



**TUGAS AKHIR - KS184822**

# **ANALISIS REGRESI UNTUK MEMPREDIKSI TAHANAN KAPAL CEPAT**

**MOHAMMAD HAIDAR ALVIN PURWANA  
NRP 062115 4000 0131**

**Dosen Pembimbing:  
R. Mohamad Atok, S.Si., M.Si., Ph.D.  
Mahendra Indriyanto, S.T., M.T.**

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





**TUGAS AKHIR - KS184822**

## **ANALISIS REGRESI UNTUK MEMPREDIKSI TAHANAN KAPAL CEPAT**

**MOHAMMAD HAIDAR ALVIN PURWANA  
NRP 062115 4000 0131**

**Dosen Pembimbing:  
R. Mohamad Atok, S.Si., M.Si., Ph.D.  
Mahendra Indriyanto, S.T., M.T.**

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





**FINAL PROJECT - KS184822**

**REGRESSION ANALYSIS FOR HIGH-SPEED CRAFT  
RESISTANCE PREDICTION**

**MOHAMMAD HAIDAR ALVIN PURWANA  
SN 062115 4000 0131**

**Supervisors:**

**R. Mohamad Atok, S.Si., M.Si., Ph.D.  
Mahendra Indriyanto, S.T., M.T.**

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



**LEMBAR PENGESAHAN**

**ANALISIS REGRESI UNTUK MEMPREDIKSI TAHANAN  
KAPAL CEPAT**

**TUGAS AKHIR**

**Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Statistika  
pada**

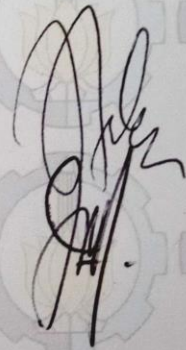
**Program Studi Sarjana Departemen Statistika  
Fakultas Sains dan Analitika Data  
Institut Teknologi Sepuluh Nopember**

Oleh :

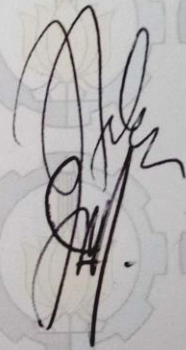
**Mohammad Haidar Alvin Purwana  
NRP. 062115 4000 0131**

Disetujui oleh Pembimbing:

R. Mohamad Atok, S.Si., M.Si., Ph.D.  
NIP. 19710915 199702 1 001

(  )

Mahendra Indriaryanto, S.T., M.T.  
NIP. 1985 0404 2009 1210 03

(  )



**Mengetahui,  
Kepala Departemen**

Dr. Kartika Fithriasari, M.Si.  
NIP. 19691212 199303 2 002

SN

**SURABAYA, JANUARI 2020**





# ANALISIS REGRESI UNTUK MEMPREDIKSI TAHANAN KAPAL CEPAT

**Nama Mahasiswa** : **Mohammad Haidar Alvin P.**  
**NRP** : **062115 4000 0131**  
**Departemen** : **Statistika**  
**Dosen Pembimbing** : **R. Mohamad Atok, S.Si., M.Si.,  
Ph. D.  
Mahendra Indriaryanto S.T.,  
M.T.**

## **Abstrak**

*Kapal merupakan bagian penting dari kehidupan manusia, terutama dalam bidang transportasi laut. Industri perkapalan di Indonesia telah berkembang dengan pesat, ditandai dengan berdirinya galangan-galangan kapal di Indonesia. Tahap awal dalam proses pembuatan kapal adalah menentukan ukuran dan parameter bentuk kapal. Bersamaan dengan hal tersebut terdapat faktor penting yang berhubungan dengan performa kapal, yaitu tahanan kapal. Kesalahan perhitungan tahanan kapal dapat berdampak buruk terhadap performa kapal dan juga para stakeholder kapal. Salah satu metode statistika yang memungkinkan untuk memprediksi tahanan kapal berdasarkan pengujian tahanan model kapal, yaitu analisis regresi. Pemilihan metode regresi sangat diperlukan karena metode yang tepat akan menghasilkan prediksi yang tepat juga. Pemilihan metode regresi dipengaruhi oleh kondisi data penelitian, seperti terpenuhi atau tidaknya asumsi-asumsi klasik meliputi multikolinieritas, distribusi normal, autokorelasi, dan homoskedastisitas. Hasil pengujian asumsi menunjukkan asumsi multikolinieritas dan distribusi normal tidak dapat terpenuhi yang mengarahkan pada penggunaan metode regresi robust ridge.*

***Kata kunci:*** *Autokorelasi, Homoskedastisitas, Multikolinieritas, Regresi Robust Ridge.*

*(Halaman ini sengaja dikosongkan)*

# REGRESSION ANALYSIS FOR HIGH-SPEED CRAFT RESISTANCE PREDICTION

**Name** : Mohammad Haidar Alvin P.  
**Student Number** : 062115 4000 0131  
**Department** : Statistics  
**Supervisor** : R. Mohamad Atok, S.Si., M.Si.,  
Ph. D.  
Mahendra Indriaryanto S.T.,  
M.T.

## **Abstract**

*Ships are an important part of human life, especially in the field of sea transportation. The shipping industry in Indonesia has grown rapidly, marked by the massive establishment of shipyards in Indonesia. The initial stage in the shipbuilding process is to determine the size and parameters of the ship's shape. Along with this, there is an important factor related to the performance of the ship, which is ship's resistance force. Mistakes in calculating ship's resistance force can adversely affect the ship performance and the clients. Statistical method that allows to predict resistance based on resistance testing data in towing tank, is defined as regression analysis. The selection of regression methods is very necessary because the right method will produce the right predictions too. The selection of proper regression method is in accordance with the conditions of the study, such as being fulfilled to classics regression assumption, like collinearity, normal distribution, autocorrelation, and homoskedasticity. The assumption test results show that collinearity and normal distribution cannot be fulfilled, yet using the robust ridge regression method is the best choice to handle it.*

**Keywords:** *Autocorrelation, Colinearity, Homoskedasticity, Robust Ridge Regression.*

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Segala puji dan syukur dipanjatkan kehadiran Tuhan Yang Maha Esa, Allah SWT, atas segala kasih, karunia, serta rahmatNya sehingga laporan Tugas Akhir yang berjudul “**Analisis Regresi Untuk Memprediksi Tahanan Kapal Cepat**” ini dapat diselesaikan dengan sebaik-baiknya. Penulisan Tugas Akhir ini merupakan salah satu syarat untuk mencapai gelar Sarjana pada Departemen Statistika di Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember (ITS).

Banyak pihak telah memberikan dukungan dan bantuan yang memungkinkan Tugas Akhir ini dapat diselesaikan sebagaimana mestinya. Untuk itu dalam sebuah kesempatan yang berharga ini, dengan kerendahan hati ijinkan penulis menyampaikan rasa hormat dan terima kasih kepada:

1. Ibu Dr. Kartika Fithriasari, M.Si. selaku Kepala Departemen Statistika Fakultas Sains dan Analitika Data ITS.
2. Ibu Santi Wulan Purnami, M.Si., Ph.D. selaku Sekretaris Departemen Statistika Fakultas Sains dan Analitika Data ITS.
3. Bapak R. Mohamad Atok, M.Si., Ph.D., selaku dosen pembimbing Tugas Akhir yang telah meluangkan waktu dan dengan sabar memberikan bimbingan, saran, motivasi, dan dukungan selama proses penyelesaian Tugas Akhir.
4. Bapak Mahendra Indriyanto, S.T., M.T., selaku pembimbing Tugas Akhir yang telah memberikan bimbingan selama proses pengerjaan hingga penulisan Tugas Akhir terselesaikan dengan baik.
5. Bapak Dr. Ir. Setiawan, M.S., selaku dosen penguji Tugas Akhir yang telah memberikan bimbingan dan saran selama proses penyusunan laporan Tugas Akhir.
6. Bapak Dr. Suhartono selaku dosen penguji Tugas Akhir yang turut memberikan arahan dan saran selama penyelesaian Tugas Akhir.
7. Bapak Dr. Muhammad Mashuri, M.T., selaku dosen wali yang telah banyak memberikan motivasi, pandangan, dan dukungan selama proses belajar di Departemen Statistika.

8. Seluruh dosen dan tenaga pendidikan Program Studi Sarjana Departemen Statistika Fakultas Sains dan Analitika Data Institut Teknologi Sepuluh Nopember yang telah membantu penulis selama menempuh masa kuliah hingga penyelesaian Tugas Akhir.
9. Kepala Balai Teknologi Hidrodinamika-Badan Pengkajian dan Penerapan Teknologi (BTH-BPPT), beserta jajarannya yang memberikan kesempatan untuk membantu dalam menganalisis data tahanan kapal cepat.
10. Kedua orang tua penulis, Dani Gumelar Daneswara dan Rosmiyati, saudara perempuan penulis, Nadhira Nurannisa, serta segenap keluarga penulis yang telah memberikan doa, motivasi, kasih sayang, dan perjuangan agar penulis dapat menyelesaikan proses perkuliahan dengan baik dan lancar.
11. Bapak Kresnayana Yahya, M.Sc., yang selalu memberikan perhatian, wawasan, motivasi, dan bimbingan selama penulis menjalankan proses perkuliahan sampai laporan Tugas Akhir ini dapat diselesaikan dengan baik.
12. Teman-teman Statistika ITS angkatan 2015 seperjuangan, yang selalu memberikan perhatian, motivasi, dan dukungan kepada penulis sehingga penulis dapat menyelesaikan Tugas Akhir.
13. Mutia Pandan Sari dan Tria Ayu Lestari selaku mahasiswa bimbingan Tugas Akhir Bapak Atok yang selalu memberikan informasi penting yang berkaitan dalam proses penyelesaian Tugas Akhir.
14. Segenap anggota Tim e-Bonni Indonesia, yang telah memberikan pengertian serta dukungan kepada penulis dari awal penulis memulai perkuliahan sebagai mahasiswa hingga penulis dapat menyelesaikan laporan Tugas Akhir dan masa studi selama berkuliah di ITS.
15. Ciko Pramuliasari, yang selalu sabar dan senantiasa memberikan dukungan moral dan motivasi kepada penulis mulai dari proses penyusunan proposal sampai proses penyelesaian laporan Tugas Akhir.

16. Semua teman, relasi, dan berbagai pihak yang tidak bisa penulis sebutkan namanya satu persatu yang telah membantu dalam keberhasilan proses yang panjang ini, baik secara langsung maupun tidak langsung, semoga Tuhan Yang Maha Esa selalu melimpahkan berkahNya.

Penulis menyadari bahwa dalam penulisan dan penyusunan laporan Tugas Akhir ini masih banyak kekurangan dan kelemahan. Oleh karena itu, kritik dan saran membangun akan sangat membantu penulis dalam memperbaikinya di masa yang akan datang. Penulis berharap semoga laporan Tugas Akhir ini dapat bermanfaat bagi masyarakat dan bagi ilmu pengetahuan.

Surabaya, 27 Mei 2019

Penulis

*(Halaman ini sengaja dikosongkan)*



## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>xi</b>
<b>DAFTAR ISI</b> .....	<b>xv</b>
<b>DAFTAR GAMBAR</b> .....	<b>xvii</b>
<b>DAFTAR TABEL</b> .....	<b>xix</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xxi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat .....	6
1.5 Batasan Masalah.....	7
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>9</b>
2.1 Analisis Regresi Berganda .....	9
2.1.1 Uji Linieritas Data .....	11
2.1.2 Homoskedastisitas .....	12
2.1.3 Uji Normalitas Data.....	14
2.1.4 Multikolinieritas.....	14
2.2 Regresi <i>Ridge</i> .....	15
2.3 Regresi <i>Robust</i> .....	18
2.4 Regresi <i>Ridge</i> dengan Parameter <i>Robust</i> .....	20
2.4.1 Estimasi Parameter <i>Robust Ridge</i> .....	21
2.4.2 Uji Signifikansi Parameter .....	22
2.5 Validasi dan Kriteria Kebaikan Model .....	23
2.6 Tinjauan Non Statistik .....	24
2.6.1 Kapal Cepat.....	24
2.6.2 Tahanan Kapal.....	26
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>27</b>
3.1 Sumber Data.....	27
3.2 Variabel Penelitian.....	27
3.3 Definisi Operasional .....	28
3.4 Struktur Data .....	28
3.5 Langkah Penelitian .....	29
<b>BAB IV ANALISIS DAN PEMBAHASAN</b> .....	<b>33</b>
4.1 Statistika Deskriptif .....	33
4.2 Uji Asumsi Linieritas Data.....	40

4.3 Uji Asumsi Homoskedastisitas.....	43
4.4 Uji Asumsi Distribusi Normal.....	44
4.5 Pemenuhan Asumsi Multikolinieritas.....	45
4.6 Pemodelan Menggunakan Regresi <i>Robust</i> .....	46
4.7 Pemodelan Menggunakan Regresi <i>Robust Ridge</i> .....	48
4.8 Kriteria Kebaikan Model .....	51
<b>BAB V KESIMPULAN DAN SARAN.....</b>	<b>53</b>
5.1 Kesimpulan .....	53
5.2 Saran .....	54
<b>DAFTAR PUSTAKA.....</b>	<b>55</b>
<b>LAMPIRAN.....</b>	<b>59</b>

## DAFTAR GAMBAR

<b>Gambar 3.1</b>	Diagram Alir Penelitian .....	30
<b>Gambar 4.1</b>	Matriks Korelasi Variabel Penelitian .....	34
<b>Gambar 4.2</b>	<i>Boxplot</i> Variabel $X_1 - X_8$ dan Y .....	36
<b>Gambar 4.3</b>	<i>Scatterplot</i> Variabel Y dengan X .....	38
<b>Gambar 4.4</b>	<i>Scatterplot</i> Variabel Y dengan Transformasi $X_4,$ $X_6, X_7$ .....	42

*(Halaman ini sengaja dikosongkan)*

## DAFTAR TABEL

<b>Tabel 2.1</b> <i>Analysis of Variance (ANOVA)</i> .....	22
<b>Tabel 3.1</b> Variabel Prediktor Penelitian.....	27
<b>Tabel 3.2</b> Variabel Respon Penelitian.....	28
<b>Tabel 3.3</b> Struktur Data Penelitian.....	28
<b>Tabel 4.1</b> Statistika Deskriptif Variabel Penelitian .....	33
<b>Tabel 4.2</b> Hasil Uji Linieritas Data.....	41
<b>Tabel 4.3</b> Hasil Uji Linieritas Data Transformasi .....	41
<b>Tabel 4.4</b> Hasil Uji Linieritas Data Transformasi .....	42
<b>Tabel 4.5</b> Hasil Uji Asumsi Homoskedastisitas .....	44
<b>Tabel 4.6</b> Hasil Uji Asumsi Distribusi Normal .....	44
<b>Tabel 4.7</b> Hasil Uji Asumsi Multikolinieritas .....	45
<b>Tabel 4.8</b> Hasil RMSE dan $R^2$ Model <i>Robust</i> .....	47
<b>Tabel 4.9</b> Koefisien Regresi <i>Robust MM-Estimator</i> .....	47
<b>Tabel 4.10</b> Koefisien Regresi <i>Robust Ridge</i> .....	49
<b>Tabel 4.11</b> $R^2$ Regresi Robust Ridge .....	50
<b>Tabel 4.12</b> Koefisien Regresi <i>Robust Ridge</i> $c=0,2$ .....	50
<b>Tabel 4.13</b> Hasil Prediksi Tahanan Kapal.....	51
<b>Tabel 4.14</b> Perbandingan RMSE dan $R^2$ Model Regresi ....	51
<b>Tabel 4.15</b> Batasan Variabel Prediktor .....	52

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

- Lampiran 1** Data Penelitian Kapal Cepat
- Lampiran 2** Data Penelitian Setelah Transformasi
- Lampiran 3** Gambar Kapal Cepat TNI-AL KRI Albakora
- Lampiran 4** Syntax Analisis Tugas Akhir
- Lampiran 5** Model Regresi Linier Sebelum Transformasi
- Lampiran 6** Model Regresi Linier Setelah Transformasi
- Lampiran 7** Model Regresi *Robust LMS-Estimator*
- Lampiran 8** Model Regresi *Robust LTS-Estimator*
- Lampiran 9** Model Regresi *Robust S-Estimator*
- Lampiran 10** Nilai Aktual & Prediksi Tahanan Kapal Cepat
- Lampiran 11** Surat Keterangan Data
- Lampiran 12** Biodata Penulis

*(Halaman ini sengaja dikosongkan)*



# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Sejarah peradaban manusia telah menjadikan kapal sebagai alat pendorong utama untuk eksplorasi wilayah teritorial dan juga pengembangan ilmu pengetahuan. Perkembangan kapal berjalan pesat seiring dengan perkembangan umat manusia. Melihat proses perkembangan teknologi kapal hingga saat ini, dapat dikatakan bahwa penggunaan kapal telah memiliki sumbangsih yang besar terjadinya perdagangan antar negara di dunia dan juga menjadikan kapal sebagai alat pertahanan suatu negara. Revolusi industri, metode mekanis baru dari propulsi, dan kemampuan untuk membangun kapal dari logam memicu terobosan dalam desain kapal (Marine Connector, 2018). Berakhirnya perang dunia yang berjalan lama dan boros, meningkatnya permintaan pasar terhadap kapal yang lebih efisien, dan munculnya era revolusi industri mendorong terciptanya produksi kapal penumpang dan kapal yang lebih khusus secara masal. Kapal-kapal khusus dibangun untuk fungsi yang sepenuhnya baru, seperti kapal pemadam kebakaran, penyelamatan, dan kapal penelitian. Indonesia sebagai negara kepulauan tentunya sangat bergantung terhadap kapal. Pemanfaatan kapal di Indonesia telah mencakup dalam berbagai bidang, seperti logistik, transportasi massal, dan juga militer.

Pada dasarnya kapal dapat dibedakan ke dalam dua kategori, yaitu menurut bentuk lambung kapal dan menurut fungsinya. Pengelompokan tipe kapal menurut bentuk lambung dapat dibedakan menjadi empat jenis, yaitu kapal *Aerostatic*, kapal *Hydrostatic*, kapal *Hydrodynamic*, dan kapal Multi Lambung. Menurut Sumaryanto (2013) pengelompokan kapal berdasarkan fungsinya terbagi menjadi tiga, yaitu kapal niaga dan komersil, kapal perang, dan kapal khusus. Macam-macam kapal yang termasuk sebagai kapal niaga dan komersil antara lain adalah kapal angkut, kapal penumpang, kapal ikan dan kapal tunda. Cakupan kapal khusus, diantaranya adalah kapal observasi dan kapal penelitian. Lain halnya dengan kedua jenis kapal yang telah disebutkan, kapal perang adalah jenis kapal yang dirancang khusus untuk penggunaan kekuatan angkatan laut. Kapal perang berbeda

dari kapal dagang dan sipil dalam hal desain, konstruksi, dan penggunaan. Kapal-kapal perang biasanya terbuat dari paduan baja khusus yang membuat mereka merusuk tangguh selama serangan musuh. Kapal perang dipersenjatai dengan sistem senjata canggih dengan pengecualian untuk kapal pengangkut pasukan dimana persenjataannya ringan atau tidak dipersenjatakan sama sekali.

Kapal perang diklasifikasi menjadi beberapa macam antara lain kapal tempur, patroli, dan kapal pendukung atau pengangkut pasukan (Sumaryanto, 2013). Kapal perang adalah kapal paling modern dalam aspek teknologi, sementara aspek ekonomis menjadi pertimbangan kedua. Kapal perang khususnya kapal patroli dan kapal tempur mengutamakan kecepatan dan manuver yang baik agar mampu beroperasi sesuai peruntukannya. Untuk itu lambung kapal perang jenis ini berbentuk *planing hull* untuk mendukung kebutuhan operasionalnya. Pertimbangan utama dari kapal perang ada pada berat, stabilitas, kecepatan, tenaga dan ketahanan, ruang, daya angkut, rencana umum (pengaturan ruang dan peralatan), kamar-kamar personel dan gerak kapal. Seluruh pertimbangan tersebut sudah harus didefinisikan sebelum desain awal kapal dibuat.

Kecepatan menjadi aspek penting karena kapal perang harus mampu bersaing dalam pengejaran maupun melepaskan diri dari musuh sedangkan manuver diperlukan dalam hal kontak senjata. Seiring berjalannya waktu, muncul penggolongan kapal menurut bilangan *Froude* yang berhubungan dengan kecepatan dan tahanan kapal. Bilangan *Froude* merupakan bilangan tak berdimensi yang sebanding dengan akar kuadrat dari rasio gaya inersia terhadap berat fluida (Chanson, 2004). Bilangan *Froude* adalah bilangan yang sangat signifikan yang digunakan untuk menentukan tahanan objek terendam yang bergerak melalui air. Dinamika kapal yang memiliki bilangan *Froude* yang sama, mudah dibandingkan karena menghasilkan bangun yang sama dan bahkan jika ukuran atau geometrinya berbeda. Menurut Siswanto (1988) kapal cepat merupakan kapal yang berlayar dengan bilangan *Froude* lebih dari sama dengan 0,35. Maka dari banyak hal tersebut, kapal perang dapat digolongkan sebagai kapal cepat. Hingga saat ini, bilangan *Froude* tetap menjadi rujukan di dunia perkapalan dalam menggolongkan kapal menurut kecepatannya.

