TUGAS AKHIR – KS184822

# REGRESI *RIDGE ROBUST* UNTUK PEMODELAN LUAS PANEN PADI DI KABUPATEN GRESIK DENGAN INDIKATOR IKLIM

ANGGUN NURFITRIANI HANDOKO
NRP 062116 4000 0106

Dosen Pembimbing
Dr. Sutikno S.Si., M.Si.

PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020

TUGAS AKHIR – KS184822

# REGRESI *RIDGE ROBUST* UNTUK PEMODELAN LUAS PANEN PADI DI KABUPATEN GRESIK DENGAN INDIKATOR IKLIM

**ANGGUN NURFITRIANI HANDOKO**
**NRP 062116 4000 0106**

**Dosen Pembimbing**
**Dr. Sutikno S.Si., M.Si.**

**PROGRAM STUDI SARJANA**
**DEPARTEMEN STATISTIKA**
**FAKULTAS SAINS DAN ANALITIKA DATA**
**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**
**SURABAYA 2020**

(*Halaman ini sengaja dikosongkan*)

FINAL PROJECT – KS184822

# ROBUST RIDGE REGRESSION FOR MODELING THE HARVESTED AREA OF PADDY IN GRESIK USING CLIMATE INDICATOR

**ANGGUN NURFITRIANI HANDOKO**
**SN 062116 4000 0106**

**Supervisor**
**Dr. Sutikno S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME**
**DEPARTMENT OF STATISTICS**
**FACULTY OF SCIENCE AND DATA ANALYTICS**
**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**
**SURABAYA 2020**

(*Halaman ini sengaja dikosongkan*)

## LEMBAR PENGESAHAN

# REGRESI *RIDGE ROBUST* UNTUK PEMODELAN LUAS PANEN PADI DI KABUPATEN GRESIK DENGAN INDIKATOR IKLIM

**TUGAS AKHIR**
Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh:
**Anggun Nurfitriani Handoko**
NRP. 062116 4000 0106

Disetujui oleh Pembimbing Tugas Akhir:
Dr. Sutikno, S.Si., M.Si.
NIP. 19710313 199702 1 001 ( )

Mengetahui,
Kepala Departemen Statistika

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN INSTITUT TEKNOLOGI SEPULUH NOPEMBER FAKULTAS SAINS DAN ANALITIKA DATA DEPARTEMEN STATISTIKA

**Dr. Dra. Kartika Fithriasari, M.Si**
NIP. 19691212 199303 2 002
SURABAYA, Agustus 2020

(*Halaman ini sengaja dikosongkan*)

# REGRESI *RIDGE ROBUST* UNTUK PEMODELAN LUAS PANEN PADI DI KABUPATEN GRESIK DENGAN INDIKATOR IKLIM

**Nama Mahasiswa : Anggun Nurfitriani Handoko**
**NRP : 062116 4000 0106**
**Departemen : Statistika-FSAD-ITS**
**Dosen Pembimbing : Dr. Sutikno, S.Si, M.Si**

## Abstrak

*Kabupaten Gresik memiliki produktivitas tanaman padi tertinggi kedua setelah Kota Malang. Pada sektor pertanian, khususnya padi, iklim berpengaruh pada luas area panen, produktivitas, dan kualitas hasil panen. Identifikasi hubungan antara luas panen dan indikator iklim dibutuhkan untuk memodelkan hubungan dari kedua variabel ini. Adanya keragaman curah hujan di wilayah Kabupaten Gresik dan luas panen padi serta dependensi antar data curah hujan menyebabkan amatan outlier dan multikolinearitas. Penelitian ini digunakan metode robust untuk mengatasi amatan outlier dan metode ridge untuk mengatasi masalah multikolinearitas. Hasil penelitian menunjukkan bahwa luas panen padi pada periode pertama terdapat pengamatan outlier, sehingga model optimum adalah regresi robust S-Estimation. Sedangkan, untuk periode kedua dan ketiga yang terdapat amatan outlier dan multikolinearitas, model terbaik diperoleh melalui penggabungan antara metode ridge dan robust, yaitu Regresi Ridge Robust S-Estimation untuk periode kedua dan Regresi Ridge Robust M-Estimation untuk periode ketiga. Dari model yang optimum diperoleh hasil estimasi untuk luas panen padi tahun 2018 dengan akurasi 99,17 persen.*

***Kata Kunci: Luas Panen Padi, Multikolinieritas, Outlier, Ridge, Robust, Ridge Robust***

(*Halaman ini sengaja dikosongkan*)

# *ROBUST RIDGE* REGRESSION FOR MODELING THE HARVESTED AREA OF PADDY IN GRESIK USING CLIMATE INDICATOR

**Name : Anggun Nurfitriani Handoko**
**Student Number : 062116 4000 0106**
**Department : Statistics**
**Supervisor : Dr. Sutikno, S.Si, M.Si**

## Abstract

*Gresik has the second highest productivity in paddy after Malang. In the agricultural sector, especially paddy, the climate affects the harvested area, productivity, and the quality of the crop. Identification the relation between harvests area and climate indicator is required for modeling the relationship of both variables. However, the rainfall diversity in Gresik and harvested area data as well as the correlations between rainfall data caused outlier cases and multicollinearity cases. In this study, robust regression applied to resolve the outliers and ridge regression to overcome the multicollinearity. The result of study indicates that the harvested area in the first period that have outliers in it, was better to estimated by the robust S-Estimation regression method. Whereas, for the second and third periods that contain outliers and multicollinearity case, the best found from combining the method between ridge and robust regression, these are Ridge Robust S-Estimation Regression for the second period and Ridge Robust M-Estimation Regression for the third period. From the best model the estimation results obtained for the harvested area in 2018 with an accuracy of 99.17 percent.*

***Keywords: Harvested Area of Paddy, Multicolinierity, Ridge, Robust, Ridge Robust***

(*Halaman ini sengaja dikosongkan*)

# KATA PENGANTAR

Puji syukur kehadirat Tuhan Yang Maha Esa, atas segala berkat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **"Regresi *Ridge Robust* untuk Pemodelan Luas Panen Padi di Kabupaten Gresik dengan Indikator Iklim"** dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Allah SWT yang selalu memberikan penulis begitu banyak kesempatan dan rahmat.
2. Kedua orang tua dan adik penulis yang selalu memberikan dukungan dalam bentuk semangat maupun doa sehingga penulis dapat terus termotivasi.
3. Dr. Sutikno, S.Si., M.Si. selaku dosen pembimbing Tugas Akhir yang senantiasa memberikan bimbingan, pengarahan, serta selalu sabar dan percaya kepada kemampuan penulis selama proses menyelesaikan Tugas Akhir ini.
4. Dr. Bambang W. Otok, M.Si. dan Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D. selaku dosen penguji yang telah menginspirasi serta banyak memberi masukan kepada penulis dalam proses menyelesaikan Tugas Akhir ini.
5. Dr. Dra. Kartika Fithriasari, M.Si. selaku Kepala Departemen Statistika dan Dr. Santi Wulan Purnami, S.Si, M.Si. selaku Sekretaris Departemen Statistika Bidang Akademik dan Kemahasiswaan yang telah memberikan fasilitas, sarana, dan prasaranan.
6. Dr.rer.pol. Dedy Dwi Prastyo, S.Si., M.Si. selaku dosen wali dan seluruh dosen serta karyawan Departemen Statistika ITS yang dengan ikhlas memberikan pengetahuan baru yang bermanfaat, pengalaman, dan bantuan kepada penulis selama proses perkuliahan.
7. Teman-teman Statistika ITS angkatan 2016, TR16GER, yang dari awal menjadi keluarga yang baik, saling

menyemangati dan berjuang bersama menjalani dunia perkuliahan serta memberikan banyak bantuan maupun dukungan pada penulis selama ini.

8. Sahabat *squad* penulis (Bhut Jolokia dan Siwoceh) yang selalu menjadi pendengar setia dan memberikan pandangan serta motivasi untuk penulis dalam bentuk apapun.
9. Indri dan Cinto, selaku teman seperjuangan penulis yang selalu memberikan masukan dalam proses kuliah maupun penulisan Tugas Akhir.
10. Abid, Naufal, Jefry, Sofi, Reza, Fadhli, yang senantiasa memberikan saran dan banyak pengetahuan baru yang berguna bagi penulis selama masa perkuliahan terutama selama proses pengerjaan Tugas Akhir ini mengenai konsep dasar pertanian maupun metode yang penulis gunakan.
11. Diri sendiri, yang selalu berani untuk mengambil resiko dan mencoba hal baru tanpa berpikir untuk menyerah dan berhenti.
12. Semua pihak yang telah membantu dalam penyelesaian Tugas Akhir ini yang tidak bisa penulis sebutkan satu persatu.

Terlepas dari semua itu, penulis menyadari bahwa Tugas Akhir ini masih banyak kekurangan. Oleh karena itu, dengan terbuka penulis menerima segala kritik dan saran dari pembaca sebagai bahan perbaikan bagi penulisan selanjutnya. Selain itu, penulis berharap laporan Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juni 2020

Penulis

## DAFTAR ISI

Halaman


**HALAMAN JUDUL**…… i
**LEMBAR PENGESAHAN**……..……..... v
**ABSTRAK**……..……..... vii
**ABSTRACT**……..……..... ix
**KATA PENGANTAR**……...……... xi
**DAFTAR GAMBAR**……..... xv
**DAFTAR TABEL**……...……... xvii
**DAFTAR LAMPIRAN**…….. xix
**BAB I PENDAHULUAN…….. 1**
1.1 Latar Belakang…… 1
1.2 Rumusan Masalah…… 3
1.3 Tujuan……..... 3
1.4 Manfaat……...…… 4
1.5 Batasan Masalah…….. 4
**BAB II TINJAUAN PUSTAKA……..... 5**
2.1 Statistika Deskriptif……. 5
2.2 *Boxplot*……..…….. 7
2.3 *Outlier*……. 8
2.3.1 *Difference in Fitted Value* (*DfFITS* )……..…… 9
2.4 Analisis Regresi……...…… 9
2.4.1 Uji Parameter Regresi secara Serentak……..…. 10
2.4.2 Uji Parameter Regresi Partial……..…… 11
2.5 Regresi *Robust*……..…….. 11
2.5.1 *Efficiency*……..… 12
2.5.2 *Breakdown Point*……..……... 12
2.6 Metode Estimasi Regresi *Robust*…… 13
2.6.1 *M-Estimation*……..…… 13
2.6.2 *S-Estimation*……..…… 16
2.6.3 *MM-Estimation*…… 18
2.7 Multikolinearitas……..…… 19
2.8 Regresi *Ridge*……..…… 19
2.8.1 Metode *Centering* dan *Rescaling*…… 20

| | | |
|---|---|---|
| 2.8.2 | Konstanta Regresi *Ridge*…… | 21 |
| 2.9 | Regresi *Ridge Robust*……..…… | 21 |
| 2.10 | Uji Normalitas……..…… | 22 |
| 2.11 | Uji Independensi…….…… | 23 |
| 2.12 | Curah Hujan…….…… | 24 |
| 2.13 | Luas Panen Padi……..…… | 24 |
| 2.14 | Indeks Performansi……..…… | 25 |
| **BAB III** | **METODE PENELITIAN……** | **27** |
| 3.1 | Sumber Data…….…… | 27 |
| 3.2 | Variabel Penelitian……..…… | 27 |
| 3.3 | Langkah Analisis…….…… | 28 |
| **BAB IV** | **ANALISIS DAN PEMBAHASAN……** | **31** |
| 4.1. | Karakteristik Data dan Identifikasi Outlier……. | 31 |
| 4.1.1 | Deskripsi Luas Panen di Kabupaten Gresik……. | 31 |
| 4.1.2 | Deskripsi Curah Hujan di Kabupaten Gresik…..… | 33 |
| 4.1.3 | Identifikasi Pola Hubungan Variabel Prediktor terhadap Respon……...…… | 35 |
| 4.1.4 | Identifikasi *Outlier*…….. | 38 |
| 4.2 | Pemodelan Luas Panen Padi Kabupaten Gresik….. | 40 |
| 4.2.1 | Pemodelan *Ordinary Least Square* (OLS)…….. | 40 |
| 4.2.2 | Pemodelan Regresi *Robust*…….. | 44 |
| 4.2.3 | Pemodelan *Ridge Robust*…… | 48 |
| 4.3 | Metode Estimasi Terbaik…… | 50 |
| 4.3.1 | Pemilihan Metode Estimasi Terbaik…… | 50 |
| 4.3.2 | Penggunaan Model Estimasi…… | 51 |
| **BAB V** | **KESIMPULAN DAN SARAN……** | **55** |
| 5.1 | Kesimpulan……..……... | 55 |
| 5.2 | Saran……..……..……... | 56 |
| **DAFTAR PUSTAKA……..……..……..** | | **57** |
| **LAMPIRAN……..……...** | | **59** |

# DAFTAR GAMBAR

Halaman


| | | |
|---|---|---|
| **Gambar 2.1** | Contoh *Boxplot*……..… | 7 |
| **Gambar 3.1** | Langkah Analisis Data……...…… | 29 |
| **Gambar 4.1** | Luas Panen Padi Kabupaten Gresik tiap Periode…… | 32 |
| **Gambar 4.2** | Rata-Rata Curah Hujan tiap bulan di Kabupaten Gresik Tahun 2006-2017……. | 33 |
| **Gambar 4.3** | *Scatterplot* antara Curah Hujan dengan Luas Panen di Kabupaten Gresik…….. | 36 |
| **Gambar 4.4** | *Boxplot* Luas Panen Padi per Periode di Kabupaten Gresik…… | 38 |
| **Gambar 4.5** | *Boxplot* Curah Hujan tiap Bulan di Kabupaten Gresik…… | 39 |
| **Gambar 4.6** | Perbandingan Data Aktual dan Prediksi Luas Panen Padi Tahun 2018 di Kabupaten Gresik……... | 53 |

(*Halaman ini sengaja dikosongkan*)

## DAFTAR TABEL

Halaman


| | | |
|---|---|---|
| **Tabel 2.1** | Estimasi M-Tukey *Bisquare*…… | 15 |
| **Tabel 3.1** | Variabel Penelitian…….….. | 27 |
| **Tabel 3.2** | Struktur Data Penelitian Periode 1……..... | 28 |
| **Tabel 4.1** | Mean, Standar Deviasi, Minimum, dan Maksimum Luas Panen Padi Kabupaten Gresik Tahun 2006-2017 (dalam hektar)…….……..... | 33 |
| **Tabel 4.2** | Mean, Standar Deviasi, Minimum, dan Maksimum Curah Hujan tiap Bulan di Kabupaten Gresik Tahun 2006-2017..... | 34 |
| **Tabel 4.3** | Koefisien dan P-value Korelasi Pearson antara Luas Panen Padi dengan Curah Hujan di Kabupaten Gresik Tahun 2006-2017 ……. | 37 |
| **Tabel 4.4** | Nilai *DFFITS* Pengamatan *Outlier* Tahun 2006-2017…….. | 40 |
| **Tabel 4.5** | Statistik $F_{hitung}$ dari Model Hasil Regresi OLS tiap Periode di Kabupaten Gresik Tahun 2006-2017…. | 41 |
| **Tabel 4.6** | Estimasi Parameter Regresi OLS di Kabupaten Gresik…….…….. | 41 |
| **Tabel 4.7** | Nilai p-value Pengujian Asumsi Normalitas Residual……..... | 42 |
| **Tabel 4.8** | Nilai VIF Variabel dengan Curah Hijan…….. | 43 |
| **Tabel 4.9** | Estimasi Parameter Regresi *Robust M-Estimation*…… | 44 |
| **Tabel 4.10** | Estimasi Parameter Regresi *Robust MM-Estimation*…… | 45 |
| **Tabel 4.11** | Estimasi Parameter Regresi *Robust S-Estimation*…….……..……. | 47 |
| **Tabel 4.12** | Perbandingan Nilai Konstanta Regresi *Ridge 3* Estimasi…….…..…… | 48 |
| **Tabel 4.13** | Perbandingan *Ridge Robust* dengan 3 Metode Estimasi Tahun 2006-2017…….. | 49 |
| **Tabel 4.14** | Perbandingan RMSE……...…… | 51 |

(*Halaman ini sengaja dikosongkan*)

## DAFTAR LAMPIRAN

Halaman


**Lampiran 1A.** Data Luas Panen dan Curah Hujan Kabupaten Gresik Periode 1……. 59

**Lampiran 1B.** Data Luas Panen dan Curah Hujan Kabupaten Gresik Periode 2……..……... 59

**Lampiran 1C.** Data Luas Panen dan Curah Hujan Kabupaten Gresik Periode 3……..…….. 60

**Lampiran 2A.** *Output Ordinary Least Square* Periode 1…. 60

**Lampiran 2B.** *Output Ordinary Least Square* Periode 2…. 61

**Lampiran 2C.** *Output Ordinary Least Square* Periode 3…. 61

**Lampiran 3.** *Output DFFITS*……...…...…..…….. 62

**Lampiran 4.** *Output* Pengujian Asumsi Normalitas Residual……...……..… 63

**Lampiran 5.** *Output* Residual Plot Hasil OLS……. 64

**Lampiran 6.** Koefisien dan P-value Korelasi Pearson antara Luas Panen dengan Curah Hujan….. 65

**Lampiran 7.** Nilai Variance Inflation Factor (VIF)…….. 66

**Lampiran 8.** Syntax Regresi *Robust* dengan R……. 66

**Lampiran 9A.** *Output* Regresi *Robust* M, MM, dan S Periode 1 dengan R……..…...…… 67

**Lampiran 9B.** *Output* Regresi *Robust* M, MM, dan S Periode 2 dengan R……..…...…...… 68

**Lampiran 9C.** *Output* Regresi *Robust* M, MM, dan S Periode 3 dengan R……...……..… 69

**Lampiran 10.** Syntax *Centering* dan *Rescaling* dengan R 70

**Lampiran 11A.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik Periode 1……...…....…… 71

**Lampiran 11B.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik Periode 2……..…...…...…… 72

**Lampiran 11C.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik

| | | |
|---|---|---|
| | Periode 3…….…...…...…… | 72 |
| **Lampiran 12.** | Syntax Regresi *Ridge*……...…...… | 72 |
| **Lampiran 13A.** | Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 1……..…...…...… | 75 |
| **Lampiran 13B.** | Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 2……..……...…...… | 77 |
| **Lampiran 13C.** | Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 3…..……...…...…… | 79 |
| **Lampiran 14A.** | Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 1……..……...…...…… | 81 |
| **Lampiran 14B.** | Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 2……..……...…...…… | 83 |
| **Lampiran 14C.** | Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 3…….……...…...…… | 85 |
| **Lampiran 15.** | Surat Pernyataan Data……...…...… | 87 |

# BAB I

# PENDAHULUAN

## 1.1 Latar Belakang

Kabupaten Gresik memiliki angka luas panen yang tinggi, walaupun jumlahnya tidak seluas tiga kabupaten penyumbang utama yaitu Kabupaten Lamongan, Jember, dan Bojonegoro. Namun produktivitas tanaman padi di Kabupaten Gresik tergolong tinggi bahkan merupakan salah satu kabupaten dengan produktivitas tertinggi di Jawa Timur. Produktivitas tanaman padi di Kabupaten Gresik menduduki urutan nomor dua setelah Kota Malang pada tahun 2017.

Pemerintah Kabupaten Gresik berupaya meningkatan produksi padi. Beberapa upaya yang dilakukan oleh pemerintah Kabupaten Gresik mulai dari mempelajari faktor yang menyebabkan keragaman luas panen padi, pengembangan bibit unggul yang sesuai dengan karakter daerah, intensifikasi pupuk organik dan penyediaan sarana produksi pertanian. Salah satu upaya lainnya adalah melakukan analisis keadaan alam seperti iklim yang berpengaruh terhadap hasil pertanian. Sebagian besar luas lahan sawah di Kabupaten Gresik adalah sawah non-irigasi, sehingga hujan menjadi salah satu sumber penting dalam sistem pengairannya. Selain itu, pada sektor pertanian, terutama padi, yang merupakan subsektor paling rentan terhadap perubahan keragaman hujan. Menurut Haryono (2011), padi merupakan tanaman pangan rentan terhadap keragaman hujan. Keragaman hujan berdampak pada luas area tanam maupun panen, produktivitas, serta kualitas hasil panen.

Analisis Statistika yang digunakan untuk mengetahui hubungan antara variabel dependen dan variabel independen adalah analisis regresi. Parameter regresi seringkali diestimasi menggunakan metode *Ordinary Least Square* (OLS), namun untuk beberapa kejadian ekstrem yang menyebabkan adanya amatan *outlier*, metode OLS tidak cukup baik untuk diaplikasikan. Hal ini karena OLS untuk regresi sensitif terhadap data *outlier* (Greene,

2002). Oleh karena itu, digunakan analisis menggunakan metode Regresi *Robust*. Regresi *Robust* merupakan solusi dari *least square regression* yang digunakan untuk asumsi yang tidak terpenuhi.

Beberapa metode estimasi parameter dikembangkan dalam penyelesaian Regresi *Robust*, diantaranya *M-Estimation* (*Maximum Likelihood Estimation*), *S-Estimation* (*Scale Estimation*), *MM-Estimation* (*Method of Moment Estimation*), *Least Median Squares* (LMS), dan *Least Trimmed Squares* (LTS). Kelima metode estimasi parameter pada Regresi *Robust* tersebut mempunyai kelemahan dan kelebihan masing-masing. Selain asumsi *outlier*, pada regresi berganda yang melibatkan beberapa variabel independen perlu dilakukan pengujian asumsi multikolinearitas. Apabila terjadi pelanggaran asumsi multikolinearitas perlu diatasi, salah satu metode yang dapat digunakan adalah Regresi *Ridge*.

Kurniati, Yundari, dan Rizky (2019) membandingkan metode OLS dengan Regresi *Robust* LMS pada produksi jeruk di Indonesia tahun 2016. Hasil penelitian menunjukkan bahwa metode OLS menunjukkan kinerja yang lebih baik dibandingkan dengan Regresi *Robust* LMS yang digunakan. Penelitian lainnya dilakukan Zahiroh (2017) membandingkan antara OLS, Regresi *Robust S-Estimation*, dan *MM-Estimation* untuk memodelkan luas panen padi dengan indikator iklim di sentra produksi padi Jawa Timur. Dari hasil penelitian tersebut menyimpulkan bahwa Regresi *Robust* menghasilkan variabel signifikan lebih banyak serta ketepatan prediksi yang lebih tinggi dibandingkan dengan metode OLS. Penelitian lain dilakukan Prabeswari (2019) untuk memodelkan konsumsi rumah tangga di Kota Surabaya berdasarkan indikator tingkat kesejahteraan menggunakan metode Regresi *Ridge Robust*. Pada penelitian tersebut dilakukan perbandingan beberapa metode estimasi diantaranya, *M-Estimation*, *S-Estimation*, dan *MM-Estimation*, dan didapatkan hasil Metode *Ridge Robust S-Estimation* merupakan metode estimasi yang paling baik digunakan dalam mengestimasi data tersebut.

Berdasarkan uraian yang telah dijelaskan, dalam penelitian ini dilakukan analisis mengenai faktor-faktor yang mempengaruhi luas

panen padi di Kabupaten Gresik menggunakan Regresi *Ridge Robust M-Estimation*, *S-Estimation*, dan *MM-Estimation* dikarenakan adanya dugaan amatan *outlier* dan multikolinearitas yang terjadi pada data. Faktor yang diduga berpengaruh adalah curah hujan rata-rata menurut bulan di Kabupaten Gresik. Hasil pemodelan luas panen padi yang terbaik diharapkan dapat digunakan sebagai pendukung keputusan dalam peningkatan luas panen padi serta jumlah produksi padi di Kabupaten Gresik.

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, permasalahan yang dibahas dalam penelitian ini adalah pemodelan luas panen padi di Kabupaten Gresik dengan faktor yang diduga memberikan pengaruh yang signifikan. Tahapan pemodelan menjadi persoalan jika terdapat amatan *outlier*. Amatan *outlier* menyebabkan estimator parameter regresi hasil metode OLS tidak efisien. Selain itu adanya multikolinearitas pada pengamatan juga dapat mempengaruhi hasil estimasi parameter regresi. Oleh karena itu, pada penelitian ini dilakukan pemodelan jumlah produksi padi dengan metode Regresi *Ridge Robust* untuk mengatasi amatan *outlier* dan multikolinearitas.

