error adalah normal, kita dapat menggunakan grafik  $Q-Q$ .

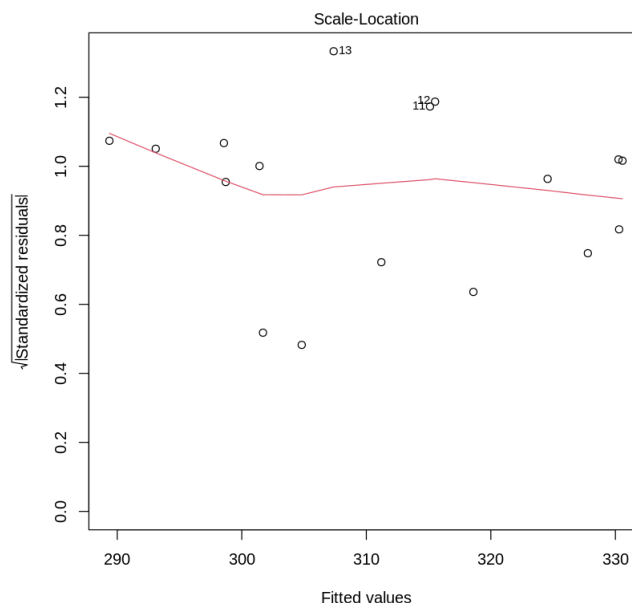


**Gambar 2.9:** Printout Grafik  $Q-Q$  untuk Model 2

Pada grafik tersebut, jika distribusi error normal, titik-titik pada grafik akan terletak pada garis. Posisi titik-titik yang mendekati garis menunjukkan bahwa distribusi dari residual mendekati distribusi normal. Ini merupakan indikasi bahwa error dari model 2 mendekati distribusi normal. Secara umum error yang tidak sepenuhnya mengikuti distribusi normal tidak terlalu memengaruhi regresi. Oleh karena itu, kami akan menganggap bahwa asumsi kedua sudah cukup terpenuhi.

Untuk memeriksa homoskedastisitas error, kita akan menggunakan grafik *Scale-Location*. Grafik ini menunjukkan nilai akar dari residual setelah standarisasi terhadap nilai  $y_i$ . Garis yang mengikuti garis lurus mengindikasikan bahwa asumsi homoskedastisitas terpenuhi.

Terakhir, kita dapat memeriksa asumsi independensi error dengan grafik *residual vs fitted* pada bagian sebelumnya. Error yang dependen akan membentuk suatu pola, umumnya sebuah periodisitas. Terlihat bahwa pada



**Gambar 2.10:** Printout Grafik *Scale-Location* untuk Model 2

grafik tersebut nilai residual tidak membentuk pola apapun. Ini adalah indikasi bahwa error independen satu dengan yang lain.

#### 2.4.4 Penilaian Kelayakan Model

Pada bagian ini, kita akan menilai performa model dengan statistik  $R^2$ . Pada gambar 2.7, kita memperoleh nilai  $R^2 = 0.9213$ . Artinya, sekitar 92% variabilitas dari nilai variabel dependen dapat dijelaskan oleh model. Nilai ini merupakan akurasi yang baik.

#### 2.4.5 Analisis Koefisien Regresi

Pada gambar 2.7 kolom *estimate* merupakan kolom estimasi parameter koefisien regresinya untuk setiap variabel *proxy*. Nilai estimasi positif menandakan bahwa kenaikan nilai variabel *proxy* tersebut akan menyebabkan kenaikan nilai dari *total column ozone*. sebaliknya, nilai estimasi yang negatif menandakan bahwa kenaikan nilai dari variabel *proxy*, akan menyebabkan penurunan nilai dari *total column ozone*. Selain itu, perlu diingat bahwa *magnitude* dari koefisien regresi tersebut tidak dapat kita bandingkan karena pada proses pemodelan, kita tidak melakukan normalisasi data.

Tabel berikut merangkum efek-efek variabel *proxy* terhadap *total column ozone* pada daerah tropis di bulan April .