Industri perkapalan di Indonesia telah berkembang dengan pesat. Hal tersebut ditandai dengan berdirinya galangan-galangan kapal di Indonesia. Berdasarkan data Kementerian Perindustrian Republik Indonesia (Kemenperin), saat ini ada sebanyak 250 galangan kapal di Indonesia dan sekitar 70 galangan kapal berada di Batam (Kementerian Perindustrian Republik Indonesia, 2019). Pada tahapan proses desain kapal, khususnya pada tahap awal penentuan *design requirement*, kondisi tinggi gelombang perairan di wilayah rute operasional kapal yang akan dilalui menjadi pertimbangan penting khususnya berkaitan dengan pertimbangan kekuatan struktur konstruksi kapal yang digunakan, khususnya berhubungan dengan elemen-elemen struktur kekuatan konstruksi memanjang yang tergantung dari dimensi panjang kapal. Prinsip dasar pendekatan dalam perhitungan desain kekuatan struktur konstruksi kapal adalah melalui pertimbangan bahwa badan kapal sebagai sebuah balok memanjang yang berada di atas tumpuan media fluida dalam hal ini gelombang laut yang mempunyai tekanan (*fluid pressure*) dan gaya dinamis (*dynamic forces*), sehingga kekuatan struktur konstruksi berdasarkan hasil perhitungan dari peraturan standar yang berlaku mampu menahan beban resultan berupa beban lengkung (*bending moment*), beban geser (*shear forces*), dan beban puntir (*torsion*) yang ditimbulkan oleh distribusi tekanan fluida dari tinggi gelombang sepanjang kapal dan distribusi muatan dan berat kapal (Hardjono, 2017).

Sahlan dkk (2016) dalam penelitiannya mengatakan bahwa tahap awal dalam proses teknis pembuatan kapal adalah menentukan ukuran dan parameter bentuk kapal, namun selain hal tersebut terdapat salah satu faktor penting lain yang berpengaruh terhadap kemampuan kapal, yaitu tahanan kapal atau *resistance force*. Besar tahanan kapal berhubungan dengan daya mesin yang dibutuhkan suatu kapal untuk mencapai kecepatan yang diinginkan. Berdasarkan hal tersebut, prediksi tahanan kapal sangatlah penting untuk dilakukan sebelum kapal tersebut dibangun (Sahlan, Ali, Nugroho, Bisri, & Arswendo, 2016). Pada proses pembuatan kapal, tentunya pengujian tahanan kapal diperlukan sebagai penentu daya mesin kapal agar kapal dapat bergerak di fluida dengan kecepatan tertentu.

Balai Teknologi Hidrodinamika-Badan Pengkajian dan Penerapan Teknologi (BTH-BPPT) merupakan salah satu lembaga di Indonesia yang dapat melakukan pengujian tahanan kapal. Pengujian tahanan kapal dilakukan dalam sebuah tangki berisikan air untuk menarik model kapal uji yang biasa dikenal dengan nama *towing tank*. Menurut BTH-BPPT pengujian tahanan kapal mulai dari membuat model kapal sampai dengan tahap pengujian pada *towing tank* memerlukan waktu sekitar satu bulan. Sedangkan waktu pengujian tahanan kapal pada *towing tank* menghabiskan waktu sekitar 7-14 hari. Sugianto dan Winarno (2017) dalam penelitiannya menyatakan bahwa biaya pengujian tahanan kapal masih tergolong mahal. Maka dari itu, penyingkatan waktu uji tentunya sangat berpengaruh bagi BTH-BPPT dari segi ekonomi. Sebagai badan yang sering melakukan pengujian tahanan kapal, diperlukan sebuah acuan tahanan kapal berdasarkan data kapal yang pernah dilakukan pengujian agar terciptanya efisiensi biaya dan waktu pengujian. Acuan tersebut memungkinkan BTH-BPPT mengetahui prediksi tahanan kapal berdasarkan kemiripan model kapal.

Salah satu metode statistika yang memungkinkan untuk memprediksi tahanan kapal berdasarkan pengujian tahanan model kapal, yaitu analisis regresi (Roh & Lee, 2018). Pemodelan menggunakan analisis regresi merupakan salah satu teknik dalam analisis data yang bertujuan untuk menjelaskan hubungan antara variabel prediktor dengan variabel respon (Rawling, Pantula, & Dickey, 1998). Terdapat berbagai metode dan model regresi yang berkembang hingga saat ini. Setiap metode memiliki keunggulan dan peruntukan masing-masing. Model regresi linier sederhana memiliki dua parameter, yaitu  $\beta_0$  dan  $\beta_1$  yang mana besar koefisiennya diperoleh dari data. Sebagai metode parametrik, tentunya diharuskan adanya pemenuhan asumsi-asumsi klasik. Walaupun demikian, teknik prediksi dengan memodelkan data menggunakan regresi tetap digunakan hingga saat ini.

Sebuah penelitian yang dilakukan oleh Mao dkk (2016) tentang memprediksi kecepatan kapal menggunakan pemodelan regresi linier berganda memberikan hasil yang kurang baik terhadap prediksi. Hal tersebut dikarenakan model regresi linier yang didapat menjadi bias karena terdapat multikolinieritas pada

data. Namun tetap saja pemodelan regresi dapat digunakan untuk memprediksi kecepatan kapal walaupun memiliki model yang kurang baik. Pada penelitian lain, prediksi kecepatan kapal menggunakan analisis regresi juga telah dilakukan oleh Brandsaeter dan Vanem (2018). Prediksi dilakukan dengan membandingkan kebaikan model yang didapat dari tiga metode regresi, yaitu regresi linier, *Generalized Additive Model* (GAM), dan *Projection Pursuit Regression* (PPR). Hasilnya baik GAM maupun model linier mampu menjelaskan hubungan antara kecepatan kapal dan poros dorong, dalam kata lain regresi mampu melakukan prediksi.

Pada kedua penelitian di atas, menyatakan bahwa dalam pengujian terdapat kondisi multikolinieritas antar variabel prediktor. Multikolinieritas adalah terjadinya hubungan linier antara variabel bebas (prediktor) dalam suatu model regresi linier berganda (Gujarati, 2003). Dampak yang dapat ditimbulkan dari multikolinieritas, yaitu dapat membuat penduga parameter *Ordinary Least Square* (OLS) kehilangan sifat *Best Linear Unbiased Estimator* atau biasa disebut penduga parameter yang BLUE (Widarjono, 2009). Sebuah penelitian tentang prediksi kebangkrutan korporat menggunakan regresi *Ridge* dan LASSO oleh Pereira, Basto, dan da Silva (2016), didasarkan karena kedua metode tersebut mampu mengatasi kasus multikolinieritas dengan baik dan menampilkan properti yang ideal untuk meminimalkan ketidakstabilan numerik yang mungkin terjadi karena overfitting. Hasil penelitian menunjukkan bahwa model LASSO dan *Ridge* cenderung memiliki hasil yang lebih memuaskan pada training set data, bila dibandingkan dengan metode Stepwise yang diterapkan dalam SPSS.

Pengujian tahanan kapal pada *towing tank* tentunya memerlukan faktor pendukung yang cukup banyak. Faktor-faktor utama dari proses tersebut adalah data hidrostatis kapal yang telah dibuat saat melakukan perancangan kapal. Pada proses ini, terdapat banyak variabel prediktor yang menentukan perhitungan tahanan kapal. Banyaknya variabel prediktor dapat mengantarkan pada kondisi multikolinieritas antar variabel prediktor. Alasan tersebut menjadikan penelitian ini untuk menggunakan analisis regresi dalam memprediksi tahanan kapal cepat.

## 1.2 Rumusan Masalah

Tahanan kapal berhubungan dengan besarnya daya mesin kapal yang dibuthkan. Untuk itu, tahanan kapal telah menjadi suatu keterikatan dan pertimbangan terpenting dalam konstruksi kapal. Tahap awal pembentukan kapal, yaitu merancang model kapal menggunakan faktor-faktor yang akan menjadi karakteristik dari kapal tersebut. Karakteristik kapal sangat menentukan besar dan kecilnya tahanan suatu kapal. Proses perhitungan yang panjang dan rumit dalam pembuatan model kapal tentu diperlukan sebagai langkah awal untuk mendapatkan besarnya tahanan kapal. Tahap akhir untuk mendapatkan tahanan kapal tentunya proses pengujian model kapal menggunakan *carriage* pada *towing tank*. Faktor-faktor pada karakteristik kapal yang diduga berpengaruh dalam pembentukan tahanan kapal dapat dimodelkan secara statistik menggunakan regresi. Model regresi yang didapatkan dari faktor-faktor yang diduga berpengaruh dapat digunakan sebagai acuan untuk memprediksi besarnya tahanan suatu kapal.

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mendeskripsikan kapal cepat berdasarkan hasil uji tahanan kapal.
2. Mendapatkan faktor-faktor yang mempengaruhi prediksi tahanan kapal cepat.
3. Mendapatkan model statistik yang dapat memprediksi tahanan kapal cepat.

## 1.4 Manfaat

Manfaat dari penelitian ini adalah sebagai berikut :

1. Dapat menjadi acuan bagi BTH-BPPT mengenai prediksi tahanan kapal cepat berdasarkan data pengujian tahanan kapal.
2. Diperoleh model statistik yang dapat menjadi rekomendasi bagi pembaca khususnya BTH-BPPT untuk memprediksi tahanan kapal cepat.

### **1.5 Batasan Masalah**

Pembatasan masalah dilakukan untuk memfokuskan penelitian sehingga penelitian yang dilakukan menjadi spesifik. Pada penelitian ini, data yang akan digunakan dibatasi hanya menggunakan data yang diperoleh berdasarkan uji tahanan kapal cepat yang dilakukan oleh BTH-BPPT dari tahun 2011-2018.

*(Halaman ini sengaja dikosongkan)*



## BAB II TINJAUAN PUSTAKA

Dalam bab ini akan dibahas beberapa teori statistika maupun non statistika yang mendasari penelitian tentang pemodelan tahanan kapal cepat menggunakan metode analisis regresi. Penelitian ini diawali dengan penjelasan mengenai teori statistik yang kemudian dilanjutkan dengan penjelasan mengenai kapal cepat dan tahanan kapal. Teori statistik yang mendasari penelitian ini dimulai dari regresi linier berganda beserta asumsi yang harus dipenuhi, regresi *Ridge*, regresi *Robust* kemudian dilanjutkan dengan *Robust Ridge Regression*.

### 2.1 Analisis Regresi Berganda

Istilah regresi pertama kali dikemukakan oleh Francis Galton (1886) dalam makalahnya yang berjudul “*Regression Towards Mediocrity in Hereditary Stature*”. Regresi dapat menganalisis hubungan antara beberapa variabel dengan terlebih dahulu menentukan variabel pengujian, yaitu variabel dependen (respon) dan variabel independen atau prediktor.

Analisis hubungan atau pengaruh satu variabel prediktor terhadap satu variabel respon dapat menggunakan model regresi linier sederhana. Sedangkan untuk menganalisis hubungan atau pengaruh dua atau lebih variabel prediktor terhadap variabel respon, maka model regresi yang digunakan adalah model regresi linier berganda (Draper & Smith, 1998). Kemudian untuk mendapatkan model regresi linier sederhana maupun model regresi linier berganda dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya menggunakan metode tertentu. Bentuk umum persamaan dari regresi linier berganda ditunjukkan sebagai berikut :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ji} + \varepsilon_i \quad (2.1)$$

dengan,  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, p$ ;  $k = 0, 1, \dots, q$

Berdasarkan bentuk umum persamaan dari regresi linier berganda pada persamaan (2.1), maka persamaan dapat disederhanakan seperti persamaan (2.2) (Rawling, Pantula, & Dickey, 1998).

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.2)$$

atau

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & x_{23} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} & \cdots & x_{np} \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

dimana :

$\mathbf{y}$  : vektor kolom  $n \times 1$  dari pengamatan pada variabel respon  $y_i$

$\mathbf{X}$  : matriks  $n \times (p + 1)$  terdiri dari satu kolom yang diberi label **1**, diikuti oleh vektor-vektor kolom  $p$  dari pengamatan pada variabel prediktor

$\boldsymbol{\beta}$  : vektor  $q \times 1$  dari parameter yang akan dilakukan estimasi

$\boldsymbol{\varepsilon}$  : vektor  $n \times 1$  dari *error*

Estimasi parameter  $\boldsymbol{\beta}$  bertujuan untuk mendapatkan model regresi linier berganda yang akan digunakan dalam analisis. Metode yang digunakan untuk mengestimasi parameter model regresi linier berganda berdasarkan asumsi normalitas data adalah metode *Ordinary Least Square* atau OLS (Gujarati, 2003). Metode OLS bertujuan meminimumkan jumlah kuadrat dari residual. Penaksir OLS merupakan penaksir yang memiliki sifat *Best Linier Unbiased Estimator* (BLUE). Notasi matriks pada persamaan normal ditunjukkan pada persamaan (2.3) berikut :

$$\begin{aligned} \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y}) \end{aligned} \quad (2.3)$$

dimana matriks  $(\mathbf{X}'\mathbf{X})$  dan  $(\mathbf{X}'\mathbf{y})$  dijabarkan pada persamaan (2.4) dan (2.5).

$$(\mathbf{X}'\mathbf{X}) = \begin{bmatrix} n & \sum x_{i1} & \sum x_{i2} & \cdots & \sum X_{ip} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \cdots & \sum x_{i1}x_{ip} \\ \sum x_{i2} & \sum x_{i1}x_{i2} & \sum x_{i2}^2 & \cdots & \sum x_{i2}x_{ip} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_{ip} & \sum x_{i1}x_{ip} & \sum x_{i2}x_{ip} & \cdots & \sum x_{ip}^2 \end{bmatrix} \quad (2.4)$$

$$(\mathbf{X}'\mathbf{y}) = \begin{pmatrix} \sum y_i \\ \sum x_{i1}y_i \\ \sum x_{i2}y_i \\ \vdots \\ \sum x_{ip}y_i \end{pmatrix} \quad (2.5)$$

Vektor estimasi variabel respon  $\mathbf{y}$  yang didapat dari hasil pemodelan nilai variabel prediktor  $\mathbf{X}$  disajikan pada persamaan (2.6).

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (2.6)$$

Vektor residual  $\mathbf{e}$  yang menjelaskan ketidaksesuaian antara data pengamatan  $\mathbf{y}$  dan data estimasi  $\hat{\mathbf{y}}$  ditunjukkan pada persamaan (2.7).

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \quad (2.7)$$

Regresi sebagai metode parametrik tentunya mempunyai asumsi-asumsi yang harus dipenuhi agar model regresi tidak *bias*. Asumsi pada regresi linier berganda akan disebutkan sebagai berikut dalam buku (Gujarati, 2003) :

1. Data linier dalam parameter.
2. Kemiripan variansi residual data (homoskedasitas).
3. Residual berdistribusi normal.
4. Tidak terdapat kolinearitas antar variabel prediktor (multikolinearitas).

Penjelasan asumsi-asumsi tersebut akan dibahas satu per satu sebagai berikut :

### 2.1.1 Uji Linieritas Data

Uji linieritas data digunakan untuk mengetahui apakah dua variabel mempunyai hubungan yang linier atau tidak secara signifikan. Pengujian linieritas data digunakan sebagai syarat utama dalam analisis regresi linier. Tentu saja hubungan umum yang alamiah antara variabel respon dan variabel prediktor tidak ada yang benar-benar mengikuti garis lurus yang linier. Pola hubungan antara variabel respon dan variabel prediktor akan memiliki bentuk yang sedikit melengkung atau bahkan murni berupa lengkungan alias tidak linier (Johnson & Bhattacharyya, 2010). Uji linieritas data dilakukan dengan mencari persamaan

regresi antara variabel prediktor terhadap variabel respon (Sonnberger & Kraemer, 1986). Salah satu metode yang dapat digunakan dalam pengujian linieritas data adalah *Ramsey RESET Test*. Hasil pengujian linieritas data ditentukan berdasarkan hipotesis berikut :

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

$$H_1 : \mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$$

dimana,

$\mathbf{R}$  : Matriks elemen variabel prediktor

$\boldsymbol{\beta}$  : Vektor parameter model

$\mathbf{r}$  : Matriks *rank* dari  $\mathbf{R}$

Keputusan akan menolak  $H_0$  apabila nilai *RESET* >  $F_{0,05,df_1,df_2}$  dengan taraf signifikan  $\alpha=5\%$ . Jika hasil pengujian menghasilkan keputusan untuk tolak  $H_0$  maka dapat dikatakan bahwa hubungan antar variabel bersifat tidak linier.

### 2.1.2 Homoskedastisitas

Homoskedastisitas adalah kondisi di mana residual model regresi memiliki variansi yang sama (homogen) satu sama lain (Gujarati, 2003). Homoskedastisitas merupakan asumsi yang harus dipenuhi dalam kasus regresi linier. Asumsi ini secara tersirat menjelaskan bahwa variabel respon pada setiap observasi memberikan informasi yang sama. Akibat dari kondisi homoskedastisitas, semua observasi pada OLS mendapatkan bobot yang sama (Rawling, Pantula, & Dickey, 1998) Lawan dari homoskedastisitas adalah heteroskedastisitas, yaitu sebuah kondisi di mana residual model regresi memiliki variansi yang berbeda (heterogen) satu sama lain (Widarjono, 2009). Dampak adanya heteroskedastisitas dalam model regresi adalah penduga OLS tidak lagi mempunyai variansi yang minimum dan menyebabkan perhitungan *standard error* metode OLS tidak dapat dipercaya kredibilitasnya meskipun penduga OLS akan tetap linier dalam parameter.

Metode yang dapat dilakukan untuk mendeteksi apakah residual bersifat homogen atau heterogen adalah Uji *Glejser* dan Uji *Breusch-Pagan*. Untuk mengetahui apakah pola residual mengandung heteroskedastisitas, Glejser menyarankan untuk melaku-

kan regresi nilai mutlak residual dengan variabel independen. Jika hasil uji F dari model regresi yang diperoleh tidak signifikan, maka tidak ada heteroskedastisitas dalam model regresi (Widarjono, 2009). Hipotesis yang digunakan untuk pengujian homoskedastisitas adalah sebagai berikut :

$$H_0 : \text{Var}(y_i) = \sigma^2$$

$$H_1 : \text{Var}(y_i) \neq \sigma^2$$

dimana,

$\text{Var}(y_i)$  : Variansi y ke-i

$\sigma^2$  : Variansi model

Untuk melakukan Uji *Breusch-Pagan* diperlukan pemodelan regresi linear berganda OLS dan menghitung residual dari regresi tersebut. Hasil residual dari regresi OLS dan perbedaan variansi antar variabel prediktor digunakan untuk membentuk model regresi pembantu (*auxiliary regression*) yang dijabarkan pada persamaan (2.8).

$$e_i^2 = \gamma_1 + \gamma_2 z_{2i} + \dots + \gamma_j z_{ji} + \eta_i \quad (2.8)$$

dimana,

$e$  : Residual regresi OLS

$\gamma$  : Parameter regresi pembantu

$\eta$  : residual regresi pembantu

$z$  : Beda variansi antara variabel prediktor pada regresi OLS

Statistik uji yang digunakan untuk Uji *Breusch-Pagan* adalah *Lagrange Multiplier* dengan penjabaran seperti pada persamaan (2.9). Statistik uji didapatkan berdasarkan koefisien determinasi dari regresi pembantu (*auxiliary regression*) dan statistik uji pada *Breusch-Pagan* asimtotik terhadap distribusi  $\chi^2$  (Sonnberger & Kraemer, 1986).

$$LM = nR^2 \quad (2.9)$$

dimana,

$R^2$  : Koefisien determinasi model regresi pembantu

$n$  : Jumlah observasi

Keputusan untuk menolak  $H_0$  dilakukan jika nilai statistik uji lebih dari  $\chi^2$  tabel atau jika  $p$ -value lebih kecil dari taraf signifikan  $\alpha=5\%$ . Apabila  $H_0$  ditolak maka akan didapatkan kesimpulan bahwa residual model memiliki variansi yang sama antar pengamatannya atau dalam kata lain asumsi homoskedastisitas tidak terpenuhi.

### 2.1.3 Uji Normalitas Data

Uji ini merupakan pengujian terhadap normalitas residual yang digunakan untuk melihat apakah variabel prediktor dan variabel respon mengikuti ataupun mendekati pola distribusi normal. Pengujian normalitas data dapat menggunakan beberapa metode, diantaranya *Kolmogorov-Smirnov Test*, *Shaphiro-Wilk Test*, atau *Jarque-Bera Test*. Ketiga metode tersebut merupakan metode yang paling umum dan relevan untuk dipakai. Hipotesis dalam pengujian ini dijabarkan sebagai berikut :

$$H_0 : \varepsilon \cong N(\mu, \sigma)$$

$$H_1 : \varepsilon \neq N(\mu, \sigma)$$

dimana,

$N(\mu, \sigma)$  : Distribusi normal

$\varepsilon$  : Residual model

Keputusan akan menolak  $H_0$  apabila statistika uji dari salah satu metode terpilih bernilai kurang dari nilai t tabel dengan derajat bebas dan taraf signifikan yang telah disesuaikan berdasarkan data residual. Pada analisis regresi linier berganda, pengujian kenormalan data diharapkan menghasilkan keputusan gagal tolak  $H_0$  agar proses analisis dapat dilanjutkan. Apabila keputusan tolak  $H_0$  maka langkah selanjutnya yang dapat dilakukan adalah ransformasi data ataupun memilih metode estimasi parameter lain yang sesuai dengan kondisi data.