## 1.3 Tujuan

Berdasarkan rumusan masalah pada uraian di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik indikator yang mempengaruhi luas panen padi di Kabupaten Gresik.
2. Mengetahui pola hubungan curah hujan dengan luas panen padi di Kabupaten Gresik.
3. Mendapatkan model terbaik untuk pemodelan luas panen padi di Kabupaten Gresik menggunakan regresi *ridge robust*.

## 1.4 Manfaat

Adapun manfaat yang ingin diperoleh dari penelitian ini ialah sebagai berikut.

1. Bagi keilmuan statistika

Dapat menjadi referensi untuk penelitian selanjutnya dalam melakukan pemodelan luas panen padi dengan metode Regresi *Ridge Robust*. Dapat menjadi referensi, khususnya yang melakukan penelitian dalam bidang pemodelan dengan data *outlier* dan multikolinearitas.

2. Bagi dinas terkait

Dapat menjadi bahan informasi mengenai model luas panen padi sebagai bahan evaluasi dan saran dalam strategi pengembangan pertanian di Kabupaten Gresik untuk lebih tepat dalam melakukan sistem penyediaan pangan yang produktif dan berkelanjutan.

## 1.5 Batasan Masalah

Batasan permasalahan dalam penelitian ini yaitu data yang digunakan adalah data sekunder berupa data luas panen padi sawah dan curah hujan bulanan Kabupaten Gresik periode 2006 hingga 2019 yang diperoleh dari *website* Badan Pusat Statistik (BPS) Kabupaten Gresik dan situs resmi Pemerintah Kabupaten Gresik. Metode estimasi parameter yang digunakan dalam pemodelan Regresi *Ridge Robust* yaitu *S-Estmiation*, *M-Estimation*, dan *MM-Estimation.*

# BAB II

# TINJAUAN PUSTAKA

## 2.1 Statistika Deskriptif

Statistika deskriptif adalah metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna berupa tabel, grafik, diagram, dan besaran lain (Walpole, 1995). Statistika deskriptif terdapat 2 jenis ukuran data yaitu ukuran pemusatan data (rata-rata, median, dan modus) dan ukuran penyebaran data (*range*, varians, dan standar deviasi). Penelitian ini menggunakan ukuran pemusatan dan penyebaran data berupa nilai rata-rata (*mean*), standard deviasi, serta nilai minimum dan maksimum. Ukuran Penyebaran adalah untuk mengukur penyimpangan nilai-nilai amatan disekitar nilai rata-ratanya. Perhitungan deviasi didasarkan pada penyimpangan nilai-nilai data secara individu terhadap rata-ratanya, karena itu deviasi akan makin besar jika nilai-nilai data menyebar.

a. Rata-rata

Rata-rata *(mean)* adalah nilai rata-rata dari data numerik hasil pengamatan. Rata-rata hitung dari populasi diberi simbol μ. Rata-rata hitung dari sampel diberi simbol $\bar{x}$ . Nilai dari rata-rata dapat diperoleh dengan perhitungan seperti pada persamaan 2.1 (Hasan, 1999).

$$\bar{x} = \frac{x_1+x_2+x_3+\cdots+x_n}{n} \tag{2.1}$$

dengan,

$\bar{x}$ : rata-rata hitung *(mean)*
$x_1$ : nilai amatan pertama
$x_2$ : nilai amatan kedua
$x_3$ : nilai amatan ketiga
$n$ : jumlah sampel

b. Keragaman (*Variance*)

Keragaman adalah alat ukur variabilitas serangkaian data yang dihitung dengan mencari rata- rata beda kuadrat antara data observasi dengan pusat datanya, seperti pada persamaan 2.2 untuk keragaman populasi dan persamaan 2.3 untuk keragaman sampel (Hasan, 1999).

Rumus keragaman:

Keragaman untuk populasi

$$\sigma^2 = \frac{\sum_{i=1}^{N} \left( X_i - \mu_x \right)^2}{N} \tag{2.2}$$

Keragaman untuk sampel

$$s^2 = \frac{\sum_{i=1}^{n} \left( X_i - \bar{X} \right)^2}{n - 1} \tag{2.3}$$

dengan,

$\sigma^2$ : keragaman populasi

$X_i$ : data ke-i dari variabel acak X

$\mu_x$ : rata- rata populasi

N : ukuran populasi

$s^2$ : keragaman sampel

$\bar{X}$ : rata- rata sampel

c. Simpangan baku *(standard deviation)*

Simpangan baku *(standard deviation)* yang merupakan ukuran variabilitas data yang diperoleh dari akar kuadrat nilai variasi. Hal ini disebabkan bahwa variasi tidak dapat dinyatakan dalam satuan ukuran apapun seperti rupiah, kilogram, ton dan lain sebagainya. Sedangkan simpangan baku dapat dinyatakan dalam satu satuan (Hasan, 1999).

Rumus simpangan baku:

Simpangan baku untuk populasi :

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} \left( X_i - \mu_x \right)^2}{N}} \tag{2.4}$$

Simpangan baku untuk sampel :

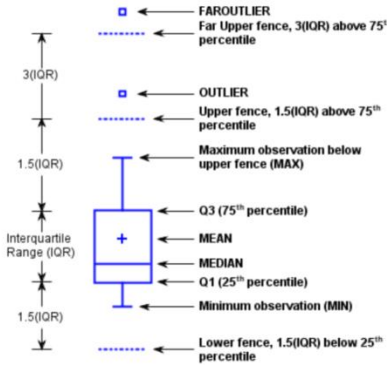
$$s = \sqrt{\frac{\sum_{i=1}^{n} \left( X_i - \bar{X} \right)^2}{n - 1}} \tag{2.5}$$

dengan,

$\sigma_x$ : simpangan baku untuk populasi
$s$ : simpangan baku untuk sampel
$\mu_x$ : rata- rata populasi
N : ukuran populasi
$X_i$ : data ke-i dari variabel acak X
$\bar{X}$ : rata- rata sampel

## 2.2 *Boxplot*

*Boxplot* adalah suatu metode statistika deskriptif dalam penyajian data secara grafis. *Boxplot* adalah ringkasan dari lima ukuran yaitu nilai observasi terkecil, kuartil pertama (Q1) yang berada pada 25% dari data terendah, median (Q2) yang berada di tengah data, kuartil ke tiga (Q3) yang berada pada 25% dari data tertinggi, dan nilai maksimum data. Pada *boxplot* juga menampilkan jika ada nilai *outlier* dalam data.




**Gambar 2.1** Contoh *Boxplot*

Garis yang memperpanjang box dinamakan dengan *whiskers* (Gambar 2.1). Jangkauan (IQR, *Interquartile Range*) didefinisikan sebagai selisih antara kuartil 1 dengan kuartil 3.

Panjang *whisker* bagian atas kurang dari atau sama dengan Q3 + (1,5 × IQR), sedangkan panjang garis *whisker* bagian bawah lebih besar atau sama dengan Q1 – (1,5 × IQR). Masing-masing *whisker* dimulai dari akhir box. Data *outlier* adalah data dengan nilai kurang dari 1.5*IQR terhadap kuartil 1 atau bernilai lebih dari 1.5*IQR terhadap kuartil 3 (Paludi, 2009).

## 2.3 *Outlier*

*Outlier* adalah suatu pengamatan individu yang berbeda dengan pengamatan mayoritas. Pengamatan *outlier* bisa merupakan pengamatan *error*, pengamatan yang dicatat dalam kondisi luar biasa, atau milik dari populasi lainnya (Rousseeuw, 2011). Kemunculan dari data yang buruk ini dapat secara signifikan mengganggu atau merusak perhitungan dari regresi. Ketika metode *least square* mengandung satu *outlier*, hal ini dapat mempengaruhi garis regresi mendekat kearah *outlier* (Hardin, 2003).

Pada penerapannya, untuk menangani *outlier* dalam regresi dapat digunakan dua cara. Metode pertama adalah *regression diagnostic*, metode klasik ini baik digunakan jikajumlah amatan *outlier* yang ada sedikit, akan tetapi apabila terdapat banyak amatan *outlier* maka metode klasik ini dapat dipengaruhi oleh *outlier* sehingga model yang dihasilkan tidak dapat mendeteksi observasi yang menyimpang. Hal ini menyebabkan data yang baik mungkin dapat dideteksi menjadi data *outlier*. Metode kedua adalah regresi *robust*, metode ini mencoba menemukan estimator yang tidak terlalu terpengaruh kuat dengan *outlier* (Rousseeuw, 1987).

### 2.3.1 *Difference in Fitted Value* (*DfFITS* )

Nilai dari pengamatan berperan penting dalam menentukan sifat-sifat model regresi, terutama titik-titik pengamatan yang lokasinya terpencil. Titik amatan dengan lokasi tersebut berpotensi menjadi *outlier* dan menyebabkan estimasi parameter tidak proporsional. Pada penerapannya, untuk mendeteksi *outlier* dalam regresi dapat digunakan beberapa cara. Salah satu metode

yang sering digunakan adalah *Difference in Fitted Value* (*DfFITS*) atau Standardized *DfFITS*. Metode *DfFITS* menampilkan perbedaan nilai prediksi dengan atau tanpa pengamatan ke-*i*. Metode ini dihitung dengan rumus sebagai berikut.

$$DfFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{S_{(i)}^2 - h_{ii}}},\ i = 1, 2, ..., n \qquad (2.6)$$

Pada persamaan 2.6, $\hat{y}_{(i)}$ merupakan hasil penduga dari $y_i$ yang diperoleh tanpa melibatkan pengamatan ke-*i*. Secara sederhana *DfFITS* merupakan nilai standar deviasi dari $\hat{y}_i$ yang mengalami perubahan ketika observasi ke-*i* dihapus.

Nilai *DfFITS* digunakan untuk mendeteksi *outlier* pada variabel bebas. Adapun ketentuan yang berlaku dalam pengambilan keputusan terhadap *outlier* adalah sebagai berikut (Rousseeuw, 1987).

Jika, $|DfFITS| > 2 * sqrt(p / n)$

dengan,

$n$ : jumlah observasi

$p$ : jumlah parameter

## 2.4 Analisis Regresi

Analisis regresi merupakan suatu analisis statistika yang digunakan untuk memodelkan pola hubungan antara variabel independen dengan variabel dependen. Persamaan yang digunakan untuk mendapatkan garis regresi pada *scatterplot* disebut persamaan regresi. Bentuk pola hubungan yang dapat diidentifikasi berdasarkan pada informasi masa lalu atau *scatterplot* data.

Diagram *scatterplot* menjelaskan bahwa hubungan x dan y berupa garis lurus. Sebuah garis lurus dapat dituliskan dalam bentuk persamaan 2.8 (Montgomery, Peck, dan Vinning, 2012).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + e \qquad (2.8)$$

dengan,

| | |
|---|---|
| $\beta_0, \beta_1, \dots, \beta_p$ | : parameter regresi |
| $x_1, x_2, \dots, x_{-1}$ | : variabel bebas |
| $e$ | : nilai error |

### 2.4.1 Uji Parameter Regresi secara Serentak

Pada analisis regresi pengujian parameter regresi secara serentak dilakukan dengan tujuan untuk mengetahui ada tidaknya variabel prediktor yang berpengaruh secara signifikan terhadap variabel respon. Pengujian regresi secara serentak atau uji-F dilakukan untuk mengetahui pengaruh variabel prediktor secara keseluruhan, pengujian ini dilakukan dengan hipotesis uji sebagai berikut.

Hipotesis

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

$H_1$ : minimal ada satu $\beta_k \neq 0$, untuk k $= 1, 2, 3, ..., p$

Statistik uji yang digunakan dalam pengujian adalah statistik F dengan rumus sebagai berikut.

$$F = \frac{(SSR / p)}{SSE / [n - (p+1)]} = \frac{MSR}{MSE} \tag{2.9}$$

Pada persamaan 2.9, $n$ merupakan ukuran sampel dan $p$ merupakan jumlah variabel independen. Daerah keputusan dari pengujian serentak ini adalah tolak $H_0$ apabila $F > F_{\alpha(p,n-(p+1))}$ atau $P - value < \alpha$, maka dapat disimpulkan bahwa secara serentak variabel prediktor memiliki pengaruh signifikan terhadap variabel respon.

### 2.4.2 Uji Parameter Regresi Partial

Pengujian parameter parsial adalah metode untuk menguji signifikansi tiap variabel prediktor terhadap variabel respon. Pengujian ini bertujuan untuk seleksi dan memastikan tiap

variabel prediktor berpengaruh. Uji yang digunakan adalah uji-t dengan hipotesa uji sebagai berikut.

Hipotesis

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0,$ untuk $k = 1, 2, 3, ..., p$

Pada pengujian parsial statistik uji yang digunakan adalah statistik uji-t dengan rumus pada persamaan 2.10.

$$t_{hitung} = \frac{\hat{\beta}_k}{s(\hat{\beta}_k)} \tag{2.10}$$

Dengan menggunakan persamaan 2.10, hipotesis nol ditolak apabila $|t_{hitung}| > t_{(\frac{\alpha}{2};n-(p+1))}$, dengan $n$ merupakan ukuran sampel.

## 2.5 Regresi *Robust*

Analisis yang umum digunakan untuk mengetahui hubungan antara variabel dependen (Y) dan variabel independen (X) adalah analisis regresi. Parameter regresi umumnya diestimasi menggunakan metode *Ordinary Least Square* (OLS), namun untuk beberapa kejadian yang menyebabkan adanya *outlier* pada pengamatan, metode OLS tidak cukup baik untuk diaplikasikan. Hal ini dikarenakan OLS untuk regresi dianggap sangat sensitif terhadap *outlier* pada data, maka dari itu perlu dilakukan analisis regresi *robust* (Greene, 2002). Regresi *robust* merupakan solusi dari *least square regression* yang digunakan untuk asumsi yang tidak terpenuhi. Regresi ini lebih baik menghasilkan koefisien regresi apabila terdapat *outlier* pada data karena pada regresi *robust* pengamatan yang mengalami *outlier* akan diberikan bobot yang lebih berat. System pembobotan ini berbeda dengan konsep yang diterapkan oleh OLS (Draper dan Smith, 1998).

Beberapa metode estimasi parameter dikembangkan dalam penyelesaian Regresi *Robust*, diantaranya *M-Estimation*

(*Maximum Likelihood Estimation*), *S-Estimation* (*Scale Estimation*), *MM-Estimation* (*Method of Moment Estimation*), *Least Median Squares* (LMS), dan *Least Trimmed Squares* (LTS). Kelima metode estimasi parameter pada Regresi *Robust* tersebut mempunyai kelemahan dan kelebihan masing-masing. Salah satu metode Regresi *Robust* yang popular dan paling umum digunakan untuk menangani *bad leverage points* karena *outlier* pada regresi adalah estimasi MM (*Method of Moment*). Estimasi MM mempunyai efisiensi yang tinggi ketika error berdistribusi normal serta memiliki nilai *breakdown* yang tinggi pula. *M-Estimation* mempunyai efisiensi yang tinggi dengan *breakdown point* bernilai 0, *S-Estimation*, LTS dan LMS mempunyai efisiensi yang rendah dengan *breakdown point* tinggi bernilai 0,5.

### 2.5.1 *Efficiency*

Regresi *Robust* akan menghasilkan hasil estimasi yang serupa dengan regresi OLS apabila data tidak mengandung *outlier* dan mengikuti seluruh asumsi. Nilai efisiensi dari estimator *robust* diperoleh dari pembagian MSE dari regresi OLS dengan MSE hasil regresi *robust*. Nilai dari perbandingan ini diharapkan mendekati satu (Montgomery, Peck, dan Vinning, 2012).

### 2.5.2 *Breakdown Point*

*Breakdown point* adalah pecahan terkecil dari data anomali yang dapat menyebabkan estimator menjadi tidak berguna. Pada awalnya *breakdown point* digunakan dalam pengestimasian rata-rata populasi. Nilai terkecil yang memungkinkan untuk *breakdown point* adalah $1/n$ sehingga sebuah pengamatan dapat mengganggu estimator. Pada regresi OLS *breakdown point* yang digunakan adalah $1/n$ (Rousseeuw, 2011). Semakin besar nilai persentase dari *breakdown point* pada suatu estimator, maka estimator tersebut semakin *robust*. Dalam perkembangannya, nilai *breakdown point* yang digunakan dalam regresi r*obust* beragam tergantung dari metode estimasi yang digunakan (Montgomery, Peck, dan Vinning, 2012).

## 2.6 Metode Estimasi Regresi *Robust*

Regresi *robust* memiliki bebberapa jenis metode dalan pengestimasian parameternya. Metode yang sering digunakan antara lain *M-Estimation* (*Maximum Likelihood Estimation*), *S-Estimation* (*Scale Estimation*), *MM-Estimation* (*Method of Moment Estimation*. Ketiga metode ini masing-masing memiliki cara kerja dan karakteristik yang berbeda. Berikut cara kerja dan karakteristik dari ketiga metode estimasi regresi *robust*.

### 2.6.1 *M-Estimation*

Regresi *Robust M-Estimation* adalah metode estimasi *Robust* yang pertama kali diperkenalkan. Metode ini mengestimasi parameter regresi dengan estimator *maximum likelihood* (Draper dan Smith, 1998). Pada metode estimasi ini diasumsikan bahwa residual secara independen memiliki distribusi dan semua megikuti distribusi yang sama. Metode ini tergolong dalam kelompok *robust* estimator yang meminimumkan fungsi $\rho$ dari residual (Montgomery, Peck, dan Vinning, 2012).

$$\underset{\beta}{\text{Min}} \sum_{i=1}^{n} \rho(u) = \underset{\beta}{\text{Min}} \sum_{i=1}^{n} \rho\left(\frac{Y_i - \mathbf{x}_i\boldsymbol{\beta}}{s}\right) \tag{2.11}$$

Fungsi $\rho$ dalam persamaan 2.11 berhubungan dengan fungsi *likelihood* dalam ketepatan pemilihan distribusi residual. Estimator dari M tidak harus memiliki skala *invariant* untuk menghadapi kasus ini nilai dari residual dibagi dengan standar deviasi (s), nilai standar deviasi diperoleh dengan formula seperti pada persamaan 2.12.

$$s = \text{median} \mid e_i - \text{median}(e_i) \mid / \beta_0 \tag{2.12}$$

Untuk meminimumkan fungsi 2.11 dengan kondisi skala dari estimator *invariant* dilakukan penyamaan turunan parsial pertama dari $\rho$ terhadap $\beta$ dengan 0, untuk menghasilkan kondisi minimum yang diperlukan. Sehingga didapatkan persamaan sebagai berikut.

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i - \mathbf{x}_i\boldsymbol{\beta}}{s}\right) = 0, \quad j = 0,1,\ldots,k \tag{2.13}$$

Dari persamaan 2.13 dapat disimpulkan bahwa $\Psi(u)$ adalah turunan dari $\rho(u)$, umumnya fungsi $\Psi(u)$ bersifat nonlinear maka harus diatasi dengan metode iterasi, salah satu metode yang umum digunakan adalah *Iteratively Reweighted Least Squares* (IRLS). Untuk melakukan IRLS diasumsikan bahwa $\hat{\beta}_0$ tersedia dan $s$ merupakan estimasi dari skala. Pada setiap tahapan digunakan nilai estimasi baru dari koefisien regresi yang didapatkan dari persamaan matrix sebagai berikut.

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{t+1}} = (\mathbf{X'W}_t\mathbf{X})^{-1}\mathbf{X'W}_t\boldsymbol{Y} \tag{2.14}$$

$\boldsymbol{W_t}$ dalam persamaan 2.14 adalah fungsi pembobot berbentuk matriks diagonal berukuran $nxn$ yang berisi bobot tiap pengamatan. Proses dari penaksiran parameter dilakukan secara berulang untuk mendapatkan hasil yang konvergen. Kelemahan dari metode *M-Estimation* adalah kurang memperhatikan distribusi data dan menggunakan median sebagai pembentuk bobot. Selain itu, *breakdown point* dari metode ini adalah 0 (Rousseeuw, 2011). Pada penelitian ini digunakan *Tukey Bisquare* sebagai metode estimasi. Fungsi objektif, pengaruh dan pembobot untuk Tukey *Bisquare* dijabarkan pada Tabel 2.1.

**Tabel 2.1** Estimasi M-Tukey *Bisquare*

| **Metode Tukey *Bisquare*** | |
|---|---|
| Fungsi Objektif | $\rho_B(u_i) = \begin{cases} \frac{u_i^2}{2} - \frac{u_i^4}{2c^2} + \frac{u_i^6}{6c^4}, & \text{untuk } \lvert u_i \rvert \le c \\ \frac{c^2}{6}, & \text{untuk } \lvert u_i \rvert > c \end{cases}$ |
| Fungsi Pengaruh | $\psi_B(u_i) = \begin{cases} u_i\left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & \text{untuk } \lvert u_i \rvert \le c \\ 0, & \text{untuk } \lvert u_i \rvert > c \end{cases}$ |
| Fungsi Pembobot | $w_B(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & \text{untuk } \lvert u_i \rvert \le c \\ 0, & \text{untuk } \lvert u_i \rvert > c \end{cases}$ |

dengan

$$u_i = \left( \frac{e_i}{s} \right) \tag{2.15}$$

Prosedur estimasi *M-Estimation* dengan *iteratively reweighted least squares* (IRLS) adalah sebagai berikut (Susanti dan Pratiwi, 2014).

1. Melakukan estimasi koefisien regresi dengan OLS.
2. Menghitung nilai residual $e_i = y_i - \hat{y}_i$.
3. Menghitung nilai $s_i$ dari residual langkah 3.

$$\hat{s} = \frac{\text{median} \mid e_i - \text{median}(e_i) \mid}{0,6745} \tag{2.16}$$

4. Menghitung nilai residual terstandarisasi $u_i = \frac{e_i}{s}$
5. Menghitung nilai bobot menggunakan fungsi pembobot *Tukey's Bisquare* dengan nilai *turning constant* $c$=4,685.

$$w_B(u_i) = \begin{cases} \left[ 1 - \left( \frac{u_i}{c} \right)^2 \right]^2, & |u_i| \le c \\ 0, & |u_i| > c \end{cases} \tag{2.17}$$

6. Mendapatkan nilai estimasi dari parameter menggunakan formula sebagai berikut.

$$\hat{\boldsymbol{\beta}}_M{}^{(t)} = (\mathbf{X'W}^{(t-1)}\mathbf{X})^{-1}\mathbf{X'W}^{(t-1)}\mathbf{Y} \tag{2.18}$$

7. Melakukan pengulangan pada langkah 2-6 untuk mendapatkan hasil penaksiran parameter yang konvergen, yaitu apabila diperoleh nilai $\left| \hat{\boldsymbol{\beta}}_j{}^{(t)} - \hat{\boldsymbol{\beta}}_j{}^{(t-1)} \right|$ cukup kecil atau sama dengan 0 untuk j = 0, 1, 2, …, k.

### 2.6.2 *S-Estimation*

*S-Estimation* adalah salah satu metode estimasi Regresi *Robust* yang masih berhubungan dengan M-*scale*. Secara konsep, *S-Estimation* bergantung pada skala residual *M-Estimation*. Pada

*S-Estimation* pembobotan didapatkan dari standar deviasi residual, metode ini mengatasi permasalahan pembobotan dengan median yang terjadi pada *M-Estimation* (Rousseeuw, 1987). *S-Estimation* akan meminimumkan jumlah kuadrat *error* pada persamaan umum regresi linear, selain itu *S-Estimation* adalah salah satu metode estimasi Regresi *Robust* dengan *high breakdown point*. *S-Estimator* didefinisikan dengan

$$\min \sum_{i=1}^{n} \rho \left( \frac{y_i - \sum_{j=1}^{k} x_{ij}\beta_j}{\hat{s}_S} \right) \tag{2.19}$$

di mana,

$$\hat{s}_S = \sqrt{\frac{1}{nK} \sum_{i=1}^{n} w_i e_i^2} \tag{2.20}$$

Persamaan 2.21 diperoleh dari penurunan persamaan 2.19 terhadap $\beta$ sehingga didapatkan fungsi minimum disamadengankan dengan 0 sebagai berikut.

$$\frac{1}{n} \sum_{i=1}^{n} x_{ij} \psi \left( \frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{\hat{s}_S} \right) = 0, \quad j = 0, 1, \ldots, k \tag{2.21}$$

Sama halnya dengan metode *M-Estimation*, pada metode *S-Estimation* untuk mendapatkan estimator yang konvergen perlu dilakukan iterasi hingga diperoleh hasil yang konvergen. Proses ini disebut sebagai *Iteratively Reweighted Least Square* (IRLS). Prosedur *S-Estimation* dengan *Iteratively Reweighted Least Square* (IRLS) adalah sebagai berikut (Susanti dan Pratiwi, 2014).

