

**ANALISIS REGRESI DATA PANEL  
PADA TINGKAT CURAH HUJAN DI PULAU SUMATERA**

(Skripsi)

Oleh

Dian Tri Diaty



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2017**

## ABSTRACT

### PANEL DATA ANALYSIS FOR LEVEL OF RAINFALL DATA IN SUMATERA

By

**Dian Tri Diaty**

This study aims to determine the best model of the level of rainfall data in Sumatera using panel data regression. The results of analysis, starting from estimating the parameters model of panel data regression, selection of the best model, diagnostic checking on the best model of panel data regression, significance checking of the parameters, and interpret the best model of panel data regression, showed that the fixed effect model is the best model with the value of  $R^2 = 0.799444$ . The estimating model is:

$$\hat{Y}_{it} = -5277.236 + C_i + 17.76998X_{1it} + 300.1607X_{2it} + \varepsilon_{it}$$

**Key words:** Panel Data Regression, *Common Effect Model*, *Fixed Effect Model*, *Random Effect Model*.

## ABSTRAK

### ANALISIS REGRESI DATA PANEL PADA TINGKAT CURAH HUJAN DI PULAU SUMATERA

Oleh

**Dian Tri Diaty**

Penelitian ini bertujuan untuk menentukan model terbaik pada tingkat curah hujan di Pulau Sumatera menggunakan metode regresi data panel. Dari tahapan analisis yang dilakukan, yaitu mengestimasi parameter model regresi data panel, melakukan uji pemilihan model terbaik, melakukan uji diagnostik model regresi data panel terbaik, menguji signifikansi parameter, dan interpretasi model regresi data panel, diperoleh model efek tetap merupakan model regresi data panel terbaik dengan nilai  $R^2 = 0.799444$ . Persamaan model hasil estimasi adalah:

$$\hat{Y}_{it} = -5277.236 + C_i + 17.76998X_{1it} + 300.1607X_{2it} + \varepsilon_{it}.$$

**Kata kunci:** Regresi Data Panel, *Common Effect Model*, *Fixed Effect Model*, *Random Effect Model*.

**ANALISIS REGRESI DATA PANEL  
PADA TINGKAT CURAH HUJAN DI PULAU SUMATERA**

**Oleh**

**DIAN TRI DIATY**

**Skripsi**

**Sebagai Salah Satu Syarat Untuk Memperoleh Gelar  
SARJANA SAINS**

**Pada**

**Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2017**

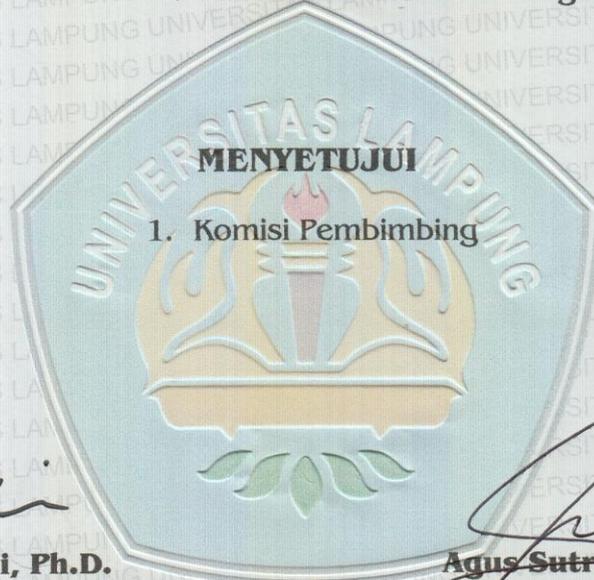
**Judul Skripsi** : **ANALISIS REGRESI DATA PANEL PADA  
TINGKAT CURAH HUJAN DI PULAU  
SUMATERA**

**Nama Mahasiswa** : **Dian Tri Diaty**

**Nomor Pokok Mahasiswa** : **1317031023**

**Jurusan** : **Matematika**

**Fakultas** : **Matematika dan Ilmu Pengetahuan Alam**

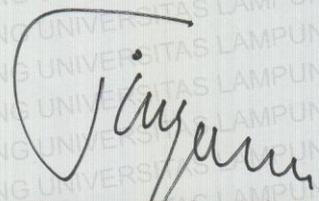


**1. Komisi Pembimbing**

  
**Netti Herawati, Ph.D.**  
NIP. 19650125 199003 2 001

  
**Agus Sutrisno, M.Si.**  
NIP. 19700831 199903 1 002

**2. Ketua Jurusan Matematika**

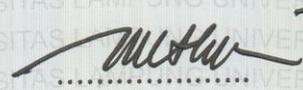
  
**Tiryono Ruby, Ph.D.**  
NIP. 19620704 198803 1 002

**MENGESAHKAN**

**1. Tim Penguji**

**Ketua**

**: Netti Herawati, Ph.D.**



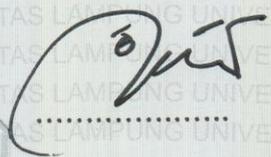
**Sekretaris**

**: Agus Sutrisno, M.Si.**



**Penguji**

**Bukan Pembimbing : Eri Setiawan, M.Si.**



**2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Prof. Warsito, S.Si., D.E.A., Ph.D.**

**NIP. 19710212 199512 1 00 1**



**Tanggal Lulus Ujian Skripsi: 8 Februari 2017**

## PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertandatangan di bawah ini :

Nama : **Dian Tri Diaty**

Nomor Pokok Mahasiswa : **1317031023**

Judul : **Analisis Data Panel pada Tingkat Curah Hujan di Pulau Sumatera**

Jurusan : **Matematika**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung.

Bandar Lampung, Februari 2017

Penulis,


Dian Tri Diaty  
NPM. 1317031023

## **RIWAYAT HIDUP**

Penulis bernama lengkap Dian Tri Diaty, dilahirkan di Ogan Komering Ulu Timur pada tanggal 12 Mei 1995 sebagai anak ketiga dari empat bersaudara pasangan Bapak Arifin Azis dan Ibu Sunasiah.

Penulis menyelesaikan pendidikan di TK Indriasana 3 Mojosari pada tahun 2001, sekolah dasar di SD Charitas 02 Mojosari pada tahun 2007, sekolah menengah pertama di SMP N 1 Belitang pada tahun 2010, dan sekolah menengah atas di SMA N 1 Belitang pada tahun 2013.

Pada tahun 2013 penulis terdaftar sebagai mahasiswa Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN). Selama menjadi mahasiswa, penulis bergabung dalam Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai anggota Bidang Keilmuan periode 2014-2015 hingga periode 2015-2016.

Pada Januari 2016 penulis melaksanakan Kerja Prakter di PT. Kereta Api Indonesia (KAI) Persero Sub. Divre 3.2 Tanjung Karang guna mengaplikasikan ilmu yang diperoleh dalam perkuliahan. Selanjutnya pada Juli 2016 penulis melaksanakan Kuliah Kerja Nyata di Desa Ngesti Rahayu, Kecamatan Punggur, Kabupaten Lampung Tengah.

## KATA INSPIRASI

*“Allah meninggikan orang-orang yang beriman di antara kamu dan orang-orang yang diberi ilmu pengetahuan beberapa derajat”*  
(QS. Al-Mujadillah: 11)

*“Ilmu adalah teman akrab dalam kesepian, sahabat dalam keterasingan, pengawas dalam kesedihan, penunjuk jalan ke arah yang benar, penolong disaat sulit, dan simpanan setelah kematian”*  
(‘Aidh Al-Qarni)

*“Ketika dunia ternyata jahat padamu, maka kamu harus menghadapinya. Karena tidak seorangpun yang akan menyelamatkanmu jika kamu tidak berusaha”*  
(Roronoa Zorro)

*“What is life without a little risk?”*  
(Dian Tri Diaty)

*Dengan mengucap Alhamdulillah,  
puji dan syukur kepada Allah SWT atas segala rahmat dan karunia-Nya, serta  
suri tauladan Nabi Muhammad SAW yang menjadi pedoman hidup dalam  
berikhtiar.*

*Kupersembahkan sebuah karya sederhana ini untuk:*

***Ayahanda Arifin Aziz & Ibunda Sunasiah***

*Terimakasih Ayah, Ibu untuk semua limpahan kasih sayang, pengorbanan, doa,  
dan dukungan selama ini. Karena atas ridho kalianlah Allah memudahkan setiap  
langkah-langkah yang aku tapaki.*

*Mungkin karya ini tak sebanding dengan pengorbanan yang telah kalian lakukan.  
Tapi percayalah ini sebuah titik awal perjuangan baktiku untuk kalian, karena  
kalian adalah motivasi terbesar dalam hidupku.*

*Kak Windy, Kak Retno, dede Ayang, dan sahabat-sahabatku yang senantiasa  
berdoa untuk keberhasilanku.*

*Serta,*

*Almamater tercinta yang turut dalam pembentukan pribadi menjadi lebih  
dewasa dalam berpikir, berucap, dan bertindak.*

## SANWACANA

Puji syukur kehadirat Allah SWT yang telah melimpahkan berkah dan rahmatNya sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Analisis Data Panel pada Tingkat Curah Hujan di Pulau Sumatera**” ini.