### 2.1.4 Multikolinieritas

Multikolinieritas adalah terjadinya hubungan linier antara variabel bebas dalam suatu model regresi linier berganda (Gujarati, 2003). Hubungan linier antara variabel bebas dapat terjadi dalam bentuk hubungan linier yang sempurna (*perfect*) dan hubungan linier yang kurang sempurna (*imperfect*). Adapun dampak adanya

multikolinieritas dalam model regresi linier berganda sebagai berikut :

1. Penaksir OLS masih bersifat BLUE, tetapi mempunyai variansi dan kovariansi yang yang besar sehingga sulit mendapatkan taksiran (estimasi) yang tepat.
2. Akibat penaksir OLS mempunyai variansi dan kovariansi yang yang besar, menyebabkan interval estimasi akan cenderung lebih lebar dan nilai hitung statistik uji t akan kecil, sehingga membuat variabel bebas secara statistik tidak signifikan mempengaruhi variabel tidak bebas.
3. Walaupun secara individu variabel bebas tidak berpengaruh terhadap variabel tidak bebas melalui uji t, tetapi nilai koefisien determinasi ( $R^2$ ) masih bisa relatif tinggi.

Selanjutnya untuk mendeteksi adanya multikolinieritas dalam model regresi linier berganda dapat digunakan nilai *Variance Inflation Factor* (VIF) dengan ketentuan jika nilai VIF melebihi angka 10, maka terjadi multikolinieritas dalam model regresi (Rawling, Pantula, & Dickey, 1998). Berikut ditunjukkan persamaan untuk mendapatkan nilai VIF pada persamaan (2.10).

$$VIF_j = \frac{1}{1-R_j^2} \quad (2.10)$$

dimana,

$R_j^2$  : Koefisien determinasi dari regresi  $X_j$  terhadap variabel prediktor yang lain

## 2.2 Regresi Ridge

Hoerl dan Kennard (1970) menyebutkan bahwa regresi *Ridge* merupakan metode regresi untuk menduga koefisien regresi dengan menambahkan konstanta *bias c* pada diagonal matriks  $(X'X)$ . Regresi *Ridge* bergantung pada sebuah transformasi normalisasi variabel prediktor. Hal tersebut dilakukan dengan mengurangi dari setiap nilai yang merupakan estimasi rata-rata prediktor dan membagi masing-masing nilai dengan estimasi standar deviasi prediktor. Melalui cara tersebut, semua nilai berada pada skala yang sama dan memiliki standar deviasi 1. Regresi *Ridge* memberikan penalti pada ukuran total koefisien, sehingga model dengan koefisien yang lebih kecil lebih diutamakan.

Pernyataan tersebut didasarkan pada intuisi bahwa prediktor yang lebih penting akan memiliki koefisien yang lebih tinggi, dan penalti akan membuat koefisien lainnya tetap rendah (Al-Obeidat, Spencer, & Alfandi, 2018).

Metode regresi *Ridge* dapat digunakan untuk mengatasi korelasi yang tinggi antar beberapa variabel independen (prediktor) atau dalam kata lain metode ini mampu mengatasi multikolinieritas yang terjadi pada kasus regresi linier berganda (Hoerl & Kennard, 1970). Meskipun metode ini akan menghasilkan parameter dugaan yang sedikit *bias* namun parameter dugaan tersebut dapat menghasilkan parameter yang memiliki nilai taksiran mendekati nilai sebenarnya. Pernyataan tersebut dapat dibuktikan dengan membandingkan *Mean Square Error* (MSE) penduga *Ridge* dengan penduga *Least Square*.

Estimasi (pendugaan) parameter *Ridge* dinotasikan sebagai  $\hat{\beta}^*$  dan untuk *Sum Square Error* (SSE) dapat ditulis seperti pada persamaan (2.11).

$$\begin{aligned}\phi &= (Y - X\hat{\beta}^*)'(Y - X\hat{\beta}^*) \\ \phi &= (Y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}^*)'(Y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}^*) \\ \phi &= (Y - X\hat{\beta} + X(\hat{\beta} - \hat{\beta}^*))'(Y - X\hat{\beta} + X(\hat{\beta} - \hat{\beta}^*)) \\ \phi &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta} - \hat{\beta}^*)'X'X(\hat{\beta} - \hat{\beta}^*)\end{aligned}\quad (2.11)$$

Untuk  $\phi$  tetap, maka dipilih penduga  $\hat{\beta}^*$  dan dibuat meminimumkan  $\hat{\beta}^{*'}\hat{\beta}^*$  dengan residual  $\phi_0 = (\hat{\beta} - \hat{\beta}^*)'X'X(\hat{\beta} - \hat{\beta}^*)$ . Hoerl dan Kennard (1970) menggunakan pertimbangan tersebut dan memberikan fungsi *Lagrange* seperti yang dituliskan pada persamaan (2.12).

$$F = \hat{\beta}^{*'}\hat{\beta}^* + \frac{1}{c} \left[ (\hat{\beta} - \hat{\beta}^*)'X'X(\hat{\beta} - \hat{\beta}^*) - \phi_0 \right] \quad (2.12)$$

dengan  $\frac{1}{c}$  merupakan pengganda *Lagrange*, sehingga untuk mendapatkan penduga parameter *Ridge* diperlukan penurunan fungsi *Lagrange* terhadap  $\hat{\beta}^*$ . Hasil akhir dari penurunan berupa penduga  $\hat{\beta}^*$  yang dinotasikan bersamaan dengan estimasi  $\hat{\beta}^*$  sebagai berikut :



$$\begin{aligned}
\frac{\partial F}{\partial \hat{\boldsymbol{\beta}}^*} &= 2\hat{\boldsymbol{\beta}}^* + \frac{1}{c} [2(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}^* - 2(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}] \\
\frac{\partial F}{\partial \hat{\boldsymbol{\beta}}^*} &= 2\hat{\boldsymbol{\beta}}^* + \frac{2}{c}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}^* - \frac{2}{c}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} \\
\frac{\partial F}{\partial \hat{\boldsymbol{\beta}}^*} &= \hat{\boldsymbol{\beta}}^* + \frac{1}{c}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}^* - \frac{1}{c}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} \\
\frac{\partial F}{\partial \hat{\boldsymbol{\beta}}^*} &= 0 \\
\hat{\boldsymbol{\beta}}^* \left[ 1 + \frac{1}{c}(\mathbf{X}'\mathbf{X}) \right] &= \frac{1}{c}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} \\
\hat{\boldsymbol{\beta}}^* [c\mathbf{I} + (\mathbf{X}'\mathbf{X})] &= (\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} \\
\hat{\boldsymbol{\beta}}^* &= [c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} \\
\hat{\boldsymbol{\beta}}^* &= \mathbf{Z}\hat{\boldsymbol{\beta}} \\
\hat{\boldsymbol{\beta}}^* &= [c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\mathbf{X}'\mathbf{y} \tag{2.13}
\end{aligned}$$

Berdasarkan hasil penaksiran penduga parameter *ridge* pada persamaan (2.13), maka ekspektasi dari penduga parameter *ridge* dijabarkan pada persamaan (2.14).

$$\begin{aligned}
E(\hat{\boldsymbol{\beta}}^*) &= E([c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= E([c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}[(\mathbf{X}'\mathbf{X}) + c\mathbf{I} - c\mathbf{I}]\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= E([c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + c[c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\hat{\boldsymbol{\beta}} \\
&\quad - c[c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= E([c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}[\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + c\mathbf{I}\hat{\boldsymbol{\beta}}] \\
&\quad - c[c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= E([c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}[\mathbf{X}'\mathbf{X} + c\mathbf{I}]\hat{\boldsymbol{\beta}} \\
&\quad - c[c\mathbf{I} + (\mathbf{X}'\mathbf{X})]^{-1}\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= E(\hat{\boldsymbol{\beta}} - c(\mathbf{X}'\mathbf{X} + c\mathbf{I})^{-1}\hat{\boldsymbol{\beta}}) \\
E(\hat{\boldsymbol{\beta}}^*) &= \boldsymbol{\beta} - c(\mathbf{X}'\mathbf{X} + c\mathbf{I})^{-1}\boldsymbol{\beta} \tag{2.14}
\end{aligned}$$

Sesuai dengan persamaan estimasi parameter *Ridge* pada persamaa (2.14) maka penduga *Ridge* merupakan penduga yang *bias* dengan besaran  $c(\mathbf{X}'\mathbf{X} + c\mathbf{I})^{-1}\boldsymbol{\beta}$  (Hoerl & Kennard, 1970). Untuk memperoleh variansi dari penduga *Ridge* adalah sebagai berikut. Misalkan  $\lambda_1, \lambda_2, \dots, \lambda_{k+1}$  adalah nilai-nilai eigen dari  $(\mathbf{X}'\mathbf{X})$  dan  $\hat{\beta}_{(c),0}^*, \hat{\beta}_{(c),1}^*, \hat{\beta}_{(c),2}^*, \dots, \hat{\beta}_{(c),k}^*$  adalah penduga koefisien

regresi *Ridge*. Jika diketahui bahwa  $\hat{\beta}^* = \mathbf{Z}\hat{\beta}$ , maka variansi dari  $\hat{\beta}^*$  disajikan pada persamaan (2.15).

$$\begin{aligned}\text{var}[\hat{\beta}^*] &= \text{var}[\mathbf{Z}\hat{\beta}] \\ \text{var}[\hat{\beta}^*] &= \mathbf{Z} \text{var}[\hat{\beta}]\mathbf{Z}' \\ \text{var}[\hat{\beta}^*] &= \mathbf{Z}\sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}' \\ \text{var}[\hat{\beta}^*] &= \sigma^2\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\end{aligned}\quad (2.15)$$

### 2.3 Regresi *Robust*

Salah satu ide untuk mengatasi masalah dalam regresi yang umum muncul adalah mengidentifikasi *outlier*, menghapusnya, dan kemudian melanjutkan seperti langkah awal, dan diasumsikan bahwa set data telah sesuai untuk metode standar. Jika koefisien sebenarnya diketahui, maka *outlier* tidak akan sulit dideteksi. Bidang diagnosa regresi mencoba untuk mengatasi masalah bagaimana mengidentifikasi titik-titik dan *outlier* yang berpengaruh, yang mana dalam kasus umum ketika kita tidak mengetahui nilai-nilai koefisien sebenarnya.

Ketika hanya ada satu *outlier*, beberapa metode diagnostik bekerja dengan sangat baik dengan melihat efek penghapusan satu per satu titik data. Sayangnya jauh lebih sulit untuk mendiagnosis *outlier* ketika data yang terdiagnosis sebagai *outlier* terbilang banyak (Lukman, Osowole, & Ayinde, 2015). Dalam situasi ini, perlu untuk berurusan dengan fenomena “*outliers masking*”. *Outlier masking* terjadi ketika satu set *outlier* tidak terdeteksi karena kehadiran set *outlier* lain. Seringkali ketika *outlier* digunakan agar sesuai dengan nilai parameter, estimasi menjadi *bias* dan sangat buruk. Setelah terdapat banyak *outlier*, metode penghapusan tidak lagi layak digunakan secara komputasi. Maka perlu untuk melihat penghapusan semua himpunan bagian dari data poin di bawah jumlah maksimum *outlier* yang dipilih.

Pendekatan lain untuk berurusan dengan *outlier* adalah regresi yang kuat (*robust regression*), yang mencoba untuk menghasilkan estimator yang resisten atau setidaknya tidak sangat dipengaruhi oleh *outlier* (Wilcox, 2017). Dalam mempelajari residual dari regresi *robust*, mungkin *outlier* yang sebenarnya

dapat ditemukan. Dalam bidang ini banyak ide yang berbeda telah diusulkan, termasuk *Least Trimmed Squares* (LTS), *Least Median of Squares* (LMS), *M-estimators*, *GM-estimators*, dan *S-estimators* (Lukman, Osowole, & Ayinde, 2015).

Regresi *robust* dan metode diagnostik *outlier* akhirnya menjadi sangat mirip dalam hal penyelesaian masalah yang telah disebutkan. Keduanya melibatkan upaya untuk mencari *outlier* dan mencoba memperkirakan koefisien dengan cara yang tidak terlalu dipengaruhi oleh *outlier*. Hal yang berbeda adalah urutan di mana kedua langkah ini dilakukan. Saat menggunakan diagnostik, cari *outlier* pertama dan kemudian setelah mereka dihapus gunakan OLS pada set data yang baru untuk estimasi parameter yang lebih baik. Regresi *robust* sebagai gantinya mencari untuk menemukan estimasi kuat yang lebih baik terlebih dahulu dan melalui estimasi ini, kita dapat menemukan *outlier* dengan menganalisis residual. Model dasar dari regresi *robust* sama dengan model regresi linier berganda. Hal yang membedakan hanya pada metode estimasi parameter model.

Metode regresi *robust* memiliki banyak cara dalam mengestimasi parameter yang *robust*. *M-Estimator* merupakan metode yang sederhana dan memiliki konsep mengasumsikan bahwa sebagian besar data terkena dampak dari *outlier*. *S-Estimator* merupakan metode estimasi parameter yang menggunakan *high breakdown point*. Mengukur proporsi minimal dari banyaknya data yang terpengaruh oleh *outlier* disebut sebagai *breakdown point*.

Lukman, Osowole, dan Ayinde (2015) dalam penelitiannya menyebutkan bahwa *MM-Estimator* lebih efisien dibandingkan dua metode yang telah disebutkan. Berdasarkan penjabaran singkat mengenai metode *M-Estimator* dan *S-Estimator* yang telah disebutkan sebelumnya, estimasi parameter *robust* pada penelitian ini menggunakan *MM-Estimator*. Menurut Wilcox (2017), *MM-Estimator* merupakan perpaduan antara *M-Estimator* dan *S-Estimator* yang memiliki *breakdown point* tertinggi dan efisiensi yang tinggi terhadap normalitas. Penaksiran parameter *robust* menggunakan *MM-Estimator* dijelaskan melalui persamaan (2.16).

$$\hat{\beta}_{MM} = \sum_{i=1}^n x_{ij} \varphi_1 \left( \frac{y_i - x_i^T \hat{\beta}}{s_n} \right) = 0, \quad j = 1, 2, \dots, k \quad (2.16)$$

dimana,

$$\mathbf{S}_n = S \left( r_i(\tilde{\beta}), \dots, r_n(\tilde{\beta}) \right) \quad (2.17)$$

$$\mathbf{r}_i(\beta) = y_i - \mathbf{x}_i^T \beta \quad (2.18)$$

$$\boldsymbol{\varphi}_1(\boldsymbol{\varepsilon}) = \frac{\partial \rho_1(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \quad (2.19)$$

Melalui persamaan (2.16) di atas, maka estimasi parameter *robust* menggunakan *MM-Estimator* dapat disederhanakan seperti pada persamaan (2.20).

$$\hat{\boldsymbol{\beta}}_{MM} = (\mathbf{X}^T \mathbf{w}_i \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{w}_i \mathbf{y}) \quad (2.20)$$

diman,

$$w_i = \begin{cases} \left( 1 - \left( \frac{\varepsilon_i}{\hat{\sigma}_i(4,685)} \right)^2 \right)^2 & \text{jika } \left| \frac{\varepsilon_i}{\hat{\sigma}_i} \right| \leq 4,685 \\ 0 & \text{jika } \left| \frac{\varepsilon_i}{\hat{\sigma}_i} \right| > 4,685 \end{cases}$$

## 2.4 Regresi *Ridge* dengan Parameter *Robust*

Regresi *ridge* dengan parameter *robust* merupakan sebuah penggabungan dua metode regresi, yaitu regresi *ridge* dan regresi *robust*. Metode ini lebih dikenal dengan nama *robust ridge regression*. Tujuan dari metode ini adalah untuk mengatasi pelanggaran asumsi distribusi normal dan multikolinieritas yang tidak dapat diatasi oleh metode regresi linier berganda OLS. Konsep dasar dari regresi *robust ridge* menggunakan teknik pemodelan regresi *ridge* dengan menyisipkan parameter *robust* dalam kalkulasi parameter model *ridge*.

*Robust Ridge Regression* diprakarsai oleh Holland (1973) dan dikembangkan oleh Alpu dan Samkar (2010). Ketika kasus *outlier* dan multikolinieritas terjadi dalam dataset, akan lebih bermanfaat untuk menggabungkan metode yang dirancang untuk menangani masalah tersebut secara individual. Dengan demikian, estimator *ridge* yang *robust* akan tahan terhadap masalah multikolinieritas dan akan lebih sedikit terpengaruh oleh *outlier* dibandingkan dengan metode regresi *ridge* (Alpu & Samkar, 2010).

### 2.4.1 Estimasi Parameter *Robust Ridge*

Merujuk pada persamaan (2.13) dan persamaan (2.20) yang merupakan landasan untuk estimasi parameter *robust ridge*, diberikan persamaan (2.21) yang merupakan estimasi parameter *robust* untuk regresi *ridge*. Notasi  $\hat{\beta}^{**}$  merupakan matriks dari taksiran parameter *robust* untuk regresi *ridge*.

$$\hat{\beta}^{**} = [c^*I + (X'X)]^{-1}X'X\hat{\beta}_{MM} \quad (2.21)$$

Telah banyak metode untuk mencari nilai  $c^*$  dijelaskan dalam berbagai literatur. Hoerl dan Kennard (1970) mengatakan bahwa nilai  $c$  terbaik merupakan nilai yang mampu memberikan nilai  $MSE(\hat{\beta}^{**})$  yang kecil. Berdasarkan pada pengalaman dalam penelitian, metode terbaik untuk mendapatkan estimasi  $\hat{\beta}^{**}$  yang lebih baik adalah dengan melakukan iterasi dengan beberapa nilai  $c^*$ . Untuk itu, persamaan (2.22) disajikan untuk melakukan perhitungan  $MSE(\hat{\beta}^{**})$ .

$$MSE(\hat{\beta}^{**}) = \delta^2 \sum_{j=1}^{p+1} \frac{\lambda_i}{(\lambda_j + c_j^*)^2} + \sum_{j=1}^{p+1} \frac{(c_j^*)^2 \beta_j^2}{(\lambda_j + c_j^*)^2} \quad (2.22)$$

dimana,

$\delta^2$  : Jumlah residual kuadrat (*Sum Squared Error*)

$\lambda_i$  : Nilai eigen dari matriks  $X'X$

$\beta_j$  : Estimasi parameter *robust*

Alternatif metode perhitungan  $c^*$  lain yang telah diformulasikan oleh Hoerl, Kennard, dan Baldwin (1975) disajikan pada persamaan (2.23). Namun demikian pemilihan nilai  $c^*$  terbaik agar masalah multikolinieritas teratasi, yaitu memperhatikan korelasi antar variabel respon dan prediktor, dan juga koefisien regresi.

$$c^* = \frac{k\hat{\sigma}_{Robust}^2}{\hat{\beta}'_{MM}\hat{\beta}_{MM}} \quad (2.23)$$

dimana,

$k$  : Jumlah regresor

$\hat{\sigma}_{Robust}^2$  : Estimasi skala *robust*

### 2.4.2 Uji Signifikansi Parameter

Regresi *robust ridge* menggabungkan konsep dasar estimasi parameter regresi *ridge* dan regresi *robust*. Untuk itu dalam pengujian signifikansi parameter *robust ridge* secara serentak menggunakan uji F. Hipotesis nol yang digunakan disajikan sebagai berikut :

$$H_0 : \hat{\beta}_1^{**} = \hat{\beta}_2^{**} = \dots = \hat{\beta}_j^{**} = 0$$

$$H_1 : \text{Minimal satu } \hat{\beta}_j^{**} \neq 0, j = 1, 2, \dots, p$$

Statistik uji F didapatkan melalui tabel ANOVA yang disajikan pada Tabel 2.1. Keputusan untuk menolak  $H_0$  didapatkan apabila  $F_{hitung} > F_{\alpha, p, n-p-1}$  (Obenchain, 1977). Jika keputusan pengujian signifikansi parameter secara serentak memberikan hasil untuk menolak  $H_0$ , maka dilanjutkan pada pengujian signifikansi parameter secara parsial.