1. Melakukan estimasi koefisien regresi dengan OLS
2. Menghitung estimator awal dan nilai residual $e_i = y_i - \hat{y}_i$.
3. Menghitung nilai $s_i$ dari residual.

$$\hat{s}_S = \begin{cases} \dfrac{\text{median} \mid e_i - \text{median}(e_i) \mid}{0,6745}, & \text{iterasi=1} \\ \sqrt{\dfrac{1}{0,199n} \sum_{i=1}^{n} w_i e_i^2}, & \text{iterasi>1} \end{cases} \tag{2.22}$$

4. Menghitung nilai $u_i = \frac{e_i}{s_S}$
5. Menghitung nilai pembobot menggunakan fungsi pembobot *Tukey's Bisquare* dengan nilai *turning constant* $c = 1.547$ sehingga diperoleh *breakdown point* sebesar 50%.

$$w_B(u_i) = \begin{cases} \left[1 - \left(\dfrac{u_i}{c}\right)^2\right]^2, & |u_i| \le c \\ 0, & |u_i| > c \end{cases} \tag{2.23}$$

6. Mendapatkan nilai estimasi dari parameter digunakan formula sebagai berikut.

$$\hat{\boldsymbol{\beta}}_S = (\mathbf{X'WX})^{-1}\mathbf{X'WY} \tag{2.24}$$

7. Melakukan pengulangan pada langkah 2-6 untuk mendapatkan hasil penaksiran parameter yang konvergen.

### 2.6.3 *MM-Estimation*

*MM-Estimation* adalah salah satu metode yang mengestimasi parameter regresi dengan *S-Estimation* yang meminimumkan skala residual dari *M-Estimation* kemudian dilanjutkan dengan *M-Estimation*. Berikut perhitungan matematis dan *MM-Estimation*.

$$\hat{\boldsymbol{\beta}}_{MM} = \min \sum_{i=1}^{n} \rho\left(\frac{y_i - \mathbf{x}_i\boldsymbol{\beta}}{s_{MM}}\right) \tag{2.25}$$

Persamaan diperoleh dari penurunan terhadap $\beta$ sehingga didapatkan persamaan 2.26

$$\sum_{i=1}^{n} \rho_1 \left( \frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{s_{MM}} \right) X_{ij} = 0, \quad j = 0,1,...,k \tag{2.26}$$

*MM-Estimation* digunakan dengan tujuan memperoleh hasil estimasi dengan nilai *breakdown point* dan *efficiency* yang tinggi. Nilai dari *breakdown* merupakan ukuran yang umum digunakan untuk menghitung proporsi dari *outlier* yang bisa diatasi sebelum pengamatan tersebut mempengaruhi model. Berikut tahapan estimasi parameter dengan *MM-Estimation* (Susanti dan Pratiwi, 2014).

1. Menghitung estimator awal dengan menggunakan S-*estimation* yang memiliki *high breakdown point*
2. Menghitung estimator awal dan nilai residual $e_i = y_i - \hat{y}_i$dari *S-Estimation.*
3. Menghitung nilai $s_i$ dari residual langkah 2.

$$s_i = \frac{\text{median}|e_i - \text{median}(e_i)|}{0.6745} \tag{2.27}$$

4. Menghitung nilai bobot sesuai dengan fungsi *Tukey Bisquare* dengan nilai *turning constant c*=4,685.

$$w_B(u_i) = \begin{cases} \left[ 1 - \left( \frac{u_i}{c} \right)^2 \right]^2, & |u_i| \le c \\ 0, & |u_i| > c \end{cases} \tag{2.28}$$

5. Mendapatkan nilai estimasi dari parameter digunakan formula sebagai berikut.

$$\hat{\boldsymbol{\beta}}_{MM}^{(l)} = (\mathbf{X'W}^{(l-1)}\mathbf{X})^{-1}\mathbf{X'W}^{(l-1)}\mathbf{y} \tag{2.29}$$

6. Melakukan pengulangan pada langkah 2-5 untuk mendapatkan hasil penaksiran parameter yang konvergen, selisih antara $\hat{\boldsymbol{\beta}}_{MM}^{(l-1)}$ dengan $\hat{\boldsymbol{\beta}}_{MM}^{(l)}$ mendekati 0, dengan l merupakan banyaknya iterasi.

## 2.7 Multikolinearitas

Multikolinearitas adalah sebuah kondisi yang menunjukkan adanya hubungan linear yang kuat atau korelasi yang besar antar variabel independen. Kasus multikolinearitas dapat menyebabkan perubahan pada nilai koefisien regresi dari suatu variabel berubah secara drastis saat ada penambahan atau pengurangan variabel baru (Gujarati, 2004). Selain itu multikolinearitas juga menyebabkan nilai standard error yang besar karena ketidak tepatan pengukuran koefisien partial regresi. Multikolinearitas dapat dideteksi dengan beberapa cara antara lain pengecekan nilai korelasi antar variabel prediktor, jika korelasi antar variabel independen tinggi makan ada indikasi terjadi kasus multikolinearitas. Selain itu untuk mendeteksi multikolinearitas dapat pula dilakukan pengecekan nilai *Variance Inflation Factor* (VIF). Apabila nilai VIF lebih besar dari 10 maka terdapat multikolinearitas. Multikolinieritas dapat disebabkan karena adanya efek kombinasi dua atau lebih variabel bebas.

## 2.8 Regresi *Ridge*

Regresi *Ridge* adalah salah satu metode regresi yang digunakan saat terjadi *"ill-conditioned"* di mana terdapat dependensi antara beberapa variabel independennya. Hal ini dapat menyebabkan matrix $\boldsymbol{X'X}$ mendekati singular, sehingga meningkatkan hasil estimasi parameter tidak stabil. Hal ini menyebabkan nilai yang besar pada standar error, sehingga perlu digunakan Regresi *Ridge* karena menambahkan informasi tambahan untuk mengatasi kondisi ini. Berikut hasil modifikasi dari persamaan normal mejadi persamaan untuk Regresi *Ridge* (Montgomery, Peck, dan Vinning, 2012).

$$(\mathbf{X'X} + k\mathbf{I})\hat{\boldsymbol{\beta}}_{ridge} = \mathbf{X'y} \tag{2.30}$$

Nilai dari $k$ pada persamaan 2.30 apabilan lebih besar dari sama dengan 0 dan didapatkan dari analisis. Ketika nilai $k = 0$ maka *ridge estimator* sama dengan *least square estimator*. Pada

dasarnya regresi ini memiliki konsep seperti halnya OLS, hanya saja pada nilai variabel ditransformasi melalui prosedur *centering* dan *rescaling*. Semakin kecil konstanta bias, maka estimator yang dihasilkan akan semakin stabil, dan nilai bias dari meningkat jika $k$ mengalami peningkatan. Sebaliknya, varians akan meningkat jika nilai $k$ semakin kecil.

### 2.8.1 Metode *Centering* dan *Rescaling*

Metode *centering* dan *rescaling* bertujuan untuk standarisasi variabel. Apabila terdapat n jumlah observasi dengan $p$ prediktor, maka didapatkan persamaan regresi berganda sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip-1} + e_i \tag{2.31}$$

Pada proses *centering* persamaan 2.31 diubah menjadi persamaan 2.32

$$\begin{aligned} y_i &= \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_1\bar{x}_1 + \ldots + \beta_p(x_{ip} - \bar{x}_p) + \beta_p\bar{x}_{cp} + e_i \\ &= (\beta_0 + \beta_1\bar{x}_1 + \ldots + \beta_p\bar{x}_p) + \beta_1(x_{i1} - \bar{x}_1) + \ldots + \\ &\quad \beta_p(x_{ip} - \bar{x}_p) + e_i \end{aligned} \tag{2.32}$$

di mana,

$$\beta_0' = (\beta_0 + \beta_1\bar{x}_1 + \ldots + \beta_p\bar{x}_p) = \bar{Y} \tag{2.33}$$

Dengan adanya proses *centering* elemen dari matriks $\boldsymbol{X}$ mengalami perubahan sebagai mana pada ilustrasi 2.34.

$$\mathbf{X} = \begin{bmatrix} 1 & (X_{11} - \bar{X}_1) & \cdots & (X_{p1} - \bar{X}_p) \\ 1 & (X_{12} - \bar{X}_1) & \cdots & (X_{p2} - \bar{X}_p) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (X_{1n} - \bar{X}_1) & \cdots & (X_{pn} - \bar{X}_p) \end{bmatrix} \tag{2.34}$$

Dari proses *centering* didapatkan nialai estimasi dengan metode yang sama dengan OLS, adanya proses *centering* menyebabkan hilangnya intersep, $\beta_0$. Dari matriks 2.34 dibentuk persamaan 2.35 merupakan persamaan regresi setelah proses *rescaling* (Draper & Smith, 1998).

$$Y_i^R = \beta_1 X_1^R + \beta_2 X_2^R + \cdots + \beta_p X_p^R + e_i \tag{2.35}$$

di mana,

$$X_{1j}^{R} = \frac{x_{1j} - \bar{x}_1}{\sqrt{\sum_{j=1}^{n} (x_{1j} - \bar{x}_1)^2}} \tag{2.36}$$

Keseluruhan proses dari prosedur ini disebut prosedur *centering* dan *rescaling*.

### 2.8.2 Konstanta Regresi *Ridge*

Dari proses pemilihan konstanta pada regresi *Ridge* dapat dilakukan dengan beberapa cara. Salah satunya adalah dengan menggunakan *Ridge trace*. *Ridge trace* adalah suatu metode yang menggunakan *plot* dari hasil estimasi regresi *Ridge* dengan berbagai nilai kemungkinan konstanta secara bersama. Nilai konstanta, $k$, dalam regresi *Ridge* menggambarkan bias dalam estimator. Umumnya nilai $k$ terletak pada interval 0 sampai dengan 1. Pemilikan nilai $k$ yang tepat sangat diperlukan untuk mendapatkan hasil estimasi yang stabil, sehingga kemungkinanan didapatkan nilai MSE yang lebih kecil dibandingkan dengan estimasi menggunakan *least square*.

## 2.9 Regresi *Ridge Robust*

Regresi *Ridge Robust* merupakan penggabungan dari metode regresi *robust* dan regresi *ridge*. Tujuan dari penggabungan dua metode ini adalah untuk menangani masalah *outlier* dan multikolinearitas yang kerap terjadi pada regresi berganda. Kedua permasalahan ini menyebabkan nilai taksiran parameter yang dihasilkan oleh OLS tidak optimum, sehingga perlu ditangani dengan gabungan kedua metode ini. Hasil dari metode estimasi ini diharapkan menjadi parameter regresi yang stabil dan resistant terhadap kasus *outlier* maupun multikolinearitas. Berikut adalah estimator regresi *ridge robust* menggunakan estimasi *robust*.

$$\hat{\boldsymbol{\beta}}_{\text{ridge}} = \left( \boldsymbol{X}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{R} + k\boldsymbol{I} \right)^{-1} \boldsymbol{X}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{Y} \qquad (2.37)$$

Salah satu metode untuk menentukan $k$ dikemukakan oleh Hoerldan Kennard (1970b) dan Gibbons (1981).

$$k = \frac{p s_{robust}{}^{2}}{\boldsymbol{\beta}^{T}\boldsymbol{\beta}} \qquad (2.38)$$

di mana,

$$s^{2}_{robust} = \frac{\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{robust}}\right)^{\mathrm{T}}\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{robust}}\right)}{n - p} \qquad (2.39)$$

## 2.10 Uji Normalitas

Uji normalitas adalah salah satu pengujian untuk mengetahui distribusi dari data. Pengujian normalitas dalam regresi menggunakan nilai dari variabel residual. Jika pada pengujian didapatkan bahwa variabel residual tidak berdistribusi normal, maka dapat dikatakan bahwa asusmsi tidak terpenuhi. Pengujian normalitas dapat dilakukan dengan beberapa metode antara lain *Kolmogorov Smirnov, Lilliefors, Chi-Square, Shapiro Wilk* dan lain-lain. Data yang berdistribusi normal akan memperkecil kemungkinan terjadinya bias.

Uji normalitas yang digunakan adalah *Kolmogorov-Smirnov*. Uji ini dilakukan karena lebih sederhana dan mengurangi perbedaan keputusan karena adanya persepsi yang terjadi apabila dilakukan pengujian menggunakan metode grafik. Pada pengujian ini menggunakan *Test* dengan *level of significant* 5%, sehingga apabila nilai *Asymp. Sig.* suatu variabel lebih besar dari *level of significant* 5% (>0,05) maka variabel tersebut berdistribusi normal, sedangkan jika nilai *Asymp. Sig.* suatu variabel lebih kecil dari *level of significant* 5% (<0,05) maka variabel tersebut tidak berdistribusi normal.

Hipotesis yang digunakan adalah sebagai berikut:

$H_0$ :$F_0(x) = F_n(x)$untuk semua nilai x

$H_1$ : $F_0(x) \neq F_n(x)$untuk semua nilai x

$F_0(x)$ merupakan fungsi ditribusi yang dihipotesiskan, sementara $F(x)$ adalah fungsi ditribusi yang belum diketahui. Nilai statistik uji:

$$D = \underset{\varepsilon}{Sup}\left|F_n(\varepsilon) - F_0(\varepsilon)\right| \tag{2.40}$$

$F_n(x)$ adalah fungsi peluang komulatif data. Apabila nilai $D > 1 - \alpha$ atau *p-value* < $\alpha$, maka kesimpulan didapatkan adalah tolak $H_0$.

## 2.11 Uji Independensi

Uji asumsi residual independen pada penelitian ini bertujuan untuk mengetahui apakah terdapat hubungan antar residual. Pengujian asumsi residual independen juga dikenal dengan uji autokorelasi residual. Salah satu pengujian yang dapat dilakukan untuk mendeteksi autokorelasi adalah uji *Durbin-Watson* dengan hipotesa sebagai berikut (Draper & Smith, 1998).

$H_0$ :$\rho = 0$ ; tidak terjadi autokorelasi

$H_1$ : $\rho \neq 0$ ; terjadi autokorelasi

Statistik uji:

$$d = \frac{\sum_{j=2}^{n} (e_j - e_{j-1})^2}{\sum_{j=i}^{n} e_j^2} \tag{2.41}$$

Keputusan diambil dengan membandingkan nilai statistik uji dengan nilai $d_u$. Tolak $H_0$ apabila nilai d<$d_L$ (batas bawah dari tabel *Durbin-Watson*) atau 4-d<$d_L$ yang berarti terdapat autokorelasi atau asumsi residual independen tidak terpenuhi. Gagal tolak $H_0$ apabila d>$d_u$ (batas atas dari tabel *Durbin-Watson*).

## 2.12 Curah Hujan

Curah hujan adalah indikator yang menunjukkan jumlah air hujan yang jatuh pada suatu daerah dalam periode waktu tertentu. Jumlah curah hujan merupakan hasil penghitungan volume air

yang terkumpul di permukaan tanah horizontal yang diasumsikan tidak terjadi infiltrasi, run off ataupun evaporasi. Curah hujan diukur dengan cara menghitung tebal atau tingginya air hujan yang terkumpul pada bidang datar tertentu. Satuan curah hujan yang umumnya dipakai oleh Badan Meteorologi, Klimatologi dan Geofisika (BMKG) adalah millimeter (mm), curah hujan 20 milimeter artinya dalam luasan satu meter persegi pada tempat yang datar tertampung air setinggi 20 millimeter. Pengukuran ini digunakan untuk menentukan tingkat curah hujan di suatu daerah. Jika curah hujan di suatu daerah berada pada rentang 0-100 mm maka tergolong ke dalam daerah dengan curah hujan rendah, jika rentang curah hujan 100-300 mm maka tergolong daerah curah hujan sedang, dan jika curah hujan diatas 300 mm maka tergolong dalam daerah dengan curah hujan tinggi.

## 2.13 Luas Panen Padi

Luas panen padi adalah suatu pengukuran total luas lahan yang padinya dipanen saat cukup umur. Data luas lahan panen merupakan data yang penting, hal ini karena data luas lahan dapat digunakan untuk mengukur produktivitas dari suatu sistem pertanian di daerah. Selain itu data luas panen juga digunakan untuk evaluasi dan analisis dari kondisi cuaca serta alam yang terkadang mengganggu kestabilan produksi. Seperti halnya badai, curah hujan ekstrem, serangan hama, dan beberapa gangguan lainnya. Data luas panen padi dicatat setiap bulan dan dikumpulkan untuk bahan hitung setiap 4 bulan sekali oleh Badan Pusat Statistik (BPS) yang bekerjasama dengan Kementrian Pertanian. Luas dari panen padi tiap daerah beragam. Keberagaman ini disebabkan oleh banyak faktor, mulai dari iklim hingga kondisi dari daerah tersebut.

## 2.14 Indeks Performansi

Indeks performansi merupakan ukuran akurasi yang digunakan untuk mengetahui ketepatan suatu metode peramalan dalam memodelkan data deret waktu. Jika semakin rendah tingkat error maka performansi model semakin baik, namun

sebaliknya jika tingkat error semakin besar maka performansi model semakin buruk. *Root Mean Square Error* (RMSE) adalah metode alternatif untuk mengevaluasi teknik peramalan yang digunakan untuk mengukur tingkat akurasi hasil prakiraan suatu model. RMSE merupakan nilai rata-rata dari jumlah kuadrat kesalahan, juga dapat menyatakan ukuran besarnya kesalahan yang dihasilkan oleh suatu model prakiraan. Nilai RMSE rendah menunjukkan bahwa variasi nilai yang dihasilkan oleh suatu model perkiraan mendekati variasi nilai obeservasinya. Selain RMSE, indeks performansi lain yang kerap digunakan adalah *Mean Absolute Percentage Error* (MAPE). MAPE menyatakan persentase kesalahan hasil peramalan terhadap permintaan aktual selama periode tertentu yang akan memberikan informasi persentase kesalahan terlalu tinggi atau terlalu rendah.

(*Halaman ini sengaja dikosongkan*)

# BAB III

# METODE PENELITIAN

## 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder diperoleh dari publikasi Gresik dalam Angka mulai dari publikasi tahun 2006 hingga 2019. Publikasi diakses melalui *website* Badan Pusat Statistik (BPS) Kabupaten Gresik dan situs resmi pemerintah Kabupaten Gresik.

## 3.2 Variabel Penelitian

Variabel Penelitian yang digunakan dalam penelitian ini adalah luas panen padi (Ha) per periode dengan variabel yang diduga berpengaruh secara signifikan yaitu curah hujan rata-rata (mm). Pada penelitian ini dalam satu tahun dibagi menjadi 3 periode, tiap periode terdiri dari 4 bulan dimulai dari bulan Januari hingga Desember.

**Tabel 3.1** Variabel Penelitian

| Periode | Variabel | Keterangan |
|---|---|---|
| 1 (Jan-Apr) | $LP_1$ | Luas Panen Padi Periode 1 |
| | $CH_1$ | Curah Hujan Bulan Januari |
| | $CH_2$ | Curah Hujan Bulan Februari |
| | $CH_3$ | Curah Hujan Bulan Maret |
| | $CH_4$ | Curah Hujan Bulan April |
| 2 (Mei-Agust) | $LP_2$ | Luas Panen Padi Periode 2 |
| | $CH_5$ | Curah Hujan Bulan Mei |
| | $CH_6$ | Curah Hujan Bulan Juni |
| | $CH_7$ | Curah Hujan Bulan Juli |
| | $CH_8$ | Curah Hujan Bulan Agustus |
| 3 (Sep-Des) | $LP_3$ | Luas Panen Padi Periode 3 |
| | $CH_9$ | Curah Hujan Bulan September |
| | $CH_{10}$ | Curah Hujan Bulan Oktober |
| | $CH_{11}$ | Curah Hujan Bulan November |
| | $CH_{12}$ | Curah Hujan Bulan December |

Dari variabel yang digunakan berikut sktruktur data penelitian pada periode 1 seperti pada Tabel 3.1. Data selengkapnya disajikan pada Lampiran 1.

**Tabel 3.2** Struktur Data Penelitian Periode 1

| **Tahun** | $LP_1$ | $CH_1$ | **…** | $CH_4$ |
|---|---|---|---|---|
| 2006 | $LP_{1,1}$ | $CH_{1,1}$ | … | $CH_{1,4}$ |
| 2007 | $LP_{2,1}$ | $CH_{2,1}$ | … | $CH_{2,4}$ |
| 2008 | $LP_{3,1}$ | $CH_{3,1}$ | … | $CH_{3,4}$ |
| ⋮ | ⋮ | ⋮ | … | ⋮ |
| 2018 | $LP_{12,1}$ | $CH_{12,1}$ | … | $CH_{12,4}$ |

Model prediksi luas panen padi yang disusun terdiri atas tiga model sesuai dengan periode yang disusun oleh BPS sebagai berikut.

a. Model Prediksi Luas Panen Padi Periode 1 (Januari-April)
$LP_1 = \beta_0 + \beta_1 CH_1 + \beta_2 CH_2 + \beta_3 CH_3 + \beta_4 CH_4$
b. Model Prediksi Luas Panen Padi Periode 2 (Mei-Agustus)
$LP_2 = \beta_0 + \beta_1 CH_5 + \beta_2 CH_6 + \beta_3 CH_7 + \beta_4 CH_8$
c. Model Prediksi Luas Panen Padi Periode 3 (September-Desember)
$LP_3 = \beta_0 + \beta_1 CH_9 + \beta_2 CH_{10} + \beta_3 CH_{11} + \beta_4 CH_{12}$

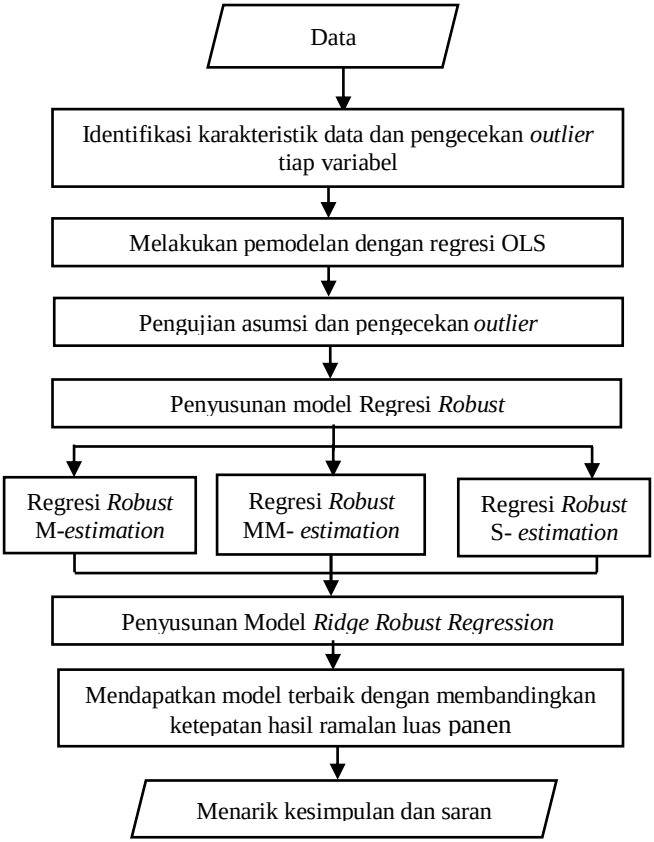
## 3.3 Langkah Analisis Data

Tahapan yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Melakukan eksplorasi data untuk mengetahui karakteristik dari data tiap variabel.
2. Melakukan identifikasi *outlier* secara visual pada setiap variabel dengan *boxplot* .
3. Melakukan regresi OLS dan pengecekan *outlier* secara statistik dengan residual regresi OLS menggunakan nilai *DFFITS* .