Variabel proxy	Efek variabel terhadap Total Column Ozone
EESC	Berbanding terbalik
QBO-30	Tidak signifikan
QBO-50	Berbanding lurus
Solar Flux	Berbanding lurus
Aerosol backscatter ratio	Berbanding lurus

## 2.5 Hasil Pemodelan

### 2.5.1 Rangkuman Hasil Pemodelan

Bagian berikut adalah rangkuman (berupa tabel) mengenai hasil pemodelan untuk setiap daerah observasi (tabel dibagi per daerah observasi, terdapat 5 tabel). Kolom dari tabel menunjukkan variabel *proxy* dan baris dari tabel menunjukkan bulan (Januari hingga Desember). sebuah ceklis pada tabel menandakan bahwa variabel *proxy* tersebut lolos uji-t (uji signifikansi dengan distribusi t) untuk bulan yang bersesuaian. Variabel *proxy* yang lolos uji t berarti variabel *proxy* tersebut dapat menjelaskan variabilitas dari *total column ozone* pada daerah observasinya untuk bulan yang bersesuaiannya. Halaman berikut merupakan rangkuman dari tabel yang dimaksud.

### 2.5.2 Interpretasi

Pada daerah tropis, variabel *proxy* EESC adalah variabel terbaik yang menjelaskan variabilitas *total column ozone* karena dapat terlihat bahwa pada tabel 2.1, variabel EESC dapat menjelaskan variabilitas *total column ozone* 7 dari 12 bulan. Selanjutnya, variabel *proxy aerosol backscatter ratio* dapat menjelaskan variabilitas *total column ozone* 5 dari 12 bulan. Variabel *proxy* QBO-30, QBO-50, Solar Flux tidak dapat menjelaskan variabilitas *total column ozone*.

Pada Daerah subtropis Selatan, variabel *proxy* QBO-30 dan EESC adalah variabel *proxy* terbaik untuk menjelaskan variabilitas *total column ozone* (dengan total 5 dari 12 bulan terjelaskan dengan kedua variabel *proxy*). Namun, pada daerah subtropis utara, tidak ada variabel *proxy* yang menjelaskan dengan baik variabilitas *total column ozone*.

Pada daerah kutub, secara general tidak ada variabel *proxy* yang menjelaskan variabilitas *total column ozone* diatas dua bulan. Temuan unik adalah variabel interaksi antara *aerosol backscatter ratio* dengan EESC (orde 2 regresi) dapat

Tropis	EESC	QBO-30	QBO-50	SolarFlux	Aerosol backscatter
Jan	✓	x	x	x	x
Feb	x	x	x	x	x
Mar	✓	x	x	x	✓
Apr	✓	x	✓	✓	✓
May	✓	x	x	x	✓
Jun	✓	x	x	x	✓
Jul	✓	x	x	x	✓
Aug	✓	x	x	x	x
Sep	x	x	x	x	x
Oct	x	x	x	x	x
Nov	x	x	x	x	x
Dec	x	x	x	x	x

**Tabel 2.1:** Rangkuman Hasil Pemodelan Pada Daerah Tropis

Subtropis (S)	EESC	QBO-30	QBO-50	SolarFlux	Aerosol backscatter
Jan	x	x	x	x	x
Feb	x	x	x	x	x
Mar	x	x	x	x	x
Apr	x	x	x	x	x
May	x	x	x	x	x
Jun	✓	✓	x	x	✓
Jul	✓	✓	x	x	x
Aug	✓	✓	x	x	x
Sep	✓	✓	x	x	x
Oct	✓	✓	x	x	x
Nov	x	x	x	x	x
Dec	x	x	x	x	x

**Tabel 2.2:** Rangkuman Pemodelan Daerah Subtropis Selatan

Subtropis (N)	EESC	QBO-30	QBO-50	SolarFlux	Aerosol backscatter
Jan	x	✓	x	x	x
Feb	x	x	x	x	x
Mar	x	x	x	x	x
Apr	x	x	x	x	x
May	x	x	x	x	x
Jun	x	x	x	x	x
Jul	x	x	x	x	x
Aug	x	x	x	x	x
Sep	x	x	x	x	x
Oct	x	x	x	x	x
Nov	x	x	x	x	x
Dec	x	x	x	x	x

**Tabel 2.3:** Rangkuman Pemodelan Daerah Subtropis Utara

Polar (N)	EESC	QBO-30	QBO-50	SolarFlux	Aerosol backscatter	AB : EESC
Jan	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Feb	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Mar	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Apr	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
May	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Jun	✓	<i>x</i>	✓	<i>x</i>	✓	✓
Jul	✓	<i>x</i>	✓	<i>x</i>	✓	<i>x</i>
Aug	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Sep	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Oct	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Nov	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Dec	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>

**Tabel 2.4:** Rangkuman Pemodelan Daerah Kutub Utara

Polar (S)	EESC	QBO-30	QBO-50	SolarFlux	Aerosol backscatter
Jan	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Feb	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Mar	✓	<b>x</b>	<b>x</b>	<b>x</b>	✓
Apr	<b>x</b>	<b>x</b>	✓	<b>x</b>	✓
May	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Jun	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Jul	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Aug	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Sep	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Oct	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Nov	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
Dec	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>

**Tabel 2.5:** Rangkuman Pemodelan Daerah Kutub Selatan



menjelaskan variabilitas *total column ozone* di kutub selatan pada bulan Juni, dan Juli.