**Tabel 2.1** *Analysis of Variance (ANOVA)*

Sumber Variasi	df	Sum of Square	Mean Square	F
Regresi	p	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{\text{Regresi}}}{df_{\text{Regresi}}}$	$\frac{MS_{\text{Regresi}}}{MS_{\text{Residual}}}$
Residual	n-p-1	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{\text{Residual}}}{df_{\text{Residual}}}$	
Total	n-1	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Untuk melakukan uji signifikansi parameter secara parsial dapat digunakan uji Wald dengan modifikasi yang menyesuaikan dengan estimasi parameter model yang digunakan (Cule, Vineis, & De Iorio, 2011). Hipotesis yang digunakan untuk uji parsial disajikan sebagai berikut :

$$H_0 : \hat{\beta}_j^{**} = 0$$

$$H_1 : \hat{\beta}_j^{**} \neq 0, k = 1, 2, \dots, p$$

Statistik uji Wald disajikan pada persamaan (2.24). Keputusan untuk menolak  $H_0$  dari pengujian parsial terjadi jika statistik uji  $t_\lambda > t_{\frac{\alpha}{2}, n-p-1}$ .

$$t_\lambda = \frac{\hat{\beta}_j^{**}}{SE(\hat{\beta}_j^{**})} \quad (2.24)$$

dimana,

$SE(\hat{\beta}_j^{**})$  : Residual standar parameter *robust ridge* ke-j

## 2.5 Validasi dan Kriteria Kebaikan Model

Pada permasalahan regresi, estimasi parameter regresi dan kesalahan standar (*standard error*) yang dihasilkan oleh model mengarahkan para peneliti untuk melakukan evaluasi terhadap model. Evaluasi kebaikan model regresi dalam memprediksi variabel respon telah dikembangkan dan memiliki berbagai macam cara (Gujarati, 2003). Validasi silang (*cross validation*) merupakan metode untuk mengevaluasi model regresi yang akan digunakan untuk memprediksi tahanan kapal. Prinsip dasar validasi silang, yaitu membagi data penelitian menjadi dua subset data atau lebih (Burnham & Anderson, 2002). Jika menggunakan dua subset data, maka salah satu subset tersebut akan digunakan untuk membentuk model sementara subset lain digunakan untuk validasi hasil prediksi. Validasi silang memiliki berbagai jenis metode yang salah satunya adalah *hold out cross validation*.

$R^2$  merupakan salah satu metode yang dapat digunakan untuk melihat kebaikan model regresi berdasarkan kriteria informasi yang dihasilkan oleh variabel prediktor untuk memprediksi variabel respon. Secara singkat, Gujarati (2003) menjelaskan jika  $R^2$  digunakan untuk melihat seberapa baik garis regresi untuk mencocokkan terhadap data. Nilai  $R^2$  memiliki interval antara 0 sampai 1 dan tidak bernilai negatif. Perhitungan  $R^2$  didefinisikan sebagai berikut :

$$R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.25)$$

dimana :

$R^2$  : Koefisien determinasi

$y_i$  : Data observasi aktual ke-i  
 $n$  : Banyak pengamatan

Metode lain untuk melakukan penilaian kebaikan model regresi adalah *Root Mean Square Error* (RMSE). RMSE adalah aturan penilaian kuadrat yang mengukur besarnya rata-rata kesalahan (*error*) dari model. RMSE merupakan akar kuadrat dari rata-rata perbedaan kuadrat antara data hasil prediksi dan observasi aktual (Chai & Draxler, 2014). Rumusan RMSE didefinisikan pada persamaan (2.26).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2.26)$$

dimana :

$\hat{y}_i$  : Data taksiran ke-i  
 $y_i$  : Data observasi aktual ke-i  
 $n$  : Banyak pengamatan

## 2.6 Tinjauan Non Statistik

Tinjauan non statistik pada penelitian meliputi pemahaman tentang Kapal Cepat dan Tahanan Kapal.

### 2.6.1 Kapal Cepat

Kapal cepat merupakan kapal yang memiliki bilangan *Froude* lebih dari sama dengan 0,35 saat berlayar (Siswanto, 1988). Kapal cepat dapat ditemui dengan mudah di lingkungan pesisir dan juga laut lepas. Terdapat banyak desain untuk berbagai jenis kapal cepat, namun satu desain yang pasti adalah bentuk lambung kapal yang menyerupai huruf ‘V’. Kapal cepat lebih banyak diperuntukkan sebagai kapal perang dan patroli. Kapal perang dan patroli difungsikan sebagai kapal pemantau dari kegiatan-kegiatan yang terjadi di laut khususnya kegiatan yang sekiranya berpotensi atau bahkan telah melanggar peraturan yang telah ditetapkan (Sumaryanto, 2013).

Umumnya kapal patroli ditemukan dalam berbagai peran perlindungan perbatasan, termasuk anti-penyelundupan, anti-pembajakan, patroli perikanan, dan penegakan hukum imigrasi. Kapal patroli dapat dioperasikan oleh angkatan laut suatu negara, penjaga pantai, atau kepolisian. Kapal patroli juga sering



dipergunakan untuk berpartisipasi dalam operasi penyelamatan. Untuk memperjelas gambaran tentang bentuk lambung kapal patroli yang tergolong ke dalam kapal cepat dapat dilihat pada Lampiran 3.

Kapal cepat lain yang umum digunakan di berbagai negara yang memiliki perairan adalah kapal perang. Kapal perang adalah kapal tentara angkatan laut yang dibangun dan terutama ditujukan untuk perang angkatan laut. Kehadiran kapal perang dimulai ketika banyak kerajaan atau pemerintahan membutuhkan atau merasa perlu menegaskan posisinya di perairan sekaligus memberikan jaminan keamanan di perairan untuk melindungi negaranya dan aktivitasnya seperti nelayan dan perdagangan. Banyak gangguan keamanan di perairan yang harus dicegah, termasuk adanya serangan dari negara-negara lain yang lebih aman bila langsung ditangkal dari laut. Pada masa sekarang ini, khususnya negara-negara yang memiliki kawasan perairan, kebutuhan membangun angkatan laut dan kapal-kapal perang adalah penting. Kemudian disesuaikan dengan kemampuan dan kesanggupan masing-masing negara. Sebagaimana membangun armada angkatan udara, pembangunan angkatan laut bergantung pada perkembangan ilmu pengetahuan dan teknologi khususnya di dunia perairan. Proses perancangan kapal-kapal perang dengan kecepatan dan kelengkapan teknologi yang mumpuni akan sangat dipertimbangkan. Selain itu, membangun angkatan laut membutuhkan biaya dan sumber daya yang besar tidak hanya untuk membangun jumlah armada yang dibutuhkan, namun juga untuk menjaga agar armada yang ada dapat beroperasi dengan baik.

Umumnya, kemampuan negara-negara dalam menjaga kelangsungan operasi angkatan laut juga bergantung pada kemampuan perekonomian suatu negara. Namun daripada itu, hal tersebut juga dapat berarti bahwa mengoperasikan angkatan laut yang kuat juga dapat dianggap atau dipandang sebagai investasi untuk menjaga perekonomiannya. Maka dari hal tersebut, kelangsungan kehidupan angkatan laut bergantung penuh kepada kebijakan politik dan ekonomi suatu negara. Pengaruh pandangan suatu negara terhadap perkembangan politik di kawasan negara tersebut atau yang lebih dikenal sebagai geopolitik, juga menjadi salah satu faktor dalam kelangsungan hidup angkatan laut.

### 2.6.2 Tahanan Kapal

Resistensi atau tahanan kapal pada kecepatan tertentu adalah gaya yang diperlukan untuk menarik kapal pada kecepatan itu dalam air yang halus, dengan asumsi tidak ada gangguan dari kapal yang ditarik. Jika lambung tidak memiliki pelengkap, ini disebut resistensi *bare-hull*. Kapal yang bergerak dengan kecepatan tertentu pasti akan mengalami tahanan (*resistance*) yang besarnya tergantung dari banyak faktor dari kapal itu sendiri, namun hal yang paling utama karena arahnya yang berlawanan dengan arah gerak kapal itu sendiri (Burhanudin, 2012). Rumus untuk tahanan total berdasarkan metode Guldhammer dan Harvald disajikan pada persamaan (2.27)

$$R_T = \frac{1}{2} \rho C_T S V^2 \quad (2.27)$$

dimana,

- $R_T$  : Total Tahanan Kapal (kN)
- $\rho$  : Masa Jenis Air
- $C_T$  : Koefisien Total Kapal
- $S$  : Area Dibasahi (m<sup>2</sup>)
- $V$  : Kecepatan kapal (knot atau m/s)

Kapal memerlukan kekuatan tertentu untuk mencapai kecepatan yang diinginkan. Kekuatan tersebut harus menyesuaikan dengan besarnya total tahanan kapal agar kapal mampu melawan tahanan kapal dan dapat melaju dengan baik. Kekuatan yang diperlukan untuk mengatasi perlawanan ini disebut sebagai *towrope* atau kekuatan efektif dan dijabarkan pada persamaan (2.28) berikut ini (Lewis, 1988).

$$P_E = R_T V \quad (2.28)$$

dimana,

- $P_E$  : Daya Efisien Mesin Kapal (kW)

## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari pengujian yang dilakukan oleh Balai Teknologi Hidrodinamika Badan Pengkajian dan Penerapan Teknologi (BTH-BPPT). Data berasal dari hasil pengujian tahanan kapal patroli yang dilakukan pada towing tank dari tahun 2011 sampai 2018. Proses pengujian berlangsung selama 1-2 bulan mulai dari pembuatan purwarupa, pengujian tahanan kapal, hingga pembuatan laporan pengujian. Setiap pengamatan uji tahanan kapal dilakukan selama 7-14 hari.

### 3.2 Variabel Penelitian

Variabel yang digunakan pada penelitian ini terdiri dari delapan variabel prediktor yang diduga berpengaruh terhadap prediksi tahanan kapal cepat. Pemilihan delapan variabel prediktor tersebut merujuk dari persamaan perhitungan tahanan total yang ada pada buku karya Roh dan Lee (2018) dan juga hasil diskusi dengan pihak BTH-BPPT. Variabel prediktor yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.1 sebagai berikut.

**Tabel 3.1** Variabel Prediktor Penelitian

Variabel	Simbol	Deskripsi	Jenis data
$X_1$	$L_{WL}$	Panjang kapal tercelup air (m)	Rasio
$X_2$	$B$	Lebar kapal tercelup air (m)	Rasio
$X_3$	$T_F$	Sarat tercelup air bagian depan (m)	Rasio
$X_4$	$Disp/\Delta$	Total keseluruhan berat kapal (Ton)	Rasio
$X_5$	$C_B$	Koefisien Blok	Rasio
$X_6$	$C_M$	Koefisien <i>Midship</i>	Rasio
$X_7$	$S$	Area Dibasahi (m <sup>2</sup> )	Rasio
$X_8$	$V^2$	Kecepatan Kapal (knot)	Rasio

Variabel respon merupakan variabel yang bersifat dependen terhadap variabel prediktor (Sugiyono, 2011). Variabel respon yang digunakan pada penelitian ini disajikan pada Tabel 3.2 seperti berikut ini.

**Tabel 3.2** Variabel Respon Penelitian

	Variabel	Deskripsi	Jenis data
Y	$R_T$	Tahanan Total Kapal (kN)	Rasio

### 3.3 Definisi Operasional

Variabel yang digunakan dalam penelitian ini terdiri dari dua variabel respon dan delapan variabel prediktor. Adapun variabel-variabel tersebut adalah sebagai berikut.

1. Panjang kapal menjelaskan ukuran panjang keseluruhan kapal dari depan hingga buritan kapal yang tercelup air dengan satuan meter.
2. Lebar kapal merupakan ukuran lebar lambung kapal yang tercelup air saat dioperasikan dengan satuan meter.
3. Sarat tercelup air adalah jarak vertikal antara garis air sampai dengan lunas kapal dan diukur ketika kapal bermuatan.
4. Koefisien blok adalah perbandingan antara isi *carena* suatu balok yang dibatasi oleh ukuran panjang, lebar, dan tinggi kapal. Koefisien blok tidak memiliki satuan tertentu.
5. Koefisien *midship* adalah koefisien tanpa satuan yang menjelaskan perbandingan antara luas penampang gading besar yang terendam air dengan luas suatu penampang.
6. Area dibasahi atau *Wetted Surface Area* adalah luas lambung kapal yang tercelup di dalam air.
7. Tahanan total kapal merupakan gaya tahanan total dari kapal setelah melakukan pengujian dan memiliki satuan kilo newton.

### 3.4 Struktur Data

Struktur data yang akan dianalisis menggunakan regresi *Ridge* ditunjukkan pada Tabel 3.3 sebagai berikut.

**Tabel 3.3** Struktur Data Penelitian

Observasi	Variabel			
	$X_1$	...	$X_8$	Y
1	$X_{1,1}$	...	$X_{8,1}$	$Y_1$
2	$X_{1,2}$	...	$X_{8,2}$	$Y_2$
3	$X_{1,3}$	...	$X_{8,3}$	$Y_3$
⋮	⋮	⋮	⋮	⋮
n	$X_{1,n}$	...	$X_{8,n}$	$Y_n$

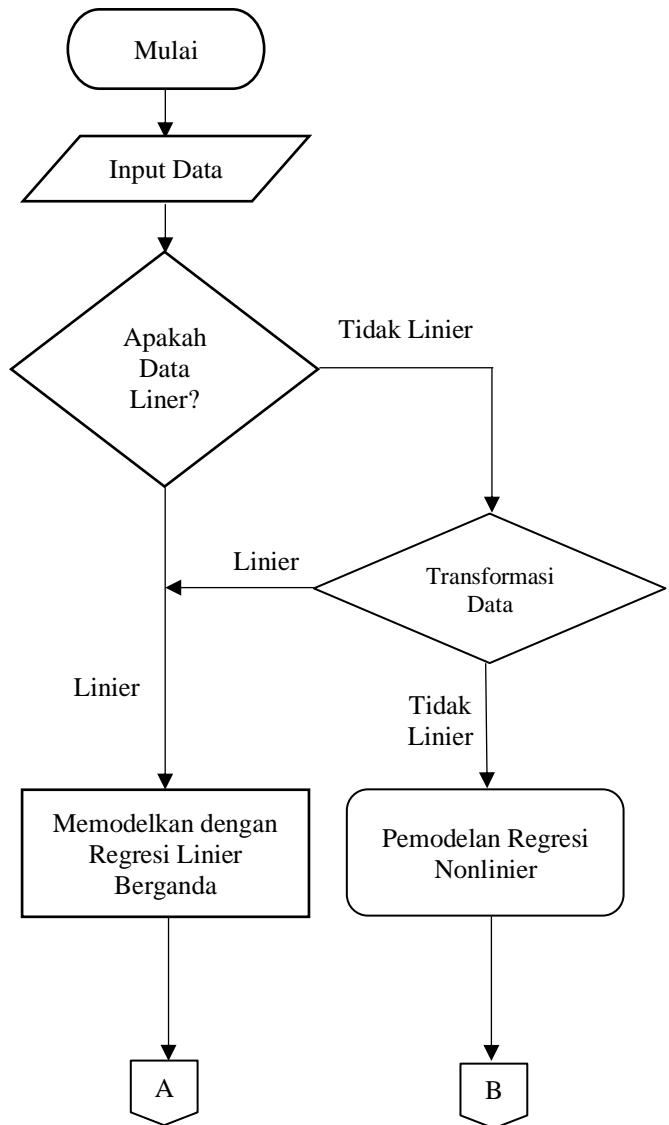
dengan n merupakan banyaknya data pengamatan.

### 3.5 Langkah Penelitian

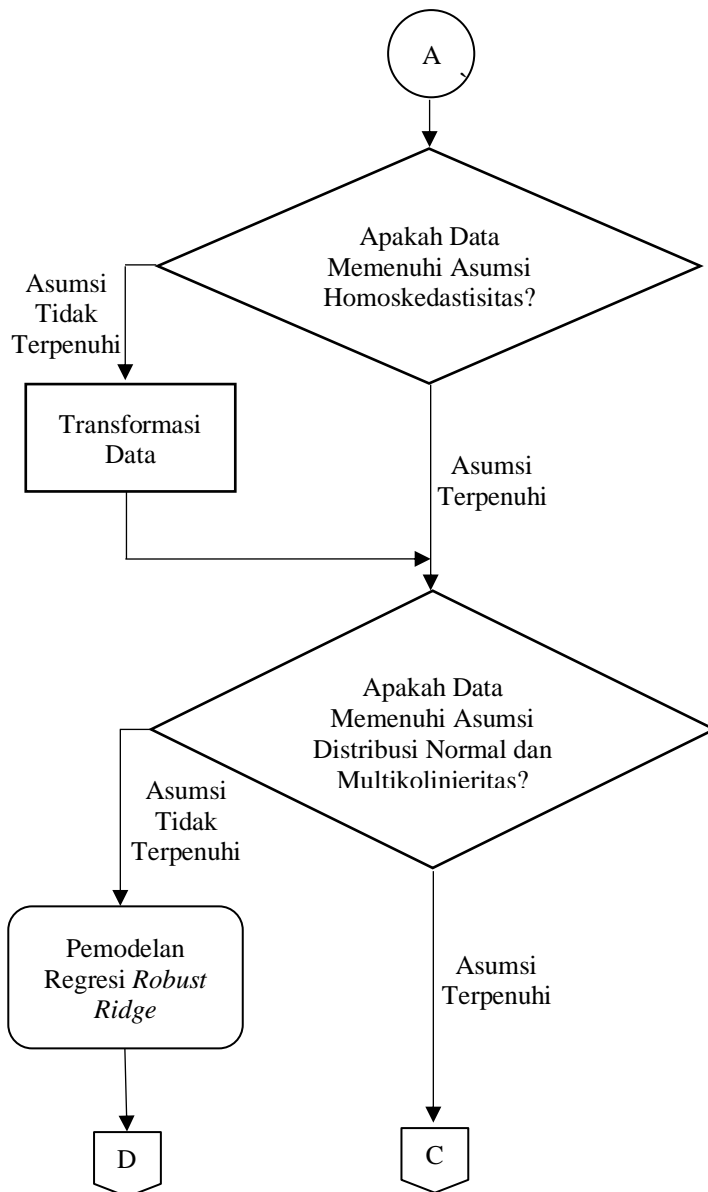
Langkah-langkah penelitian yang akan dilakukan berdasarkan dengan tujuan pada penelitian ini adalah sebagai berikut.

1. Mengumpulkan data penelitian yang berupa data sekunder.
2. Menentukan variabel yang akan digunakan dalam penelitian.
3. Eksplorasi data dan mendeskripsikan faktor-faktor yang diduga mempengaruhi pemodelan tahanan kapal.
4. Melakukan uji linieritas data menggunakan *Ramsey RESET Test*.
5. Melakukan pemodelan menggunakan analisis regresi dan pendugaan parameter.
6. Melakukan uji asumsi klasik regresi linier berganda, yaitu :
  - a. Homoskedastisitas
  - b. Autokorelasi
  - c. Distribusi Normal
  - d. Multikolinieritas
7. Menentukan ulang metode regresi yang sesuai dengan hasil uji asumsi dan memodelkan kembali.
8. Melakukan uji signifikansi parameter.
9. Melakukan perhitungan RMSE dan  $R^2$  dari model.
10. Mengintepretasi model terpilih.
11. Menarik kesimpulan dan saran dari hasil analisis.

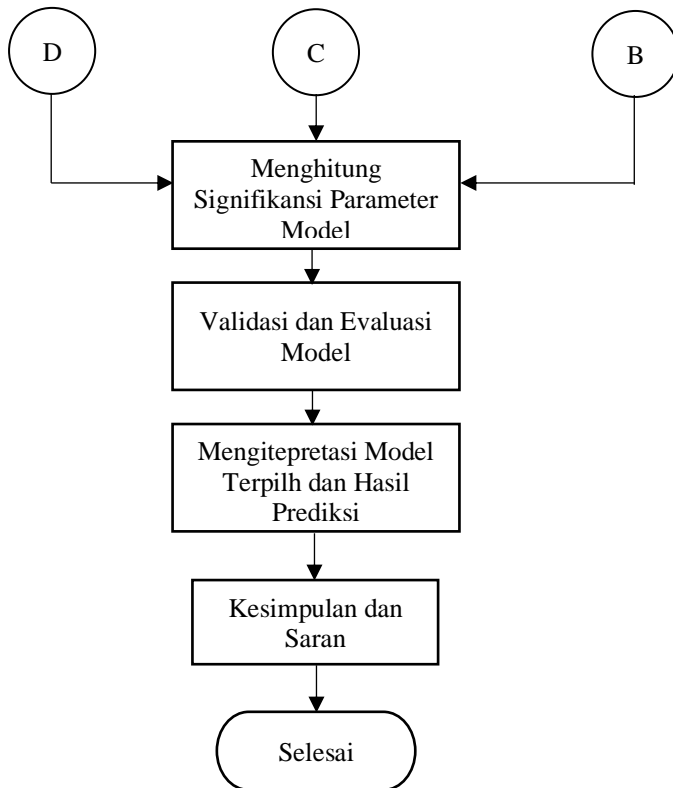
Berdasarkan langkah-langkah penelitian yang akan dilakukan dan penjelasan sebelumnya, pada penelitian ini akan menggunakan 1 diagram alir. Berikut akan diberikan diagram alir yang dapat dilihat pada Gambar 3.1.



**Gambar 3.1** Diagram Alir Penelitian



**Gambar 3.1** Diagram Alir Penelitian (lanjutan)



**Gambar 3.1** Diagram Alir Penelitian (lanjutan)



## BAB IV ANALISIS DAN PEMBAHASAN

### 4.1 Statistika Deskriptif

Statistika deskriptif adalah statistika yang berkaitan dengan metode atau cara mendeskripsikan, menggambarkan, menjabarkan atau menguraikan data. Statistika deskriptif didasarkan tentang eksplorasi dan menyajikan data dengan baik. Statistika deskriptif yang dapat dilakukan misalnya dengan menentukan nilai rata-rata hitung, median, modus, standar deviasi. Cara lain dalam mendeskripsikan sebuah data, yaitu dengan membuat plot atau grafik distribusi data. Informasi mengenai seluruh variabel penelitian disajikan dalam Tabel 4.1.