4. Melakukan pemodelan luas panen padi dengan pendekatan regresi Regresi *Robust* dengan langkah sebagai berikut.
   a. Menyusun model regresi luas panen padi (Y) dengan curah hujan dan jumlah hari hujan (X) dari data menggunakan Regresi *Robust S-Estimation*.
   b. Menyusun model regresi luas panen padi (Y) dengan curah hujan dan jumlah hari hujan (X) dari data menggunakan Regresi *Robust M-Estimation*.
   c. Menyusun model regresi luas panen padi (Y) dengan curah hujan dan jumlah hari hujan (X) dari data menggunakan Regresi *Robust MM-Estimation*.
5. Melakukan pengecekan multikolinearitas. Jika tidak memenuhi asumsi multikolinieritas, maka dilakukan metode Regresi *Ridge Robust* dengan 3 estimasi.
6. Melakukan pemodelan luas panen padi dengan pendekatan Regresi *Ridge Robust*.
7. Melakukan prediksi luas panen padi dan menghitung persentase ketepatan hasil prediksi.
8. Menginterpretasi hasil dari analisis.
9. Menarik kesimpulan dan memberikan saran.

Tahapan analisis data di atas disajikan dalam diagram alir seperti pada Gambar 3.1.




**Gambar 3.1** Langkah Analisis Data

# BAB IV

# ANALISIS DAN PEMBAHASAN

Pada bab analisis dan pembahasan ini dijelaskan beberapa hal dengan tujuan menjawab dari rumusan masalah. Hal-hal yang dibahas meliputi analisis mengenai faktor yang berpengaruh terhadap luas panen padi di Kabupaten Gresik. Pada tahap pertama yaitu melakukan analisis karakteristik pada luas panen padi dan faktor yang mempengaruhi. Selanjutnya dilakukan identifikasi *outlier* menggunakan *boxplot* dan *DfFITS* dan pola hubungan antara luas panen dengan curah hujan menggunakan regresi *robust* dengan beberapa metode estimasi parameter *M, S* dan *MM*. Setelah itu dilakukan analisis lanjutan menggunakan regresi *ridge* menggunakan hasil estimasi parameter regresi *robust M, S,* dan juga *MM*. Dari hasil analisis regresi dengan metode regresi *ridge robust* selanjutnya perbandingan untuk menentukan metode terbaik dalam yang mempengaruhi luas panen padi di Kabupaten Gresik.

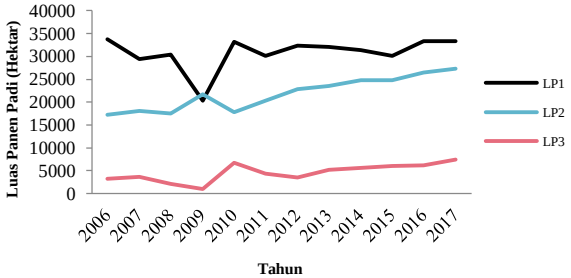
## 4.1. Karakteristik Data dan Identifikasi *Outlier*

Data luas panen padi pada Kabupatn Gresik tahun 2006 hingga 2017 tiap periode pada subbab ini dideskripsikan dengan *line chart*, dengan demikian dapat dilihat perbandingan jumlah luas panen tiap periodenya. Selain itu, dilakukan juga perbandingan nilai statistika deskriptif berupa nilai *mean*, simpangan baku, serta nilai minimum dan maksimum dari tiap periode luas panen dan curah hujan tiap bulan. Proses selanjutnya adalah identifikasi menggunakan *boxplot* dan penghitungan *DfFITS* untuk menemukan amatan *outlier*.

### 4.1.1 Deskripsi Luas Panen di Kabupaten Gresik

Kabupaten Gresik sebagai salah satu kabupaten dengan fokusan pada sektor pertanian di Jawa Timur. Hal ini karena

Kabupaten Gresik menyumbang angka jumlah luas panen di Jawa Timur sebesar 3,18 persen pada tahun 2017. Selain itu produktivitas dari padi di Kabupaten Gresik sebesar 65,55 kwintal per hektar angka ini diatas dari rata-rata produktivitas Jawa Timur yang hanya 58,19 kwintal per hektar. Luas panen pada tiap periode di Kabupaten Gresik memiliki nilai yang berbeda. Berikut *plot* luas panen padi di Kabupaten Gresik tiap periode dalam kurun waktu 12 tahun.




**Gambar 4. 1** Luas Panen Padi Kabupaten Gresik Tahun 2006-20017

Gambar 4.1 menunjukkan bahwa dari tahun 2006 hingga 2017 luas panen padi periode 1 (Januari-April) lebih tinggi dibandingkan dengan luas panen pada periode 2 maupun 3 kecuali pada tahun 2009. Pada tahun 2009 luas panen periode pertama mengalami penurunan hingga jumlahnya lebih kecil dibandingkan dengan luas panen pada periode 2. Hal ini disebabkan oleh adanya bencana banjir yang menggenangi sebagian besar lahan sawah di Kabupaten Gresik pada awal tahun 2009, sehingga menyebabkan rusaknya tanaman padi yang baru ditanam. Pada periode kedua luas panen cenderung mengalami peningkatan dari tahun ke tahun. Hal ini mengindikasikan bahwa periode kedua memiliki potensi untuk menjadi periode yang alternatif yang efektif untuk menanam padi. Gambaran umum dari data luas panen padi tiap periodenya di Kabupaten Gresik sebagaimana disajikan pada Tabel 4.1.

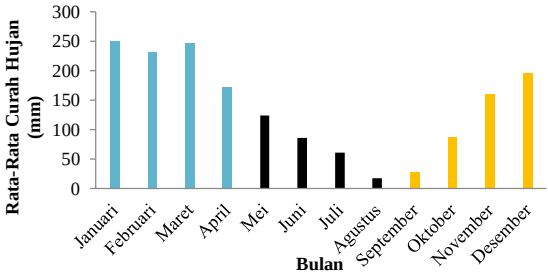
**Tabel 4.1** Mean, Standar Deviasi, Minimum, dan Maksimum Luas Panen Padi Kabupaten Gresik Tahun 2006-2017 (dalam hektar)

| Luas Panen | *Mean* | *Stdev* | *Min* | *Max* |
|---|---|---|---|---|
| Periode 1 | 30.828,57 | 3.626,43 | 20.303,21 | 33.770,00 |
| Periode 2 | 21.866,98 | 3.635,83 | 17.290,00 | 27.266,40 |
| Periode 3 | 4.588,12 | 1.946,99 | 941,00 | 7.417,60 |

Periode pertama menjadi periode tanam yang paling efektif, hal ini dikarenakan periode pertama memiliki rata-rata luas panen padi tertinggi. Nilai simpangan baku luas panen padi yang besar menunjukkan ketidakstabilan luas panen tiap tahun pada tiap periode. Pada periode pertama dan kedua memiliki simpangan baku besar mengindikasikan luas panen padi yang beragam. Perbedaan nilai pada luas panen pada ketiga periode disebabkan adanya perbedaan curah hujan setiap bulannya di Indonesia.

### 4.1.2 Deskripsi Curah Hujan di Kabupaten Gresik

Data curah hujan suatu wilayah menggambarkan rata-rata curah hujan dalam satu bulan. Kondisi sawah di Kabupaten Gresik yang mayoritas merupakan sawah non-irigasi menyebabkan hujan menjadi faktor penting dalam pengairan sawah. Berikut pola curah hujan di Kabupaten Gresik.




**Gambar 4.2** Rata-Rata Curah Hujan tiap bulan di Kabupaten Gresik tahun 2006-2017

Dari Gambar 4.2 menunjukan pola penurunan rata-rata curah hujan pada periode pertama dan kedua sedangkan pada periode ketiga mengalami pola kenaikan. Pada sistem pertainian terutama padi, kecukupan jumlah air pada awal masa tanam atau masa vegetatif sangat penting untuk menunjang pertumbuhan benih. Pada bulan menjelang masa panen, jumlah air yang dibutuhkan tidak terlalu banyak. Pada bulan Januari dan Mei pada masa tanam jumlah curah hujan tinggi sedangkan pada masa panen bulan April dan Agustus curah hujan lebih rendah, hal ini membantu mencegahnya terjadi pembusukan pada tumbuhan dan gagal panen. Berikut statistika deskriptif dari luas panen padi tiap periodenya di Kabupaten Gresik.

**Tabel 4.2** *Mean*, Standar Deviasi, Minimum, dan Maksimum Curah Hujan tiap Bulan di Kabupaten Gresik Tahun 2006-2017 (dalam mm)

| **Variabel** | ***Mean*** | ***Stdev*** | ***Min*** | ***Max*** |
|---|---|---|---|---|
| $CH_1$ | 246.7 | 94.8 | 68.0 | 394.7 |
| $CH_2$ | 228.2 | 78.0 | 133.8 | 348.8 |
| $CH_3$ | 246.1 | 68.1 | 112.1 | 319.6 |
| $CH_4$ | 171.6 | 79.2 | 61.1 | 316.1 |
| $CH_5$ | 131.3 | 81.1 | 4.5 | 266.0 |
| $CH_6$ | 81.6 | 86.9 | 0.0 | 238.0 |
| $CH_7$ | 65.9 | 87.0 | 0.0 | 224.0 |
| $CH_8$ | 19.3 | 41.6 | 0.0 | 140.0 |
| $CH_9$ | 30.7 | 79.7 | 0.0 | 266.0 |
| $CH_{10}$ | 82.1 | 111.8 | 0.0 | 322.0 |
| $CH_{11}$ | 147.5 | 108.7 | 0.0 | 370.0 |
| $CH_{12}$ | 187.0 | 106.5 | 0.0 | 350.3 |

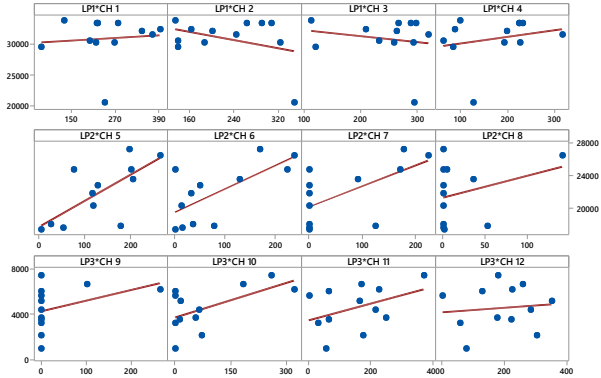
Tabel 4.2 menunjukkan rata-rata curah hujan di Kabupaten Gresik masuk dalam kelompok intensitas curah hujan sedang dan rendah, hal ini berdasarkan penggolongan yang ditetapkan oleh BMKG. Pada bulan Juni hingga Oktober, rata-rata curah hujan berada pada interval 0 hingga 100 mm. Hal ini mengindikasikan adanya curah hujan yang rendah. Sedangkan pada bulan November hingga Mei dengan curah hujan berada pada selang 100mm hingga 300mm

mengindikasikan curah hujan sedang. Di Kabupaten Gresik dalam periode tahun 2006 hingga 2017 hanya beberapa kali mengalami hujan dengan curah tinggi, lebih dari 300mm. berdasarkan nilai simpangan baku, tiga bulan terakhir yaitu Oktober, November dan Desember memiliki simpangan baku yang besar, hal ini mengindikasikan curah hujan yang terjadi pada tiap bulan beragam.

Berdasarkan keragaman curah hujan dan luas panen padi dapat dilihat bahwa periode penanaman padi optimum ada pada periode pertama dan kedua karena memiliki curah hujan yang menurun seiring dengan berlanjutnya masa tanam. Peningkatan curah hujan yang terjadi setiap bulan pada periode ketiga menyebabkan periode ketiga memiliki luas panen yang tidak cukup besar. Penanaman padi pada periode ketiga memerlukan bantuan sistem pengairan lain karena curah hujan yang rendah pada masa awal tanam dan perlakuan khusus pada masa panen karena tingginya curah hujan dapat memicu kegagalan panen. Pola yang serupa juga ditunjukkan dalam simpulan Zahiroh (2017) yang memodelkan luas panen padi di sentra produksi padi Jawa Timur.

### 4.1.3 Identifikasi Pola Hubungan Variabel Prediktor terhadap Respon

Pada analisis korelasi secara visual menggunakan *scatterplot*, garis linear yang menunjukkan pola penurunan pada periode pertama bulan Februari dan Maret menunjukkan korelasi negatif antara curah hujan dengan luas panen, seperti pada Gambar 4.3. Sedangkan untuk periode kedua terlihat bahwa gradien garis yang terbentuk besar hal ini menunjukkan besarnya pengaruh curah hujan terhadap luas panen padi.




**Gambar 4.3** *Scatterplot* antara Curah Hujan dengan Luas Panen di Kabupaten Gresik

Data *outlier* pada *scatterplot* sangat berpengaruh pada arah dan kuatnya hubungan antar kedua variabel. Data diduga sebagai amatan *outlier* apabila letak titik berada jauh dari sekumpulan data pada *scatterplot*. Pada *scatterplot* antara luas panen periode ketiga dengan curah hujan bulan ke-9 atau bulan September terdapat dua amatan yang berada diluar sekumpulan data yang diduga sebagai amatan *outlier*. Jika amatan yang diduga *outlier* dihilangkan, maka hubungan antar variabel bisa menjadi lebih kuat atau lemah dan memungkinkan terjadi perubahan pola keterkaitan antar variabel dari hubungan linier positif menjadi hubungan linier negatif maupun sebaliknya.

Selain dilakukan analisis secara visual dengan *scatterplot*, untuk melihat besarnya korelasi dapat dilakukan degan cara matematis menggunakan koefisien korelasi P*earson*, di mana nilai koefisien yang rendah pada sebagian besar pasangan variabel menunjukkan hubungan antar variabel curah hujan dengan luas panen padi yang lemah.

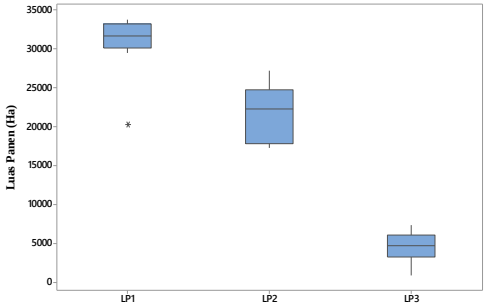
**Tabel 4.3** Koefisien dan P-value Korelasi Pearson antara Luas Panen Padi dengan Curah Hujan di Kabupaten Gresik Tahun 2006-2017

| Variabel | | Pearson | |
|---|---|---|---|
| | | **Koefisien** | ***P-value*** |
| $LP_1$ | $CH_1$ | 0,091 | 0,778 |
| | $CH_2$ | -0,361 | 0,249 |
| | $CH_3$ | -0,178 | 0,581 |
| | $CH_4$ | 0,241 | 0,450 |
| $LP_2$ | $CH_5$ | **0,715** | **0,009** |
| | $CH_6$ | **0,690** | **0,013** |
| | $CH_7$ | **0,615** | **0,033** |
| | $CH_8$ | 0,305 | 0,336 |
| $LP_3$ | $CH_9$ | 0,376 | 0,228 |
| | $CH_{10}$ | **0,585** | **0,046** |
| | $CH_{11}$ | 0,408 | 0,188 |
| | $CH_{12}$ | 0,107 | 0,741 |

Dari hasil korelasi pada Tabel 4.3 menunjukkan beberapa variabel memiliki *p-value* < α (0,05) mengindikasikan bahwa terdapat hubungan yang linier diantara variabel prediktor tersebut terhadap variabel respon hal ini menandakan bahwa adanya korelasi antara kedua variabel. Secara keseluruhan pada periode kedua memiliki curah hujan yang berhubungan linear terhadap luas panen kecuali pada bulan ke-8 atau bulan Agustus. Hal ini berbanding terbalik dengan dua periode lainnya, pada periode ketiga hanya curah hujan pada bulan Oktober yang memiliki hubungan linear dengan luas panen sedangkan pada periode pertama tidak ada bulan yang memiliki curah hujan linear dengan luas panen pada taraf signifikansi 5%. Pada tahapan selanjutnya dilaukan analisa hubungan seluruh variabel prediktor yaitu curah hujan tiap bulan terhadap variabel respon yaitu luas panen padi di Kabupaten Gresik.

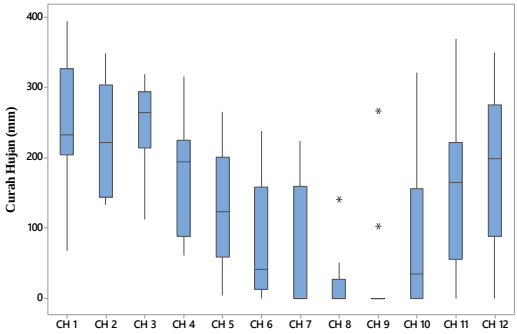
### 4.1.4 Identifikasi *Outlier*

Pada identifikasi pengamatan *outlier* dalam data luas panen padi dan curah hujan dilakukan dengan dua cara yaitu cara visual dengan *boxplot* dan pemeriksaan statistik menggunakan nilai *DfFITS* dari residual regresi OLS. Pengidentifikasian ini dilakukan sebagai pertimbangan dalam penggunaan regresi *robust*. Berikut hasil *boxplot* dari luas panen setiap periodenya.




**Gambar 4.4** *Boxplot* Luas Panen Padi per Periode di Kabupaten Gresik

Gambar 4.4 menunjukkan bahwa secara visual hanya satu *outlier* yang terjadi yaitu pada luas panen periode pertama tepatnya pada tahun 2009. Luas panen padi tiap periode memiliki perbedaan yang cukup signifikan terutama antara luas panen periode kedua dengan ketiga. Pada periode kedua memilik rentang data yang cukup luas dibandingkan dengan dua periode lainnya. Selanjutnya, dilakukan identifikasi *outlier* pada data curah hujan. Berikut merupakan tampilan *boxplot* untuk curah hujan setiap bulan.




**Gambar 4.5** *Boxplot* Curah Hujan tiap Bulan di Kabupaten Gresik

Secara visual dari Gambar 4.5 terdapat tiga pengamatan *outlier* yaitu satu pengamatan pada bulan Agustus, dan dua pengamatan pada bulan September. Hal ini mengindikasikan adanya ketidak stabilan pada curah hujan yang terjadi pada bulan tersebut. Selain itu, data curah hujan setiap bulan juga memiliki rentang yang luas dan persebaran data yang tidak simetris, hal ini dikarenakan nilai dari garis median yang tidak berlokasi di tengah box dan salah satu dari *whisker* lebih panjang dari yang lainnya.

Setelah melakukan identifikasi *outlier* secara visual dan didapatkan adanya dua periode yang memiliki amatan *outlier*, selanjutnya dilakukan identifikasi *outlier* secara matematis dengan menghitung nilai *DfFITS* yang didapatkan dari residual hasil OLS. Pengamatan digolongkan menjadi amatan *outlier* apabila nilai $|DfFITS| > 2\sqrt{\frac{k+1}{n}}$, diperoleh nilai 1,291 sebagai *cutoff*. Tabel 4.4 menunjukkan nilai *DfFITS* residual yang menunjukkan adanya *outlier*.

**Tabel 4.4** Nilai *DFFITS* Pengamatan *Outlier* Tahun 2006-2017

| Periode | Pengamatan ke- | *DFFITS* |
|---|---|---|
| 1 | 4 | -5.10072 |
| | 7 | 1.470019 |
| | 9 | -2.00524 |
| 2 | 5 | -3.23011 |
| | 10 | 1.343738 |
| | 11 | 6.685189 |
| | 12 | 1.860642 |
| 3 | 11 | -1.6073 |
| | 12 | 1.99233 |

Berdasarkan Tabel 4.4 didapatkan 9 pengamatan dari total 36 data yang memiliki nilai *DFFITS* lebih besar dibandingkan dengan nilai *cutoff*nya. Setelah dilakukan deskripsi dan pendeteksian *outlier*, ditemukan adanya data *outlier* pada residual setiap periode secara visual maupun matematis. Adanya *outlier* pada setiap periode dapat menjadi pertimbangan untuk dilakukan pemodelan luas panen padi dengan analisis regresi *robust* yang dapat menghasilkan koefisien regresi yang tidak terpengaruh walaupun terdapat residual *outlier*.

## 4.2 Pemodelan Luas Panen Padi Kabupaten Gresik

Tahap selanjutnya adalah analisis perbandingan untuk mendapatkan metode estimasi terbaik menggunakan *Ordinary Least Square (OLS)*, Regresi *Robust M-Estimation*, Regresi *Robust S-Estimation*, Regresi *Robust MM-Estimation*, dan Regresi *Ridge Robust.*

### 4.2.1 Pemodelan *Ordinary Least Square* (OLS)

Pemodelan menggunakan Ordinary Least Square (OLS) bertujuan untuk menaksirkan parameter dengan meminimumankan jumlah kuadrat residual. Pemodelan awal menggunakan OLS ditujukan sebagai kontrol, selain itu untuk pengecekan *outlier* dan multikolinearitas sebelum nantinya dilakukan pemodelan dengan pendekatan menggunakan Regresi *ridge robust.*

Pada pemodelan luas panen padi di Kabupaten Gresik dilakukan uji signifikansi parameter secara serentak dengan taraf signifikan α = 5%.

**Tabel 4.5** Statistik $F_{hitung}$ dari Model Hasil Regresi OLS tiap Periode di Kabupaten Gresik Tahun 2006-2017

| Periode | $F_{hitung}$ | Keputusan |
|---|---|---|
| 1 | 1,92 | Gagal tolak $H_0$ |
| 2 | 2,62 | Gagal tolak $H_0$ |
| 3 | 0,84 | Gagal tolak $H_0$ |

Hasil pengujian signifikansi parameter secara serentak didapatkan bahwa semua model menghasilkan keputusan gagal tolak $H_0$. Hal ini dikarenakan $F_{hitung}$ lebih kecil dibandingkan dengan $F_{0,05(4,7)}$ senilai 4,12. Hasil pemodelan untuk ketiga periode luas panen padi tidak signifikan secara serentak berdasarkan pengolahan menggunakan OLS atau dengan kata lain diartikan bahwa tidak ada valiabel curah hujan yang berpengaruh signifikan terhadap luas panen.

**Tabel 4.6** Estimasi Parameter Regresi OLS di Kabupaten Gresik

| Periode | Parameter | Koef | R-sq |
|---|---|---|---|
| | $\beta_0$ | 35328 | |
| | $\beta_1$ | 5.1 | |
| $LP_1$ | $\beta_2$ | -26.7 | 52,30% |
| | $\beta_3$ | -26.9 | |
| | $\beta_4$ | 40.5 | |
| | $\beta_0$ | 17226 | |
| | $\beta_1$ | 30.3 | |
| $LP_2$ | $\beta_2$ | 28.3 | 59,92% |
| | $\beta_3$ | -18.8 | |
| | $\beta_4$ | -21.1 | |
| | $\beta_0$ | 4342 | |
| | $\beta_1$ | -9.8 | |
| $LP_3$ | $\beta_2$ | 19.5 | 35,79% |
| | $\beta_3$ | -6.3 | |
| | $\beta_4$ | -0.26 | |

Nilai koefisien determinasi (R-Sq) pada Tabel 4.6 dapat menjelaskan bahwa variabel luas panen padi sebagai variabel dependen ($Y$) mampu dijelaskan oleh curah hujan sebagai variabel prediktor ($X$) sebesar 52,3%, 59,92%, dan 35,79% untuk periode pertama, kedua, dan ketiga sedangkan sisanya dijelaskan oleh variabel lain.

Model regresi yang diperoleh dari *OLS* dapat dikatakan memiliki estimator yang tidak bias apabila memenuhi uji asumsi klasik. Model regresi yang baik adalah model yang memiliki nilai residual berdistribusi normal. *Kolmogorov-Smirnov Test* dapat digunakan untuk pengecekan asumsi normalitas pada residual. Berdasarkan *plot* dari residual hasil pengujian menggunakan *Kolmogorov-Smirnov Test* pada Lampiran 4, didapatkan nilai *p-value* sebagai berikut.