Penulis menyadari bahwa tanpa bimbingan, bantuan, dan doa dari berbagai pihak skripsi ini tidak akan dapat diselesaikan. Oleh karena itu, dalam kesempatan ini penulis menyampaikan terimakasih kepada:

1. Ibu Netti Herawati, Ph.D. selaku pembimbing utama atas kesediaan waktu, tenaga, pemikiran, dan pengarahan dalam proses penyusunan skripsi.
2. Bapak Agus Sutrisno, M.Si. selaku pembimbing kedua atas kesediaan waktu, tenaga, pemikiran, dan pengarahan yang telah diberikan.
3. Bapak Eri Setiawan, M.Si. selaku pembahas atas kesediaan waktu dan pemikiran dalam memberikan evaluasi, arahan, dan saran yang membangun dalam proses penyusunan skripsi ini.
4. Bapak Warsono, Ph.D. selaku pembimbing akademik yang telah memberi arahan, nasihat, dan meluangkan waktunya kepada penulis selama proses perkuliahan.
5. Bapak Tiryono Ruby, Ph.D. selaku Ketua Jurusan Matematika atas izin dan bantuan selama masa pendidikan.

6. Bapak Warsito, S.Si., D.E.A., Ph.D. selaku Dekan FMIPA Universitas Lampung.
7. Seluruh dosen Jurusan Matematika atas bimbingan, nasihat, dan ilmu yang diberikan selama masa studi.
8. Ibu, ayah, kak Windy, kak Retno, dan Ayang yang telah memberikan doa, dorongan, semangat, dan kasih sayang yang tulus kepada penulis.
9. Kak Deri, Annisa, Debi, Della, Eno, Yucky, Heni, Tina, Widya, Mila, Omyzha, dan Ici atas kebersamaan, keceriaan, dan dukungannya selama ini. Semoga akan terus berlanjut sampai kapanpun.
10. Teman-teman satu bimbingan atas bantuan, semangat, dan dukungannya dalam menyelesaikan skripsi ini. Jangan menyerah, kita pasti bisa!
11. Teman-teman seperjuangan Matematika angkatan 2013 atas keakraban dan kebersamaan selama ini.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Oleh karena itu, kritik dan saran dari pembaca akan sangat bermanfaat bagi penulis. Semoga skripsi ini dapat bermanfaat bagi semua pihak yang membaca.

Bandar Lampung, Februari 2017

Penulis,

**Dian Tri Diaty**

## DAFTAR ISI

Halaman

<b>DAFTAR TABEL</b> .....	iii
---------------------------	-----

### I. PENDAHULUAN

1.1 Latar Belakang dan Masalah .....	1
1.2 Tujuan Penelitian .....	3
1.3 Manfaat Penelitian .....	3

### II. TINJAUAN PUSTAKA

2.1 Model Regresi Linear .....	4
2.1.1 Model Regresi Linear Sederhana .....	4
2.1.2 Model Regresi Linear Berganda .....	5
2.2 Regresi Data Panel .....	5
2.3 Estimasi Regresi Data Panel .....	8
2.3.1 <i>Common Effect Model</i> (CEM) .....	8
2.3.1.1 <i>Ordinary Least Square</i> (OLS) .....	9
2.3.2 <i>Fixed Effect Model</i> (FEM) .....	10
2.3.2.1 <i>Least Square Dummy Variable</i> .....	11
2.3.3 <i>Random Effect Model</i> (REM) .....	12
2.3.3.1 <i>Generalized Least Square</i> .....	14
2.4 Pemilihan Model Estimasi Regresi Data Panel .....	15
2.4.1 Uji <i>Chow</i> .....	15
2.4.2 Uji <i>Hausman</i> .....	16
2.5 Pemeriksaan Persamaan Regresi .....	18
2.5.1 Galat Baku .....	18
2.5.2 Uji Hipotesis .....	19
2.5.2.1 Uji Serentak (Uji F) .....	20
2.5.2.1 Uji Parsial (Uji t) .....	20
2.5.3 Koefisien Determinasi .....	22

2.6 Uji Asumsi Model Regresi Data Panel .....	22
2.6.1 Uji Normalitas .....	23
2.6.2 Uji Multikolinearitas .....	24
2.6.3 Uji Heteroskedastisitas .....	25
2.6.4 Uji Autokorelasi .....	26
2.7 Curah Hujan .....	27
2.8 Faktor-faktor yang Mempengaruhi Curah Hujan .....	27

### **III. METODOLOGI PENELITIAN**

3.1 Waktu dan Tempat Penelitian .....	30
3.2 Data Penelitian .....	30
3.3 Metode Analisis Data .....	32

### **IV. HASIL DAN PEMBAHAAN**

4.1 Estimasi Parameter Model Regresi Data Panel .....	33
4.1.1 <i>Common Effect Model</i> (CEM) .....	33
4.1.2 <i>Fixed Effect Model</i> (FEM) .....	34
4.1.3 <i>Random Effect Model</i> (REM) .....	35
4.2 Pemilihan Model Estimasi Regresi Data Panel .....	37
4.2.1 Uji <i>Chow</i> .....	37
4.2.2 Uji <i>Hausman</i> .....	37
4.2.3 Pemilihan Model Akhir .....	38
4.3 Uji Asumsi .....	39
4.3.1 Uji Normalitas .....	40
4.3.2 Uji Heteroskedastisitas .....	40
4.3.3 Uji Autokorelasi .....	41
4.3.4 Uji Multikolinearitas .....	42
4.4 Pemeriksaan Persamaan Regresi .....	43
4.4.1 Uji F .....	43
4.4.2 Uji-t .....	43
4.4.3 Koefisien Determinasi .....	46

<b>V. KESIMPULAN</b> .....	47
----------------------------	----

### **DAFTAR PUSTAKA**

### **LAMPIRAN**

## DAFTAR TABEL

Tabel	Halaman
1. Aturan keputusan uji Durbin-Watson .....	26
2. Data Tingkat Curah di Sumatera tahun 2010-2013.....	30
3. Estimasi parameter <i>Common Effect Model</i> .....	33
4. Estimasi parameter <i>Fixed Effect Model</i> .....	34
5. Nilai variabel <i>dummy</i> dengan metode pengaruh tetap .....	35
6. Estimasi parameter <i>Random Effect Model</i> .....	36
7. Nilai galat acak dengan metode pengaruh acak .....	36
8. Hasil uji <i>Chow</i> .....	37
9. Hasil uji <i>Hausman</i> .....	38
10. Perbandingan nilai statistik REM dan FEM.....	39
11. Hasil uji normalitas dengan metode Kolmogorov-Smirnov .....	40
12. Hasil uji Park.....	41
13. Hasil uji Durbin-Watson .....	42
14. Hasil uji multikolinearitas .....	42
15. Hasil uji-F .....	43
16. Hasil uji-t.....	44
17. Nilai <i>R-square</i> model regresi data panel dengan FEM .....	46

## I. PENDAHULUAN

### 1.1 Latar Belakang dan Masalah

Regresi pertama kali diperkenalkan oleh Sir Francis Galton pada tahun 1886. Analisis regresi dipakai secara luas untuk melakukan prediksi dan ramalan. Analisis ini juga digunakan untuk memahami variabel bebas mana saja yang berhubungan dengan variabel terikat dan untuk mengetahui bentuk-bentuk hubungan tersebut.

Data panel merupakan salah satu perkembangan ilmu dari regresi. Data panel adalah gabungan antara data runtun waktu dan data silang. Data panel diperoleh dari data silang yang disurvei berulang kali pada objek atau unit individu yang sama dan pada waktu yang berlainan dengan tujuan untuk memperoleh gambaran tentang perilaku objek tersebut selama periode tertentu. Karena data panel merupakan gabungan dari data silang dan data runtun waktu maka akan mempunyai observasi lebih banyak dibanding data silang atau runtun waktu saja. Akibatnya, ketika digabungkan menjadi pool data guna membuat regresi maka hasilnya cenderung akan lebih baik dibandingkan regresi yang hanya menggunakan data silang atau runtun waktu saja (Nachrowi & Usman, 2006).