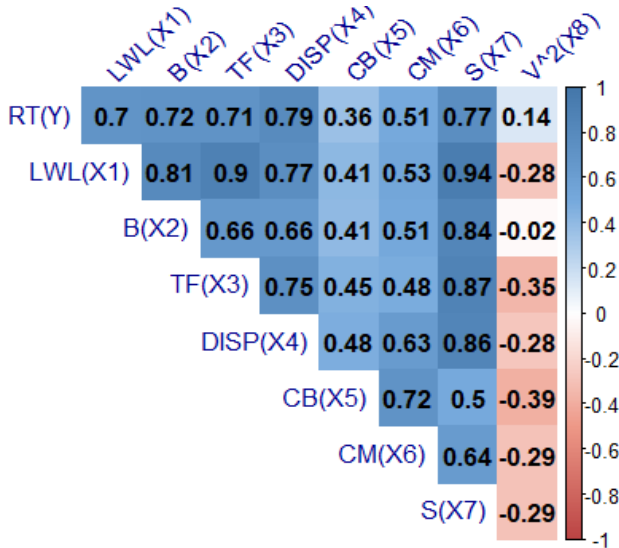
**Tabel 4.1** Statistika Deskriptif Variabel Penelitian

Variabel	Symbol	Min	Max	Med	Mean	Var
Y	$R_T$	6,80	461,00	141,50	158,80	7610,76
X <sub>1</sub>	$L_{WL}$	8,89	76,15	42,80	42,99	276,91
X <sub>2</sub>	$B$	2,80	16,00	7,80	7,73	6,71
X <sub>3</sub>	$T_F$	0,45	3,00	1,80	1,82	0,42
X <sub>4</sub>	$Disp/\Delta$	4,00	1626,00	235,40	327,60	112062,70
X <sub>5</sub>	$C_B$	0,23	0,54	0,40	0,42	0,01
X <sub>6</sub>	$C_M$	0,40	0,93	0,60	0,63	0,01
X <sub>7</sub>	$S$	19,35	1041,88	303,47	355,65	5309,00
X <sub>8</sub>	$V^2$	289,00	784,00	484,00	509,80	21511,88

Melalui Tabel 4.1 telah disajikan informasi sederhana dari variabel penelitian. Informasi sederhana seperti tabel di atas didapatkan dari total 37 kapal cepat yang menjadi objek penelitian. Diketahui berdasarkan informasi pada Tabel 4.1, tahanan kapal terkecil dari objek penelitian berada pada angka 6,8 kN, sedangkan tahanan kapal terbesar berada pada angka 461 kN. Rata-rata tahanan kapal cepat didapatkan sebesar 158,8 kN dengan variansi data sebesar 7610,76. Selain itu, kecepatan kapal berdasarkan data pengujian berada pada nilai 17 knot untuk yang terkecil dan 28 knot untuk yang terbesar. Angka 17 knot dan 28 knot didapatkan dari akar nilai yang ada pada Tabel 4.1 variabel X<sub>8</sub>.

Statistika deskriptif lain yang dapat digunakan adalah melihat hubungan antar variabel penelitian. Menghitung besaran korelasi antar variabel penelitian menjadi solusi yang tepat untuk men-

deteksi seberapa kuat hubungan antar dua variabel. Visualisasi matriks korelasi disajikan pada Gambar 4.1 untuk mempermudah pembaca dalam melihat besaran korelasi antar variabel.



**Gambar 4.1** Matriks Korelasi Variabel Penelitian

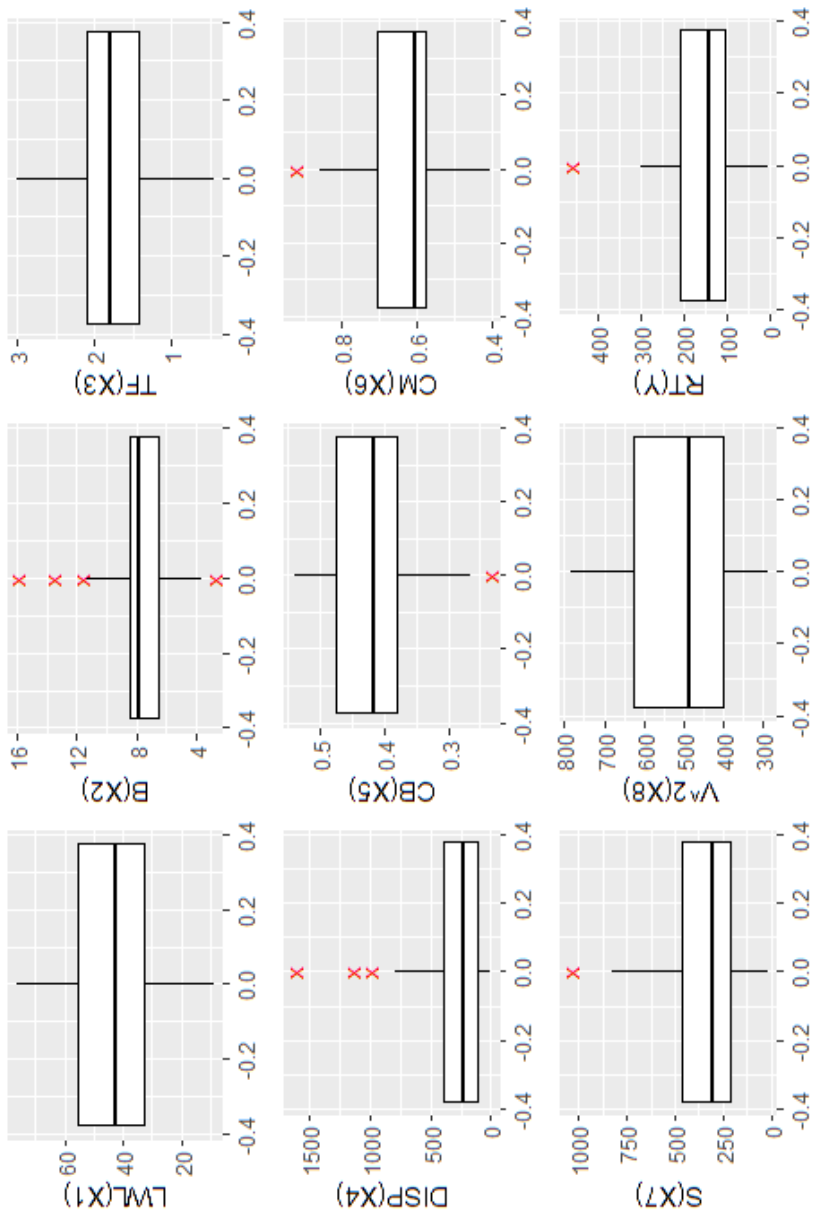
Melalui Gambar 4.1 dapat dilihat hubungan antar variabel penelitian mulai dari variabel respon hingga masing-masing variabel prediktor. Besaran korelasi bernilai antara -1 sampai dengan 1. Jika besaran korelasi mendekati angka 1 atau -1 menandakan bahwa korelasi yang dimiliki kedua variabel tersebut sangatlah kuat. Sebaliknya, apabila besaran korelasi mendekati angka 0 maka dapat dinyatakan bahwa antar kedua variabel memiliki korelasi yang lemah. Pada Gambar 4.1 semakin gelap warna biru maka menunjukkan kedua variabel memiliki korelasi positif yang semakin kuat. Sementara semakin gelap warna merah maka menunjukkan kedua variabel memiliki korelasi negatif yang semakin kuat. Jika warna pada Gambar 4.1 semakin muda atau mendekati ke arah warna putih, maka dapat dinyatakan kedua variabel memiliki korelasi yang sangat lemah.

Informasi lain yang disajikan pada matriks korelasi di atas, yaitu hubungan antara variabel  $R_T$  (Y) dengan variabel  $\text{Disp}/\Delta$  ( $X_4$ )

merupakan korelasi terkuat yang terjadi antara variabel respon dengan variabel prediktor dengan nilai korelasi sebesar 0,79. Sementara korelasi terendah antara variabel respon dan prediktor terjadi pada variabel  $R_T$  (Y) dengan variabel  $V^2$  ( $X_8$ ) dengan nilai korelasi sebesar 0,14. Variabel prediktor yang memiliki korelasi positif yang kuat terhadap variabel  $R_T$  (Y), diantaranya  $L_{WL}$  ( $X_1$ ),  $B$  ( $X_2$ ),  $T_F$  ( $X_3$ ),  $Disp/\Delta$  ( $X_4$ ),  $C_M$  ( $X_6$ ), dan  $S$  ( $X_7$ ). Sementara variabel  $C_B$  ( $X_5$ ) dan  $V^2$  ( $X_8$ ) memiliki nilai korelasi positif yang lemah terhadap variabel respon. Nilai korelasi yang paling tinggi diantara semua variabel dimiliki oleh variabel  $L_{WL}$  ( $X_1$ ) dengan variabel  $S$  ( $X_7$ ). Nilai korelasi kedua variabel tersebut sebesar 0,94 yang menunjukkan kedua variabel memiliki hubungan yang sangat kuat. Sedangkan nilai korelasi terkecil dimiliki oleh variabel  $B$  ( $X_2$ ) dengan variabel  $V^2$  ( $X_8$ ). Nilai korelasi kedua variabel tersebut sebesar -0,02 yang berarti hubungan kedua variabel sangat lemah atau bahkan hampir tidak terdapat hubungan karena hampir mendekati angka 0.

Hubungan antara variabel  $R_T$  (Y) dengan masing-masing variabel prediktor berkorelasi positif. Hubungan positif yang kuat antara variabel  $R_T$  (Y) dengan masing-masing variabel prediktor mengindikasikan bahwa parameter masing-masing variabel prediktor akan memiliki tanda yang positif juga. Apabila terdapat parameter yang bertanda negatif, maka dapat dipastikan bahwa multikolinieritas terjadi di dalam model.

Statistika deskriptif yang diperuntukkan untuk melihat bagaimana persebaran data dari masing-masing variabel prediktor dan juga variabel respon, dapat digunakan *boxplot* yang disajikan pada Gambar 4.2. Selain hal tersebut, *boxplot* juga berguna untuk mendeteksi ada atau tidaknya *outlier* dari masing-masing variabel. *Outlier* merupakan data observasi yang memiliki nilai yang berbeda jauh diantara data observasi lainnya baik secara univariat ataupun multivariat. *Outlier* pada *boxplot* digambarkan seperti tanda “x” berwarna merah pada masing-masing *boxplot*. Setiap *boxplot* terdapat tiga garis yang menjadikan batas-batas atau penanda bagi data. Garis atas pada *boxplot* merupakan kuantil 3 dari data, sedangkan garis tengah dan bawah merupakan kuantil 2 dan kuantil 1 dari data.

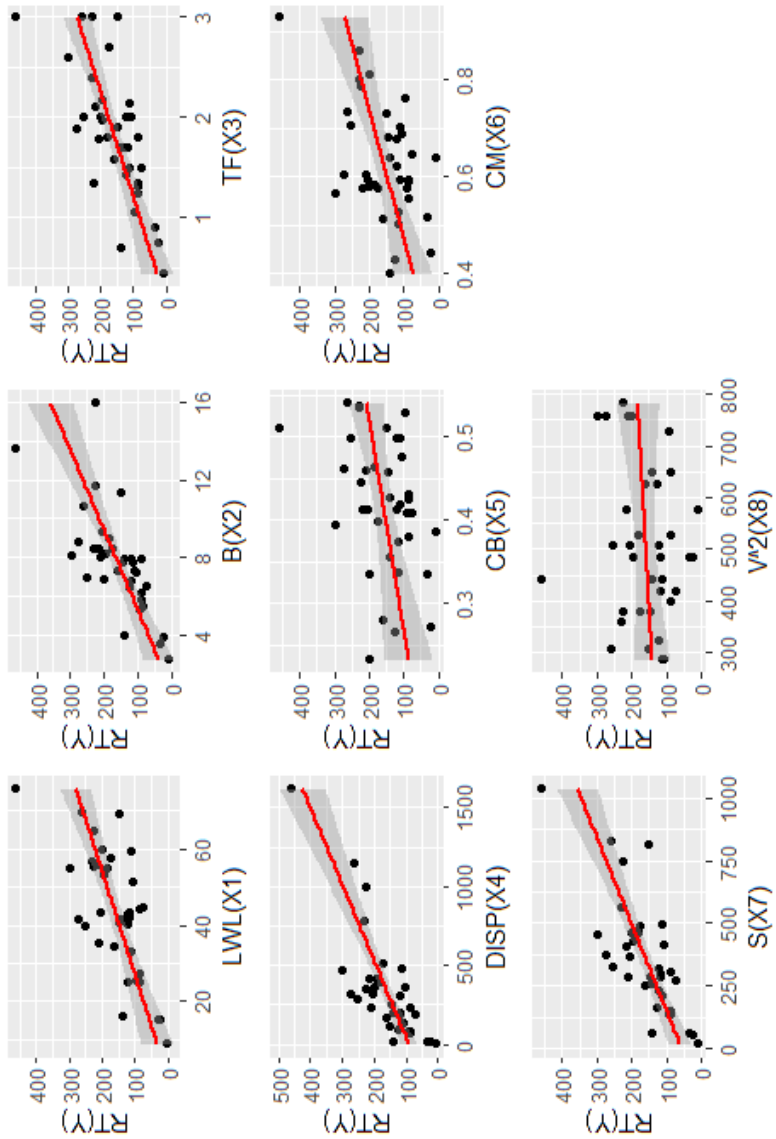


**Gambar 4.2** *Boxplot* Variabel  $X_1 - X_8$  dan  $Y$

Berdasarkan Gambar 4.2, dapat dilihat bahwa variabel  $L_{WL}$  ( $X_1$ ) memiliki nilai tengah (median) di angka 42,8 meter. Sedangkan variabel  $B$  ( $X_2$ ) memiliki nilai tengah sebesar 7,8. Variabel  $B$  ( $X_2$ ) memiliki data yang bersifat *outlier* sebanyak empat pengamatan. *Outlier* yang terdapat pada variabel  $Disp/\Delta$  ( $X_4$ ) sebanyak tiga pengamatan. Sementara pada variabel  $T_F$  ( $X_3$ ) tidak ditemukan *outlier* yang dibuktikan dengan tidak adanya tanda silang berwarna merah pada *boxplot* variabel  $T_F$  ( $X_3$ ).

Pada variabel  $C_B$  ( $X_5$ ) dan  $C_M$  ( $X_6$ ), *outlier* yang terdeteksi berjumlah sebanyak satu pengamatan. Nilai *outlier* yang terdapat pada variabel  $C_B$  ( $X_5$ ) bernilai pada kisaran 0,2 yang juga merupakan data terkecil pada variabel tersebut. Nilai *outlier* yang terdapat pada variabel  $C_M$  ( $X_6$ ) bernilai pada kisaran 0,9 yang juga merupakan data terbesar pada variabel  $C_M$  ( $X_6$ ). Melalui Gambar 4.2 juga dapat diketahui bahwa variabel kecepatan atau  $V^2$  ( $X_8$ ) tidak memiliki data yang bersifat *outlier*. Sedangkan pada variabel  $S$  ( $X_7$ ) terdapat satu data *outlier* yang merupakan data terbesar dari variabel  $S$  ( $X_7$ ). Setelah dilakukan visualisasi karakteristik data variabel prediktor menggunakan *boxplot*, dapat dirangkum bahwa lima dari delapan variabel prediktor terdapat data *outlier*. Variabel-variabel yang terdapat *outlier* diantaranya, yaitu variabel  $B$  ( $X_2$ ),  $Disp/\Delta$  ( $X_4$ ),  $C_B$  ( $X_5$ ),  $C_M$  ( $X_6$ ), dan  $S$  ( $X_7$ ).

Untuk mendeteksi *outlier* pada variabel respon dapat juga dilakukan pada Gambar 4.2 menunjukkan bahwa variabel respon memiliki satu *outlier* pada data yang merupakan data terbesar pada variabel respon. Data *outlier* tersebut merupakan data terbesar yang dimiliki oleh variabel  $R_T$  ( $Y$ ). Statistika deskriptif lain yang dapat digunakan dalam penelitian ini adalah memvisualisasikan kecenderungan hubungan antara variabel respon dan variabel prediktor. Teknik visualisasi yang digunakan adalah *scatterplot* dengan tambahan *fit line* di dalamnya untuk membantu proses pendugaan apakah antar variabel tersebut memiliki kecenderungan hubungan linier atau tidak. Semakin dekat sebaran data pada garis linier maka hubungan linier pada pengamatan tersebut semakin kuat. *Scatterplot* seperti yang telah disebutkan disajikan dalam Gambar 4.3.



**Gambar 4.3** Scatterplot Variabel Y dengan X

Hasil visualisasi pola hubungan dari masing-masing variabel seperti pada Gambar 4.3 akan digunakan sebagai pertimbangan deteksi hubungan linieritas antara variabel respon dan variabel prediktor. *Scatterplot* antara variabel  $R_T$  (Y) dengan variabel  $L_{WL}$  ( $X_1$ ) yang disajikan pada Gambar 4.3 mengindikasikan hubungan yang linier. Hal ini tergambarkan melalui sebaran data yang mengikuti garis linier (*fit line*). Pola hubungan antara variabel  $R_T$  (Y) dengan variabel B ( $X_2$ ) memiliki kecenderungan hubungan yang linier. Meskipun begitu, hubungan linier antara  $R_T$  (Y) dengan B ( $X_2$ ) belum dapat dipastikan melalui visualisasi karena terdapat kecenderungan penumpukan data pada satu titik meskipun pesebarannya mengikuti garis linier (*fit line*). Identifikasi hubungan linier melalui visualisasi *scatterplot* juga dilakukan terhadap variabel penelitian lain. Melalui visualisasi pada Gambar 4.3 dapat diduga bahwa antara variabel  $R_T$  (Y) dengan  $T_F$  ( $X_3$ ) memiliki hubungan yang linier. Hal tersebut terlihat dari pesebaran data yang mengikuti garis linier.

Hubungan antara variabel  $R_T$  (Y) dengan  $Disp/\Delta$  ( $X_4$ ) diduga tidak memiliki hubungan yang linier berdasarkan visualisasi *scatterplot* yang disajikan pada Gambar 4.3. Pola sebaran data cenderung berkumpul pada satu bagian dan tidak mengikuti garis linier. Untuk kasus seperti ini, diperlukan pengujian linieritas data agar didapat-kan keyakinan penuh apakah hubungan antara variabel  $R_T$  (Y) dengan  $Disp/\Delta$  ( $X_4$ ) bersifat linier atau tidak.

Terlihat melalui visualisasi Gambar 4.3, variabel  $C_B$  ( $X_5$ ) diduga memiliki hubungan linier dengan variabel  $R_T$  (Y). Sebaran data variabel  $C_B$  ( $X_5$ ) dengan variabel  $R_T$  (Y) telah mengikuti garis linier sehingga timbul dugaan bahwa kedua variabel tersebut memiliki hubungan yang linier. Meskipun demikian, terdapat salah satu pengamatan yang berada cukup jauh jaraknya dari garis linier. Pengamatan tersebut berjarak jauh karena adanya nilai pengamatan variabel  $R_S$  (Y) yang cukup ekstrem. Pola hubungan antara variabel  $R_T$  (Y) dan  $C_M$  ( $X_6$ ). Berdasarkan Gambar 4.3, diketahui variabel  $C_M$  ( $X_6$ ) memiliki kecenderungan hubungan yang linier terhadap variabel  $R_T$  (Y) dikarenakan sebaran data antar pengamatan saling berdekatan dan cenderung mengikuti garis linier.

Selain itu, berdasarkan Gambar 4.3 juga didapatkan informasi bahwa antara variabel  $R_T$  (Y) dan S ( $X_7$ ) cenderung tidak memiliki hubungan yang linier. Hal tersebut dapat dilihat dengan pola sebaran data yang cenderung berkumpul pada satu daerah dan sebaran data tidak terlalu mengikuti garis linier. Pola hubungan yang dimiliki antara variabel respon  $R_T$  (Y) dan variabel prediktor  $V^2$  ( $X_8$ ) diduga bersifat linier. Dugaan tersebut dikarenakan pola sebaran data yang memiliki kecenderungan untuk mengikuti garis linier. Informasi lain yang didapat yaitu, sebaran data antara variabel respon  $R_T$  (Y) dan variabel prediktor  $V^2$  ( $X_8$ ) pada Gambar 4.3 memiliki jarak yang relatif berjauhan antar data pengamatan. Meskipun demikian, *scatterplot* seperti Gambar 4.3 tidak dapat dijadikan dasar pengambilan keputusan hubungan linier antar dua variabel.