**Tabel 4.7** Nilai p-value Pengujian Asumsi Normalitas Residual

| **Periode** | ***p-value*** | **Keputusan** |
|---|---|---|
| 1 | >0,150 | Tolak $H_0$ |
| 2 | >0,150 | Tolak $H_0$ |
| 3 | >0,150 | Tolak $H_0$ |

Berdasarkan pengujian normalitas menggunakan *Kolmogorov - Smirnov Test* didapatkan *p-value* (>0,15) > $\alpha$ (0,05) untuk semua periode, sehingga diambil keputusan tolak $H_0$ artinya residual data berdistribusi normal, maka dapat disimpilkan bahwa asumsi terpenuhi untuk semua periode. Selain pengujian distribusi normal dilakukan pengujian asumsi lain yaitu asumsi identik menggunakan pendekatan analisis grafik pada Lampiran 5. Berdasarkan grafik sebaran pada *versus fits*, titik residual dalam grafik menyebar secara acak, sehingga menunjukkan tidak terdapat heteroskedastisitas dalam model atau residual memiliki varians yang homogen maka asumsi identik terpenuhi untuk ketiga periode.

Pengujian asumsi lainnya adalah asumsi independen yang memiliki tujuan untuk megetahui ada tidaknya korelasi antar pengamatan satu dengan pengamatan sebelumnya. Pengujian independensi menggunakan pengujian autokorelasi

dengan pendekatan analisis grafik pada Lampiran 5. Berdasarkan sebaran titik-titik dalam grafik *versus order* tidak membentuk suatu pola tertentu sehingga dapat ditarik kesimpulan tidak terdapat autokorelasi atau residual independen pada setiap periode luas panen. Asumsi lainnya yang penting untuk dilakukan pengecekan adalah asumsi multikolinearitas.. Multikolinearitas dapat dideteksi dengan beberapa cara antara lain pengecekan nilai korelasi antar variabel prediktor, seperti pada Tabel 4.8. Apabila nilai VIF lebih besar dari 10 maka mengindikasikan multikolinearitas.

**Tabel 4.8** Nilai VIF Variabel Curah Hujan

| Periode | Variabel | VIF |
|---|---|---|
| 1 | $CH_1$ | 1.7 |
| | $CH_2$ | 2.11 |
| | $CH_3$ | 3.53 |
| | $CH_4$ | 2.2 |
| 2 | $CH_5$ | 4.62 |
| | $CH_6$ | **11.19** |
| | $CH_7$ | **11.91** |
| | $CH_8$ | 1.94 |
| 3 | $CH_9$ | 5.93 |
| | $CH_{10}$ | **12.25** |
| | $CH_{11}$ | 7.11 |
| | $CH_{12}$ | 1.84 |

Dari 3 periode hanya periode pertama yang memenuhi asumsi multikolinearitas sedangkan untuk periode kedua dan tiga tidak. Hal ini karena nilai VIF yang lebih dari 10 untuk variabel $CH_6$, $CH_7$, dan $CH_{10}$. Adapun hasil terhadap identifikasi multikolinieritas secara lengkap terdapat dalam Lampiran 7. Adanya multikolinearitas pada data dapat dijadikan sebagai bahan pertimbangan untuk selanjutnya dilakukan pemodelan luas panen padi dengan analisis regresi *ridge* untuk menangani multikolinearitas.

### 4.2.2 Pemodelan Regresi *Robust*

Regresi *robust* digunakan sebagai salah satu cara untuk menangani pemodelan dengan pengamatan yang memiliki residual *outlier*. Pemodelan regresi *robust* yang digunakan untuk memodelkan luas panen padi di Kabupaten Gresik menggunakan 3 metode estimasi. Ketiga metode estimasi tersebut antara lain *M-Estimation*, *S-Estimation*, dan *MM-Estimation*.

**a. Regresi *Robust M-Estimation***

Analisis regresi *robust M-Estimation* digunakan dengan pembobot Tukey *Bisquare* untuk mendapatkan estimasi parameter terbaik. Berikut merupakan nilai koefisien hasil estimasi menggunakan regresi *robust M-Estimation* pada masing-masing periode luas panen padi di Kabupaten Gresik tahun 2006-2017.

**Tabel 4.9** Estimasi Parameter Regresi *Robust M-Estimation*

| Periode | Parameter | Koef | R-sq |
|---|---|---|---|
| | $\beta_0$ | 34933.570 | |
| | $\beta_1$ | 5.811 | |
| $LP_1$ | $\beta_2$ | -23.664 | 3,56% |
| | $\beta_3$ | -25.693 | |
| | $\beta_4$ | 36.614 | |
| | $\beta_0$ | 15984.840 | |
| | $\beta_1$ | 25.898 | |
| $LP_2$ | $\beta_2$ | 68.260 | 64,88% |
| | $\beta_3$ | -70.572 | |
| | $\beta_4$ | 20.596 | |
| | $\beta_0$ | 4251.667 | |
| | $\beta_1$ | -10.945 | |
| $LP_3$ | $\beta_2$ | 20.115 | 33,26% |
| | $\beta_3$ | -6.614 | |
| | $\beta_4$ | 0.682 | |

Pada Tabel 4.9 nampak bahwa pemodelan luas panen yang dilakukan menggunakan regresi *robust M-Estimation*

menghasilkan model estimasi yang baik pada periode kedua. Hal ini dapat dilihat dari nilai koefisien determinasi yang dihasilkan besar, selain itu nilai koefisien deteminasi dari regresi *robust M-estimation* lebih besar dibandingkan dengan OLS. Pada periode pertama dan ketiga estimasi menggunakan metode ini menghasilkan model estimasi yang kurang baik karena nilai koefisien determinasi kecil, bahkan lebih kecil dibandingkan dengan model yang dihasilkan dari OLS.

**b. Regresi *Robust MM-Estimation***

Pada regresi *robust* dengan *MM-Estimation* digunakan dengan pembobot Tukey *Bisquare* dan estimasi awal *S-Estimation* untuk mendapatkan estimasi parameter terbaik. Pemodelan regresi *robust* luas panen padi di Kabupaten Gresik menggunakan *MM-Estimation* karena memiliki *breakdown point* yang cukup besar yaitu 50%. Berikut merupakan nilai koefisien dan koefisien determinasi hasil estimasi menggunakan regresi *robust MM-Estimation* pada setiap periode luas panen padi di Kabupaten Gresik tahun 2006-2017.

**Tabel 4.10** Estimasi Parameter Regresi *Robust MM-Estimation*

| Periode | Parameter | Koef | R-sq |
|---|---|---|---|
| | $\beta_0$ | 30563.390 | |
| | $\beta_1$ | 5.067 | |
| $LP_1$ | $\beta_2$ | 12.094 | 18,63% |
| | $\beta_3$ | -11.185 | |
| | $\beta_4$ | 0.697 | |
| | $\beta_0$ | 16622.900 | |
| | $\beta_1$ | 29.060 | |
| $LP_2$ | $\beta_2$ | 36.190 | 71,85% |
| | $\beta_3$ | -23.240 | |
| | $\beta_4$ | -17.330 | |

**Tabel 4.10** Estimasi Parameter Regresi *Robust MM-Estimation* (Lanjutan)

| Periode | Parameter | Koef | R-sq |
|---|---|---|---|
| | $\beta_0$ | 4100.828 | |
| | $\beta_1$ | -9.887 | |
| $LP_3$ | $\beta_2$ | 19.495 | 34,43% |
| | $\beta_3$ | -5.886 | |
| | $\beta_4$ | 0.452 | |

Pada estimasi menggunakan metode *robust* dengan *MM-Estimation*, menghasilkan model pada periode pertama yang konvergen setelah melalui 17 iterasi sedangkan untuk periode kedua membutuhkan 25 iterasi, dan periode ketiga konvergen setelah 8 kali proses iterasi. Pemodelan menggunakan metode ini menghasilkan model estimasi yang baik pada periode kedua dan ketiga dengan koefisien determinasi yang besar. Pada periode kedua nilai koefiseien determinasi senilai 71,85%, hal ini menunjukkan besarnya variabilitas curah hujan yang mampu menjelaskan model luas panen padi. Sedangkan untuk periode pertama nilai koefisien determinasi kecil yaitu 18,63%, yang artinya model luas panen pada periode pertama hanya mampu dijelaskan 18,63% oleh variabel curah hujan, sedangkan sisahnya dijelaskan dengan variabel lain. Jika dibandingkan dengan hasil estimasi dengan metode *M-Estimation*, estimasi dengan menggunakan *MM-Estimation* pada data ini memberikan hasil model yang lebih baik walaupun memiliki pola yang sama dengan *M-Estimation* yaitu model pada periode kedua paling baik dan model pada periode pertama paling buruk ditinjau dari besarnya koefisien determinasi yang dihasilkan.

### c. Regresi *Robust S-Estimation*

Metode estimasi ketiga yang digunakan adalah *S-Estimation*. Metode ini digunakan dalam analisis regresi *robust* karena memiliki *breakdown point* yang cukup tinggi yaitu 50%. Pada analisa ini fungsi pembobot yang digunakan adalah Tukey *Bisquare* untuk mendapatkan estimasi parameter terbaik. Nilai koefisien hasil estimasi dan koefisien

determinasi pada setiap periode luas panen padi di Kabupaten Gresik tahun 2006-2017 menggunakan regresi *robust S-Estimation* adalah sebagai berikut.

**Tabel 4.11** Estimasi Parameter Regresi *Robust S-Estimation*

| Periode | Parameter | Koef | R-sq |
|---|---|---|---|
| | $\beta_0$ | 25.854,730 | |
| | $\beta_1$ | 7,750 | |
| $LP_1$ | $\beta_2$ | 30,835 | 94,79% |
| | $\beta_3$ | -1,628 | |
| | $\beta_4$ | -12,902 | |
| | $\beta_0$ | 15.826,965 | |
| | $\beta_1$ | 27,208 | |
| $LP_2$ | $\beta_2$ | 69,060 | 95,41% |
| | $\beta_3$ | -72,127 | |
| | $\beta_4$ | 20,526 | |
| | $\beta_0$ | 4.847,355 | |
| | $\beta_1$ | -14,854 | |
| $LP_3$ | $\beta_2$ | 22,528 | 54,27% |
| | $\beta_3$ | -10,272 | |
| | $\beta_4$ | 2,188 | |

Pada periode pertama dan kedua koefisien determinasi dari model yang dihassilkan besar, mencapai 94,79% untuk periode pertama dan 95,41% untuk periode kedua. Nilai koefisien determinasi ini mengindikasikan besarnya variabilitas dari curah hujan yang mampu menggambarkan luas panen padi di Kabupaten Gresik. Pada periode pertama luas panen mampu dijelaskan oleh curah hujan sebesar 94,79%, sedangkan sisanya dijelaskan oleh variabel lain. Periode ketiga memiliki nilai koefisien determinasi yang kecil dibandingkan dengan dua periode lainnya, namun nilainya berada diatas nilai koefisien determinasi menggunakan regresi OLS.

Performa dari model estimasi yang dihasilkan dengan menggunakan *robust estimation* secara keseluruhan lebih baik dibandingkan OLS terutama untuk model *S-Estimation* karena memiliki model dengan koefisien determinasi yang besar pada tiap periodenya. Zahiroh (2017) dalam penelitiannya membuktikan hal yang serupa yaitu model estimasi menggunakan regresi *robust* memberikan performa yang lebih baik dibandingkan dengan OLS. Hasil dari estimasi ini juga menunjukkan bahwa metode *MM-Estimation* yang merupakan penggabungan dari dua metode estimasi yaitu *S-Estimation* dan *M-Estimation* tidak menghasilkan model yang lebih baik dibandingkan dengan metode estimasi penyusunnya ketika berdiri sendiri. Penelitian ini mendukung simpulan dari Savitri (2015) yang juga menunjukkan hasil serupa yaitu model estimasi dengan metode *MM-Estimation* yang tidak lebih baik dibandingkan dengan metode estimasi penyusunnya, *S-Estimation* dan *M-Estimation.*

### 4.2.3 Pemodelan *Ridge Robust*

Setelah dilakukan pemodelan menggunakan regresi *robust* untuk menangani amatan *outlier*, tahap selanjutnya adalah melakukan analisis regresi *ridge* untuk mengatasi multikolinearitas pada data yang digunakan. Tahapan pertama yang dilakukan adalah *centering* dan *rescaling* data menggunakan persamaan pada 2.35. proses ini dilakukan untuk menangani data *ill-conditioned.* Dari proses didapatkan bentuk matriks $Y_i^R$ dan $X_i^R$ yang merupakan hasil *centering* dan *rescaling* data pada Lampiran 11. Proses selanjutnya perhitungan nilai konstanta regresi *ridge*, *k*, menggunakan formula yang ada pada 2.38 untuk setiap model.

**Tabel 4.12** Perbandingan Nilai Konstanta Regresi *Ridge* 3 Estimasi

| Periode | *Robust M* | *Robust MM* | *Robust* S |
|---|---|---|---|
| $LP_1$ | 0,0409 | 0,1356 | 0,3176 |
| $LP_2$ | 0,2848 | 0,1579 | 0,2945 |
| $LP_3$ | 1,0424 | 1,1101 | 0,9383 |

Nilai *k* atau konstanta regresi *ridge* umumya berada pada rentang 0 hingga 1. Namun pada Tabel 4.12 dapat dilihat bahwa model luas panen periode ketiga dengan menggunakan *robust* estimasi M dan MM menghasilkan nilai *k* yang lebih dari rentang tersebut. Dari nilai *k* yang didapat selanjutnya dilakukan estimasi parameter regresi *ridge robust*. Pada analisis regresi *ridge robust* digunakan tiga metode estimasi *robust* yaitu *M-Estimation, MM-Estimation,* dan *S-Estimation*. Berikut hasil estimasi untuk setiap periode luas panen padi di Kabupaten Gresik tahun 2006 smpai 2017 menggunakan regresi *ridge robust M, S,* dan *MM-Estimation.*

**Tabel 4.13** Perbandingan Regresi *Ridge Robust* dengan 3 Metode Estimasi Tahun 2006-2017

| Periode | Parameter | ***Ridge Robust M*** | ***Ridge Robust MM*** | ***Ridge Robust* S** |
|---|---|---|---|---|
| $LP_1$ | $\beta_0$ | 34.897,818 | 34.196,310 | 33.416,561 |
| | $\beta_1$ | 4,631 | 4,038 | 3,408 |
| | $\beta_2$ | -25,107 | -21,863 | -17,387 |
| | $\beta_3$ | -22,970 | -17,324 | -11,970 |
| | $\beta_4$ | 35,940 | 28,474 | 20,297 |
| $LP_2$ | $\beta_0$ | 18.685,350 | 18.316,207 | 18.707,404 |
| | $\beta_1$ | 16,753 | 20,211 | 16,557 |
| | $\beta_2$ | 11,784 | 13,685 | 11,683 |
| | $\beta_3$ | 3,577 | 1,145 | 3,683 |
| | $\beta_4$ | -11,176 | -15,321 | -10,915 |
| $LP_3$ | $\beta_0$ | 3.964,315 | 3.972,201 | 3.952,222 |
| | $\beta_1$ | 2,277 | 2,259 | 2,298 |
| | $\beta_2$ | 3,699 | 3,593 | 3,879 |
| | $\beta_3$ | 2,113 | 2,076 | 2,173 |
| | $\beta_4$ | -0,329 | -0,291 | -0,394 |

Pada Tabel 4.13 dapat terlihat hasil estimasi parameter *Ridge Robust* yang terbentuk dari tiga metode estimasi yang digunakan pada tiap periode luas panen padi. Model yang terbentuk memiliki perbedaan nilai koefisien dengan estimasi

yang lainnya namun perbedaan yang muncul tidak terlalu jauh.

Setelah mendapatkan model estimasi dengan metode regresi *ridge robust*, maka tahapan selanjutnya adalah menentukan model estimasi terbaik. Pemilihan model hasil estimasi terbaik dapat dilakukan dengan beberapa cara antara lain menghitung nilai indeks performansi kemudian membandingkan nilai tersebut disetiap estimasi yang digunakan. Model terbaik adalah model yang memiliki nilai indeks performansi terkecil.

## 4.3 Metode Estimasi Terbaik

Peritungan indeks performansi menggunakan *Root Mean Square* Error (RMSE) untuk menentukan estimasi terbaik pada masing-masing periode luas panen padi dilakukan kerena RMSE memuat dua unsur penting yang berguna dalam menentukan model estimasi terbaik, yaitu unsur presisi yang menggambarkan jarak antara titik satu pengamatan dengan pengamatan lain dan unsur akurasi yang menunjukkan kedekatan antara nilai prediksi dengan kenyataan. Model terbaik adalah model dengan nilai RMSE terkecil. Teknik yang digunakan dalam pemilihan model terbail adalah *Leave One Out Cross Validation* (LOOCV).

### 4.3.1 Pemilihan Metode Estimasi Terbaik

Nilai RMSE yang didapatkan dari proses *Leave One Out Cross Validation* (LOOCV) dibandingkan untuk setiap metode estimasi yang digunakan untuk menentukan metode estimasi terbaik. Teknik LOOCV digunakan karena data pengamatan sedikit sehingga proses dari LOOCV yang membutuhkan pengulangan tidak memerlukan banyak waktu. Selain itu, keuntungan dari teknik ini adalah mengurangi potensi bias karena melibatkan semua data. Berikut adalah tabel perbandingan nilai RMSE dari proses LOOCV untuk masing-masing metode regresi tiap periode.

**Tabel 4.14** Perbandingan RMSE

| Model | Periode 1 | Periode 2 | Periode 3 |
|---|---|---|---|
| OLS | 5.078,702 | 6.131,694 | 2.554,191 |
| *Robust* M | 3.594,240 | 6.704,887 | 2.464,390 |
| *Robust* MM | 3.457,740 | 6.348,529 | 2.196,351 |
| *Robust* S | **3.308,176** | 7.139,868 | 2.497,996 |
| *Ridge Robust* M | 3.990,461 | 5.035,982 | **2.036,689** |
| *Ridge Robust* MM | 3.912,151 | 5.009,441 | 2.108,786 |
| *Ridge Robust* S | 3.812,321 | **4.821,427** | 2.038,727 |

Bedasarkan Tabel 4.14 dapat dilihat bahwa nilai RMSE terkecil dimiliki oleh estimasi regresi *Robust S-Estimation* untuk periode pertama, regresi *Ridge Robust S-Estimation* untuk periode kedua, dan regresi *Ridge Robust M-Estimation* untuk periode ketiga.

Pada periode pertama metode regresi *Robust* lebih baik dibandingkan dengan *Ridge Robust* karena berdasarkan pada pengujian multikolinearitas menggunakan nilai VIF menunjukkan bahwa tidak terjadi kejadian multikolinearitas di periode pertama. Di lain sisi pada periode kedua dan ketiga dengan adanya kejadian multikolinearitas secara optimum dimodelkan dengan metode *Ridge Robust*. Maka dapat disimpulan bahwa metode estimasi yang optimum digunakan dalam mengestimasi luas panen padi pada di Kabupaten Gresik yang mengandung *outlier* dan multikolinearitas beragam bergantung dari karakteristik data. Hal ini memperkuat simpulan dari penelitian yang dilakukan oleh Savitri (2015) yang menunjukkan bahwa untuk memodelkan luas panen padi di Kabupaten Bondowoso membutuhkan metode estimasi yang berbeda disetiap periodenya.

### 4.3.2 Penggunaan Model Estimasi

Model estimasi optimum dari setiap periode berbeda pada setiap periodenya. Pada periode pertama, model estimasi terbaik yang digunakan adalah regresi *robust S-Estimation*, berikut merupakan model estimasi periode pertama.

$$\hat{LP}_1 = 25854,73 + 7,75CH_1 + 30,853CH_2 - 1,628CH_3 - 12,90CH_4$$

Luas panen pada periode pertama akan mengalami peningkatan 7,75 hektar apabila curah hujan pada bulan Januari meningkat satu satuan atau meningkat 30,85 hektar apabila curah hujan pada bulan Februari meningkat satu satuan dengan syarat nilai variabel lainnya tetap. Sebaliknya, luas panen pada periode pertama akan mengalami penurunan sebesar 1,62 hektar apabila curah hujan pada bulan Maret meningkat satu satuan atau akan menurun sebesar 12,9 hektar apabila curah hujan bulan April meningkat satu satua, dengan syarat curah hujan lainnya tetap.

Model estimasi terbaik pada periode kedua adalah regresi *ridge robust S-Estimation*. Berikut merupakan model estimasi periode kedua.

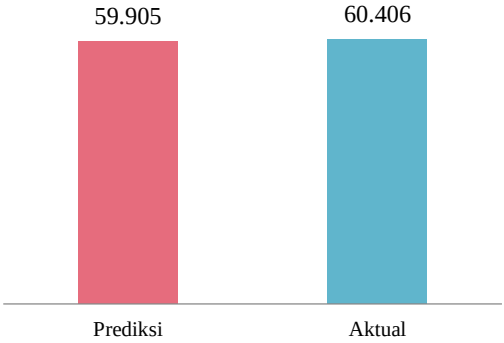
$$\hat{LP}_2 = 18707,40 + 16,56\ CH_5 + 11,68CH_6 + 3,68\ CH_7 - 10,92CH_8$$

Pada periode kedua, luas panen akan mengalami penurunan senilai 10,92 hektar apabila terjadi kenakan curah hujan pada bulan Agustus sebesar satu satuan, di mana variabel lainnya konstan. Peningkatan satu satuan pada curah hujan bulan Mei, Juni, dan Juli masing masing akan meningkatkan total luas panen sebesar 16,56 hektar, 11,68 hektar, dan 3,68 ketar dengan syarat curah hujan pada bulan lainnya konstan. Pada luas panen periode ketiga, model estimasi terbaik adalah regresi *ridge robust M-Estimation*, dengan model estimasi sebagai berikut.

$$\hat{LP}_3 = 3964,32 + 2,28CH_9 + 3,70CH_{10} + 2,11\ CH_{11} - 0.33CH_{12}$$

Pada periode ketiga luas panen mengalami penurunan jika terjadi peningkatan curah hujan pada bulan panennya, dalam hal ini adalah bulan Desember, sama halnya seperti pada periode pertama dan kedua. Peningkatan curah hujan pada bulan Desember akan menurunkan luas panen sebsar 0,33 hektar dengan syarat variabel lainnya konstan.

Dari model optimum tiap periode luas panen padi digunakan untuk memprediksi luas panen padi pada tahun 2018 di Kabupaten Gresik menggunakan data curah hujan pada tahun 2018. Sehingga dari proses perhitungan didapatkan nilai luas panen total sebesar 59.905,45 hektar dengan rincian 33.899,71 hektar pada periode pertama, 20.948,55 hektar pada periode kedua, dan 5.057,19 hektar pada periode ketiga. Prediksi luas panen periode pertama yang lebih tinggi dibandingkan dengan dua periode lainnya menunjukkan bahwa periode pertama menjadi periode penanaman padi yang lebih produktif. Berikut perbandingan antara total luas panen di Kabupaten Gresik dibandingkan dengan hasil dari estimasi tersebut.




**Gambar 4.6** Perbandingan Data Aktual dan Prediksi Luas Panen Padi Tahun 2018 di Kabupaten Gresik

Berdasarkan dari Gambar 4.6 dapat dilihat bahwa hasil prediksi total luas panen lebih kecil jika dibandingkan dengan data aktualnya. Jumlah luas panen aktual mencapai 60.406 hektar atau lebih besar 500,55 hektar dibandingkan dengan nilai prediksinya. Jika dilihat dari persentase ketepatan prediksi, nilai prediksi berbeda 0,83% dibandingkan dengan nilai aktual maka dari itu dapat dikatakan bahwa nilai prediksi

hampir mendekati data aktual. Nilai tersebut menunjukkan kinerja regresi *robust* maupun *ridge robust* yang digunakan cukup baik dibandingkan dengan regresi OLS untuk data dengan amatan *outlier* maupun kejadian multikolinearitas.

# BAB V

# KESIMPULAN DAN SARAN

## 5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan, dapat diambil beberapa kesimpulan sebagai berikut.

1. Data luas panen padi di Kabupaten Gresik tahun 2006 hingga 2017 mengandung data *outlier*, begitu pula dengan data curah hujan perbulannya. Pola hubungan antara luas panen dengan curah hujan setiap bulannya sebagian besar berhubungan secara linear positif. Selain itu terdapat juga korelasi yang tinggi antara variabel prediktor pada periode kedua dan ketiga yang mengindikasikan terjadinya multikolinearitas.
2. Hasil pemodelan luas panen padi menggunakan regresi *robust M-Estimation*, *S-Estimation*, dan *MM-Estimation* serta *ridge robust M-Estimation*, *S-Estimation*, dan *MM-Estimation* menghasilkan model estimasi yang lebih baik daripada OLS. Penentuan model terbaik dari regresi *ridge robust* dilakukan dengan membandingkan nilai *Root Mean Square Error* (RMSE) dari masing-masing metode estimasi. Model luas panen pada periode pertama secara optimal dimodelkan menggunakan regresi *Robust S-Estimation*, sedangkan untuk periode kedua dan ketiga masing masing lebih optimal dimodelkan menggunakan metode Regresi *Ridge Robust S-Estimation* dan Regresi *Ridge Robust M-Estimation*.
3. Berdasarkan dari model terbaik dari regresi *ridge robust* setiap periode didapatkan jumlah luas panen hasil prediksi lebih kecil jika dibandingkan dengan data aktualnya. Perbedaan jumlah luas panen sebesar 500,55 hektar atau berbeda sekitar 0,83% dibandingkan dengan data aktualnya.