Ada tiga macam model data panel yaitu *Common Effect Model*, *Fixed Effect Model*, dan *Random Effect Model*. Model yang cocok untuk analisis regresi data

panel pada tingkat curah hujan diperoleh dari uji spesifikasi model data panel. Uji-uji tersebut yaitu uji Chow dan uji Hausman.

Negara Indonesia merupakan negara dengan iklim tropis yang memiliki dua musim, yaitu musim kemarau dan musim penghujan. Musim penghujan berperan dalam menunjang berlangsungnya proses kehidupan masyarakat Indonesia seperti produksi pertanian, perkebunan, perikanan, hingga penerbangan. Informasi tentang banyaknya curah hujan adalah salah satu unsur penting dan besar pengaruhnya terhadap segala macam aktifitas tersebut.

Curah hujan adalah endapan atau deposit air dalam bentuk cair maupun padat yang berasal dari atmosfer. Curah hujan mencakup tetes hujan, salju, batu es, dan embun. Curah hujan dipengaruhi oleh beberapa faktor antara lain kelembaban udara, temperatur, tekanan udara, intensitas sinar matahari, kecepatan angin, dan lain-lain. Prakiraan curah hujan dengan segala bentuk analisis dan informasi yang dihasilkan besar dampaknya guna membantu dan menunjang kegiatan sosial ekonomi di Indonesia.

Berdasarkan uraian di atas penulis akan menganalisis tingkat curah hujan di Pulau Sumatera berdasarkan faktor-faktor yang mempengaruhinya sehingga didapatkan model terbaik yang dapat digunakan untuk memprediksi tingkat curah hujan yang akan datang dan dapat memperkecil resiko yang tidak diinginkan. Untuk itu penulis mengambil judul “Analisis Regresi Data Panel pada Tingkat Curah Hujan di Pulau Sumatera”.

## **1.2 Tujuan Penelitian**

Penelitian ini bertujuan untuk menentukan model terbaik dalam menganalisis atau memprediksi data panel pada tingkat curah hujan di Pulau Sumatera.

## **1.3 Manfaat Penelitian**

Manfaat yang diperoleh dari hasil penelitian ini antara lain:

- a. Membantu penulis dalam mengaplikasikan ilmu yang telah didapat di bangku perkuliahan sehingga menunjang kesiapan untuk terjun ke dunia kerja.
- b. Dapat memperoleh model yang sesuai untuk memprediksi tingkat curah hujan di Pulau Sumatera.
- c. Sebagai bahan pertimbangan pihak berwenang dalam mengambil kebijakan yang tepat dan dapat mengetahui seberapa besar tingkat curah hujan yang akan datang.

## II. TINJAUAN PUSTAKA

### 2.1 Model Regresi Linear

#### 2.1.1 Model Regresi Linear Sederhana

Analisis regresi adalah salah satu metode statistik yang dapat digunakan untuk menyelidiki atau membangun model hubungan antara beberapa variabel. Dalam regresi sederhana, bentuk hubungan fungsi (keterkaitan antarvariabel) yang dipelajari adalah bentuk hubungan fungsi antara dua variabel (variabel bebas dan variabel terikat). Model yang dibuat pada regresi sederhana dapat berbentuk garis lurus atau bukan garis lurus. Apabila model yang dibuat tidak garis lurus maka model yang tidak terbentuk garis lurus tersebut sedapatnya ditransformasikan.

Model regresi linear sederhana secara matematis dapat dirumuskan sebagai berikut:

$$Y_i = \alpha + \beta X_i + \varepsilon_i; \quad i = 1, 2, \dots, n$$

$Y$  disebut variabel terikat,  $X$  adalah variabel bebas atau variabel penjelas,  $\varepsilon$  adalah variabel gangguan,  $\alpha$  dan  $\beta$  adalah parameter-parameter regresi. Subskrip  $i$  menunjukkan pengamatan ke- $i$ . Parameter  $\alpha$  dan  $\beta$  ditaksir atas dasar data yang tersedia untuk variabel  $X$  dan  $Y$ .

### 2.1.2 Model Regresi Berganda

Analisis regresi berganda merupakan perluasan dari analisis regresi sederhana yang melibatkan lebih dari satu variabel bebas. Bentuk umum model regresi berganda dengan  $k$  variabel bebas adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i$$

dengan:

$Y$  = variabel terikat  
 $\beta_0$  = intersep  
 $\beta_1, \beta_2, \dots, \beta_k$  = koefisien kemiringan  
 $X$  = variabel bebas  
 $\varepsilon$  = galat  
 $i$  = observasi ke- $i$

Beberapa asumsi yang penting dalam regresi linear berganda (Widarjono, 2005)

antara lain:

- a. Hubungan antara  $Y$  (variabel dependen) dan  $X$  (variabel independen) adalah linear dalam parameter.
- b. Tidak ada hubungan linear antara variabel independen atau tidak ada multikolinieritas antara variabel independen.
- c. Nilai rata-rata dari galat adalah nol,  $E(\varepsilon) = 0$ .
- d. Tidak ada korelasi antara  $(\varepsilon_i)$  dan  $(\varepsilon_j)$ .  $E(\varepsilon_i \varepsilon_j) = 0, i \neq j$ .
- e. Variansi setiap galat sama (homoskedastisitas).

### 2.2 Regresi Data Panel

Data panel adalah data yang merupakan hasil dari pengamatan pada beberapa individu (unit tabel silang) yang masing-masing diamati dalam beberapa periode

waktu yang berurutan (unit waktu) (Baltagi, 2005). Menurut Wanner dan Pevalin sebagaimana dikutip oleh Sembodo (2013), menyebutkan bahwa regresi panel merupakan sekumpulan teknik untuk memodelkan pengaruh peubah penjelas terhadap peubah respon pada data panel.

Ada beberapa model regresi panel, salah satunya adalah model dengan kemiringan konstan dan intersep bervariasi. Model regresi panel yang hanya dipengaruhi oleh salah satu unit saja (unit tabel silang atau unit waktu) disebut model komponen satu arah, sedangkan model regresi panel yang dipengaruhi oleh kedua unit (unit tabel silang dan unit waktu) disebut model komponen dua arah. Secara umum terdapat dua pendekatan yang digunakan dalam menduga model dari data panel yaitu model tanpa pengaruh individu (*common effect*) dan model dengan pengaruh individu (*fixed effect* dan *random effect*).

Data panel diperkenalkan pertama kali oleh Howles pada tahun 1950. Data panel merupakan gabungan dari data tabel silang dan data runtun waktu. Data tabel silang merupakan data yang terdiri dari sejumlah individu yang dikumpulkan pada suatu waktu tertentu. Sedangkan data runtun waktu merupakan data yang terdiri dari satu individu tetapi meliputi beberapa periode waktu tertentu.

Mengingat data panel merupakan gabungan dari data tabel silang dan data runtun waktu, maka bentuk umum model regresi data panel adalah sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

dengan:

$$Y_{it} = \text{pengamatan unit data tabel silang ke-}i \text{ waktu ke-}t$$

$$\alpha = \text{intersep}$$

$\beta$  = koefisien kemiringan untuk semua unit  
 $X_{it}$  = variabel bebas untuk unit data tabel silang ke- $i$  dan waktu ke- $t$   
 $\varepsilon_{it}$  = nilai galat pada unit data tabel silang ke- $i$  dan waktu ke- $t$

Secara umum dengan menggunakan data panel akan dihasilkan koefisien intersep dan kemiringan yang berbeda pada setiap individu dan setiap periode waktu.

Oleh karena itu, dalam mengestimasi model regresi data panel akan sangat bergantung pada asumsi yang dibuat tentang intersep, koefisien kemiringan, dan variabel gangguannya. Ada beberapa kemungkinan yang akan muncul, yaitu:

1. Semua koefisien baik intersep maupun kemiringan konstan.

$$Y_{it} = \beta_1 + \sum_{k=2}^{\kappa} \beta_k X_{kit} + \varepsilon_{it}$$

2. Koefisien kemiringan konstan, tetapi intersep berbeda untuk semua individu.

$$Y_{it} = \beta_{1i} + \sum_{k=2}^{\kappa} \beta_k X_{kit} + \varepsilon_{it}$$

3. Koefisien kemiringan konstan, tetapi intersep berbeda baik sepanjang waktu maupun antarindividu.

$$Y_{it} = \beta_{1it} + \sum_{k=2}^{\kappa} \beta_k X_{kit} + \varepsilon_{it}$$

4. Intersep dan koefisien kemiringan berbeda untuk semua individu.

$$Y_{it} = \beta_{1i} + \sum_{k=2}^{\kappa} \beta_{ki} X_{kit} + \varepsilon_{it}$$

5. Intersep dan koefisien kemiringan berbeda sepanjang waktu dan untuk semua individu.

$$Y_{it} = \beta_{1it} + \sum_{k=2}^{\kappa} \beta_{kit} X_{kit} + \varepsilon_{it}$$

dengan:

$$i = 1, 2, \dots, N$$

$$t = 1, 2, \dots, T$$

$N$  = banyak unit data silang

$T$  = banyak data runtun waktu

$Y_{it}$  = nilai variabel terikat data silang ke- $i$  dan runtun waktu ke- $t$

$X_{it}$  = nilai variabel bebas ke- $k$  untuk data silang ke- $i$  dan runtun waktu ke- $t$

$\beta_{it}$  = parameter yang ditaksir

$\varepsilon_{it}$  = unsur gangguan populasi

$K$  = banyak parameter regresi yang ditaksir

### 2.3 Estimasi Regresi Data Panel

Berdasarkan asumsi pengaruh yang digunakan dalam regresi data panel, model regresi data panel dibagi menjadi 3, yaitu *Common Effect model*, *Fixed Effect model*, dan *Random Effect Model*.