Melalui visualisasi *scatterplot* yang telah dilakukan pada Gambar 4.3 dapat dikerucutkan bahwa secara visual variabel prediktor yang tidak memiliki hubungan linier terhadap variabel respon adalah variabel B ( $X_2$ ), Disp/ $\Delta$  ( $X_4$ ), dan S ( $X_7$ ). Visualisasi hubungan linier antara variabel respon dengan prediktor menggunakan *scatterplot* tidak dapat dijadikan acuan utama dalam pengambilan keputusan. Maka dari itu, pemantauan hubungan linier antara variabel respon dengan variabel prediktor tidak cukup jika hanya melalui visualisasi *scatterplot*. Untuk mengetahui hubungan linier antar dua variabel secara objektif diperlukan pengujian linieritas data. Seperti yang telah diketahui bahwa metode regresi linier berganda merupakan salah satu metode statistika parametrik sehingga diperlukan pengujian asumsi-asumsi klasik yang salah satunya adalah uji asumsi linieritas data. Langkah selanjutnya yang harus dilakukan adalah melakukan pengujian hubungan linieritas data yang pada penelitian ini menggunakan metode *Ramsey RESET Test*.

## 4.2 Uji Asumsi Linieritas Data

Uji linieritas data diperlukan sebagai salah satu syarat pemenuhan asumsi jika menggunakan analisis regresi linier berganda. Metode pengujian yang digunakan dalam penelitian adalah *Ramsey RESET Test*.  $H_0$  pada pengujian ini adalah data memiliki hubungan yang linier. Jika hasil pengujian memberikan keputusan untuk



menolak  $H_0$ , maka data tidak memiliki hubungan linier. Keputusan untuk menolak  $H_0$  dilakukan jika besaran  $p$ -value yang didapat melebihi dari taraf signifikan  $\alpha = 5\%$  atau sebesar 0,05. Hasil dari Ramsey RESET Test dapat dilihat pada Tabel 4.2.

**Tabel 4.2** Hasil Uji Linieritas Data

Variabel	RESET Statistik	df <sub>1</sub>	df <sub>2</sub>	$p$ -value
Y ~ X <sub>1</sub>	3,0298	2	33	0,0619
Y ~ X <sub>2</sub>	1,9798	2	33	0,1542
Y ~ X <sub>3</sub>	0,0027	2	33	0,9973
Y ~ X <sub>4</sub>	5,8913	2	33	0,0064
Y ~ X <sub>5</sub>	1,2285	2	33	0,3058
Y ~ X <sub>6</sub>	4,5989	2	33	0,0173
Y ~ X <sub>7</sub>	4,2261	2	33	0,0232
Y ~ X <sub>8</sub>	2,4597	2	33	0,1010

Tabel 4.2 di atas memperlihatkan hasil pengujian linieritas data antara variabel respon dengan variabel prediktor. Melihat besaran  $p$ -value, diketahui bahwa variabel yang memiliki hubungan linier terhadap variabel R<sub>T</sub> (Y) adalah variabel L<sub>WL</sub> (X<sub>1</sub>), B (X<sub>2</sub>), T<sub>F</sub> (X<sub>3</sub>), C<sub>B</sub> (X<sub>5</sub>), dan V<sup>2</sup> (X<sub>8</sub>). Untuk itu, dapat diputuskan variabel-variabel yang tidak memiliki hubungan linier terhadap variabel R<sub>T</sub> (Y), yaitu Disp/Δ (X<sub>4</sub>), C<sub>M</sub> (X<sub>6</sub>), dan S (X<sub>7</sub>). Seperti yang telah dibahas pada bagian sebelumnya, bahwa asumsi linieritas antara variabel respon dengan masing-masing variabel prediktor harus dipenuhi agar dapat menggunakan regresi linier berganda. Langkah selanjutnya yang dapat dilakukan adalah melakukan transformasi data variabel prediktor hingga asumsi linieritas dapat terpenuhi. Metode transformasi data yang dilakukan harus berlandaskan pada pola sebaran data sehingga data hasil transformasi menjadi linier. Metode transformasi data penelitian yang dilakukan disajikan pada Tabel 4.3.

**Tabel 4.3** Hasil Uji Linieritas Data Transformasi

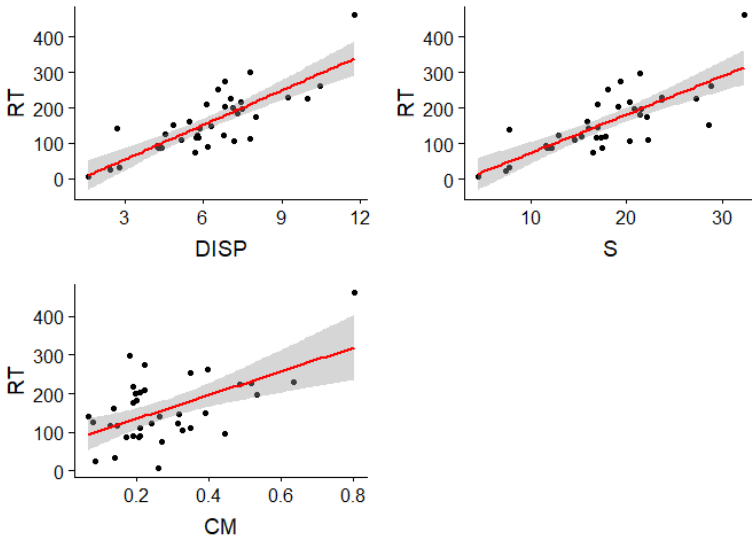
Variabel	Metode Transformasi	Variabel	Metode Transformasi
Y	Y	X <sub>5</sub>	X <sub>5</sub>
X <sub>1</sub>	X <sub>1</sub>	X <sub>6</sub>	(X <sub>6</sub> ) <sup>3</sup>
X <sub>2</sub>	X <sub>2</sub>	X <sub>7</sub>	(X <sub>7</sub> ) <sup>1/2</sup>
X <sub>3</sub>	X <sub>3</sub>	X <sub>8</sub>	X <sub>8</sub>
X <sub>4</sub>	(X <sub>4</sub> ) <sup>1/3</sup>		

Transformasi hanya dilakukan pada variabel prediktor yang tidak memiliki hubungan linier terhadap variabel respon. Setelah proses transformasi data dilakukan, uji linieritas data yang telah ditransformasi perlu dilakukan kembali untuk memastikan variabel tersebut telah linier. Tabel 4.4 berisikan hasil *Ramsey RESET Test* dari variabel yang telah dilakukan transformasi data.

**Tabel 4.4** Hasil Uji Linieritas Data Transformasi

Variabel	RESET Statistik	df <sub>1</sub>	df <sub>2</sub>	<i>p-value</i>
$Y \sim (X_4)^{1/3}$	1,8751	2	33	0,1693
$Y \sim (X_6)^3$	3,2638	2	33	0,0508
$Y \sim (X_7)^{1/2}$	1,3255	2	33	0,2794

Berdasarkan hasil Tabel 4.3 terlihat jelas bahwa variabel  $\text{Disp}/\Delta$  ( $X_4$ ),  $C_M$  ( $X_6$ ), dan  $S$  ( $X_7$ ) telah linier terhadap variabel  $R_T$  ( $Y$ ). Keputusan tersebut didapat berdasarkan *p-value* yang bernilai lebih besar dari  $\alpha = 5\%$ . Visualisasi hasil uji linieritas pada masing-masing variabel prediktor yang telah dilakukan transformasi data disajikan pada Gambar 4.4.



**Gambar 4.4** Scatterplot Variabel Y dengan Transformasi  $X_4$ ,  $X_6$ ,  $X_7$

Gambar 4.4 merupakan visualisasi hubungan antara variabel  $R_T$  ( $Y$ ) dengan variabel prediktor yang telah dilakukan transfor-

masi. Terlihat melalui Gambar 4.4, sebaran data antara variabel  $R_T$  (Y) dengan  $Disp/\Delta$  ( $X_4$ ) telah mengikuti garis linier yang mana pada Gambar 4.3 data cenderung berkumpul pada satu bagian dan tidak mengikuti garis linier. Selain itu, visualisasi pola hubungan antara variabel  $R_T$  (Y) dan  $C_M$  ( $X_6$ ) yang telah dilakukan transformasi memberikan hasil bahwa data telah mengikuti garis linier dan sebaran data berjarak semakin mendekati garis linier. Hasil dari transformasi data S ( $X_7$ ) menjadikan sebaran data variabel  $R_T$  (Y) dan S ( $X_7$ ) mengikuti garis linier. Pada visualisasi Gambar 4.3 terlihat sebaran data variabel  $R_T$  (Y) dan S ( $X_7$ ) cenderung berkumpul pada satu bagian dan tidak mengikuti garis linier.

Untuk perbandingan antara data pengamatan sebelum dan sesudah ditransformasi dapat dilihat pada Lampiran 1 dan Lampiran 2. Langkah selanjutnya yang dapat dilakukan adalah memodelkan dataset yang telah ditransformasi menggunakan regresi linier berganda untuk melakukan pengujian asumsi lain seperti homoskedastisitas, independensi, dan distribusi normal yang didasarkan pada residual model. Model regresi linier berganda yang didapat disajikan pada persamaan (4.1).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4^{1/3} + \beta_5 X_5 + \beta_6 X_6^3 + \beta_7 X_7^{1/2} + \beta_8 X_8 \quad (4.1)$$

### 4.3 Uji Asumsi Homoskedastisitas

Sebelum masuk ke dalam pengujian asumsi klasik seperti yang telah disebutkan sebelumnya, diperlukan pemodelan regresi linier berganda untuk didapatkan model linier. Tujuan dari pembentukan model untuk memeriksa jika asumsi klasik regresi linier berganda telah dipenuhi oleh model linier. Bentuk model regresi linier berganda tersebut disajikan pada Lampiran.

Berdasarkan model linier terlampir, uji asumsi pertama yang dapat dilakukan setelah melakukan uji linieritas data adalah uji asumsi homoskedastisitas.  $H_0$  pada pengujian ini adalah data bersifat homoskedastisitas. Metode yang digunakan untuk memeriksa pemenuhan asumsi homoskedastisitas adalah *Breusch-Pagan Test*. Hasil dari uji homoskedastisitas disajikan pada Tabel 4.5.

**Tabel 4.5** Hasil Uji Asumsi Homoskedastisitas

Metode	Statistik Uji	df <sub>1</sub>	<i>p-value</i>	Keterangan
<i>Breusch-Pagan</i>	0,0003	1	0,9856	Gagal Tolak $H_0$

Tabel 4.5 menunjukkan bahwa hasil pengujian asumsi homoskedastisitas menggunakan metode *Breusch-Pagan* memberikan keputusan gagal menolak  $H_0$ . Berdasarkan keputusan tersebut dapat ditarik kesimpulan bahwa model telah memenuhi asumsi homoskedastisitas. Maka dari hal tersebut, analisis selanjutnya yang dapat dilakukan adalah melakukan uji asumsi distribusi normal.

#### 4.4 Uji Asumsi Distribusi Normal

Uji asumsi selanjutnya adalah uji asumsi distribusi normal dengan menggunakan metode *Jarque-Bera Test*.  $H_0$  yang digunakan pada pengujian ini, yaitu data mengikuti pola distribusi normal. Keputusan menolak  $H_0$  dilakukan jika besaran *p-value* yang didapat melebihi dari taraf signifikan  $\alpha = 5\%$  atau sebesar 0,05. Hasil pengujian disajikan pada Tabel 4.6.

**Tabel 4.6** Hasil Uji Asumsi Distribusi Normal

Metode	Statistik Uji	df <sub>1</sub>	<i>p-value</i>	Keterangan
<i>Jarque-Bera</i>	95,427	18	0,0000	Tolak $H_0$

Hasil pengujian didapatkan besaran *p-value* sebesar  $0,000 \cong 0$  yang memberikan keputusan menolak  $H_0$  pada taraf signifikan  $\alpha$  sebesar 5%. Kesimpulan yang didapat berdasarkan hasil pengujian, yaitu model tidak mengikuti distribusi normal. Melalui proses analisis statistika deskriptif pada Gambar 4.2, dapat diduga dengan kuat bahwa *outlier* merupakan faktor utama penyebab pelanggaran asumsi distribusi normal oleh model. *Outlier* dapat diatasi dengan membuang (menghilangkan) pengamatan yang terdapat *outlier* dari proses penelitian. Meskipun begitu, pada kasus ini terlalu beresiko untuk menghilangkan pengamatan yang memiliki *outlier* mengingat jumlah pengamatan pada data penelitian yang relatif sedikit dan jumlah pengamatan yang mengandung *outlier* relatif banyak.

Regresi *robust* merupakan salah satu solusi dari permasalahan tidak terpenuhinya asumsi distribusi normal dan dapat digunakan jika hanya asumsi distribusi normal yang tidak terpenuhi. Maka dari hal tersebut langkah selanjutnya yang perlu dilakukan adalah membiarkan data tidak memenuhi asumsi distribusi normal dan

memeriksa satu asumsi terakhir yaitu, pemenuhan asumsi multikolinieritas. Keputusan metode regresi yang akan digunakan akan terjadi setelah semua pemeriksaan asumsi dilakukan.

#### 4.5 Pemenuhan Asumsi Multikolinieritas

Terdapat satu asumsi lain yang perlu dipenuhi dalam kasus regresi linier berganda, yaitu asumsi multikolinieritas. Pemeriksaan terhadap asumsi ini dapat menggunakan metode *Variance Inflation Factor* (VIF). Besaran nilai VIF didapatkan melalui koefisien determinasi dari regresi masing-masing variabel prediktor terhadap variabel prediktor lainnya. Keputusan bahwa terdapat multikolinieritas pada data terjadi apabila minimal satu variabel prediktor memiliki nilai VIF yang lebih besar dari atau sama dengan 10. Hasil perhitungan nilai VIF masing-masing variabel prediktor disajikan pada Tabel 4.7.

**Tabel 4.7** Hasil Uji Asumsi Multikolinieritas

Variabel	VIF <sub>i</sub>
X <sub>1</sub>	37,0689
X <sub>2</sub>	9,3290
X <sub>3</sub>	11,7261
X <sub>4</sub>	6,3864
X <sub>5</sub>	2,0662
X <sub>6</sub>	2,3992
X <sub>7</sub>	95,5422
X <sub>8</sub>	1,5171

Melalui Tabel 4.7 dapat dilihat bahwa terdapat tiga variabel prediktor dengan nilai VIF lebih dari 10. Sesuai dengan ketentuan mengenai pemenuhan asumsi multikolinieritas menggunakan VIF, maka diputuskan bahwa variabel X<sub>1</sub>, X<sub>3</sub>, dan X<sub>7</sub> tidak memenuhi asumsi multikolinieritas. Kesimpulan yang didapat berdasarkan hasil pengujian asumsi klasik regresi linier berganda, yaitu data (model) hanya memenuhi asumsi homoskedastisitas. Namun daripada itu, asumsi distribusi normal dan juga multikolinieritas tidak dapat terpenuhi sehingga metode regresi linier berganda tidak relevan untuk digunakan sebagai metode pemodelan dalam analisis lebih lanjut. Mengarah pada hasil pengujian dan pemenuhan asumsi regresi linier berganda, maka diputuskan untuk melakukan

pemodelan data penelitian meng-gunakan metode regresi *ridge* dengan parameter *robust* (*robust ridge regression*).

#### 4.6 Pemodelan Menggunakan Regresi *Robust*

Pada sub-bab sebelumnya, telah dilakukan pengujian dan pemeriksaan asumsi regresi linear berganda dan didapatkan hasil sedemikian rupa sehingga memberikan kepu-tusan untuk melakukan pemodelan dengan regresi *ridge* dengan parameter *robust*. Langkah awal sebelum melakukan pemodelan menggunakan regresi *ridge*, terlebih dahulu dilakukan pemodelan menggunakan regresi *robust*. Tujuan dilakukan langkah ini adalah untuk mendapatkan parameter *robust* terbaik dan juga hasil prediksi dari model *robust* yang nantinya digunakan dalam melakukan estimasi parameter *ridge* yang *robust*.

Lukman, Osowole, dan Ayinde (2015) dalam penelitiannya menyebutkan bahwa metode penaksiran parameter *MM-Estimator* lebih efisien dibandingkan dengan metode *M-Estimator* dan *S-Estimator*. Pada penelitian ini, dilakukan pemodelan regresi *robust* dengan membandingkan empat metode estimasi parameter *robust*. Keempat metode tersebut, yaitu *MM-Estimator*, *S-Estimator*, *Least Trimmed Square* (LTS), dan *Least Median of Square* (LMS). Membandingkan keempat metode tersebut berguna untuk memperkuat pernyataan Lukman, Osowole, dan Ayinde (2015) bahwa *MM-Estimator* merupakan metode pendugaan parameter yang lebih efisien.

RMSE berguna untuk melihat seberapa besar kesalahan model dalam prediksi data. Semakin kecil nilai *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE) maka model dapat dikatakan memiliki *bias* yang semakin kecil. Sementara  $R^2$  atau koefisien determinasi merupakan statistik yang digunakan untuk mengukur seberapa baik model dalam menggambarkan atau menjelaskan informasi yang ada pada variabel prediktor untuk memprediksi variabel respon. Jika nilai  $R^2$  semakin besar, maka model dapat dikatakan semakin baik. Untuk dapat menentukan model dengan penduga terbaik, kedua aspek di atas harus dipertimbangkan. Tabel 4.8 menyajikan hasil *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan  $R^2$  dari masing-masing model.

**Tabel 4.8** Hasil RMSE dan  $R^2$  Model *Robust*

Metode	RMSE	$R^2$	MAE
MM	38,5608	0,8112	22,0453
LMS	82,5867	0,3989	26,6272
LTS	69,5980	0,5128	23,8726
S	39,7661	0,8274	37,2047

Tabel 4.8 menunjukkan model dengan nilai RMSE dan MAE terkecil dimiliki oleh model *robust* dengan *MM-Estimator*.  $R^2$  terbesar dimiliki oleh model *robust* dengan *S-Estimator* dan model *robust* dengan *MM-Estimator* yang hanya mempunyai selisih kurang dari 1%. Meskipun demikian dalam memilih metode yang memiliki estimasi terbaik, indikator yang paling utama untuk dipertimbangkan adalah nilai RMSE dan MAE terkecil. Maka dari hal tersebut, pernyataan Lukman, Osowole, dan Ayinde (2015) dalam penelitiannya dapat dibenarkan. Langkah analisis selanjutnya adalah memodelkan data menggunakan regresi *ridge* dengan parameter *robust* ( $\hat{Y}_{Robust}$ ) sebagai nilai  $Y$  terbaru. Koefisien regresi dari model *robust MM-Estimator* dapat dilihat pada Tabel 4.9.

**Tabel 4.9** Koefisien Regresi *Robust MM-Estimator*

Parameter	Koefisien	Standar Error
$\beta_0$	-220,0407	33,3516
$\beta_1$	-3,4135	1,3834
$\beta_2$	2,7333	4,4596
$\beta_3$	62,0874	20,0827
$\beta_4$	19,4446	4,3979
$\beta_5$	-4,7347	67,9376
$\beta_6$	-12,2643	36,5871
$\beta_7$	7,9731	5,9402
$\beta_8$	0,2455	0,0318

Tabel 4.9 memperlihatkan koefisien estimasi parameter regresi *robust* menggunakan *MM-Estimator* dan *standard error* masing-masing parameter. *Standard error* terkecil dimiliki oleh parameter  $\beta_8$  dengan nilai sebesar 0,0318. Sementara *standard error* terbesar dimiliki oleh parameter  $\beta_5$ . *Standard error* mendeskripsikan jarak antara data hasil prediksi dengan garis regresi. Semakin besar *standard error* maka semakin jauh jarak data prediksi dari garis regresi. Model regresi *robust* dengan *MM-Estimator* dapat dilihat pada persamaan (4.2).

$$\hat{Y}_{Robust} = -220,0407 - 3,4135X_1 + 2,7333X_2 + 62,0874X_3 + 19,4446X_4^{\frac{1}{3}} - 4,7347X_5 - 12,2643X_6^3 + 7,9731X_7^{1/2} + 0,2455X_8 \quad (4.2)$$

Model pada persamaan (4.2) digunakan untuk mendapatkan estimasi variabel respon yang nantinya digunakan untuk mengestimasi parameter *ridge* yang *robust*. Hasil estimasi variabel respon menggunakan regresi *robust* dapat dilihat pada Lampiran 10. Model regresi *robust* lain juga dapat dilihat secara lengkap pada bagian Lampiran 7 sampai Lampiran 9.

#### 4.7 Pemodelan Menggunakan Regresi *Robust Ridge*

Penduga parameter *robust* telah didapatkan, namun model *robust* yang telah diformulasikan tidak relevan untuk digunakan dikarenakan pelanggaran asumsi multikolinieritas belum teratasi dengan baik. Hal ini, dikarenakan data yang dimiliki mengandung multikolinieritas. Sementara regresi *robust* hanya mampu mengatasi asumsi distribusi normal yang tidak terpenuhi dan efek *outlier*. Pemodelan menggunakan regresi *robust ridge* diharapkan menghasilkan model dengan nilai RMSE dan  $R^2$  yang lebih baik dibandingkan dengan model regresi *robust* maupun regresi linier. Untuk memastikan bahwa model regresi *robust ridge* dapat digunakan untuk memprediksi tahanan kapal cepat diperlukan aplikasi metode validasi silang. Validasi silang dilakukan dengan mempartisi data menjadi dua bagian, yaitu bagian *training* dan *testing*. Bagian data *training* digunakan untuk melakukan pembentukan model regresi, sedangkan data *testing* digunakan untuk menguji kebaikan model dalam memprediksi tahanan kapal cepat. Validasi silang akan diterapkan setelah didapatkan konstanta *bias c* dalam model regresi *robust ridge*.