## 5.2 Saran

Berdasarkan kesimpulan yang diperoleh, berikut merupakan beberapa saran untuk penelitian selanjutnya.

1. Dapat dilakukan pendalaman metode, diantaranya efek *breakdown point* dan efisiensi terhadap hasil estimasi parameter.
2. Dapat dilakukan percobaan regresi *robust* dengan estimasi yang lain, seperti halnya *Least Median Square (LMS)*, *Least Trimmed Square (LTS)* dan lainnya.
3. Perlu dilakukan peninjauan ulang terhadap variabel lain yang memiliki potensi mempengaruhi luas panen padi.

# DAFTAR PUSTAKA


Draper, N. R., dan Smith, H. (1998). *Applied Regression Analysis Third Edition.* Canada: John Wiley & Sons, Inc.

Greene, W. (2002). *Econometric Analysis.* Prentice Hall,Inc.

Gujarati, D. N. (2004). *Basic Econometrics.* New York: Mc Gwra Hill,Inc.

Hardin, P. I. (2003). *Common Errors in Statistics (And How to Avoid Them).* Canada: John Wiley & Sons, Inc.

Haryono. (2011). *Pedoman Umum: Adaptasi Perubahan Iklim Sektor Pertanian.* Jakarta: Badan Penetitian dan Pengembanyan Pertanian.

Hasan, I. (1999). *Pokok-Pokok Materi Statistik 1 (Statistika Deskriptif).* Jakarta: Bumi Aksara.

Hoerl, Arthur E., dan Kennard, Robert W. (1976). *Ridge regression iterative estimation of the biasing parameter*. *Communications in Statistics - Theory and Methods, 5:1, 77-88*

Kurniati, M., Yundari, dan Rizki, S. W. (2019). Metode *Least Median Square* (LMS) Dalam Analisis Regresi *Robust* Ketika Terdapat *Outlier*. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, 869 – 874.

Montgomery, D. C., Peck, E. A., dan Vining, G. G. (2012). *Introduction To Linear Regression Analysis 5th Edition.* New Jersey: John Wiley & Sons, Inc.

Paludi, S. (2009, Juni). Identifikasi Dan Pengaruh Keberadaan Data Pencilan (*Outlier*). *Majalah Ilmiah Panorama Nusantara, edisi VI, Januari - Juni 2009* .

Prabeswari, G. A. (2019). *Pemodelan Konsumsi Rumah Tangga di Kota Surabaya Berdasarkan Indikator Tingkat Kesejahteraan dengan Metode Two Stage Robust Ridge Regression*. Surabaya.

Pratiwi, Y. I. (2016). *Penerapan Metode Regresi Ridge Robust Estimasi M.* Jakarta.

Rousseeuw, P. J. (1987). *Robust Regression and Outlier Detection.* Canada: John Wiley & Sons, Inc.

Rousseeuw, P. J. (2011, Januari/Februari). *Robust* statistics for *outlier* detection. *Wires Data Mining Knowl Discov*, pp. 73-79.

Saputra, C. W. (2014). Pemodelan Regresi *Robust* LTS dan Uji Kesamaan Model Anomali Luas Panen Padi pada Beberapa Kabupaten di Jawa Timur. Tugas Akhir, Institut Teknologi Sepuluh Nopember, Surabaya.

Savitri, R. A. D. (2015). *Pemodelan Luas Panen Padi Dengan Indikator El-Nino Southern Oscillation (Enso) Di Kabupaten Bondowoso Menggunakan Regresi Robus.* Surabaya.

Susanti, Y., & Pratiwi, H. (2014). *M estimation, S estimation, and MM estimation in robust regression*. *International Journal of Pure and Applied Mathematics*, *91*(3), 349-360.

Tim SUTAS 2018. (2018). *Hasil Survei Pertanian antar Sensus (SUTAS).* Jakarta: Badan Pusat Statistik.

Walpole, R. E. (1995). *Pengantar Statistika.* Jakarta: PT Gramedia Pustaka Utama.

Zairoh, N. (2017). *Regresi Robust S-Estimation Dan MM-Estimation Untuk Pemodelan Luas Panen Padi Dengan Indikator Iklim Di Sentra Produksi Padi Jawa Timur.* Surabaya.

# LAMPIRAN

**Lampiran 1A.** Data Luas Panen dan Curah Hujan Kab. Gresik Periode 1 (Januari, Februari, Maret, April)

| Tahun | LP1 | CH1 | CH2 | CH3 | CH4 |
|---|---|---|---|---|---|
| 2006 | 33770 | 130.11 | 133.83 | 112.05 | 96.66 |
| 2007 | 29467.00 | 68 | 139 | 118.79 | 81.93 |
| 2008 | 30447.00 | 200.1 | 138.9 | 231.6 | 61.1 |
| 2009 | 20303.21 | 241.3 | 348.8 | 295.1 | 125.7 |
| 2010 | 33259.00 | 277.9 | 262.1 | 288.7 | 230.9 |
| 2011 | 30184.00 | 216.9 | 187 | 292.4 | 192.1 |
| 2012 | 32328.00 | 394.7 | 161.7 | 208.7 | 85.9 |
| 2013 | 32058.00 | 343.7 | 201.1 | 262.8 | 197.2 |
| 2014 | 31401.41 | 373.7 | 243.4 | 319.6 | 316.1 |
| 2015 | 30109.00 | 270 | 324 | 259 | 225 |
| 2016 | 33336.60 | 224 | 308 | 266 | 224 |
| 2017 | 33279.60 | 220 | 290 | 298 | 223 |

**Lampiran 1B.** Data Luas Panen dan Curah Hujan Kab. Gresik Periode 2 (Mei, Juni, Juli, Agustus)

| Tahun | LP2 | CH5 | CH6 | CH7 | CH8 |
|---|---|---|---|---|---|
| 2006 | 17290 | 4.55 | 0 | 0 | 0.55 |
| 2007 | 18017.00 | 25.42 | 34.74 | 0 | 0 |
| 2008 | 17513.00 | 53 | 14.1 | 0 | 0 |
| 2009 | 21768.00 | 117.3 | 32.1 | 0 | 0 |
| 2010 | 17765.00 | 178.5 | 77.8 | 125.1 | 51.5 |
| 2011 | 20314.00 | 118.5 | 12 | 0 | 0 |
| 2012 | 22826.00 | 127.6 | 49.1 | 0 | 0 |
| 2013 | 23593.00 | 206.7 | 128 | 91.7 | 35 |
| 2014 | 24768.11 | 202 | 224.4 | 171.5 | 4 |
| 2015 | 24784.00 | 77 | 0 | 0 | 0 |
| 2016 | 26499.20 | 266 | 238 | 224 | 140 |
| 2017 | 27266.40 | 199 | 169 | 178 | 0 |

**Lampiran 1C.** Data Luas Panen dan Curah Hujan Kab. Gresik Periode 3 (September, Oktober, November, Desember)

| Tahun | LP3 | CH9 | CH10 | CH11 | CH12 |
|---|---|---|---|---|---|
| 2006 | 3245 | 0 | 0.38 | 29.66 | 55.78 |
| 2007 | 3672.00 | 0 | 54 | 247.03 | 173.87 |
| 2008 | 2126.00 | 0 | 71.5 | 174 | 301.3 |
| 2009 | 941.00 | 0 | 0 | 53.7 | 76.5 |
| 2010 | 6697.00 | 102.6 | 184.4 | 169.1 | 256.2 |
| 2011 | 4414.00 | 0 | 64.8 | 215.3 | 282.5 |
| 2012 | 3524.00 | 0 | 11.6 | 62.9 | 220.9 |
| 2013 | 5160.00 | 0 | 15.06 | 161.8 | 350.3 |
| 2014 | 5646.44 | 0 | 0 | 0 | 0 |
| 2015 | 6025.00 | 0 | 0 | 63 | 125 |
| 2016 | 6189.39 | 266 | 322 | 224 | 224 |
| 2017 | 7417.60 | 0 | 261 | 370 | 178 |

**Lampiran 2A.** *Output Ordinary Least Square* (OLS) Periode 1

**Regression Analysis: LP1 versus CH 1, CH 2, CH 3, CH 4**

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Regression | 4 | 75650685 | 18912671 | 1.92 | 0.212 |
| CH 1 | 1 | 1502306 | 1502306 | 0.15 | 0.708 |
| CH 2 | 1 | 22642918 | 22642918 | 2.30 | 0.173 |
| CH 3 | 1 | 10479234 | 10479234 | 1.06 | 0.337 |
| CH 4 | 1 | 51428598 | 51428598 | 5.22 | 0.056 |
| Error | 7 | 69010019 | 9858574 | | |
| Total | 11 | 144660704 | | | |

**Model Summary**

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 3139.84 | 52.30% | 25.04% | 0.00% |

**Coefficients**

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 35328 | 3744 | 9.44 | 0.000 | |
| CH 1 | 5.1 | 13.0 | 0.39 | 0.708 | 1.70 |
| CH 2 | -26.7 | 17.6 | -1.52 | 0.173 | 2.11 |
| CH 3 | -26.9 | 26.1 | -1.03 | 0.337 | 3.53 |
| CH 4 | 40.5 | 17.7 | 2.28 | 0.056 | 2.20 |

**Regression Equation**

LP1 = 35328 + 5.1 CH 1 - 26.7 CH 2 - 26.9 CH 3 + 40.5 CH 4

**Lampiran 2B.** *Output Ordinary Least Square* (OLS) Periode 2

**Regression Analysis: LP2 versus CH 5, CH 6, CH 7, CH 8**

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Regression | 4 | 87131519 | 21782880 | 2.62 | 0.126 |
| CH 5 | 1 | 14384218 | 14384218 | 1.73 | 0.230 |
| CH 6 | 1 | 5933652 | 5933652 | 0.71 | 0.426 |
| CH 7 | 1 | 2468297 | 2468297 | 0.30 | 0.603 |
| CH 8 | 1 | 4347918 | 4347918 | 0.52 | 0.493 |
| Error | 7 | 58280122 | 8325732 | | |
| Total | 11 | 145411641 | | | |

**Model Summary**

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 2885.43 | 59.92% | 37.02% | 0.00% |

**Coefficients**

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 17226 | 2020 | 8.53 | 0.000 | |
| CH 5 | 30.3 | 23.0 | 1.31 | 0.230 | 4.62 |
| CH 6 | 28.3 | 33.5 | 0.84 | 0.426 | 11.19 |
| CH 7 | -18.8 | 34.5 | -0.54 | 0.603 | 11.91 |
| CH 8 | -21.1 | 29.2 | -0.72 | 0.493 | 1.94 |

**Regression Equation**

LP2 = 17226 + 30.3 CH 5 + 28.3 CH 6 - 18.8 CH 7 - 21.1 CH 8

**Lampiran 2C.** *Output Ordinary Least Square* (OLS) Periode 3

**Regression Analysis: LP3 versus CH 9, CH 10, CH 11, CH 12**

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Regression | 4 | 14220700 | 3555175 | 0.84 | 0.549 |
| CH 9 | 1 | 1139944 | 1139944 | 0.27 | 0.623 |
| CH 10 | 1 | 4264841 | 4264841 | 1.00 | 0.355 |
| CH 11 | 1 | 715308 | 715308 | 0.17 | 0.696 |
| CH 12 | 1 | 4480 | 4480 | 0.00 | 0.975 |
| Error | 6 | 25509879 | 4251646 | | |
| Total | 10 | 39730579 | | | |

**Lampiran 2C.** *Output Ordinary Least Square* (OLS) Periode 3 (Lanjutan)

**Model Summary**

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 2061.95 | 35.79% | 0.00% | 0.00% |

**Coefficients**

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 4342 | 1489 | 2.92 | 0.027 | |
| CH 9 | -9.8 | 19.0 | -0.52 | 0.623 | 5.85 |
| CH 10 | 19.5 | 19.5 | 1.00 | 0.355 | 11.61 |
| CH 11 | -6.3 | 15.3 | -0.41 | 0.696 | 6.34 |
| CH 12 | -0.26 | 8.10 | -0.03 | 0.975 | 1.63 |

**Regression Equation**

LP3 = 4342 - 9.8 CH 9 + 19.5 CH 10 - 6.3 CH 11 - 0.26 CH 12

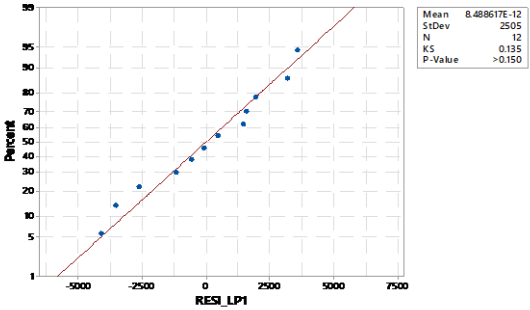
**Lampiran 3.** *Output DFFITS*

| Tahun | DFFIT LP1 | DFFIT LP2 | DFFIT LP3 |
|---|---|---|---|
| 2006 | 0.176489 | -0.02448 | -0.21704 |
| 2007 | -1.21356 | -0.32834 | -0.13385 |
| 2008 | 0.720605 | -0.28634 | -0.98905 |
| 2009 | -5.10072* | 0.015973 | -0.99767 |
| 2010 | 0.251211 | -3.23011* | 0.365294 |
| 2011 | -0.42598 | -0.18766 | 0.062386 |
| 2012 | 1.470019* | 0.093617 | -0.16452 |
| 2013 | -0.1018 | -0.24511 | 0.917397 |
| 2014 | -2.00524* | -1.19198 | 0.865277 |
| 2015 | -0.02474 | 1.343738* | 0.566751 |
| 2016 | 0.649471 | 6.685189* | -1.6073* |
| 2017 | 0.714162 | 1.860642* | 1.99233* |

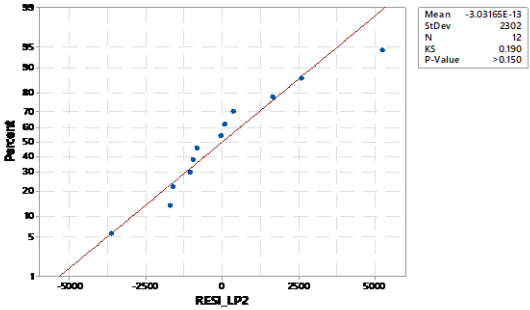
*) Keterangan : Data *Outlier*

**Lampiran 4.** *Output* Pengujian Asumsi Normalitas Residual

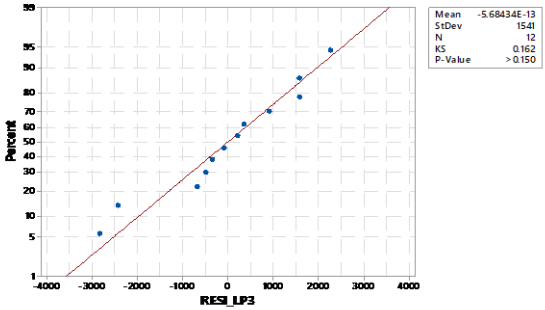
*Probability Plot* Luas Panen 1




*Probability Plot* Luas Panen 2

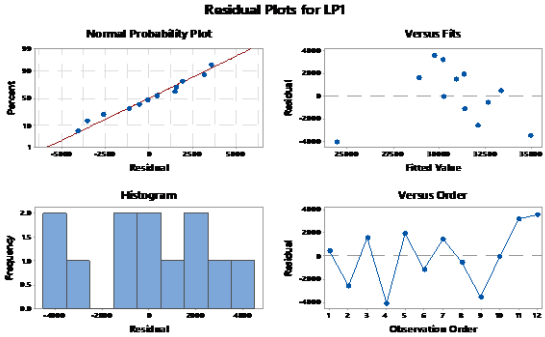



*Probability Plot* Luas Panen 3

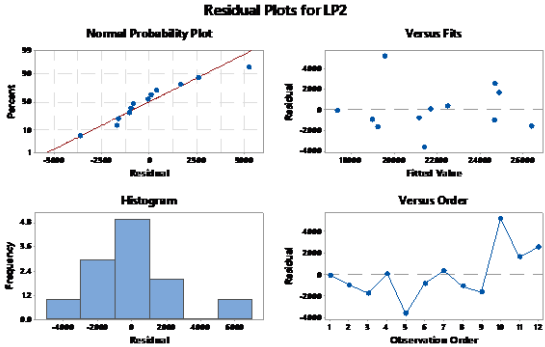


**Lampiran 5.** *Output* Residual *Plot* Hasil OLS

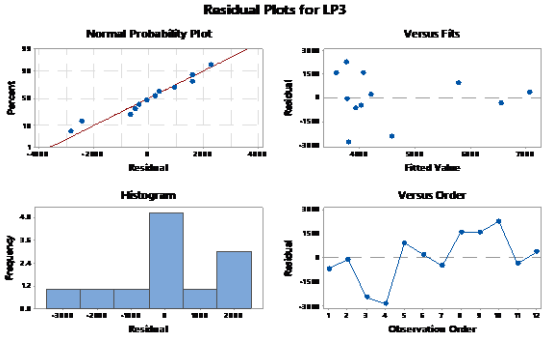
Residual *Plot* Luas Panen 1




Residual *Plot* Luas Panen 2




Residual *Plot* Luas Panen 3



**Lampiran 6.** Koefisien dan P-value Korelasi Pearson antara Luas Panen dengan Curah Hujan

| | LP1 | CH 1 | CH 2 | CH 3 |
|---|---|---|---|---|
| CH 1 | 0.091 | | | |
| | 0.778 | | | |
| CH 2 | -0.361 | 0.258 | | |
| | 0.249 | 0.419 | | |
| CH 3 | -0.178 | 0.600 | 0.683 | |
| | 0.581 | 0.039 | 0.014 | |
| CH 4 | 0.241 | 0.463 | 0.592 | 0.723 |
| | 0.450 | 0.130 | 0.043 | 0.008 |

*Cell Contents*
*Pearson correlation*
*P-Value*

| | LP2 | CH 5 | CH 6 | CH 7 |
|---|---|---|---|---|
| CH 5 | 0.715 | | | |
| | 0.009 | | | |
| CH 6 | 0.690 | 0.866 | | |
| | 0.013 | 0.000 | | |
| CH 7 | 0.615 | 0.865 | 0.948 | |
| | 0.033 | 0.000 | 0.000 | |
| CH 8 | 0.305 | 0.650 | 0.600 | 0.665 |
| | 0.336 | 0.022 | 0.039 | 0.018 |

*Cell Contents*
*Pearson correlation*
*P-Value*

| | LP3 | CH 9 | CH 10 | CH 11 |
|---|---|---|---|---|
| CH 9 | 0.376 | | | |
| | 0.228 | | | |
| CH 10 | 0.585 | 0.758 | | |
| | 0.046 | 0.004 | | |
| CH 11 | 0.408 | 0.236 | 0.729 | |
| | 0.188 | 0.459 | 0.007 | |
| CH 12 | 0.107 | 0.181 | 0.296 | 0.539 |
| | 0.741 | 0.573 | 0.350 | 0.070 |

*Cell Contents*
*Pearson correlation*
*P-Value*

**Lampiran 7.** Nilai *Variance Inflation Factor* (VIF)

Luas Panen Periode 1

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 35328 | 3744 | 9.44 | 0 | |
| CH 1 | 5.1 | 13 | 0.39 | 0.708 | 1.7 |
| CH 2 | -26.7 | 17.6 | -1.52 | 0.173 | 2.11 |
| CH 3 | -26.9 | 26.1 | -1.03 | 0.337 | 3.53 |
| CH 4 | 40.5 | 17.7 | 2.28 | 0.056 | 2.2 |

Luas Panen Periode 2

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 17226 | 2020 | 8.53 | 0 | |
| CH 5 | 30.3 | 23 | 1.31 | 0.23 | 4.62 |
| CH 6 | 28.3 | 33.5 | 0.84 | 0.426 | 11.19 |
| CH 7 | -18.8 | 34.5 | -0.54 | 0.603 | 11.91 |
| CH 8 | -21.1 | 29.2 | -0.72 | 0.493 | 1.94 |

Luas Panen Periode 3

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 4067 | 1219 | 3.34 | 0.012 | |
| CH 9 | -9.6 | 17.8 | -0.54 | 0.607 | 5.93 |
| CH 10 | 19.3 | 18.2 | 1.06 | 0.325 | 12.25 |
| CH 11 | -5.7 | 14.3 | -0.4 | 0.701 | 7.11 |
| CH 12 | 0.4 | 7.41 | 0.05 | 0.959 | 1.84 |

**Lampiran 8.** *Syntax* Regresi *Robust* dengan *R*

Estimasi M

```
 library(MASS)
 model.M=rlm(y~.,periode.x,psi=psi.bisquare,method="M",          scale.est=
"MAD")
 print(summary(model.M))
```

Estimasi MM

```
library(robustbase)
model.MM<-lmrob(y~., periode.x, method="MM")
print( summary(model.MM))
```

Estimasi S

```
library(robustbase)
model.MM<-lmrob(y~., periode.x, method="S")
print( summary(model.S))
```

**Lampiran 9A.** *Output* Regresi *Robust* M, MM, dan S Periode 1 dengan *R*

Estimasi M

```
Call: rlm(formula = y ~ ., data = periode.x, psi = psi.bisquar
e, scale.est = "MAD",
    method = "M")
Residuals:
    Min      1Q  Median      3Q     Max
-4798.9 -1459.3   138.5  1482.3  3421.8

Coefficients:
             Value      Std. Error t value
(Intercept) 34933.5678  4329.4157     8.0689
x1              5.8111    15.0414     0.3863
x2            -23.6644    20.3615    -1.1622
x3            -25.6931    30.1813    -0.8513
x4             36.6144    20.5216     1.7842

Residual standard error: 2392 on 7 degrees of freedom
```

Estimasi MM

```
lmrob(formula = y ~ ., data = periode.x, method = "MM")
 \--> method = "MM"
Residuals:
      Min        1Q    Median        3Q       Max
-12488.08   -946.21    -89.41    845.75   2114.72

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 30563.3987  5858.2652   5.217  0.00123 **
x1              5.0671     5.5108   0.919  0.38843
x2             12.0935    33.2194   0.364  0.72657
x3            -11.1849    16.4005  -0.682  0.51717
x4              0.6971    17.0937   0.041  0.96861
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 1376
Multiple R-squared:  0.1863, Adjusted R-squared:  -0.2787
Convergence in 17 IRWLS iterations
```

Estimasi S

```
Call:
lmrob(formula = y ~ ., data = periode.x, method = "S")
 \--> method = "S"
Residuals:
      Min        1Q    Median        3Q       Max
-16074.83   -299.33    -99.94    182.52   4209.72

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 25854.724    863.995  29.925  1.2e-08 ***
x1              7.750      1.897   4.086 0.004654 **
x2             30.835      4.318   7.141 0.000187 ***
x3             -1.628      4.901  -0.332 0.749491
x4            -12.902      3.931  -3.282 0.013448 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 1376
Multiple R-squared:  0.9479, Adjusted R-squared:  0.9181
Convergence in IRWLS iterations
```

**Lampiran 9B.** *Output* Regresi *Robust* M, MM, dan S Periode 2 dengan *R*

Estimasi M

```
Call: rlm(formula = y ~ ., data = periode.x, psi = psi.bisquar
e, scale.est = "MAD",
    method = "M")
Residuals:
   Min     1Q Median     3Q    Max
-997.6 -471.9  279.7  709.6 7153.8

Coefficients:
            Value      Std. Error t value
(Intercept) 15984.9216   772.1474    20.7019
x1             25.8968     8.8093     2.9397
x2             68.2604    12.8019     5.3320
x3            -70.5718    13.1904    -5.3503
x4             20.5958    11.1445     1.8481

Residual standard error: 953 on 7 degrees of freedom
```