#### 2.3.1 *Common Effect Model (CEM)*

Menurut Baltagi (2005) model tanpa pengaruh individu adalah pendugaan yang menggabungkan seluruh data runtun waktu dan tabel silang dan menggunakan pendekatan OLS (*Ordinary Least Square*) untuk menduga parameteranya. Metode OLS merupakan salah satu metode populer untuk menduga nilai parameter dalam persamaan regresi linear. Secara umum, persamaan modelnya dituliskan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}$$

dengan:

- $Y_{it}$  = variabel respon pada unit observasi ke- $i$  dan waktu ke- $t$
- $X_{it}$  = variabel prediktor pada unit observasi ke- $i$  dan waktu ke- $t$
- $\beta$  = koefisien kemiringan atau koefisien arah
- $\alpha$  = intersep model regresi
- $\varepsilon_{it}$  = galat atau komponen galat pada unit observasi ke- $i$  dan waktu ke- $t$

### 2.3.1.1 Ordinary Least Square (OLS)

Menurut Nachrowi & Usman (2006) data panel tentunya akan mempunyai observasi lebih banyak dibanding data tabel silang atau runtun waktu saja. Akibatnya, ketika data digabungkan guna membuat regresi data panel maka hasilnya cenderung akan lebih baik dibanding regresi yang hanya menggunakan data tabel silang atau runtun waktu saja.

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

Bila  $cov(\varepsilon_{it}, \varepsilon_{jt}) = 0$ ;  $cov(\varepsilon_{it}, \varepsilon_{it-1}) = 0$ ;  $E(\varepsilon_{it}) = 0$ ; dan  $var(\varepsilon_{it}) = \sigma^2$ , kita dapat mengestimasi model tersebut dengan memisahkan waktunya sehingga ada  $T$  regresi dengan  $N$  pengamatan. Atau dapat dituliskan dengan:

$$\begin{aligned} Y_{i1} &= \alpha + \beta X_{i1} + \varepsilon_{i1}; \quad i = 1, 2, \dots, N \\ Y_{i2} &= \alpha + \beta X_{i2} + \varepsilon_{i2} \\ &\vdots \\ Y_{iT} &= \alpha + \beta X_{iT} + \varepsilon_{iT} \end{aligned}$$

Model juga dapat diestimasi dengan memisahkan tabel silang-nya sehingga didapat  $N$  regresi dengan masing-masing  $T$  pengamatan. Atau dapat ditulis dengan:

$$\begin{aligned} i = 1; & Y_{1t} = \alpha + \beta X_{1t} + \varepsilon_{1t}; \quad t = 1, 2, \dots, T \\ i = 2; & Y_{2t} = \alpha + \beta X_{2t} + \varepsilon_{2t} \\ &\vdots \\ i = N; & Y_{Nt} = \alpha + \beta X_{Nt} + \varepsilon_{Nt} \end{aligned}$$

Jika dipunyai asumsi bahwa  $\alpha$  dan  $\beta$  akan sama (konstan) untuk setiap data runtun waktu dan tabel silang, maka  $\alpha$  dan  $\beta$  dapat diestimasi dengan model berikut dengan menggunakan  $N \times T$  pengamatan.

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

### 2.3.2 *Fixed Effect Model* (FEM)

*Fixed Effect Model* adalah metode regresi yang mengestimasi data panel dengan menambahkan variabel boneka. Model ini mengasumsi bahwa terdapat efek yang berbeda antarindividu. Perbedaan itu dapat diakomodasi melalui perbedaan pada intersepnya. Oleh karena itu dalam FEM, setiap individu merupakan parameter yang tidak diketahui dan akan diestimasi dengan menggunakan teknik variabel boneka sehingga metode ini seringkali disebut dengan *Least Square Dummy Variable* model. Persamaan regresi pada *Fixed Effect Model* adalah:

$$Y_{it} = \alpha_i + \sum_{k=2}^N a_k D_{ki} + \beta_i X_{it} + \varepsilon_{it}$$

dengan:

- $Y_{it}$  = variabel terikat untuk individu ke-i pada waktu ke-t
- $X_{it}$  = variabel bebas ke-j untuk individu ke-i pada waktu ke-t
- $D_{ki}$  = variabel boneka
- $\varepsilon_{it}$  = komponen galat untuk individu ke-i pada waktu ke-t
- $\beta_i$  = parameter untuk variabel ke-i

Gujarati (2004) mengatakan bahwa pada *Fixed Effect Model* diasumsikan bahwa koefisien kemiringan bernilai konstan tetapi intersep bersifat tidak konstan.

### 2.3.2.1 Least Square Dummy Variable (LSDV)

Menurut Greene (2007), secara umum pendugaan parameter model efek tetap dilakukan dengan LSDV (*Least Square Dummy Variable*), di mana LSDV merupakan suatu metode yang dipakai dalam pendugaan parameter regresi linear dengan menggunakan Metode Kuadrat Terkecil (MKT) pada model yang melibatkan variabel boneka sebagai salah satu variabel prediktornya. MKT merupakan teknik pengepasan garis lurus terbaik untuk menghubungkan variabel prediktor dan variabel respon. Berikut adalah prinsip dasar MKT:

$$u = Y - X\beta$$

Sehingga didapatkan Jumlah Kuadrat Galat sebagai berikut:

$$\begin{aligned} u'u &= (Y - X\beta)'(Y - X\beta) \\ &= Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \end{aligned}$$

Di mana, jika matriks transpos  $(X\beta)' = \beta'X'$ , maka skalar  $\beta'X'Y = Y'X\beta$ . Untuk mendapatkan penduga parameter yang menyebabkan jumlah kuadrat galat minimum, yaitu dengan cara menurunkan persamaan terhadap parameter yang kemudian hasil turunan tersebut disamakan dengan nol atau  $\frac{\partial(u'u)}{\partial\beta} = 0$ , sehingga diperoleh:

$$\begin{aligned} \frac{\partial(Y'Y - 2\beta'X'Y + \beta'X'X\beta)}{\partial\beta} &= 0 \\ \Leftrightarrow -2X'Y + 2X'X\hat{\beta} &= 0 \\ \Leftrightarrow 2X'X\hat{\beta} &= 2X'Y \\ \Leftrightarrow X'X\hat{\beta} &= X'Y \end{aligned}$$

$$\Leftrightarrow (X'X)^{-1}(X'X)\hat{\beta} = (X'X)^{-1}X'Y$$

$$\Leftrightarrow \hat{\beta} = (X'X)^{-1}X'Y$$

Pada pemodelan efek tetap, variabel boneka yang dibentuk adalah sebanyak  $N - 1$ , sehingga model yang akan diduga dalam pemodelan efek tetap adalah sebagai berikut:

$$Y_{it} = \alpha + \alpha_1 D_{1it} + \dots + \alpha_N D_{(N-1)it} \beta X_{it} + u_{it}$$

Sedangkan untuk pemodelan efek tetap waktu, variabel boneka yang dibentuk berdasarkan unit waktu, di mana variabel boneka yang terbentuk sebanyak  $T - 1$ , sehingga model yang akan diduga dalam pemodelan efek tetap waktu adalah sebagai berikut:

$$Y_{it} = \alpha + \alpha_1 D_{1it} + \dots + \alpha_T D_{(T-1)it} \beta X_{it} + u_{it}$$

dengan:

$D_{jit}$  = peubah boneka ke- $j$  ( $j = 1, 2, \dots, (N - 1)$ ) unit data silang ke- $i$  dan unit waktu ke- $t$ .  $D_{jit}$  bernilai satu jika  $j = i$  dan bernilai nol jika  $j \neq i$ .