Regresi *ridge* memiliki konstanta *bias c* yang digunakan sebagai regularisasi dalam model agar model *ridge* menjadi efisien. Untuk memilih nilai *c* optimum diperlukan simulasi memodelkan data menggunakan regresi *robust ridge* dengan berbagai nilai *c*. Pemeriksaan  $R^2$  dan koefisien regresi pada setiap nilai *c* digunakan untuk memilih model terbaik. Hasil estimasi parameter dan  $R^2$



pemodelan regresi *robust ridge* pada berbagai nilai  $c$  disajikan pada Tabel 4.10 dan Tabel 4.11.

**Tabel 4.10** Koefisien Regresi *Robust Ridge*

$c$	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$
$c=0,05$	-207,743	-0,835	4,461	54,303	15,994	17,165	0,628	1,908	0,234
$c=0,1$	-194,278	-0,308	4,174	45,707	14,049	19,580	4,155	1,825	0,221
$c=0,15$	-182,203	-0,045	4,134	40,404	12,761	18,924	6,715	1,853	0,208
$c=0,2$	-171,516	0,113	4,169	36,741	11,821	17,688	8,747	1,882	0,197
$c=0,25$	-161,999	0,216	4,221	34,019	11,093	16,501	10,419	1,901	0,187
$c=0,3$	-153,451	0,288	4,270	31,895	10,505	15,527	11,826	1,912	0,177
$c=0,35$	-145,706	0,341	4,309	30,177	10,017	14,788	13,026	1,916	0,169
$c=0,4$	-138,632	0,380	4,337	28,751	9,602	14,262	14,063	1,915	0,161
$c=0,45$	-132,128	0,410	4,355	27,542	9,243	13,916	14,965	1,911	0,154
$c=0,5$	-126,111	0,434	4,364	26,500	8,928	13,717	15,757	1,904	0,148
$c=0,55$	-120,515	0,452	4,366	25,588	8,647	13,637	16,455	1,895	0,142
$c=0,6$	-115,287	0,466	4,361	24,781	8,396	13,651	17,073	1,885	0,137
$c=0,65$	-110,381	0,477	4,351	24,059	8,168	13,738	17,623	1,873	0,132
$c=0,7$	-105,762	0,486	4,336	23,409	7,960	13,881	18,113	1,860	0,127
$c=0,75$	-101,399	0,493	4,318	22,818	7,769	14,068	18,552	1,847	0,123
$c=0,8$	-97,264	0,499	4,298	22,277	7,592	14,288	18,945	1,834	0,119
$c=0,85$	-93,337	0,503	4,274	21,780	7,429	14,532	19,297	1,820	0,115
$c=0,9$	-89,597	0,506	4,249	21,320	7,276	14,793	19,613	1,806	0,111
$c=0,95$	-86,027	0,509	4,222	20,892	7,133	15,065	19,896	1,792	0,108
$c=1$	-82,614	0,510	4,194	20,492	6,999	15,344	20,151	1,778	0,104

Pada Tabel 4.10 dapat dilihat bahwa koefisien regresi untuk variabel  $X_1$  hingga  $X_8$  memiliki tanda yang positif pada saat  $c=0,2$ . Cara mendeteksi bahwa tidak terdapat multikolinieritas pada model, yaitu dengan membandingkan tanda koefisien regresi dengan hasil korelasi antara variabel  $Y$  dan  $X$  seperti pada Gambar 4.1. Jika dilihat kembali pada Gambar 4.1, korelasi antara variabel  $R_T$  ( $Y$ ) dengan masing-masing variabel  $X$  berkorelasi positif dan dengan hal tersebut maka model dikatakan tidak terdapat

multikolinieritas jika koefisien regresi masing-masing variabel X bertanda positif.

**Tabel 4.11**  $R^2$  Regresi Robust Ridge

c	$R^2$	c	$R^2$
c=0,05	0,919	c=0,55	0,647
c=0,1	0,871	c=0,6	0,630
c=0,15	0,834	c=0,65	0,615
c=0,2	0,801	c=0,7	0,600
c=0,25	0,773	c=0,75	0,587
c=0,3	0,747	c=0,8	0,573
c=0,35	0,724	c=0,85	0,561
c=0,4	0,703	c=0,9	0,549
c=0,45	0,683	c=0,95	0,537
c=0,5	0,664		

Melalui Tabel 4.11 dapat dilihat bahwa besar  $R^2$  model untuk masing-masing nilai  $c$  berbeda satu sama lain. Semakin besar nilai  $c$  yang digunakan maka semakin kecil besaran  $R^2$  model. Melalui pertimbangan yang telah disajikan pada Tabel 4.10 dan 4.11, serta koefisien regresi pada bagian lampiran maka model regresi *robust ridge* dengan nilai  $c=0,2$  merupakan model yang terbaik.

Validasi silang dilakukan dengan membagi 80% data (32 dari total 37 pengamatan) sebagai data *training* yang akan digunakan untuk membentuk model regresi. Sementara 20% sisa dari data digunakan sebagai data *testing* yang akan digunakan untuk mengevaluasi model yang terbentuk. Data testing terdiri dari data pengamatan ke-2, ke-5, ke-26, ke-27, dan data pengamatan ke-36. Estimasi parameter hasil data *training* disajikan dalam Tabel 4.12.

**Tabel 4.12** Koefisien Regresi *Robust Ridge*  $c=0,2$

Parameter	Koefisien	$p$ -value
$\beta_0$	-162,175	0,000
$\beta_1$	0,126	0,394
$\beta_2$	4,428	0,000
$\beta_3$	33,984	0,000
$\beta_4$	11,801	0,000
$\beta_5$	13,337	0,719
$\beta_6$	10,927	0,568
$\beta_7$	1,879	0,000
$\beta_8$	0,185	0,000

Setelah melalui proses pemodelan, didapatkan besaran *p-value* pada Tabel 4.12 yang dapat digunakan untuk melakukan pengujian signifikansi parameter model secara parsial. Terlihat bahwa parameter yang memiliki pengaruh signifikan, diantaranya  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ ,  $\beta_7$ , dan  $\beta_8$ . Lima parameter tersebut memiliki *p-value* kurang dari taraf signifikan  $\alpha$  sebesar 5%. Model akhir dari regresi *robust ridge* disajikan pada persamaan (4.3).

$$\begin{aligned} \hat{Y}_{RR} = & -162,175 + 0,126 X_1 + 4,428 X_2 + \\ & 33,984 X_3 + 11,801 X_4^{1/3} + 13,337 X_5 + \\ & 10,927 X_6^3 + 1,879 X_7^{1/2} + 0,185 X_8 \end{aligned} \quad (4.3)$$

Model regresi pada persamaan (4.3) dapat digunakan untuk memprediksi tahanan kapal cepat dengan memasukkan nilai dari faktor-faktor yang berpengaruh. Hasil prediksi tahanan kapal dan data asli tahanan kapal disajikan pada Tabel 4.13.

**Tabel 4.13** Hasil Prediksi Tahanan Kapal

Obs	$Y$	$\hat{Y}_{RobustRidge}$
2	174,214	157,218
5	279,055	248,260
26	120,745	129,844
27	114,317	121,221
36	112,994	130,424

Hasil prediksi pada Tabel 4.13 terdapat perbedaan, namun perbedaan tersebut tidak terlalu jauh dari nilai sebenarnya. Selain itu, hasil prediksi menggunakan regresi *robust ridge* tidak menghasilkan nilai negatif yang menandakan bahwa kasus multikolinieritas pada model telah teratasi dengan baik

#### 4.8 Kriteria Keباikan Model

Untuk membuktikan apakah regresi *robust ridge* lebih baik untuk mengatasi dua kasus pelanggaran asumsi, dibandingkan besaran RMSE, MAE, dan  $R^2$  dari masing-masing model yang disajikan pada Tabel 4.14.

**Tabel 4.14** Perbandingan RMSE dan  $R^2$  Model Regresi

Metode	RMSE	$R^2$	MAE
<i>Linear</i>	33,179	0,940	26,087
<i>Robust</i>	36,509	0,870	21,933
<i>Robust Ridge</i>	18,284	0,983	16,245

Hasil kriteria kebaikan model *robust ridge* pada Tabel 4.14 didapatkan RMSE sebesar 18,284 dan MAE sebesar 16,245. Nilai RMSE dan MAE model *robust ridge* tersebut lebih kecil jika dibandingkan dengan model regresi *robust* dan juga regresi linier.  $R^2$  model regresi *robust ridge* juga memiliki nilai yang lebih besar dibandingkan dengan  $R^2$  model regresi *robust* dan linier, yaitu sebesar  $0,983 \approx 98,3\%$ . Melalui hasil tersebut dapat dikatakan bahwa kebaikan parameter *robust ridge* dalam memprediksi nilai tahanan kapal cepat pada kasus ini dapat dipercaya karena model mampu memprediksi dengan kebaikan sebesar 98,3%.

Model regresi pada persamaan (4.3) dapat digunakan oleh BTH-BPPT untuk memprediksi tahanan kapal cepat dengan memasukkan besaran variabel prediktor yang sesuai dengan batasan-batasan tertentu. Pembatasan besaran variabel prediktor ditujukan untuk mendapatkan hasil prediksi tahanan kapal cepat yang mendekati hasil sebenarnya. Batas-batas untuk besaran variabel prediktor disajikan pada Tabel 4.15.

**Tabel 4.15** Batasan Variabel Prediktor

Variabel	Simbol	Batas Minimum	Batas Maksimum
$X_1$	$L_{WL}$	8,89	76,15
$X_2$	$B$	2,80	16,00
$X_3$	$T_F$	0,45	3,00
$X_4$	$\Delta$	4,00	1626,00
$X_5$	$C_B$	0,23	0,54
$X_6$	$C_M$	0,40	0,93
$X_7$	$S$	19,35	1041,88
$X_8$	$V$	17,00	28,00

## BAB V KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Setelah dilakukan analisis dan pembahasan pada BAB IV, maka kesimpulan yang dapat diambil disajikan sebagai berikut :

1. Sebanyak delapan variabel prediktor dilakukan analisis deskriptif dan didapatkan hasil bahwa variabel B ( $X_2$ ),  $\text{Disp}/\Delta$  ( $X_4$ ),  $C_M$  ( $X_6$ ), dan S ( $X_7$ ) terdapat *outlier*. Selain itu melalui *scatterplot* antara variabel prediktor dan variabel respon, didapatkan dugaan awal jika variabel B ( $X_2$ ),  $\text{Disp}/\Delta$  ( $X_4$ ),  $C_M$  ( $X_6$ ), dan S ( $X_7$ ) tidak linier terhadap variabel  $R_T$  (Y). Hasil pengujian linieritas data menguatkan bahwa variabel  $\text{Disp}/\Delta$  ( $X_4$ ),  $C_M$  ( $X_6$ ), dan S ( $X_7$ ) tidak linier terhadap variabel  $R_T$  (Y). Untuk itu, proses transformasi data dilakukan untuk mendapatkan data yang linier.
2. Melalui tahap pemodelan regresi *ridge* dengan menggunakan parameter yang *robust* mampu menjawab permasalahan multikolinieritas dan ketidaknormalan data. Melalui interpretasi model terdapat faktor-faktor yang paling berpengaruh terhadap pembentukan tahanan kapal, yaitu variabel B ( $X_2$ ),  $T_F$  ( $X_3$ ),  $\text{Disp}/\Delta$  ( $X_4$ ), S ( $X_7$ ), dan  $V^2$  ( $X_8$ ). Lima faktor tersebut secara statistik memiliki pengaruh yang signifikan terhadap prediksi tahanan kapal cepat pada taraf signifikan sebesar 5%.
3. Model akhir *robust ridge* dengan  $R^2$  model sebesar 0,983 merupakan model terbaik yang didapatkan melalui data pengujian tahanan kapal cepat.  $R^2$  model sebesar 0,983 memiliki arti bahwa banyaknya informasi yang mampu dijelaskan oleh model sebesar 98,3%. Sekitar 2,7% informasi lainnya dijelaskan oleh variabel yang tidak berada di dalam model tersebut. Model regresi terbaik dapat digunakan untuk memprediksi tahanan kapal cepat dengan memasukkan besaran variabel prediktor ke dalam model pada persamaan (4.3).

## 5.2 Saran

Setelah dilakukan analisis dan penarikan kesimpulan, adapula masukan berupa saran bagi pembaca untuk dipertimbangkan pada penelitian selanjutnya.

1. Karena keterbatasan waktu yang dimiliki untuk menganalisis topik penelitian ini, disarankan bagi pembaca/penelitian selanjutnya untuk melakukan analisis menggunakan regresi *ridge* dengan metode penduga parameter yang lebih baik.
2. Sebaiknya dilakukan penyeleksian secara mendalam tentang pemilihan variabel-variabel pendukung agar didapatkan model regresi dengan kebaikan yang tinggi.
3. Untuk penelitian serupa dengan jumlah data yang lebih banyak, disarankan untuk menggunakan metode *in sample out sample* yang lebih mendukung agar didapatkan model prediksi dengan akurasi dan presisi yang lebih tinggi.

## DAFTAR PUSTAKA

- Al-Obeidat, F., Spencer, B., & Alfandi, O. (2018). Consistently accurate forecasts of temperature within buildings from sensor data using ridge and lasso regression. *Future Generation Computer Systems*.
- Alpu, O., & Samkar, H. (2010). Ridge Regression Based on Some Robust Estimators. *Journal of Modern Applied Statistical Methods*, Vol. 9 pp. 495-501.
- Brandsaeter, A., & Vanem, E. (2018). Ship speed prediction based on full scale sensor measurements of shaft thrust and environmental conditions. *Journal of Ocean Engineering*, Vol. 162 pp. 316-330.
- Burhanudin, I. (2012). *Kajian Teknis Kapal Trimaran Dengan Sistem Propulsi Water Jet Berpenggerak Motor Listrik*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference : A Practical Information-Theoretical Approach*. New York: Springer.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*.
- Chanson, H. (2004). *Hydraulics of Open Channel Flow*. Burlington: Elsevier.
- Cule, E., Vineis, P., & De lorio, M. (2011). Significance Testing in Ridge Regression for Genetic Data. *Bioinformatics*, Vol. 12 pp. 372.
- Draper, N., & Smith, H. (1998). *Applied Regression Analysis*. Danvers: John Wiley & Sons, Inc.
- Furqon. (1999). *Statistika Terapan Untuk Penelitian*. Bandung: Alfabeta.
- Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *Journal of the Anthropological Institute*, Vol. 15 pp. 246-263.
- Gujarati, D. N. (2003). *Basic Econometrics*. Boston: McGraw-Hill.
- Hardjono, S. (2017). *Kajian Kapal Cepat Rudal (KCR) 60M Terhadap Kondisi SEASTATE Perairan Kawasan Barat*

- dan Timur Indonesia*. Tangerang Selatan: Pusat Teknologi Industri Rekayasa Maritim - BPPT.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, Vol. 12 pp. 55-67.
- Holland, P. W. (n.d.). Weighted ridge regression: Combining Ridge and Robust Regression Methods. *NBER Working Paper Series*.
- Johnson, R. A., & Bhattacharyya, G. K. (2010). *Statistics Principle & Methods*. Danver: John Wiley & Sons Inc.
- Kementrian Perindustrian Republik Indonesia. (2019, Maret 3). *Berita Industri*. Retrieved from Kementrian Perindustrian Republik Indonesia Web Site: <http://www.kemenperin.go.id/artikel/2908/Galangan-Kapal-Tumbuh>
- Lewis, E. V. (1988). *Principles of Naval Architecture Second Revision*. Jersey City: The Society of Naval Architects and Marine Engineers.
- Lukman, A. F., Osowole, O. I., & Ayinde, K. (2015). Two Stage Robust Ridge Method in a Linear Regression Model. *Journal of Modern Applied Statistical Methods*, Vol. 14.
- Mao, W., Rychlik, I., Wallin, J., & Storhaug, J. (2016). Statistical Models for The Speed Prediction of a Container Ship. *Journal of Ocean Engineering*, Vol. 126 pp. 152-162.
- Marine Connector. (2018, February 13). *Maritime Encyclopedia*. Retrieved from Marine Connector: <http://maritime-connector.com/wiki-home>
- Mendenhall, W. (1987). *Introduction To Probability and Statistics*. Boston: PWS-Kent.
- Obenchain, R. (1977). Classical F-Tests and Confidence Regions for Ridge Regression. *Technometrics*, Vol. 19 pp. 429-439.
- Pereira, J. M., Basto, M., & da Silva, A. F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Economics and Finance*, Vol. 39 pp. 634-641.
- Rawling, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied Regression Analysis*. New York: Springer.



- Roh, M.-I., & Lee, K.-Y. (2018). *Computational Ship Design*. Singapore: Springer Nature.
- Sahlan, Ali, B., Nugroho, W., Bisri, A., & Arswendo, B. A. (2016). Pengaruh Bentuk Lambung Kapal Terhadap Pola Aliran dan Powering Pada Kapal Perairan Sungai dan Laut. *Kapal*, 1-6.
- Siswanto, D. (1988). *Teori Tahanan Kapal I*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Sonnberger, H., & Kraemer, W. (1986). *Linear Regression Model under Test*. Heildeberg: Physica-Verlag.
- Sugianto, E., & Winarno, A. (2017). Computational Model Tahanan Kapal untuk Menentukan Kebutuhan Daya Kapal Bulk Carrier 8664 DWT. *Kelautan*, 10(2), 1-2.
- Sugiyono. (2011). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Afabeta.
- Suliyanto. (2005). *Analisis Data dalam Aplikasi Pemasaran*. Bogor: Ghalisa Indonesia.
- Sumaryanto. (2013). *Konsep Dasar Kapal*. Jakarta: Kementerian Pendidikan dan Kebudayaan Indonesia.
- Widarjono, A. (2009). *Ekonometrika Pengantar dan Aplikasinya*. Yogyakarta: Ekonisia.
- Wilcox, R. (2017). Robust Regression. In R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing* (pp. 517–583). Elsevier Inc.

*(Halaman ini sengaja dikosongkan)*

## LAMPIRAN

### Lampiran 1 Data Penelitian Kapal Cepat

	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>
1	230,00	56,60	8,50	3,00	786,80	0,54	0,86	561,11	361,00
2	253,00	39,81	7,00	2,00	282,25	0,50	0,70	323,35	506,25
3	199,14	60,00	6,91	1,98	361,18	0,23	0,58	460,54	756,25
4	203,50	43,50	9,30	2,00	315,80	0,41	0,59	364,25	506,25
5	299,00	54,82	8,10	2,60	471,55	0,40	0,56	457,30	756,25
6	74,25	44,60	6,51	1,50	184,06	0,41	0,64	269,82	420,25
7	461,00	76,15	13,60	3,00	1626,00	0,51	0,92	1041,88	441,00
8	225,00	55,60	16,00	1,35	350,00	0,45	0,78	562,94	784,00
9	117,00	43,34	7,80	1,70	186,55	0,34	0,52	283,00	484,00
10	209,50	35,56	8,00	1,78	230,00	0,46	0,60	287,62	756,25
11	275,00	41,69	8,80	1,89	316,25	0,46	0,60	372,95	756,25
12	217,00	56,04	8,50	2,10	414,10	0,41	0,57	411,00	576,00
13	125,00	25,20	6,40	1,44	95,00	0,27	0,42	165,99	625,00
14	162,00	34,30	7,30	1,59	165,00	0,28	0,51	253,60	625,00
15	197,00	53,41	8,20	2,17	420,00	0,34	0,81	430,06	484,00
16	175,00	57,83	8,50	2,70	515,94	0,40	0,57	489,38	380,25
17	141,50	41,42	7,90	1,70	202,54	0,36	0,40	256,00	650,25
18	147,50	40,55	7,80	1,90	250,51	0,46	0,68	287,60	380,25
19	94,90	25,44	5,29	1,05	77,19	0,53	0,76	132,92	729,00
20	86,70	25,14	5,48	1,32	84,93	0,43	0,55	145,47	529,00
21	32,40	15,24	3,57	0,90	22,05	0,34	0,51	60,10	484,00
22	141,00	15,92	4,00	0,70	19,48	0,43	0,64	60,00	441,00
23	227,00	64,81	11,67	2,40	995,95	0,53	0,80	745,38	380,25
24	151,00	69,34	11,35	3,00	113,62	0,51	0,73	817,09	306,25
25	121,50	42,80	6,78	2,00	311,87	0,50	0,67	317,71	324,00
26	121,50	40,31	6,83	1,70	194,00	0,41	0,62	234,00	506,25
27	117,50	42,20	7,90	1,70	196,20	0,37	0,50	300,20	420,25
28	87,85	27,11	6,20	1,35	79,07	0,38	0,59	147,90	650,25
29	183,00	55,22	9,00	1,80	392,78	0,46	0,58	459,04	529,00
30	111,00	32,91	7,30	1,50	138,93	0,42	0,59	211,86	441,00
31	6,80	8,90	2,80	0,45	4,00	0,39	0,63	19,35	576,00
32	88,60	25,08	5,85	1,25	75,80	0,42	0,59	137,11	529,00
33	24,10	15,13	3,90	0,75	14,85	0,27	0,44	54,24	484,00
34	88,70	44,24	7,90	1,80	235,43	0,41	0,57	303,47	400,00
35	262,00	69,76	10,63	3,00	1147,32	0,54	0,73	829,09	306,25
36	106,00	51,57	7,25	2,00	366,44	0,48	0,68	411,69	289,00
37	112,00	59,25	7,31	2,15	475,95	0,50	0,70	493,95	289,00