Estimasi MM

```
lmrob(formula = y ~ ., data = periode.x, method = "MM")
 \--> method = "MM"
Residuals:
    Min      1Q  Median      3Q     Max
-3059.8  -987.5   179.1   830.9  5923.8
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 16622.90    1491.68  11.144 1.04e-05 ***
x1             29.06      12.27   2.368   0.0498 *
x2             36.19      68.49   0.528   0.6136
x3            -23.24      68.18  -0.341   0.7432
x4            -17.33      34.20  -0.507   0.6278
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 1812
Multiple R-squared:  0.7185, Adjusted R-squared:  0.5577
Convergence in 25 IRWLS iterations
```

Estimasi S

```
lmrob(formula = y ~ ., data = periode.x, method = "S")
 \--> method = "S"
Residuals:
   Min     1Q Median     3Q    Max
-900.8 -426.5  258.6  731.5 7192.5

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 15826.965    929.051  17.036 5.89e-07 ***
x1             27.208      9.816   2.772  0.02762 *
x2             69.060     14.608   4.727  0.00214 **
x3            -72.127     17.263  -4.178  0.00415 **
x4             20.526     14.932   1.375  0.21163
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 1812
Multiple R-squared:  0.9541, Adjusted R-squared:  0.9279
Convergence in IRWLS iterations
```

**Lampiran 9C.** *Output* Regresi *Robust* M, MM, dan S Periode 3 dengan *R*

Estimasi M

```
Call: rlm(formula = y ~ ., data = periode.x, psi =
psi.bisquare, scale.est = "MAD",
    method = "M")
Residuals:
     Min       1Q   Median       3Q      Max
-3007.72  -735.87   -30.39   950.66  2104.69

Coefficients:
            Value     Std. Error t value
(Intercept) 4251.6666 1516.6015     2.8034
x1           -10.9452   22.1318    -0.4945
x2            20.1150   22.6818     0.8868
x3            -6.6136   17.7798    -0.3720
x4             0.6824    9.2216     0.0740

Residual standard error: 1230 on 7 degrees of freedom
```

Estimasi MM

```
lmrob(formula = y ~ ., data = periode.x, method = "MM")
 \--> method = "MM"
Residuals:
     Min       1Q   Median       3Q      Max
-2878.32  -577.90    41.73  1057.84  2238.50

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 4100.8283  1197.2440   3.425   0.0111 *
x1            -9.8866     6.9038  -1.432   0.1952
x2            19.4948     7.0675   2.758   0.0282 *
x3            -5.8863     5.3483  -1.101   0.3075
x4             0.4521     5.8847   0.077   0.9409
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 2431
Multiple R-squared:  0.3443, Adjusted R-squared:  -0.03032
Convergence in 8 IRWLS iterations
```

Estimasi S

```
lmrob(formula = y ~ ., data = periode.x, method = "S")
 \--> method = "S"
Residuals:
    Min      1Q  Median      3Q     Max
-3522.1 -1423.4  -192.3   496.8  1551.3

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 4847.355   1804.739   2.686   0.0313 *
x1           -14.854     22.939  -0.648   0.5379
x2            22.528     23.301   0.967   0.3658
x3           -10.272     18.487  -0.556   0.5957
x4             2.188     10.695   0.205   0.8437
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 2431
Multiple R-squared:  0.5427, Adjusted R-squared:  0.2814
Convergence in IRWLS iterations
```

**Lampiran 10.** *Syntax Centering* dan *Rescaling* dengan *R*

```
av1=mean(x.1[,2])
av2=mean(x.1[,3])
av3=mean(x.1[,4])
av4=mean(x.1[,5])
x1.baru=as.matrix(0)
x2.baru=as.matrix(0)
x3.baru=as.matrix(0)
x4.baru=as.matrix(0)
```

**Lampiran 10.** *Syntax Centering* dan *Rescaling* dengan *R* (Lanjutan)

```
for (i in 1:n1){
   x1.baru[i]=x.1[i,2]-av1;
   x2.baru[i]=x.1[i,3]-av2;
   x3.baru[i]=x.1[i,4]-av3;
   x4.baru[i]=x.1[i,5]-av4
  }
  x.baru=cbind(x1.baru,x2.baru,x3.baru,x4.baru)
  jumx=matrix(0)
  for (i in 1:4){
   jumx[i]=sum((x.baru[,i])^2)
  }

x.r=cbind(x.baru[,1]/sqrt(jumx[1]),x.baru[,2]/sqrt(jumx[2]),x.baru[,3]/sqrt(jumx[3]
),x.baru[,4]/sqrt(jumx[4]))
  x.r

  avy=mean(Y.1)
  y.baru=Y.1-avy
  y.r=y.baru/sqrt(sum((y.baru)^2))
```

**Lampiran 11A.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik Periode 1

| Tahun | Y | CH1 | CH2 | CH3 | CH4 |
|---|---|---|---|---|---|
| 2006 | 0.2446 | -0.3709 | -0.3645 | -0.5929 | -0.2854 |
| 2007 | -0.1132 | -0.5685 | -0.3445 | -0.5631 | -0.3414 |
| 2008 | -0.0317 | -0.1483 | -0.3449 | -0.0640 | -0.4207 |
| 2009 | -0.8751 | -0.0172 | 0.4662 | 0.2170 | -0.1748 |
| 2010 | 0.2021 | 0.0993 | 0.1312 | 0.1887 | 0.2256 |
| 2011 | -0.0536 | -0.0948 | -0.1590 | 0.2050 | 0.0779 |
| 2012 | 0.1247 | 0.4708 | -0.2568 | -0.1653 | -0.3263 |
| 2013 | 0.1022 | 0.3086 | -0.1045 | 0.0741 | 0.0973 |
| 2014 | 0.0476 | 0.4040 | 0.0589 | 0.3254 | 0.5499 |
| 2015 | -0.0598 | 0.0741 | 0.3703 | 0.0572 | 0.2031 |
| 2016 | 0.2085 | -0.0722 | 0.3085 | 0.0882 | 0.1993 |
| 2017 | 0.2038 | -0.0849 | 0.2390 | 0.2298 | 0.1955 |

**Lampiran 11B.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik Periode 2

| Tahun | Y | CH5 | CH6 | CH7 | CH8 |
|---|---|---|---|---|---|
| 2006 | -0.3796 | -0.4712 | -0.2831 | -0.2282 | -0.1355 |
| 2007 | -0.3193 | -0.3936 | -0.1626 | -0.2282 | -0.1395 |
| 2008 | -0.3611 | -0.2911 | -0.2342 | -0.2282 | -0.1395 |
| 2009 | -0.0082 | -0.0520 | -0.1717 | -0.2282 | -0.1395 |
| 2010 | -0.3402 | 0.1755 | -0.0132 | 0.2053 | 0.2337 |
| 2011 | -0.1288 | -0.0476 | -0.2415 | -0.2282 | -0.1395 |
| 2012 | 0.0795 | -0.0137 | -0.1128 | -0.2282 | -0.1395 |
| 2013 | 0.1431 | 0.2803 | 0.1610 | 0.0895 | 0.1141 |
| 2014 | 0.2406 | 0.2628 | 0.4954 | 0.3661 | -0.1105 |
| 2015 | 0.2419 | -0.2018 | -0.2831 | -0.2282 | -0.1395 |
| 2016 | 0.3841 | 0.5007 | 0.5426 | 0.5480 | 0.8749 |
| 2017 | 0.4478 | 0.2517 | 0.3032 | 0.3886 | -0.1395 |

**Lampiran 11C.** Hasil *Centering* dan *Rescaling* Luas Panen dan Curah Hujan Kabupaten Gresik Periode 3

| Tahun | Y | CH9 | CH10 | CH11 | CH12 |
|---|---|---|---|---|---|
| 2006 | -0.2080 | -0.1161 | -0.2202 | -0.3270 | -0.3717 |
| 2007 | -0.1419 | -0.1161 | -0.0757 | 0.2760 | -0.0373 |
| 2008 | -0.3813 | -0.1161 | -0.0285 | 0.0734 | 0.3237 |
| 2009 | -0.5648 | -0.1161 | -0.2212 | -0.2603 | -0.3131 |
| 2010 | 0.3266 | 0.2718 | 0.2759 | 0.0598 | 0.1959 |
| 2011 | -0.0270 | -0.1161 | -0.0465 | 0.1879 | 0.2704 |
| 2012 | -0.1648 | -0.1161 | -0.1900 | -0.2348 | 0.0959 |
| 2013 | 0.0886 | -0.1161 | -0.1806 | 0.0396 | 0.4624 |
| 2014 | 0.1639 | -0.1161 | -0.2212 | -0.4092 | -0.5297 |
| 2015 | 0.2225 | -0.1161 | -0.2212 | -0.2345 | -0.1757 |
| 2016 | 0.2480 | 0.8895 | 0.6469 | 0.2121 | 0.1047 |
| 2017 | 0.4382 | -0.1161 | 0.4824 | 0.6171 | -0.0256 |

**Lampiran 12.** *Syntax* Regresi *Ridge*

Estimasi M

```
coef.m=coefficients(model.M) #nilai koefisien model
coef.m <-as.matrix(coef.m)
p1=nrow(coef.m)
```

```
Z1.m=(Y.1-(x.1%*%coef.m))
Z1.m=as.matrix(Z1.m)
Skuadrat1.m=(t(Z1.m)%*%Z1.m)/(n1-p1)
print(c("Skuadrat1.m=",Skuadrat1.m))
B1.m=t(coef.m)%*%coef.m #pembagi b transpose x b
konstanta.m=(p1*Skuadrat1.m)/B1.m #nilai konstanta ridge
print(c("konstanta.m=",konstanta.m))
konstanta.m=as.numeric(konstanta.m)
KI=diag(konstanta.m,np.1-1,np.1-1) #matriks identitas
BetaRidgeM=solve((t(x.r)%*%x.r)+(KI))%*%(t(x.r)%*%y.r)
colnames(BetaRidgeM)="BetaRidgeM"
print(BetaRidgeM)
BetaRidgeRobustM=0
for(i in 1:4){
  BetaRidgeRobustM[i]=(BetaRidgeM[i]*sqrt(sum((y.baru)^2)))/sqrt(jumx[i])
}
BetaRidgeRobustM=as.matrix(BetaRidgeRobustM)
colnames(BetaRidgeRobustM)="BetaRidgeRobustM"
print(BetaRidgeRobustM)
b0RR.m=avy-(av1*BetaRidgeRobustM[1])-(av2*BetaRidgeRobustM[2])-
                (av3*BetaRidgeRobustM[3])-(av4*BetaRidgeRobustM[4])
print(c("b0.M=",b0RR.m))
```

Estimasi MM

```
#Mencari nilai K
coef.mm=coefficients(model.MM) #nilai koefisien model
coef.mm <-as.matrix(coef.mm)
p1=nrow(coef.mm)
Z1.mm=(Y.1-(x.1%*%coef.mm))
Z1.mm=as.matrix(Z1.mm)
Skuadrat1.mm=(t(Z1.mm)%*%Z1.mm)/(n1-p1)
print(c("Skuadrat1.mm=",Skuadrat1.mm))
B1.mm=t(coef.mm)%*%coef.mm #pembagi b transpose x b
konstanta.mm=(p1*Skuadrat1.mm)/B1.mm #nilai konstanta ridge
print(c("konstanta.mm=", konstanta.mm))
konstanta.mm=as.numeric(konstanta.mm)
KI.mm=diag(konstanta.mm,np.1-1,np.1-1) #matriks identitas
BetaRidgeMM=solve((t(x.r)%*%x.r)+(KI.mm))%*%(t(x.r)%*%y.r)
colnames(BetaRidgeMM)="BetaRidgeMM"
print(BetaRidgeMM)
BetaRidgeRobustMM=0
```

```
for(i in 1:4){
 BetaRidgeRobustMM[i]=(BetaRidgeMM[i]*sqrt(sum((y.baru)^2)))/sqrt(jumx[i])
}
BetaRidgeRobustMM=as.matrix(BetaRidgeRobustMM)
colnames(BetaRidgeRobustMM)=("BetaRRMM")
print(BetaRidgeRobustMM)

b0RR.mm=avy-(av1*BetaRidgeRobustMM[1])-(av2*BetaRidgeRobustMM[2])-
                    (av3*BetaRidgeRobustMM[3])-(av4*BetaRidgeRobustMM[4])
print(c("b0.MM=",b0RR.mm))
```

Estimasi S

```
#Mencari nilai K
 coef.s=coefficients(model.S) #nilai koefisien model
 coef.s <-as.matrix(coef.s)
 p1=nrow(coef.s)
 Z1.s=(Y.1-(x.1%*%coef.s))
 Z1.s=as.matrix(Z1.s)
 Skuadrat1.s=(t(Z1.s)%*%Z1.s)/(n1-p1)
 print(c("Skuadrat1.s=", Skuadrat1.s))
 B1.s=t(coef.s)%*%coef.s #pembagi b transpose x b
 konstanta.s=(p1*Skuadrat1.s)/B1.s #nilai konstanta ridge
 print(c("konstanta.s=", konstanta.s))
 konstanta.s=as.numeric(konstanta.s)

 KI.s=diag(konstanta.s,np.1-1,np.1-1) #matriks identitas
 BetaRidgeS=solve((t(x.r)%*%x.r)+(KI.s))%*%(t(x.r)%*%y.r)
 colnames(BetaRidgeS)="BetaRidgeS"
 print(BetaRidgeS)
 BetaRidgeRobustS=0
 for(i in 1:4){
  BetaRidgeRobustS[i]=(BetaRidgeS[i]*sqrt(sum((y.baru)^2)))/sqrt(jumx[i])
 }
 BetaRidgeRobustS=as.matrix(BetaRidgeRobustS)
 colnames(BetaRidgeRobustS)="BetaRidgeRobustS"
 print(BetaRidgeRobustS)
 b0RR.S=avy-(av1*BetaRidgeRobustS[1])-(av2*BetaRidgeRobustS[2])-
                    (av3*BetaRidgeRobustS[3])-(av4*BetaRidgeRobustS[4])
 print(c("b0.S=",b0RR.S))
```

**Lampiran 13A.** Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 1

Estimasi M

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 33957,190 | 5,060 | -24,665 | -19,484 | 34,589 | 32474,925 | 33770 | 1677218,596 |
| 2 | 38875,940 | -1,050 | -26,755 | -33,374 | 40,446 | 35328,690 | 29467 | 34359404,538 |
| 3 | 34446,710 | 6,014 | -21,436 | -30,551 | 41,444 | 32943,168 | 30447 | 6230854,997 |
| 4 | 30956,910 | 3,831 | 6,331 | -7,706 | 2,095 | 31641,675 | 20303,21 | 128560784,430 |
| 5 | 34435,990 | 4,192 | -21,868 | -18,562 | 27,467 | 32630,074 | 33259 | 395548,035 |
| 6 | 34932,360 | 1,993 | -29,342 | -15,724 | 35,723 | 32955,956 | 30184 | 7683739,910 |
| 7 | 34628,480 | -7,782 | -26,010 | -16,093 | 44,501 | 32633,375 | 32328 | 93253,664 |
| 8 | 34844,440 | 5,183 | -25,318 | -22,703 | 35,613 | 33029,113 | 32058 | 943059,931 |
| 9 | 33993,230 | 9,049 | -34,797 | -26,045 | 55,400 | 32950,230 | 31401,41 | 2398842,433 |
| 10 | 34863,980 | 4,523 | -25,178 | -22,387 | 35,562 | 33011,840 | 30109 | 8426480,426 |
| 11 | 35071,250 | 6,177 | -31,204 | -19,231 | 33,455 | 32777,789 | 33336,6 | 312269,932 |
| 12 | 35476,310 | 9,075 | -26,456 | -29,681 | 35,453 | 33217,444 | 33279,6 | 3863,409 |
| | | | | | | | **RMSE** | 3990,461 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 32024,150 | 4,601 | -16,577 | -8,844 | 19,957 | 31342,366 | 33770 | 5893408,855 |
| 2 | 37802,400 | -0,502 | -23,822 | -27,391 | 33,187 | 34687,809 | 29467 | 27256845,349 |
| 3 | 34447,280 | 6,017 | -21,435 | -30,566 | 41,456 | 32943,715 | 30447 | 6233587,795 |
| 4 | 30957,360 | 3,841 | 6,345 | -7,735 | 2,101 | 31642,628 | 20303,21 | 128582409,731 |
| 5 | 35087,420 | 4,853 | -24,783 | -23,905 | 34,310 | 33040,063 | 33259 | 47933,421 |
| 6 | 34957,140 | 1,975 | -29,515 | -15,833 | 35,995 | 32969,320 | 30184 | 7758009,615 |
| 7 | 34628,260 | -7,780 | -26,008 | -16,092 | 44,497 | 32633,279 | 32328 | 93195,371 |
| 8 | 34847,770 | 5,188 | -25,332 | -22,732 | 35,649 | 33031,284 | 32058 | 947281,225 |
| 9 | 33993,230 | 9,049 | -34,797 | -26,045 | 55,400 | 32950,230 | 31401,41 | 2398844,811 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 10 | 33438,700 | 3,254 | -18,013 | -11,249 | 20,187 | 32142,264 | 30109 | 4134163,636 |
| 11 | 35075,240 | 6,181 | -31,226 | -19,261 | 33,497 | 32780,166 | 33336,6 | 309618,307 |
| 12 | 35476,840 | 9,076 | -26,458 | -29,687 | 35,458 | 33217,799 | 33279,6 | 3819,388 |
| | | | | | | | **RMSE** | 3912,151 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 32029,160 | 4,604 | -16,604 | -8,869 | 19,998 | 31345,194 | 33770 | 5879685,444 |
| 2 | 37232,040 | -0,215 | -22,090 | -24,394 | 29,371 | 34353,375 | 29467 | 23876657,624 |
| 3 | 33596,980 | 3,688 | -19,377 | -15,588 | 26,283 | 32277,452 | 30447 | 3350553,254 |
| 4 | 30956,980 | 3,833 | 6,333 | -7,710 | 2,096 | 31641,822 | 20303,21 | 128564122,486 |
| 5 | 33989,890 | 3,770 | -19,512 | -15,332 | 22,877 | 32362,381 | 33259 | 803926,377 |
| 6 | 33840,470 | 2,482 | -21,454 | -10,988 | 23,916 | 32372,674 | 30184 | 4790293,434 |
| 7 | 33419,300 | -1,583 | -18,431 | -11,315 | 26,648 | 32054,602 | 32328 | 74746,587 |
| 8 | 33697,080 | 3,752 | -19,165 | -13,892 | 23,295 | 32315,488 | 32058 | 66299,851 |
| 9 | 33159,120 | 6,305 | -24,459 | -16,897 | 36,127 | 32304,758 | 31401,41 | 816037,640 |
| 10 | 33426,820 | 3,244 | -17,936 | -11,175 | 20,063 | 32135,697 | 30109 | 4107500,025 |
| 11 | 33835,860 | 4,942 | -23,254 | -11,463 | 20,805 | 32093,377 | 33336,6 | 1545602,321 |
| 12 | 34423,820 | 6,895 | -22,344 | -19,777 | 25,211 | 32551,561 | 33279,6 | 530040,914 |
| | | | | | | | **RMSE** | 3812,321 |

**Lampiran 13B.** Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 2

Estimasi M

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 19359.420 | 12.986 | 10.167 | 4.523 | -7.687 | 19414.280 | 17290 | 4512567.235 |
| 2 | 19055.230 | 14.945 | 12.383 | 3.057 | -10.822 | 19117.279 | 18017 | 1210613.401 |
| 3 | 19132.560 | 15.447 | 11.475 | 3.034 | -11.079 | 19196.752 | 17513 | 2835019.157 |
| 4 | 18717.950 | 15.243 | 11.664 | 4.713 | -9.793 | 18781.921 | 21768 | 8916668.327 |
| 5 | 18250.820 | 28.892 | -2.915 | 12.458 | -11.944 | 18375.710 | 17765 | 372966.606 |
| 6 | 18521.620 | 19.126 | 12.602 | 1.890 | -13.846 | 18601.029 | 20314 | 2934268.535 |
| 7 | 18361.910 | 17.589 | 12.826 | 4.212 | -13.546 | 18434.488 | 22826 | 19285381.156 |
| 8 | 18941.680 | 14.624 | 10.347 | 4.521 | -7.813 | 19003.921 | 23593 | 21059650.645 |
| 9 | 18385.160 | 18.266 | 18.005 | 1.615 | -18.851 | 18457.901 | 24768.11 | 39818733.591 |
| 10 | 17519.870 | 19.211 | 18.732 | 1.505 | -14.807 | 17599.138 | 24784 | 51622247.530 |
| 11 | 17822.170 | 33.424 | 4.627 | 3.366 | -101.552 | 17918.394 | 26499.2 | 73630225.667 |
| 12 | 18339.710 | 19.358 | 14.201 | -7.219 | -1.438 | 18426.996 | 27266.4 | 78135059.159 |
| | | | | | | | **RMSE** | 5032,982 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 19307.280 | 13.348 | 10.396 | 4.405 | -8.235 | 19363.486 | 17290 | 4299346.085 |
| 2 | 18754.700 | 17.343 | 14.456 | 1.020 | -14.111 | 18825.851 | 18017 | 654240.089 |
| 3 | 18765.640 | 18.623 | 13.591 | 0.579 | -14.986 | 18842.135 | 17513 | 1766599.490 |
| 4 | 18336.930 | 18.792 | 13.663 | 2.424 | -14.416 | 18414.504 | 21768 | 11245937.171 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 5 | 18167.450 | 30.355 | -4.785 | 13.555 | -12.896 | 18298.475 | 17765 | 284595.050 |
| 6 | 18393.160 | 20.465 | 13.195 | 0.886 | -15.218 | 18477.904 | 20314 | 3371249.848 |
| 7 | 18322.860 | 17.978 | 13.042 | 3.929 | -14.034 | 18396.943 | 22826 | 19616548.435 |
| 8 | 18317.710 | 20.532 | 13.273 | 1.185 | -14.874 | 18402.949 | 23593 | 26936627.846 |
| 9 | 19328.480 | 11.024 | 10.248 | 4.882 | -3.858 | 19376.516 | 24768.11 | 29069286.742 |
| 10 | 17513.300 | 19.262 | 18.796 | 1.441 | -14.875 | 17592.762 | 24784 | 51713899.557 |
| 11 | 17802.620 | 33.732 | 4.466 | 3.357 | -102.261 | 17899.859 | 26499.2 | 73948659.238 |
| 12 | 18334.270 | 19.411 | 14.255 | -7.310 | -1.446 | 18421.794 | 27266.4 | 78227056.527 |
| | | | | | | | **RMSE** | 5009,441 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 19330.070 | 13.189 | 10.295 | 4.458 | -7.995 | 19385.684 | 17290 | 4391892.406 |
| 2 | 19048.660 | 14.996 | 12.424 | 3.023 | -10.896 | 19110.897 | 18017 | 1196610.223 |
| 3 | 19162.260 | 15.203 | 11.326 | 3.175 | -10.745 | 19225.525 | 17513 | 2932742.359 |
| 4 | 18725.500 | 15.178 | 11.628 | 4.743 | -9.699 | 18789.228 | 21768 | 8873084.687 |
| 5 | 18734.390 | 21.929 | 3.109 | 9.555 | -6.877 | 18830.383 | 17765 | 1135041.371 |
| 6 | 18850.350 | 15.945 | 11.221 | 3.769 | -10.077 | 18917.356 | 20314 | 1950613.642 |
| 7 | 18687.870 | 14.582 | 11.228 | 5.814 | -9.268 | 18749.121 | 22826 | 16620940.873 |
| 8 | 18786.950 | 15.977 | 11.010 | 4.005 | -9.623 | 18854.354 | 23593 | 22454769.978 |
| 9 | 19242.460 | 11.620 | 10.778 | 4.831 | -4.923 | 19292.624 | 24768.11 | 29980945.450 |
| 10 | 17693.700 | 17.879 | 17.182 | 2.985 | -12.936 | 17767.937 | 24784 | 49225146.725 |
| 11 | 18328.100 | 25.874 | 7.619 | 4.110 | -83.061 | 18400.141 | 26499.2 | 65594752.546 |
| 12 | 18551.230 | 17.290 | 12.331 | -4.113 | -0.868 | 18629.424 | 27266.4 | 74597360.372 |
| | | | | | | | **RMSE** | 4821,427 |