$D_{kit}$  = peubah boneka ke- $k$  ( $k = 1, 2, \dots, (T - 1)$ ) unit data silang ke- $i$  dan unit waktu ke- $t$ .  $D_{kit}$  bernilai satu jika  $k = i$  dan bernilai nol jika  $k \neq i$ .

$\alpha_j$  = rata-rata peubah respon jika peubah boneka ke- $j$  bernilai satu dan peubah penjelas bernilai nol.

$\alpha_k$  = rata-rata nilai peubah respon jika peubah boneka ke- $k$  bernilai satu dan peubah penjelas bernilai nol.

### 2.3.3 *Random Effect Model (REM)*

Menurut Nachrowi & Usman (2006), sebagaimana telah diketahui bahwa pada model dengan pengaruh tetap (FEM), perbedaan karakteristik-karakteristik individu dan waktu diakomodasikan pada intersep sehingga intersepanya berubah

antar waktu. Sementara *Random Effect Model* (REM) perbedaan karakteristik individu dan waktu diakomodasikan pada galat dari model. Mengingat ada dua komponen yang mempunyai kontribusi pada pembentukan galat, yaitu individu dan waktu, maka galat acak pada REM juga perlu diurai menjadi galat untuk komponen waktu dan galat gabungan. Dengan demikian persamaan REM diformulasikan sebagai berikut:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it};$$

$$\varepsilon_{it} = u_i + v_t + w_{it}$$

dengan:

$u_i$ : komponen galat tabel silang  
 $v_t$ : komponen galat runtun waktu  
 $w_{it}$ : komponen galat gabungan

Adapun asumsi yang digunakan untuk komponen galat tersebut adalah:

$$u_i \sim N(0, \sigma_u^2)$$

$$v_t \sim N(0, \sigma_v^2)$$

$$w_{it} \sim N(0, \sigma_w^2)$$

Melihat persamaan di atas, maka dapat dinyatakan bahwa REM menganggap efek rata-rata dari data tabel silang dan runtun waktu direpresentasikan dalam intersep.

Sedangkan deviasi efek secara acak untuk data runtun waktu direpresentasikan dalam  $v_t$  dan deviasi untuk data tabel silang dinyatakan dalam  $u_i$ .

$$\varepsilon_{it} = u_i + v_t + w_{it}$$

Dengan demikian varians dari galat tersebut dapat dituliskan dengan:

$$\text{Var}(\varepsilon_{it}) = \sigma_u^2 + \sigma_v^2 + \sigma_w^2$$

Hal ini tentunya berbeda dengan model OLS yang diterapkan pada data panel (*pooled data*), yang mempunyai varian galat sebesar:  $Var(\varepsilon_{it}) = \sigma_w^2$ .

Dengan demikian, REM bisa diestimasi dengan OLS bila  $\sigma_u^2 = \sigma_v^2 = 0$ . Jika tidak demikian, REM perlu diestimasi dengan metode lain. Adapun metode estimasi yang digunakan adalah *Generalized Least Square* (GLS).

### 2.3.3.1 Generalized Least Square (GLS)

Untuk *Random Effect Model* (REM), pendugaan parameternya dilakukan menggunakan *Generalized Least Square* jika matriks diketahui. Pada REM ketidaklengkapan informasi untuk setiap unit tabel silang dipandang sebagai galat sehingga  $\mu_i$  adalah bagian dari unsur gangguan. Model REM dapat dituliskan sebagai berikut.

$$Y_{it} = \alpha + \sum_{k=1}^k \beta_k X_{kit} + (\mu_i + e_{it})$$

Asumsi:

$$\mu_i \sim N(0, \sigma_\mu^2) E(\mu_i, \mu_j) = 0; \quad i \neq j$$

$$e_{it} \sim N(0, \sigma_v^2) E(\mu_i, e_{it}) = 0$$

$$E(e_{it}, e_{is}) = E(e_{it}, e_{jt}) = E(e_{it}, e_{js}) = 0; \quad i \neq j; \quad t \neq s$$

Untuk data tabel silang ke- $i$  persamaan di atas dapat ditulis  $y_i = X_i \beta + (\mu_i \mathbf{1} + e_i)$ . Varians komponen dari unsur gangguan  $(\mu_i \mathbf{1} + e_i)$  untuk unit tabel silang ke- $i$  adalah:

$$\Omega = \begin{bmatrix} \sigma_{\mu}^2 + \sigma_e^2 & \sigma_{\mu}^2 & \cdots & \sigma_{\mu}^2 \\ \sigma_{\mu}^2 & \sigma_{\mu}^2 + \sigma_e^2 & \cdots & \sigma_{\mu}^2 \\ \cdots & \cdots & \ddots & \cdots \\ \sigma_{\mu}^2 & \sigma_{\mu}^2 & \cdots & \sigma_{\mu}^2 + \sigma_e^2 \end{bmatrix}$$

Varians komponen identik untuk setiap unit tabel silang. Sehingga varians komponen untuk seluruh observasi dapat dituliskan:

$$W = \begin{bmatrix} \Omega & 0 & \cdots & 0 \\ 0 & \Omega & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \Omega \end{bmatrix}$$

Jika nilai diketahui maka persamaan dapat diduga menggunakan *Generalized Least Square* (GLS) dengan  $\hat{\beta} = (X'W^{-1}X)^{-1}(X'W^{-1}y)$ . Jika tidak diketahui maka perlu diduga dengan menduga  $\hat{\sigma}_{\mu}^2$  dan  $\hat{\sigma}_e^2$ , sehingga persamaan di atas diduga dengan  $\hat{\beta} = (X'\hat{W}^{-1}X)^{-1}(X'\hat{W}^{-1}y)$  di mana  $\hat{\sigma}_e^2 = \frac{\hat{e}'\hat{e}}{NT-N-K}$  dengan  $\hat{e} = y - X\hat{\beta}$  adalah residu dari *Least Square Dummy Variable* (LSDV).

Sedangkan  $\hat{\sigma}_{\mu}^2 = \frac{\hat{\sigma}_1^2 - \hat{\sigma}_e^2}{T}$ .

## 2.4 Pemilihan Model Estimasi Regresi Data Panel

### 2.4.1 Uji Chow

Uji ini digunakan untuk memilih salah satu model pada regresi data panel, yaitu antara model efek tetap (FEM) dengan model koefisien tetap (CEM). Prosedur pengujiannya sebagai berikut (Baltagi, 2005).

Hipotesis:

$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$  (efek unit data silang secara keseluruhan tidak berarti)

$H_1$ : Minimal terdapat satu  $\alpha_i \neq 0$ ;  $i = 1, 2, \dots, n$  (efek wilayah berarti)

Statistik uji yang digunakan merupakan uji F, yaitu:

$$F_{hitung} = \frac{[RRSS - URSS]/(n - 1)}{URSS/(nT - n - K)}$$

dengan:

$n$  = jumlah individu

$T$  = jumlah periode waktu

$K$  = jumlah variabel penjelas

$RRSS$  = *restricted residual sums of squares* yang berasal dari model koefisien tetap

$URSS$  = *unrestricted residual sums of squares* yang berasal dari model efek tetap

Jika nilai atau  $F_{hitung} > F_{(n-1, nT-n-K)}$  atau  $p\text{-value} < (\text{taraf signifikansi})$  maka tolak hipotesis awal sehingga model yang terpilih adalah model efek tetap.

#### 2.4.2 Uji Hausman

Uji ini digunakan untuk memilih model efek acak (*REM*) dengan model efek tetap (*FEM*). Uji ini bekerja dengan menguji apakah terdapat hubungan antara galat pada model (galat komposit) dengan satu atau lebih variabel penjelas (independen) dalam model. Hipotesis awalnya adalah tidak terdapat hubungan antara galat model dengan satu atau lebih variabel penjelas. Prosedur pengujiannya sebagai berikut (Baltagi, 2008).

Hipotesis:

$H_0$ : Korelasi  $(X_{it}, \varepsilon_{it}) = 0$  (efek data silang tidak berhubungan dengan regresor lain)

$H_1$ : Korelasi  $(X_{it}, \varepsilon_{it}) \neq 0$  (efek data silang berhubungan dengan regresor lain)

Statistik uji yang digunakan adalah uji *chi-squared* berdasarkan kriteria *Wald*, yaitu:

$$W = \hat{q}' [\text{var}(\hat{q}')]^{-1} \hat{q}$$

$$\Leftrightarrow W = (\hat{\beta}_{MET} - \hat{\beta}_{MEA})' [\text{var}(\hat{\beta}_{MET} - \hat{\beta}_{MEA})]^{-1} (\hat{\beta}_{MET} - \hat{\beta}_{MEA})$$

dengan:

$$\hat{\beta}_{MET} = \text{vektor estimasi kemiringan model efek tetap}$$

$$\hat{\beta}_{MEA} = \text{vektor estimasi kemiringan model efek acak}$$

Jika nilai  $W > \chi^2_{(\alpha, K)}$  atau nilai *p-value* kurang dari taraf signifikansi yang ditentukan maka tolak hipotesis awal sehingga model yang terpilih adalah model efek tetap.