## **BAB 3**

### **PENUTUP**

#### **3.1 Kesimpulan**

Dari hasil pengujian bab 2.5, dapat dilihat bahwa tidak semua variabel *proxy* dapat menjelaskan variabilitas *total column ozone* di setiap garis daerah (tropis, subtropis selatan, subtropis utara, kutub utara, kutub selatan). Pada daerah tropis, ditemukan bahwa variabel *proxy* EESC dan *aerosol backscatter ratio* adalah variabel *proxy* yang dapat menjelaskan variabilitas TCO dengan baik. Di lain sisi, pada daerah subtropis selatan, EESC dan QBO-30 lah merupakan variabel penjelas yang baik. Namun, pada daerah subtropis utara, kutub selatan, kutub utara, tidak ditemukan variabel *proxy* yang dapat menjelaskan variabilitas total column ozone dengan baik.

#### **3.2 Saran**

Proses pembentukan dan penghancuran lapisan ozon merupakan proses yang dinamik dan rumit. Kelima variabel *proxy* yang digunakan belum dapat meng-*capture* semua proses dinamik tersebut secara global. Oleh karena itu, pada penelitian selanjutnya, kami dapat mempertimbangkan variabel-variabel [*proxy* lainnya, sehingga memperluas ukuran model. Dengan begitu, kami harapkan bahwa performa model akan jauh lebih baik. Selain itu, rentang waktu tahun (derajat bebas) dari model juga dapat diperpanjang, sehingga kami harapkan juga performa model yang lebih baik.

## DAFTAR REFERENSI

- [1] Mendenhall, William Sincich, Terry. (1997). A Second Course in Statistics: Regression Analysis. Journal of the American Statistical Association. 92. 10.2307/2965740.
- [2] WMO (World Meteorological Organization), Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project – Report No. 58, 588 pp., Geneva, Switzerland, 2018.
- [3] Wohltmann, I., Rex, M., Brunner, D., Mäder, J.. (2005). Integrated equivalent latitude as a proxy for dynamical changes in ozone column. Geophysical Research Letters. 3210. 1029/2005GL022497.
- [4] Newman, P. A., Daniel, J. S., Waugh, D. W., and Nash, E. R.. A new formulation of equivalent effective stratospheric chlorine (EESC), Atmos. Chem. Phys., 7, 4537–4552, <https://doi.org/10.5194/acp-7-4537-2007>, 2007.
- [5] Bojkov, R. D., L. Bishop, W. J. Hill, G. C. Reinsel, and G. C. Tiao (1990), A statistical trend analysis of revised Dobson total ozone data over the northern hemisphere, J. Geophys. Res., 95, 9785–9807.
- [6] Ziemke, J., S. Chandra, R. McPeters, and P. Newman (1997), Dynamical proxies of column ozone with applications to global trend models, J. Geophys. Res., 102(D5), 6117– 6129.
- [7] CLUE: Chemistry, Life, the Universe, and Everything. (2022, April 3). Michigan State University and UC Bolder.<https://chem.libretexts.org/@go/page/354112>
- [8] Mäder, J. A., Staehelin, J., Brunner, D., Stahel, W. A., Wohltmann, I., and Peter, T. (2007), Statistical modeling of total ozone: Selection of appropriate explanatory variables, J. Geophys. Res., 112, D11108, doi:10.1029/2006JD007694.
- [9] Halogen Chemistry dataset reference:  
<https://gml.noaa.gov/dv/data.html>

<https://agage.mit.edu/data/agage-data>  
<https://gml.noaa.gov/hats/odgi.html>

[10] QBO dataset reference:

<https://www.geo.fuberlin.de/en/met/ag/strat/produkte/qbo/index.html>

[11] Heterogeneous Polar Chemistry dataset reference

<https://openbooks.lib.msu.edu/clue/chapter/chapter-4-heterogeneous-compounds/>

[12] Aerosol dataset reference:

<https://data.giss.nasa.gov/modelforce/strataer//tau.line2012.12.txt>

[13] Solar Cycle dataset reference:

<https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-radio/>