**Lampiran 13C.** Hasil LOOCV Regresi *Ridge Robust* Luas Panen Padi Periode 2

Estimasi M

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4143.680 | 2.220 | 4.073 | 2.067 | -1.002 | 4150.622 | 3245 | -905.622 |
| 2 | 4042.478 | 2.051 | 3.179 | 2.643 | -0.398 | 4099.891 | 3672 | -427.891 |
| 3 | 3928.861 | 1.535 | 3.898 | 2.219 | 1.053 | 4054.894 | 2126 | -1928.894 |
| 4 | 4718.775 | 2.039 | 4.261 | 1.638 | -2.560 | 4626.173 | 941 | -3685.173 |
| 5 | 3933.008 | 1.635 | 3.201 | 2.177 | -0.697 | 3959.913 | 6697 | 2737.087 |
| 6 | 3981.828 | 2.209 | 3.414 | 2.115 | -0.204 | 4034.469 | 4414 | 379.531 |
| 7 | 4033.728 | 2.150 | 3.500 | 1.903 | -0.141 | 4083.618 | 3524 | -559.618 |
| 8 | 4006.195 | 2.477 | 4.576 | 2.257 | -1.664 | 3982.058 | 5160 | 1177.942 |
| 9 | 4101.824 | 0.901 | 1.108 | 1.009 | 0.482 | 4159.057 | 5646.44 | 1487.383 |
| 10 | 4185.235 | 0.825 | 1.018 | 0.747 | 0.195 | 4218.648 | 6025 | 1806.352 |
| 11 | 3931.298 | 8.229 | 4.694 | 1.729 | -0.497 | 3956.619 | 6189.39 | 2232.771 |
| 12 | 3976.744 | 3.511 | 2.427 | -0.015 | 0.419 | 4000.589 | 7417.6 | 3417.011 |
| | | | | | | | **RMSE** | 2036,689 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4163.726 | 2.182 | 3.621 | 1.947 | -0.798 | 4178.337 | 3245 | 871117.648 |
| 2 | 4050.089 | 2.032 | 3.109 | 2.588 | -0.363 | 4107.805 | 3672 | 189925.945 |
| 3 | 3949.533 | 1.568 | 3.719 | 2.182 | 1.044 | 4073.914 | 2126 | 3794368.953 |
| 4 | 4719.816 | 2.035 | 4.286 | 1.640 | -2.578 | 4626.312 | 941 | 13581525.518 |
| 5 | 3940.896 | 1.624 | 3.081 | 2.108 | -0.635 | 3969.155 | 6697 | 7441140.345 |
| 6 | 3990.469 | 2.185 | 3.318 | 2.075 | -0.171 | 4043.714 | 4414 | 137112.062 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 7 | 4042.344 | 2.131 | 3.407 | 1.872 | -0.114 | 4092.819 | 3524 | 323554.575 |
| 8 | 4006.451 | 2.471 | 4.475 | 2.229 | -1.589 | 3985.617 | 5160 | 1379175.545 |
| 9 | 3571.468 | 1.974 | 2.763 | 2.445 | 1.046 | 3703.378 | 5646.44 | 3775491.308 |
| 10 | 3644.684 | 2.315 | 3.777 | 2.437 | 0.098 | 3723.861 | 6025 | 5295239.947 |
| 11 | 3939.302 | 7.961 | 4.517 | 1.692 | -0.441 | 3966.589 | 6189.39 | 4940844.397 |
| 12 | 3982.105 | 3.381 | 2.351 | 0.037 | 0.405 | 4006.697 | 7417.6 | 11634257.730 |
| | | | | | | | **RMSE** | 2108,786 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4148.037 | 2.217 | 3.955 | 2.039 | -0.949 | 4157.107 | 3245 | 831938.417 |
| 2 | 4016.382 | 2.107 | 3.450 | 2.850 | -0.543 | 4071.925 | 3672 | 159939.966 |
| 3 | 3894.848 | 1.442 | 4.239 | 2.264 | 1.066 | 4023.060 | 2126 | 3598836.450 |
| 4 | 4705.527 | 2.081 | 3.926 | 1.594 | -2.313 | 4625.278 | 941 | 13573906.560 |
| 5 | 3921.257 | 1.642 | 3.406 | 2.290 | -0.807 | 3945.475 | 6697 | 7570887.838 |
| 6 | 3973.158 | 2.231 | 3.517 | 2.156 | -0.240 | 4025.035 | 4414 | 151293.961 |
| 7 | 4017.485 | 2.181 | 3.692 | 1.961 | -0.200 | 4065.903 | 3524 | 293658.683 |
| 8 | 4008.149 | 2.474 | 4.964 | 2.350 | -1.953 | 3970.822 | 5160 | 1414145.457 |
| 9 | 4102.186 | 0.900 | 1.106 | 1.008 | 0.482 | 4159.367 | 5646.44 | 2211386.073 |
| 10 | 4187.216 | 0.819 | 1.010 | 0.742 | 0.194 | 4220.399 | 6025 | 3256583.073 |
| 11 | 3919.677 | 8.667 | 4.994 | 1.784 | -0.597 | 3941.192 | 6189.39 | 5054393.826 |
| 12 | 3966.462 | 3.806 | 2.594 | -0.150 | 0.454 | 3988.328 | 7417.6 | 11759905.893 |
| | | | | | | | **RMSE** | 2038,727 |

**Lampiran 14A.** Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 1

Estimasi M

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 34080.879 | 5.882 | -24.663 | -21.789 | 36.444 | 32626.676 | 33770 | 1307190.450 |
| 2 | 38869.244 | -0.417 | -25.713 | -34.965 | 40.768 | 35396.685 | 29467 | 35161163.650 |
| 3 | 34641.139 | 8.706 | -20.265 | -40.666 | 49.085 | 33249.733 | 30447 | 7855312.900 |
| 4 | 25862.213 | 7.741 | 30.847 | -1.668 | -12.915 | 29562.366 | 20303.21 | 85731961.098 |
| 5 | 31104.444 | 4.937 | 8.145 | -11.242 | 1.606 | 31732.417 | 33259 | 2330454.162 |
| 6 | 34689.652 | 2.096 | -27.945 | -14.971 | 35.484 | 32974.822 | 30184 | 7788687.973 |
| 7 | 35006.774 | -10.585 | -28.132 | -17.220 | 50.597 | 32825.791 | 32328 | 247796.133 |
| 8 | 35074.326 | 6.820 | -25.984 | -26.656 | 39.050 | 33271.986 | 32058 | 1473762.910 |
| 9 | 34236.997 | 10.059 | -37.726 | -28.969 | 60.940 | 33141.450 | 31401.41 | 3027738.980 |
| 10 | 35018.977 | 5.934 | -23.613 | -27.310 | 38.254 | 33268.336 | 30109 | 9981406.721 |
| 11 | 35543.102 | 9.678 | -31.978 | -25.295 | 34.521 | 33025.263 | 33336.6 | 96930.944 |
| 12 | 35804.174 | 10.434 | -27.680 | -33.199 | 38.120 | 33422.093 | 33279.6 | 20304.115 |
| | | | | | | | **RMSE** | 3594,240 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 25845.029 | 7.760 | 30.811 | -1.564 | -12.888 | 29556.970 | 33770 | 17749622.484 |
| 2 | 33895.260 | 0.821 | 10.286 | -21.006 | 5.707 | 33576.548 | 29467 | 16888381.354 |
| 3 | 34667.327 | 8.530 | -20.228 | -40.571 | 48.970 | 33257.436 | 30447 | 7898548.633 |
| 4 | 25957.984 | 7.751 | 30.693 | -2.085 | -12.592 | 29623.447 | 20303.21 | 86866810.819 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 5 | 35423.548 | 6.028 | -25.315 | -27.747 | 36.935 | 33281.047 | 33259 | 486.049 |
| 6 | 35241.966 | 1.910 | -32.089 | -17.026 | 40.054 | 33159.697 | 30184 | 8854774.063 |
| 7 | 35003.379 | -10.370 | -27.941 | -17.355 | 50.359 | 32837.951 | 32328 | 260049.804 |
| 8 | 35173.965 | 6.581 | -26.566 | -26.920 | 39.794 | 33305.003 | 32058 | 1555015.633 |
| 9 | 34238.143 | 10.132 | -37.774 | -28.997 | 60.926 | 33140.977 | 31401.41 | 3026094.149 |
| 10 | 25957.912 | 7.751 | 30.693 | -2.084 | -12.592 | 29623.404 | 30109 | 235803.294 |
| 11 | 35431.626 | 7.463 | -32.552 | -22.969 | 36.536 | 33004.169 | 33336.6 | 110510.376 |
| 12 | 35821.197 | 10.322 | -27.624 | -33.424 | 38.459 | 33439.384 | 33279.6 | 25530.841 |
| | | | | | | | **RMSE** | 3457,740 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 25871.105 | 7.731 | 30.849 | -1.706 | -12.932 | 29564.280 | 33770 | 17688080.551 |
| 2 | 32907.170 | 2.710 | 25.562 | -28.096 | 1.503 | 33677.775 | 29467 | 17730625.646 |
| 3 | 29349.242 | 7.187 | 27.293 | -17.917 | -3.234 | 31616.763 | 30447 | 1368346.238 |
| 4 | 25871.107 | 7.731 | 30.849 | -1.706 | -12.932 | 29564.280 | 20303.21 | 85767420.325 |
| 5 | 29498.643 | 6.883 | 25.543 | -16.381 | -4.152 | 31575.670 | 33259 | 2833600.710 |
| 6 | 29391.834 | 6.672 | 26.033 | -15.894 | -3.904 | 31585.582 | 30184 | 1964431.883 |
| 7 | 29091.108 | 11.789 | 30.588 | -19.524 | -9.069 | 31654.182 | 32328 | 454030.068 |
| 8 | 29326.660 | 6.798 | 26.965 | -16.415 | -4.248 | 31569.811 | 32058 | 238328.091 |
| 9 | 29516.993 | 8.008 | 21.442 | -18.565 | 3.613 | 31697.560 | 31401.41 | 87705.044 |
| 10 | 25871.106 | 7.731 | 30.849 | -1.706 | -12.932 | 29564.280 | 30109 | 296719.795 |
| 11 | 30011.814 | 5.918 | 33.763 | -23.148 | -4.019 | 32318.099 | 33336.6 | 1037343.685 |
| 12 | 30459.490 | 7.145 | 19.372 | -18.192 | -0.289 | 31915.169 | 33279.6 | 1861670.703 |
| | | | | | | | **RMSE** | 3308,176 |

**Lampiran 14B.** Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 2

Estimasi M

| I | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 14405.641 | 40.891 | 72.827 | -83.603 | 18.406 | 14601.820 | 33770 | 7226309.333 |
| 2 | 16878.543 | 17.519 | 75.454 | -74.784 | 25.657 | 16972.365 | 29467 | 1091261.369 |
| 3 | 17182.468 | 15.959 | 75.272 | -74.552 | 26.248 | 17269.519 | 30447 | 59283.051 |
| 4 | 15993.708 | 24.324 | 67.840 | -68.056 | 20.433 | 16115.622 | 20303.21 | 31949376.953 |
| 5 | 17150.103 | 38.269 | -16.099 | 24.745 | -19.939 | 17313.263 | 33259 | 204066.364 |
| 6 | 16294.840 | 27.015 | 45.555 | -31.342 | -12.054 | 16411.127 | 30184 | 15232421.253 |
| 7 | 16988.298 | 29.544 | 7.979 | 7.618 | -19.161 | 17112.185 | 32328 | 32647679.270 |
| 8 | 14405.611 | 40.892 | 72.828 | -83.604 | 18.406 | 14601.792 | 32058 | 80841812.821 |
| 9 | 16083.092 | 19.542 | 79.478 | -37.619 | -43.119 | 16148.291 | 31401.41 | 74301279.955 |
| 10 | 15966.503 | 26.042 | 68.462 | -70.873 | 20.629 | 16096.342 | 30109 | 75475393.167 |
| 11 | 16579.876 | 45.251 | 2.736 | 0.928 | -122.919 | 16718.164 | 33336.6 | 95668674.517 |
| 12 | 15966.577 | 26.042 | 68.462 | -70.873 | 20.630 | 16096.414 | 33279.6 | 124768590.893 |
| | | | | | | | **RMSE** | 6704,887 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 14893.773 | 34.832 | 70.259 | -77.273 | 18.846 | 15062.624 | 33770 | 4961203.727 |
| 2 | 16892.728 | 23.933 | 45.863 | -30.615 | -13.166 | 16994.381 | 29467 | 1045750.497 |
| 3 | 17252.860 | 27.707 | 30.296 | -18.430 | -17.779 | 17369.147 | 30447 | 20693.790 |
| 4 | 16830.151 | 28.445 | 32.089 | -18.699 | -18.676 | 16949.303 | 20303.21 | 23219837.583 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 5 | 17669.901 | 42.170 | -33.868 | 37.143 | -20.912 | 17850.271 | 33259 | 7271.122 |
| 6 | 16872.195 | 32.155 | 27.967 | -18.756 | -19.745 | 17007.641 | 30184 | 10932011.305 |
| 7 | 16950.061 | 27.866 | 28.087 | -14.116 | -19.186 | 17066.300 | 32328 | 33174148.773 |
| 8 | 16543.745 | 33.119 | 34.831 | -24.908 | -18.385 | 16684.326 | 32058 | 47729773.730 |
| 9 | 16205.312 | 35.633 | 41.740 | -99.842 | 94.469 | 16419.399 | 31401.41 | 69700975.829 |
| 10 | 16004.217 | 25.750 | 67.981 | -70.180 | 20.534 | 16132.671 | 30109 | 74845497.264 |
| 11 | 16660.890 | 46.143 | -0.025 | 2.556 | -126.39 | 16801.324 | 33336.6 | 94048794.479 |
| 12 | 16004.217 | 25.750 | 67.981 | -70.180 | 20.534 | 16132.671 | 33279.6 | 123959926.333 |
| | | | | | | | **RMSE** | 6348,529 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 14664.541 | 37.598 | 71.419 | -80.070 | 18.612 | 14845.848 | 33770 | 5973878.980 |
| 2 | 16947.027 | 17.052 | 75.670 | -74.903 | 25.899 | 17038.858 | 29467 | 956762.651 |
| 3 | 16555.898 | 21.815 | 69.948 | -71.229 | 22.587 | 16667.579 | 30447 | 714736.011 |
| 4 | 15943.510 | 24.701 | 68.221 | -68.666 | 20.464 | 16067.157 | 20303.21 | 32499616.583 |
| 5 | 17155.550 | 35.720 | -15.362 | 26.869 | -19.366 | 17307.423 | 33259 | 209376.863 |
| 6 | 15948.374 | 24.259 | 70.530 | -71.040 | 21.460 | 16070.554 | 30184 | 18006833.267 |
| 7 | 15979.482 | 25.444 | 67.637 | -69.135 | 20.301 | 16106.418 | 32328 | 45152787.180 |
| 8 | 15398.741 | 32.231 | 70.316 | -76.312 | 19.802 | 15556.283 | 32058 | 64588816.329 |
| 9 | 15659.907 | 35.121 | 52.701 | -96.271 | 75.142 | 15861.037 | 31401.41 | 79335953.133 |
| 10 | 14664.541 | 37.598 | 71.419 | -80.070 | 18.612 | 14845.848 | 30109 | 98766865.096 |
| 11 | 15809.065 | 34.606 | 49.464 | -52.614 | -27.559 | 15951.367 | 33336.6 | 111256774.752 |
| 12 | 14664.541 | 37.598 | 71.419 | -80.070 | 18.612 | 14845.848 | 33279.6 | 154270111.885 |
| | | | | | | | **RMSE** | 7139,868 |

**Lampiran 14C.** Hasil LOOCV Regresi *Robust* Luas Panen Padi Periode 3

Estimasi M

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 5291.614 | -11.112 | 19.592 | -7.919 | -1.816 | 4962.846 | 3245 | 2950994.762 |
| 2 | 4202.476 | -7.507 | 16.029 | -2.964 | -0.664 | 4083.618 | 3672 | 169429.083 |
| 3 | 4185.657 | -16.077 | 24.088 | -10.434 | 4.438 | 4132.863 | 2126 | 4027499.665 |
| 4 | 4951.126 | -9.334 | 17.871 | -6.006 | -1.661 | 4687.134 | 941 | 14033517.591 |
| 5 | 4242.015 | -7.865 | 15.557 | -2.849 | -1.479 | 4080.895 | 6697 | 6844005.421 |
| 6 | 4256.371 | -10.918 | 20.282 | -6.709 | 0.518 | 4093.986 | 4414 | 102409.040 |
| 7 | 4339.362 | -13.019 | 22.082 | -9.130 | 2.149 | 4196.830 | 3524 | 452699.828 |
| 8 | 4616.657 | -11.534 | 23.189 | -6.605 | -4.118 | 4199.839 | 5160 | 921908.350 |
| 9 | 1647.242 | -14.489 | 23.205 | -5.772 | 10.710 | 2082.292 | 5646.44 | 12703148.159 |
| 10 | 1652.444 | -14.490 | 23.204 | -5.779 | 10.696 | 2086.494 | 6025 | 15511831.253 |
| 11 | 4242.015 | 4.359 | 15.557 | -2.849 | -1.479 | 4080.895 | 6189.39 | 4445751.205 |
| 12 | 4131.023 | 11.190 | -0.261 | -3.444 | 2.068 | 4144.161 | 7417.6 | 10715404.842 |
| | | | | | | | **RMSE** | 2464,390 |

Estimasi MM

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4453.316 | -10.130 | 19.632 | -6.530 | -0.421 | 4243.616 | 3245 | 997232.952 |
| 2 | 4087.188 | -7.867 | 16.955 | -3.636 | -0.325 | 3967.676 | 3672 | 87424.191 |
| 3 | 3950.178 | -16.695 | 24.961 | -10.733 | 5.240 | 3933.585 | 2126 | 3267364.017 |
| 4 | 4969.718 | -7.898 | 16.571 | -5.013 | -2.389 | 4694.071 | 941 | 14085544.724 |

Estimasi MM (lanjutan)

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 5 | 4060.177 | -7.162 | 15.026 | -2.306 | -1.285 | 3925.842 | 6697 | 7679316.567 |
| 6 | 4115.665 | -9.998 | 19.876 | -6.201 | 0.315 | 3956.844 | 4414 | 208991.920 |
| 7 | 4194.473 | -11.582 | 21.074 | -7.862 | 1.506 | 4053.294 | 3524 | 280151.732 |
| 8 | 4502.821 | -11.665 | 23.538 | -6.729 | -3.771 | 4101.830 | 5160 | 1119724.385 |
| 9 | 2934.780 | -7.954 | 17.230 | -2.873 | 3.680 | 3061.365 | 5646.44 | 6682610.202 |
| 10 | 3556.804 | -10.906 | 21.291 | -5.755 | 1.440 | 3474.531 | 6025 | 6504890.811 |
| 11 | 4060.177 | 5.702 | 15.026 | -2.306 | -1.285 | 3925.842 | 6189.39 | 5123649.472 |
| 12 | 3944.732 | 10.657 | 0.774 | -3.764 | 2.541 | 3975.134 | 7417.6 | 11850572.979 |
| | | | | | | | **RMSE** | 2196,351 |

Estimasi S

| i | b0 | b1 | b2 | b3 | b4 | ytopi | y | Residual$^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 5680.280 | -16.731 | 23.777 | -13.131 | 1.438 | 5380.054 | 3245 | 4558454.380 |
| 2 | 4926.441 | -2.639 | 8.409 | 2.300 | -3.928 | 4778.724 | 3672 | 1224837.189 |
| 3 | 5027.099 | -16.783 | 24.128 | -12.279 | 3.098 | 4844.869 | 2126 | 7392246.617 |
| 4 | 5027.097 | -16.783 | 24.128 | -12.279 | 3.098 | 4844.867 | 941 | 15240179.319 |
| 5 | 4651.973 | -8.861 | 16.160 | -3.967 | -1.878 | 4435.724 | 6697 | 5113370.465 |
| 6 | 4537.926 | -12.715 | 20.920 | -8.070 | 1.462 | 4388.066 | 4414 | 672.598 |
| 7 | 5351.318 | -20.452 | 27.275 | -16.537 | 5.165 | 5159.278 | 3524 | 2674133.507 |
| 8 | 5277.833 | -11.045 | 21.675 | -6.620 | -5.705 | 4771.524 | 5160 | 150913.521 |
| 9 | 1646.140 | -14.448 | 23.150 | -5.729 | 10.702 | 2081.971 | 5646.44 | 12705438.596 |
| 10 | 1646.140 | -14.448 | 23.150 | -5.729 | 10.702 | 2081.971 | 6025 | 15547476.855 |
| 11 | 4651.973 | 2.114 | 16.160 | -3.967 | -1.878 | 4435.724 | 6189.39 | 3075345.461 |
| 12 | 4785.425 | 16.282 | -6.171 | -2.642 | 0.542 | 4734.923 | 7417.6 | 7196757.840 |
| | | | | | | | **RMSE** | 2497,996 |

**Lampiran 15.** Surat Pernyataan Data

## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FSAD ITS

Nama : Anggun Nurfitriani Handoko
NRP : 06211640000106

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ ~~Thesis~~ ini merupakan data sekunder yang diambil dari ~~penelitian~~ / ~~buku~~/ ~~Tugas Akhir~~/ ~~Thesis~~/ publikasi lainnya yaitu:

Sumber : Situs Badan Pusat Statistik Kabupaten Gresik dan Situs Resmi Pemerintahan Kabupaten Gresik

Keterangan : Data total luas lahan padi dan rata-rata curah hujan setiap bulan dalam publikasi Kabupaten Gresik dalam Angka tahun terbit 2006 sampai dengan 2019.

Surat pernyataan ini dibuat dengan sebenarnya, apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang telah ditetapkan.

Mengetahui,
Pembimbing Tugas Akhir

(Dr. Sutikno, S.Si M.Si)
NIP. 19710313 199702 1 001

Surabaya, 22 Juli 2020

(Anggun Nurfitriani Handoko)
NRP. 06211640000106

# BIODATA PENULIS

Anggun Nurfitriani Handoko atau biasa dipanggil Anggun adalah anak pertama dari dua bersaudara dan dilahirkan di Gresik pada tanggal 2 Oktober 1998. Pendidikan yang telah ditempuh oleh penulis adalah SD Muhammadiyah, Gresik (2004-2010), SMP Muhammadiyah 12, Gresik (2010-2013), dan SMAN 1 Gresik (2013-2016). Selanjutnya penullis menempuh pendidikan formal di Institut Teknologi Sepuluh Nopember Departemen Statistika. Penulis diterima di ITS pada tahun 2016 melalui jalur PKM ITS. Semasa perkuliahan, penulis aktif mengikuti beberapa organisasi diantaranya Staff Departemen Keilmiahan pada periode 2017-2018, Staff PERS HIMASTA-ITS bidang pewacanaan periode 2017-2018, Ketua bidang pewacanaan PERS HIMASTA-ITS periode 2018-2019, Staff Kementrian Kebijakan Publik BEM ITS periode 2018-2019, dan Dirjen Isu Strategis Kementrian Kebijakan Publik BEM ITS periode 2019-2020. Selain itu, penulis juga aktif dalam mengikuti kepanitiaan yang diadakan tingkat departemen dan ITS seperti Fasilitator Acara dalam kegiatan GERIGI ITS 2017, Staff ITS Expo Paper Competitio (IEPC) tahun 2017, Sie Acara Data Analysis Competition (DAC) pada *big event* Statistika ITS, Pekan Raya Statistika (PRS) 2017. Selama masa perkuliahan penulis juga pernah magang di PT Terminal Teluk Lamong (TTL) pada tahun 2019. Selain itu, penulis juga berkesempatan terpilih menjadi salah satu peserta *Study Excursion:* ITS *Goes Beyond Batch* 2 di Tokyo, Jepang pada tahun 2019. Bagi pembaca yang ingin berdiskusi,

memberikan saran, kritik, dan pertanyaan lebih lanjut mengenai Tugas Akhir ini dapat menghubungi melalui *e-mail* anggunnurfitriani@gmail.com.