Dalam perhitungan statistik Uji Hausman diperlukan asumsi bahwa banyaknya kategori tabel silang lebih besar dibandingkan jumlah variabel independen (termasuk konstanta) dalam model. Lebih lanjut, dalam estimasi statistik Uji Hausman diperlukan estimasi variansi tabel silang yang positif, yang tidak selalu dapat dipenuhi oleh model. Apabila kondisi-kondisi ini tidak dipenuhi maka hanya dapat digunakan model efek tetap.

## 2.5 Pemeriksaan Persamaan Regresi

Menurut Nachrowi & Usman (2006), baik atau buruknya regresi yang dibuat dapat dilihat berdasarkan beberapa indikator, yaitu meliputi galat baku, uji hipotesis, dan koefisien determinasi ( $R^2$ ).

### 2.5.1 Galat Baku

Metode yang digunakan untuk menduga model dilandasi pada prinsip meminimalkan galat. Oleh karena itu, ketepatan dari nilai dugaan sangat ditentukan oleh galat baku dari masing-masing penduga. Adapun galat baku dirumuskan sebagai berikut.

$$s.e(b_1) = \left\{ \frac{\sigma^2}{\sum (X_i - \bar{X})^2} \right\}^{1/2}$$

$$s.e(b_0) = \left\{ \frac{\sum X_i^2}{N \sum (X_i - \bar{X})^2} \right\}^{1/2} \sigma$$

Oleh karena  $\sigma$  merupakan penyimpangan yang terjadi dalam populasi yang nilainya tidak diketahui maka  $\sigma$  biasanya diduga berdasarkan data sampel.

Adapun penduganya adalah sebagai berikut.

$$s = \left( \frac{\sum u_i^2}{N - 2} \right)^{1/2}$$

$$u_i^2 = (Y_i - \hat{y}_i)^2$$

Terlihat hubungan galat yang minimal akan mengakibatkan galat baku koefisien minimal pula. Dengan minimalnya galat baku koefisien menunjukkan koefisien yang didapat cenderung mendekati nilai sebenarnya. Bila rasio tersebut bernilai 2

atau lebih, dapat dinyatakan bahwa nilai galat baku relatif kecil dibanding parameternya. Rasio inilah yang menjadi acuan pada Uji t.

### **2.5.2 Uji Hipotesis**

Pengujian hipotesis dalam penelitian ini dapat diukur dari *goodness of fit* fungsi regresinya. Secara statistik, analisa ini dapat diukur dari nilai statistik t, nilai statistik F, dan koefisien determinasi (Kuncoro, 2011). Analisis regresi ini bertujuan untuk mengetahui secara parsial maupun simultan pengaruh variabel independen terhadap variabel dependen serta untuk mengetahui proporsi variabel independen dalam menjelaskan perubahan variabel dependen.

Uji hipotesis berguna untuk memeriksa atau menguji apakah koefisien regresi yang didapat signifikan. Maksud dari signifikan adalah suatu nilai koefisien regresi yang secara statistik tidak sama dengan nol. Jika koefisien kemiringan sama dengan nol, berarti dapat dikatakan bahwa tidak cukup bukti untuk menyatakan variabel bebas mempunyai pengaruh terhadap variabel terikat. Untuk kepentingan tersebut, maka semua koefisien regresi harus diuji.

Ada dua jenis uji hipotesis terhadap koefisien regresi yang dapat dilakukan, yang disebut Uji F dan Uji t. Uji F digunakan untuk menguji koefisien (kemiringan) regresi secara bersama-sama, sedang Uji t untuk menguji koefisien regresi, termasuk intersep secara individu.

### 2.5.2.1 Uji Serentak (Uji F)

Uji statistik F pada dasarnya menunjukkan apakah semua variabel independen dalam model mempunyai pengaruh secara bersama-sama terhadap variabel dependen (Kuncoro, 2011). Uji-F diperuntukan guna melakukan uji hipotesis koefisien (kemiringan) regresi secara bersamaan. Dengan demikian, secara umum hipotesisnya dituliskan sebagai berikut,

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$H_1$ : Tidak demikian (paling tidak ada satu kemiringan yang  $\neq 0$ )

di mana  $k$  adalah banyaknya variabel bebas.

Statistik uji:

$$F_{hitung} = \frac{R^2/(n + K - 1)}{(1 - R^2)/(nT - n - K)}$$

dengan:

$R^2$  = koefisien determinasi  
 $n$  = jumlah data silang  
 $T$  = jumlah data runtun waktu  
 $K$  = jumlah variabel independen

Kriteria uji:  $H_0$  ditolak jika  $F_{hitung} > F_{(\alpha, n+K, nT-n-K)}$ , artinya bahwa hubungan antara semua variabel independen dan variabel dependen berpengaruh signifikan (Gujarati, 2004).

### 2.5.2.2 Uji Parsial (Uji t)

Untuk mengetahui pengaruh signifikan setiap variabel independen terhadap variabel dependen menggunakan uji t. Adapun hipotesis dalam uji ini adalah sebagai berikut,

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; \quad j = 0,1,2, \dots, k \text{ (} k \text{ adalah koefisien kemiringan)}$$

Dari hipotesis tersebut dapat terlihat arti dari pengujian yang dilakukan, yaitu berdasarkan data yang tersedia akan dilakukan pengujian terhadap koefisien regresi apakah sama dengan nol yang berarti variabel bebas tidak mempunyai pengaruh signifikan terhadap variabel terikat, atau tidak sama dengan nol yang berarti variabel bebas mempunyai pengaruh signifikan terhadap variabel terikat.

Uji t didefinisikan sebagai berikut,

$$t = \frac{b_j - \beta_j}{s.e(b_j)}$$

Tetapi karena  $\beta_j$  akan diuji apakah sama dengan nol, maka nilai  $\beta_j$  dalam persamaan harus diganti dengan nol. Maka formula uji t menjadi,

$$t = \frac{b_j}{s.e(b_j)}$$

Nilai  $t$  di atas akan dibandingkan dengan nilai  $t$  tabel. Bila ternyata setelah dihitung  $|t_{hitung}| > t_{(\alpha/2, nT-n-K)}$ , maka nilai  $t$  berada dalam daerah penolakan, sehingga hipotesis nol ditolak pada tingkat kepercayaan  $(1 - \alpha) - 100\%$ . Dalam hal ini dapat dikatakan bahwa  $\beta_j$  signifikan secara statistika.

Khusus untuk Uji t ini dapat dibuat batasan daerah penolakan secara praktis, yaitu bila derajat bebas = 20 atau lebih dan  $\alpha = 5\%$ , maka hipotesis  $\beta_j = 0$  akan ditolak jika,

$$|t| = \frac{b_j}{s.e(b_j)} > 2.$$

### 2.5.3 Koefisien Determinasi

Menurut Nachrowi & Usman (2006), koefisien determinasi (*Goodness of Fit*), yang dinotasikan dengan  $R^2$ , merupakan suatu ukuran yang penting dalam regresi karena dapat menginformasikan baik atau tidaknya model regresi yang terestimasi. Atau dengan kata lain, angka tersebut dapat mengukur seberapa dekat garis regresi yang terestimasi dengan data sesungguhnya.

Nilai koefisien determinasi ( $R^2$ ) ini mencerminkan seberapa besar variasi dari variabel terikat  $Y$  dapat diterangkan oleh variabel bebas  $X$ . Bila nilai koefisien determinasi sama dengan 0 ( $R^2 = 0$ ), artinya variasi dari  $Y$  tidak dapat diterangkan oleh sama sekali  $X$ . Sementara bila  $R^2 = 1$ , artinya variasi  $Y$  secara keseluruhan dapat diterangkan oleh  $X$ . Dengan kata lain  $R^2 = 1$ , maka semua pengamatan berada tepat pada garis regresi. Dengan demikian baik atau buruknya suatu persamaan regresi ditentukan oleh  $R^2$ -nya yang mempunyai nilai antara nol dan satu.