**Lampiran 2** Data Penelitian Setelah Transformasi

	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>
1	230,00	56,60	8,50	3,00	9,23	0,54	0,64	23,69	361,00
2	253,00	39,81	7,00	2,00	6,56	0,50	0,35	17,98	506,25
3	199,14	60,00	6,91	1,98	7,12	0,23	0,20	21,46	756,25
4	203,50	43,50	9,30	2,00	6,81	0,41	0,21	19,09	506,25
5	299,00	54,82	8,10	2,60	7,78	0,40	0,18	21,38	756,25
6	74,25	44,60	6,51	1,50	5,69	0,41	0,27	16,43	420,25
7	461,00	76,15	13,60	3,00	11,76	0,51	0,80	32,28	441,00
8	225,00	55,60	16,00	1,35	7,05	0,45	0,49	23,73	784,00
9	117,00	43,34	7,80	1,70	5,71	0,34	0,15	16,82	484,00
10	209,50	35,56	8,00	1,78	6,13	0,46	0,22	16,96	756,25
11	275,00	41,69	8,80	1,89	6,81	0,46	0,22	19,31	756,25
12	217,00	56,04	8,50	2,10	7,45	0,41	0,19	20,27	576,00
13	125,00	25,20	6,40	1,44	4,56	0,27	0,08	12,88	625,00
14	162,00	34,30	7,30	1,59	5,48	0,28	0,14	15,92	625,00
15	197,00	53,41	8,20	2,17	7,49	0,34	0,53	20,74	484,00
16	175,00	57,83	8,50	2,70	8,02	0,40	0,19	22,12	380,25
17	141,50	41,42	7,90	1,70	5,87	0,36	0,06	16,00	650,25
18	147,50	40,55	7,80	1,90	6,30	0,46	0,32	16,96	380,25
19	94,90	25,44	5,29	1,05	4,26	0,53	0,44	11,53	729,00
20	86,70	25,14	5,48	1,32	4,40	0,43	0,17	12,06	529,00
21	32,40	15,24	3,57	0,90	2,80	0,34	0,14	7,75	484,00
22	141,00	15,92	4,00	0,70	2,69	0,43	0,26	7,75	441,00
23	227,00	64,81	11,67	2,40	9,99	0,53	0,52	27,30	380,25
24	151,00	69,34	11,35	3,00	4,84	0,51	0,39	28,58	306,25
25	121,50	42,80	6,78	2,00	6,78	0,50	0,31	17,82	324,00
26	121,50	40,31	6,83	1,70	5,79	0,41	0,24	15,30	506,25
27	117,50	42,20	7,90	1,70	5,81	0,37	0,13	17,33	420,25
28	87,85	27,11	6,20	1,35	4,29	0,38	0,21	12,16	650,25
29	183,00	55,22	9,00	1,80	7,32	0,46	0,20	21,43	529,00
30	111,00	32,91	7,30	1,50	5,18	0,42	0,21	14,56	441,00
31	6,80	8,90	2,80	0,45	1,59	0,39	0,26	4,40	576,00
32	88,60	25,08	5,85	1,25	4,23	0,42	0,21	11,71	529,00
33	24,10	15,13	3,90	0,75	2,46	0,27	0,09	7,36	484,00
34	88,70	44,24	7,90	1,80	6,17	0,41	0,19	17,42	400,00
35	262,00	69,76	10,63	3,00	10,47	0,54	0,40	28,79	306,25
36	106,00	51,57	7,25	2,00	7,16	0,48	0,33	20,29	289,00
37	112,00	59,25	7,31	2,15	7,81	0,50	0,35	22,22	289,00

### Lampiran 3 Gambar Kapal Cepat TNI-AL KRI Albakora



### Lampiran 4 Syntax Analisis Tugas Akhir

```

Data1 = read.csv("KAPAL MED.csv")
head(Data1)
KapalMed=data.frame(Data1[,12],Data1[,1:2],Data1[,4:5],Data1[
,7:10])
colnames(KapalMed)=c("RS", "LWL", "B", "TF", "DISP", "CB", "C
M", "CP", "V")
head(KapalMed)

library(vars)
library(ggfortify)
library(lmtest)
library(olsrr)
library(ggplot2)
library(ggsci)
library(viridis)
library(cowplot)
library(MASS)
library(SciViews)
library(car)
library(cowplot)
  
```

```

library(PerformanceAnalytics)
library(corrplot)
library(RColorBrewer)
library(Hmisc)

summary(KapalMed)
is.na.data.frame(KapalMed)
vari<- cbind(

var(KapalMed$RS),var(KapalMed$LWL),var(KapalMed$B),var(
KapalMed$TF),

var(KapalMed$DISP),var(KapalMed$CB),var(KapalMed$CM),v
ar(KapalMed$CP),
  var(KapalMed$V)
)
colnames(vari)=c("RS","LWL","B","TF","DISP","CB","CM","C
P","V")
vari
chart.Correlation(KapalMed, histogram=TRUE, pch=19)
col <- colorRampPalette(c("#BB4444", "#EE9988", "#FFFFFF",
"#77AADD", "#4477AA"))
Matrix <- rcorr(as.matrix(KapalMed))
M <- Matrix$r
p_mat <- Matrix$P
corrplot(M, method = "color", col = col(200),
  type = "upper", order = "hclust",
  addCoef.col = "black", # Add coefficient of correlation
  tl.col = "darkblue", tl.srt = 45, #Text label color and rotation
  # hide correlation coefficient on the principal diagonal
  diag = FALSE
)
#BoxPlot
plot_grid(
  ggplot(KapalMed,aes(y=LWL))+
  geom_boxplot(col="Maroon",outlier.color="Red",outlier.shape
= "x"),

```

```

ggplot(KapalMed,aes(y=B))+
  geom_boxplot(col="Dark Green",outlier.color = "Red",
outlier.shape = "x"),
  ncol = 2, nrow = 1
)
plot_grid(
  ggplot(KapalMed,aes(y=TF))+
    geom_boxplot(col="Purple",outlier.color = "Red", outlier.shape
= "x"),
  ggplot(KapalMed,aes(y=DISP))+
    geom_boxplot(col="Dark Orange",outlier.color = "Red",
outlier.shape = "x"),
  ncol = 2, nrow = 1
)
plot_grid(
  ggplot(KapalMed,aes(y=CB))+
    geom_boxplot(col="Black",outlier.color = "Red", outlier.shape
= "x"),
  ggplot(KapalMed,aes(y=CM))+
    geom_boxplot(col="Blue",outlier.color = "Red", outlier.shape =
"x"),
  ncol = 2, nrow = 1
)
plot_grid(
  ggplot(KapalMed,aes(y=CP))+
    geom_boxplot(col="Violet",outlier.color = "Red", outlier.shape
= "x"),
  ggplot(KapalMed,aes(y=V))+
    geom_boxplot(col="Dark Green",outlier.color = "Red",
outlier.shape = "x"),
  ncol = 2, nrow = 1
)
ggplot(KapalMed,aes(y=RS))+
  geom_boxplot(col="Dark Blue",outlier.color = "Red",
outlier.shape = "x")

```

```

#Plot Per Variabel

ggplot(KapalMed,aes(x=LWL,y=RS))+geom_point(col="Maroon")
)+
  geom_smooth(method = "lm")
ggplot(KapalMed,aes(x=B,y=RS))+geom_point(col="Red")+
  geom_smooth(method = "lm")

ggplot(KapalMed,aes(x=TF,y=RS))+geom_point(col="Purple")+
  geom_smooth(method = "lm")

ggplot(KapalMed,aes(x=DISP,y=RS))+geom_point(col="Orange")
)+
  geom_smooth(method = "lm")

ggplot(KapalMed,aes(x=CB,y=RS))+geom_point(col="Black")+
  geom_smooth(method = "lm")
ggplot(KapalMed,aes(x=CM,y=RS))+geom_point(col="Blue")+
  geom_smooth(method = "lm")

ggplot(KapalMed,aes(x=CP,y=RS))+geom_point(col="Green")+
  geom_smooth(method = "lm")
ggplot(KapalMed,aes(x=V,y=RS))+geom_point(col="Green")+
  geom_smooth(method = "lm")

#Pemodelan Linier
linearMod1 <- lm(RS ~ LWL+B+TF+DISP+CB+CM+CP+V,
data=KapalMed)
print(linearMod1)
summary(linearMod1)
AIC(linearMod1)
BIC(linearMod1)
#Plot Asumsi Klasik
autoplot(linearMod1)
#Uji Linieritas Model
resettest(KapalMed$RS~KapalMed$LWL,type = "regressor")

```



```

resetest(KapalMed$RS~KapalMed$B,type = "regressor")
resetest(KapalMed$RS~KapalMed$TF,type = "regressor")
resetest(KapalMed$RS~KapalMed$DISP,type = "regressor")
resetest(KapalMed$RS~KapalMed$CB,type = "regressor")
resetest(KapalMed$RS~KapalMed$CM,type = "regressor")
resetest(KapalMed$RS~KapalMed$CP,type = "regressor")
resetest(KapalMed$RS~KapalMed$V,type = "regressor")
resetest(linearMod1, type = "regressor")
#Plot Cook's Distance for Outlier
ols_plot_cooksd_bar(linearMod1)
#Uji Autokorelasi
dwtest(linearMod1)
#Uji Heteroskedastisitas
ols_test_breusch_pagan(linearMod1)
#Uji Normalitas
mv1 <- VAR(KapalMed)
normality.test(mv1)
#Multikolinieritas
vif(linearMod1)

#Transformasi Data
KMTrans <- KapalMed
KMTrans$DISP <- (KapalMed$DISP)^(1/3)
KMTrans$CP <- (KapalMed$CP^5)
KMTrans$CM <- (KapalMed$CM)^5*log(KapalMed$CM)*(-1)
KMTrans$RS <- KapalMed$RS
head(KapalMed)
head(KMTrans)
resetest(KMTrans$RS~KMTrans$DISP,type = "regressor")
resetest(KMTrans$RS~KMTrans$CP,type = "regressor")
resetest(KMTrans$RS~KMTrans$CM,type = "regressor")

ggplot(KMTrans,aes(x=DISP,y=RS))+geom_point(col="Orange")
)+
geom_smooth(method = "lm")

```

```

ggplot(KMTrans,aes(x=CP,y=RS))+geom_point(col="Red")+
  geom_smooth(method = "lm")
ggplot(KMTrans,aes(x=CM,y=RS))+geom_point(col="Dark
Green")+
  geom_smooth(method = "lm")

#Pemodelan Trans
lmTrans <- lm(RS ~ LWL+B+TF+DISP+CB+CM+CP+V,
data=KMTrans)
print(lmTrans)
summary(lmTrans)
#Uji Autokorelasi
dwtest(lmTrans)
#Uji Heteroskedastisitas
ols_test_breusch_pagan(lmTrans)
#Uji Normalitas
mvT <- VAR(KMTrans)
normality.test(mvT)
#Multikolinieritas
vif(lmTrans)

autoplot(lmTrans)
sresid <- studres(lmTrans)
hist(sresid, freq=FALSE,
  main="Distribution of Studentized Residuals")
library(caret)
RMSE(predict(lmTrans), KMTrans$RS)
R2(predict(lmTrans), KMTrans$RS)

#Robust Regression
library(MASS)
robust <- rlm(RS ~ LWL+B+TF+DISP+CB+CM+CP+V, data =
KMTrans, method = "MM")
rLms <- lqs(RS ~ LWL+B+TF+DISP+CB+CM+CP+V, data =
KMTrans, method = "lms")

```

```

rLts <- lqs(RS ~ LWL+B+TF+DISP+CB+CM+CP+V, data =
KMTrans, method = "lts")
rS <- lqs(RS ~ LWL+B+TF+DISP+CB+CM+CP+V, data =
KMTrans, method = "S")

Baik <- rbind(cbind(RMSE(predict(robust), KMTrans$RS),
  R2(predict(robust), KMTrans$RS)),
  cbind(RMSE(predict(rLms), KMTrans$RS),
  R2(predict(rLms), KMTrans$RS)),
  cbind(RMSE(predict(rLts), KMTrans$RS),
  R2(predict(rLts), KMTrans$RS)),
  cbind(RMSE(predict(rS), KMTrans$RS),
  R2(predict(rS), KMTrans$RS)),
  deparse.level = 2
)
colnames(Baik)= c("RMSE", "R-Sqr")
rownames(Baik)= c("MM", "LMS", "LTS", "S")
Baik
summary(robust)

#Robust Ridge (RR)
library(lmridge)
library(elasticnet)
CDataRR <- cbind(robust$fitted.values,robust$x)
CDataRR
DataRR <- data.frame(CDataRR[,1],CDataRR[,3:10])
colnames(DataRR)=c("RT", "LWL", "B", "TF", "DISP", "CB", "CM", "S", "
V")
head(DataRR)
head(KMTrans)
ridge <- lmridge(RT ~ LWL+B+TF+DISP+CB+CM+S+V, data =
DataRR, K = seq(0.05,1,0.05))
cbind(print(ridge2))
rstats1(ridge2)
cbind(R2(predict(ridge2), KMTrans$RT))
kest(ridge2)
rplots.plot(ridge2)

```

```

#Cross Validation
set.seed(17)
fld = createDataPartition(y=DataRR$RT, p=0.8, list = FALSE)
fld
train1 = DataRR[fld,]
train1
test1 = DataRR[-fld,]

ridge1 <- lmridge(RT ~ LWL+B+TF+DISP+CB+CM+S+V, data
= train1,K = 0.2)
print(ridge1)
x_test1 = test1[,2:9]
y_test1 = test1[,1]
y_test0 = rbind(KMTrans[2,1],KMTrans[5,1],KMTrans[26,1],
                KMTrans[27,1],KMTrans[36,1])
pred = predict(ridge1,x_test1,interval = "prediction")

Kebaikan<-rbind(
  cbind(RMSE(pred, y_test1),
        R2(pred, y_test1),
        MAE(pred,y_test1)),
  cbind(RMSE(predict(robust,x_test1), y_test0),
        R2(predict(robust,x_test1), y_test0),
        MAE(predict(robust,x_test1),y_test0)),
  cbind(RMSE(predict(lmTrans,x_test1), y_test0),
        R2(predict(lmTrans,x_test1), y_test0),
        MAE(predict(lmTrans,x_test1), y_test0)),
  deparse.level = 2)
colnames(Kebaikan)= c("RMSE", "R-Sqr", "MAE")
rownames(Kebaikan)= c("Robust Ridge", "Robust", "Linear
Sederhana")
Kebaikan

Banding <-cbind(y_test1,pred)
colnames(Banding)=c("Y Aktual", "Y Prediksi")
Banding

```

**Lampiran 5** Model Regresi Linier Sebelum Transformasi

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8$$

$$\hat{Y} = -174,7 - 1,681X_1 + 7,313X_2 + 75,365X_3 + 0,1401X_4 + 10,64X_5 + 48,76X_6 + 0,0004X_7 + 0,254X_8$$

**Lampiran 6** Model Regresi Linier Setelah Transformasi

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4^{1/3} + \beta_5 X_5 + \beta_6 X_6^3 + \beta_7 X_7^{1/2} + \beta_8 X_8$$

$$\hat{Y} = -186,23 - 5,616X_1 - 1,713X_2 + 36,215X_3 + 21,979X_4^{1/3} - 64736X_5 + 86,405X_6^3 + 16,102X_7^{1/2} + 0,222X_8$$

**Lampiran 7** Model Regresi *Robust* LMS-Estimator

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4^{1/3} + \beta_5 X_5 + \beta_6 X_6^3 + \beta_7 X_7^{1/2} + \beta_8 X_8$$

$$\hat{Y} = -118,596 - 2,319X_1 + 16,447X_2 + 64,543X_3 + 15,29X_4^{1/3} - 52,476X_5 + 47,93X_6^3 - 1,632X_7^{1/2} + 0,109X_8$$

**Lampiran 8** Model Regresi *Robust* LTS-Estimator

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4^{1/3} + \beta_5 X_5 + \beta_6 X_6^3 + \beta_7 X_7^{1/2} + \beta_8 X_8$$

$$\hat{Y} = -116,9097 - 1,3959X_1 + 18,3091X_2 + 49,821X_3 + 16,413X_4^{1/3} - 117,667X_5 + 94,118X_6^3 - 3,015X_7^{1/2} + 0,121X_8$$

**Lampiran 9** Model Regresi *Robust* S-Estimator

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4^{1/3} + \beta_5 X_5 + \beta_6 X_6^3 + \beta_7 X_7^{1/2} + \beta_8 X_8$$

$$\hat{Y} = -173,711 - 1,858X_1 + 14,858X_2 + 125,044X_3 + 21,591X_4^{1/3} - 108,214X_5 + 45,817X_6^3 - 7,533X_7^{1/2} + 0,1889X_8$$

**Lampiran 10** Nilai Aktual & Prediksi Tahanan Kapal Cepat

Obs	$Y$	$\hat{Y}_{Linier}$	$\hat{Y}_{Robust}$
1	230.000	274.813	241.400
2	253.000	194.963	174.214
3	199.135	208.616	208.498
4	203.500	187.027	185.134
5	299.000	260.028	279.055
6	74.250	86.292	78.432
7	461.000	383.948	324.035
8	225.000	247.451	229.131
9	117.000	113.477	120.316
10	209.500	230.253	224.360
11	275.000	251.364	244.456
12	217.000	168.700	186.437
13	125.000	149.661	142.793
14	162.000	175.814	165.005
15	197.000	208.954	176.601
16	175.000	180.038	195.244
17	141.500	143.201	165.142
18	147.500	135.345	125.786
19	94.900	145.493	116.677
20	86.700	106.456	97.470
21	32.400	39.043	25.104
22	141.000	20.024	-3.366
23	227.000	270.427	234.851
24	151.000	149.131	150.265
25	121.500	137.108	123.046
26	121.500	117.656	120.745
27	117.500	111.788	114.317
28	87.850	127.777	123.493
29	183.000	164.315	165.938
30	111.000	108.111	101.033
31	6.800	6.703	-12.762
32	88.600	98.008	88.297
33	24.100	19.465	8.838
34	88.700	112.179	115.612
35	262.000	273.433	256.678
36	106.000	129.870	112.994
37	112.000	138.003	123.980

## Lampiran 11 Surat Keterangan Data

### SURAT KETERANGAN

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Mohammad Haidar Alvin Purwana

NRP : 062115 4000 0131

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : Badan Pengkajian dan Penerapan Teknologi (BPPT)

Divisi/ bagian : Balai Teknologi Hidrodinamika

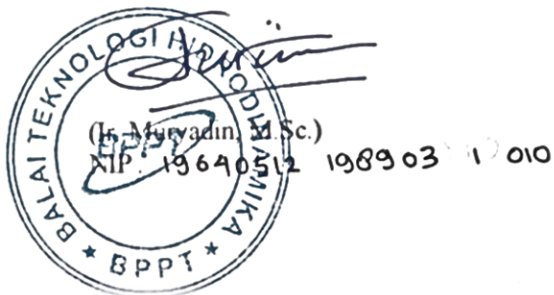
sejak tanggal 11 Maret 2019 sampai dengan 21 Juli 2019 untuk keperluan Tugas Akhir/ Thesis Semester Genap 2019.

2. Tidak Keberatan/Keberatan\* nama perusahaan dicantumkan dalam Tugas Akhir/ Thesis mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak Keberatan/Keberatan\* bahwa hasil analisis data dari perusahaan dipublikasikan dalam E-journal ITS yaitu Jurnal Sains dan Seni ITS.

Surabaya, 13 Juni 2019

Pimpinan Perusahaan

Kepala Balai Teknologi Hidrodinamika, BPPT



\*(coret yang tidak perlu)

## Lampiran 12 Biodata Penulis

Penulis yang memiliki nama lengkap Mohammad Haidar Alvin Purwana, merupakan anak pertama dari 2 bersaudara. Jakarta merupakan kota kelahiran dan juga menjadi kota pertama penulis dalam mempelajari ilmu-ilmu kehidupan. Penulis merupakan mahasiswa Departemen Statistika ITS Angkatan 2015. Menjalankan kehidupan perkuliahan dan segala bentuk kegiatan mahasiswa sedikit banyak membentuk karakter penulis seperti saat ini. Segala bentuk ilmu dan pengalaman didapatkan baik saat belajar di dalam kelas maupun saat melakukan kegiatan non akademik di luar kelas. Selama menjadi mahasiswa, penulis seringkali mengikuti pelbagai pelatihan baik pelatihan pengembangan diri hingga pelatihan keprofesional. Pada masa sebagai mahasiswa, penulis juga mencoba peruntungan dalam dunia wirausaha dengan memulai dari level mikro hingga masuk pada dunia kewirausahaan non mikro. Segala bentuk kegiatan, interaksi, pertemanan, dan hubungan baik yang telah dijalankan oleh penulis memberikan arti sendiri bagi penulis. Banyak juga pengalaman hidup yang penulis dapat melalui hubungan baik dan kemurahan hati dosen-dosen sehingga penulis merasa sangat berterimakasih kepada dosen-dosen yang senantiasa berpengaruh dalam hidup penulis.