## 2.6 Uji Asumsi Model Regresi Data Panel

Menurut Yudiatmaja (2013), model regresi data panel dapat disebut sebagai model yang baik jika model tersebut memenuhi kriteria *Best, Linear, Unbiased*, dan *Estimator* (BLUE). BLUE dapat dicapai bila memenuhi asumsi klasik. Apabila persamaan yang terbentuk tidak memenuhi kaidah BLUE, maka persamaan tersebut diragukan kemampuannya dalam menghasilkan nilai-nilai prediksi yang akurat. Tetapi bukan berarti persamaan tersebut tidak bisa digunakan untuk memprediksi. Agar suatu persamaan tersebut dapat

dikategorikan memenuhi kaidah BLUE, maka data yang digunakan harus memenuhi beberapa asumsi yang sering dikenal dengan istilah uji asumsi klasik. Uji asumsi klasik mencakup uji normalitas, uji multikolinearitas, uji heterokedastisitas, dan uji autokorelasi (Widarjono, 2007).

### **2.6.1 Uji Normalitas**

Data klasifikasi kontinu dan data kuantitatif yang termasuk dalam pengukuran data skala interval atau ratio agar dapat dilakukan uji statistik parametrik dipersyaratkan berdistribusi normal. Pembuktian data berdistribusi normal tersebut perlu dilakukan uji normalitas terhadap data. Uji normalitas berguna untuk membuktikan data dari sampel yang dimiliki berasal dari populasi berdistribusi normal. Banyak cara yang dapat dilakukan untuk membuktikan suatu data berdistribusi normal atau tidak.

Berdasarkan pengalaman empiris beberapa pakar statistik, data yang banyaknya lebih dari 30 angka ( $n > 30$ ) maka sudah dapat diasumsikan berdistribusi normal. Biasa dikatakan sebagai sampel besar. Namun untuk memberikan kepastian data yang dimiliki berdistribusi normal atau tidak, sebaiknya digunakan uji statistik normalitas karena belum tentu data yang lebih dari 30 dapat dipastikan berdistribusi normal. Demikian sebaliknya data yang banyaknya kurang dari 30 belum tentu tidak berdistribusi normal, untuk itu perlu suatu pembuktian.

Banyak jenis uji statistik normalitas yang dapat digunakan diantaranya *Kolmogorov Smirnov*, *Lilliefors*, *Chi-Square*, *Shapiro Wilk*, atau menggunakan

perangkat lunak komputer. Pada hakekatnya perangkat lunak tersebut merupakan hitungan uji statistik yang telah diprogram dalam perangkat lunak komputer.

### **2.6.2 Uji Multikolinearitas**

Asumsi multikolinearitas adalah asumsi yang menunjukkan adanya hubungan linear yang kuat di antara beberapa variabel prediktor dalam suatu model regresi linear berganda. Model regresi yang baik memiliki variabel-variabel prediktor yang independen atau tidak berkorelasi. Penyebab terjadinya kasus multikolinearitas adalah terdapat korelasi atau hubungan linear yang kuat di antara beberapa variabel prediktor yang dimasukkan kedalam model regresi.

Multikolinearitas digunakan untuk menguji suatu model apakah terjadi hubungan yang sempurna atau hampir sempurna antara variabel bebas, sehingga sulit untuk memisahkan pengaruh antara variabel-variabel itu secara individu terhadap variabel terikat. Pengujian ini untuk mengetahui apakah antar variabel bebas dalam persamaan regresi tersebut tidak saling berkorelasi.

Beberapa indikator dalam mendeteksi adanya multikolinearitas, diantaranya

(Gujarati, 2006):

- (1) Nilai  $R^2$  yang terlampau tinggi (lebih dari 0,8) tetapi tidak ada atau sedikit t-statistik yang signifikan; dan
- (2) Nilai F-statistik yang signifikan, namun t-statistik dari masing-masing variabel bebas tidak signifikan.

Untuk menguji multikolinearitas dapat melihat matriks korelasi dari variabel bebas, jika terjadi koefisien korelasi lebih dari 0,80 maka terdapat multikolinearitas (Gujarati, 2006). Atau dengan memeriksa nilai *Variance Inflation Factor* (VIF) dari masing-masing variabel bebas. Multikolinearitas terjadi ketika nilai  $VIF > 10$ .

### 2.6.3 Uji Heteroskedastisitas

Uji heteroskedastisitas digunakan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik heteroskedastisitas yaitu adanya ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi. Prasyarat yang harus terpenuhi dalam model regresi adalah tidak adanya gejala heteroskedastisitas.

Ada beberapa metode pengujian yang bisa digunakan dalam melihat ada tidaknya permasalahan heteroskedastisitas ini, salah satunya adalah uji glejser. Uji glejser adalah uji hipotesis untuk mengetahui apakah sebuah model regresi memiliki indikasi heteroskedastisitas dengan cara meregresikakn mutlak galat. Jika nilai signifikansi antara variabel bebas dengan mutlak galat lebih dari 5% maka tidak terjadi masalah heteroskedastisitas.

Kriteria pengujian sebagai berikut:

$H_0$  : tidak ada gejala heteroskedastisitas

$H_1$  : : ada gejala heteroskedastisitas

$H_0$  tidak ditolak bila  $-t_{tabel} < |t_{hitung}| < t_{tabel}$ , berarti tidak terdapat heteroskedastisitas dan  $H_0$  ditolak bila  $t_{hitung} > t_{tabel}$  atau  $-t_{hitung} < -t_{tabel}$  yang berarti terdapat heteroskedastisitas.

#### 2.6.4 Uji Autokorelasi

Uji autokorelasi digunakan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik autokorelasi yaitu korelasi yang terjadi antara residual pada satu pengamatan dengan pengamatan lain pada model regresi. Autokorelasi merupakan pelanggaran salah satu asumsi dari model regresi klasik, yaitu faktor gangguan dari setiap pengamatan yang berbeda tidak saling mempengaruhi prasyarat yang harus terpenuhi. Metode pengujian yang sering digunakan adalah uji Durbin-Watson (uji DW) dengan ketentuan sebagai berikut:

Tabel 1. Aturan keputusan uji Durbin-Watson

Hipotesis null	Keputusan	Jika
Tidak terdapat autokorelasi positif	Tolak $\frac{H_0}{H_1}$	$0 < \frac{d}{d^*} < \frac{d^*}{d^*}$
Tidak terdapat autokorelasi positif	Tidak ada keputusan	$d_L < d < d_U$
Tidak terdapat autokorelasi negatif	Tolak $\frac{H_0}{H_1}$	$(4 - \frac{d}{d^*}) < \frac{d_U}{d^*} < \frac{d^*}{d^*}$
Tidak terdapat autokorelasi negatif	Tidak ada keputusan	$(4 - d_U) < d < (4 - d_L)$
Tidak terdapat autokorelasi, positif maupun negatif	Tidak tolak $\frac{H_0}{H_1}$	$d_U < d < (4 - d_U)$

$H_0: \rho = 0$  (tidak ada korelasi residual)

$H_1: \rho \neq 0$  (terdapat korelasi residual, positif maupun negatif)

Jika  $d_U < d < (4 - d_U)$  maka tidak tolak  $H_0$ . Nilai  $d_L$  dan  $d_U$  dapat diperoleh dari tabel statistik Durbin-Watson yang bergantung pada banyaknya observasi dan banyaknya variabel yang menjelaskan. Jika hipotesis nol tidak ditolak, maka dapat disimpulkan bahwa tidak ada autokorelasi. Residual tidak saling berkorelasi, sehingga analisa regresi tidak mempunyai masalah autokorelasi.

## 2.7 Curah Hujan

Curah hujan adalah banyaknya air yang jatuh ke permukaan bumi. Derajat curah hujan dinyatakan dengan jumlah curah hujan dalam suatu satuan waktu. Biasanya satuan yang digunakan adalah mm/jam. Dalam meteorologi, butiran hujan dengan diameter lebih dari 0.5 mm disebut hujan dan diameter antara 0.5 – 0.1 mm disebut gerimis. Semakin besar ukuran butiran hujan maka semakin besar pula kecepatan jatuhnya. Ketelitian alat ukur curah hujan adalah 1/10 mm. Pembacaan dilakukan satu kali dalam sehari dan dicatat sebagai curah hujan hari terdahulu/kemarin (Suyono, 1985).

## 2.8 Faktor-faktor yang Mempengaruhi Curah Hujan

### a. Kelembaban Udara

Kelembaban adalah perbandingan antara massa uap dalam suatu satuan volume dengan massa uap yang jenuh dalam satuan volume itu pada suhu yang sama.

Secara umum kelembaban menyatakan banyaknya kadar air yang ada di udara. Banyaknya uap yang bergerak di dalam atmosfer berpengaruh terhadap besarnya hujan, lamanya hujan, dan intensitas curah hujan. Kelembaban tertinggi umumnya terjadi pada musim penghujan dan paling rendah pada musim kemarau.

Umumnya semakin tinggi suatu daerah dari permukaan laut maka kelembaban udaranya semakin tinggi. Makin tinggi kelembaban udara akan dapat menyebabkan bertambah banyak uap air yang dapat diserap awan. Uap air itu akan menghasilkan tekanan yang dinyatakan dengan satuan tinggi air raksa ( $1 \text{ mmHg} = 1,33 \text{ milibar}$ ). Tekanan yang diberikan oleh uap air disebut dengan tekanan uap air (Suyono, 1985).

#### b. Tekanan Udara

Tekanan udara merupakan tenaga yang bekerja untuk menggerakkan massa udara dalam setiap satuan luas tertentu. Diukur dengan menggunakan barometer. Satuan tekanan udara adalah milibar (mb). Garis yang menghubungkan tempat-tempat yang sama tekanan udaranya disebut sebagai isobar.

Tekanan udara dibatasi oleh ruang dan waktu. Artinya pada tempat dan waktu yang berbeda, besarnya juga berbeda. Daerah yang suhu udaranya tinggi akan bertekanan rendah dan daerah yang bersuhu udara rendah tekanannya tinggi (Soewarno, 2000).

### c. Suhu Udara

Suhu udara adalah keadaan panas atau dinginnya udara. Suhu juga disebut temperatur yang diukur dengan alat termometer. Beberapa faktor yang mempengaruhi suhu udara diantaranya tinggi tempat, daratan/lautan, radiasi matahari, indeks datang matahari, dan angin. Pengukuran biasa dinyatakan dalam skala Celsius (C), Reamur (R), dan Fahrenheit (F). Suhu udara tertinggi di permukaan bumi adalah di daerah tropis (sekitar ekuator) dan makin ke kutub makin dingin (Soewarno, 2000).

### d. Kecepatan Angin

Angin adalah udara yang bergerak akibat adanya perbedaan tekanan udara dengan arah aliran angin dari tempat yang memiliki tekanan tinggi ke tempat yang bertekanan rendah atau dari daerah yang memiliki temperatur rendah ke wilayah bersuhu tinggi. Angin memiliki hubungan yang erat dengan sinar matahari karena daerah yang terkena banyak paparan sinar matahari akan memiliki suhu yang lebih tinggi serta tekanan udara yang lebih rendah dari daerah lain di sekitarnya sehingga menyebabkan terjadinya aliran udara. Angin juga dapat disebabkan oleh pergerakan benda sehingga mendorong udara di sekitarnya untuk bergerak ke tempat lain (Soewarno, 2000).

### III. METODOLOGI PENELITIAN

#### 3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2016/2017 di jurusan Matematika fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### 3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data tingkat curah hujan di Pulau Sumatera beserta beberapa faktor yang mempengaruhinya diantaranya data kecepatan angin dan rata-rata suhu udara tahun 2010-2013 yang diperoleh dari situs resmi BPS Republik Indonesia.

Tabel 2. Data Tingkat Curah di Sumatera tahun 2010-2013

No	Provinsi	Tahun	Tingkat curah hujan (mm)	Kecepatan angin (m/det)	Suhu rata-rata (C)
1	NAD	2010	1986	4,8	27,1
		2011	1268	4,9	27,1
		2012	1098	4,8	26,9
		2013	1623	5,2	27
2	Sumatera Utara	2010	2184	1,7	27,1
		2011	2042	1,8	27,2
		2012	3175	0,6	27,3
		2013	2627	1,99	28,77

Tabel 2. Lanjutan

No	Provinsi	Tahun	Tingkat curah hujan (mm)	Kecepatan angin (m/det)	Suhu rata-rata (C)
3	Sumatera Barat	2010	5228	2	25,8
		2011	4691	0,5	25,5
		2012	4339	0,3	25,2
		2013	4627,4	1,06	25,13
4	Riau	2010	3390	6,6	27,7
		2011	2405	5,4	27
		2012	2636	5,9	27,3
		2013	2628	6,6	27,7
5	Jambi	2010	3207	4,8	27,1
		2011	2295	5,5	26,9
		2012	1874	5,1	26,7
		2013	2299	1	27,1
6	Sumatera Selatan	2010	2389	2,8	27,4
		2011	2593	2,7	27,3
		2012	3083	3	27,4
		2013	3409,2	3,6	27,3
7	Bengkulu	2010	3822	2	26,8
		2011	3850	5,1	26,5
		2012	2545	2	26,9
		2013	3980,9	3,57	26,74
8	Lampung	2010	2710	2,4	26,7
		2011	1568	4	26,8
		2012	1685	4,2	26,8
		2013	2456,7	1,5	26,7
9	Kepulauan Bangka	2010	3444	2,5	27
		2011	2921	3,4	27,4
		2012	2018	3	27
		2013	2839,2	5,67	27
10	Kepulauan Riau	2010	3283	6,9	27
		2011	3893	6,8	26,8
		2012	3253	6,9	26,8
		2013	3389,4	7	26,77

### 3.3 Metode Analisis Data

Metode yang dilakukan dalam penelitian ini ada dua yaitu analisis statistik deskriptif dan analisis regresi data panel. Adapun langkah-langkah yang akan dilakukan dalam penelitian ini adalah:

1. Mengestimasi parameter model regresi data panel pada data Tingkat Curah Hujan di Sumatera tahun 2010 - 2013 dengan pendekatan *Common Effect Model*, *Fixed Effect Model*, dan *Random Effect Model*. Metode yang digunakan adalah *Ordinary Least Square*, *Least Square Dummy Variable*, dan *Generalized Least Square*.
2. Melakukan uji pemilihan model terbaik menggunakan Chow dan Uji Hausman.
3. Melakukan uji asumsi data panel, yaitu pengujian normalitas, pengujian autokorelasi, pengujian heteroskedastisitas, dan pengujian multikolinearitas.
4. Melakukan uji parameter model regresi data panel terbaik, meliputi pemeriksaan uji hipotesis yaitu uji-F dan uji-t serta pemeriksaan koefisien determinasi.
5. Interpretasi model regresi.

## V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, diperoleh kesimpulan bahwa:

1. Model regresi data panel yang sesuai untuk pemodelan tingkat curah hujan di Pulau Sumatera tahun 2010 hingga 2013 adalah *Fixed Effect Model* (FEM) dengan model persamaan hasil estimasi adalah sebagai berikut:

$$\hat{Y}_{it} = -5277.236 + C_i + 17.76998X_{1it} + 300.1607X_{2it} + \varepsilon_{it}$$

2. Sebesar 72.07% faktor kecepatan angin dan rata-rata suhu udara bersama-sama mempengaruhi tingkat curah hujan sedangkan sisanya sebesar 27.93% dijelaskan oleh faktor lain di luar penelitian dan seluruh variabel bebas mempunyai pengaruh signifikan terhadap tingkat curah hujan secara statistik pada  $\alpha = 5\%$ .

## DAFTAR PUSTAKA

- Agustin, A. R. 2016. Penerapan Analisis Regresi Data Panel pada Ketahanan Pangan Provinsi Lampung Tahun 2010-2013. Skripsi. Jurusan Matematika FMIPA Unila, Bandar Lampung.
- Baltagi, B. H. 2005. *Econometrics Analysis of Panel Data*. Third edition. John Wiley & Sons Ltd., Chicester.
- Baltagi, B. H. 2008. *Econometrics*. Fourth edition. Springer Verlag, Berlin Heidelberg.
- Greene, W. H. 2007. *Econometric Analysis*. Sixth edition. Prentice Hall International, New Jersey.
- Gujarati, D. N. 2004. *Basic Econometrics*. Fourth edition. The McGraw-Hill Companies, New York.
- Hsiao, C. 2003. *Analysis Of Panel Data*. Cambridge University Press, Southern California.
- Nachrowi, D. N. & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan*. Lembaga Penerbit FE UI, Jakarta.
- Sembodo, H. 2013. Pemodelan Regresi Panel pada Pendapatan Asli Daerah (PAD) dan Dana Alokasi Umum (DAU) Terhadap Belanja Daerah. *Jurnal Mahasiswa Statistik*, 1(4): 297-300.
- Soewarno. 2000. *Hidrologi Operasional*. Jilid kesatu. Citra Aditya Bakti, Bandung.
- Suyono, S. 1985. *Hidrologi*. PT Pradnya Paramita, Jakarta.
- Widarjono, A. 2007. *Ekonometrika Teori dan Aplikasi Untuk Ekonomi dan Bisnis*. Edisi kedua. Ekonisia FE UII, Yogyakarta.
- Yudiatmaja, F. 2013. *Analisis Regresi dengan Menggunakan Aplikasi Komputer Statistika SPSS*. Gramedia Pustaka Utama, Jakarta.