Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Perbandingan Estimator *Robust Huber* dan *Tukey's Biweight* terhadap Berbagai Skema Pencilan dalam Regresi Linier

Linda Rassiyanti[1*], Indah Suciati[2], Vina Nurmadani[3], Yoga Aji Sukma[4]

[1,2,3,4]Program Studi Sains Data, Institut Teknologi Sumatera, Indonesia
[*]corresponding author: linda.rassiyanti@sd.itera.ac.id

Received July 29, 2025; Received in revised form July 11, 2025; Accepted July 11, 2025

Abstrak. Regresi linier secara umum menggunakan pendekatan *Ordinary Least Squares* (OLS) namun sering kali mengalami gangguan ketika data mengandung pencilan (*outlier*), yang dapat menyebabkan estimasi parameter menjadi bias dan tidak akurat. Regresi robust dikembangkan untuk mengatasi kelemahan OLS dengan menurunkan sensitivitas terhadap pencilan. Terdapat dua fungsi kerugian yang sering digunakan dalam regresi robust, yaitu Huber Loss dan Tukey's Biweight Loss. Penelitian ini bertujuan untuk membandingkan performa dua metode regresi robust, yaitu Huber Loss dan Tukey's Biweight, dalam menghadapi berbagai skema pencilan. Data simulasi dibangkitkan dengan parameter intersep dan slope masing-masing sebesar 3 dan 2, kemudian ditambahkan pencilan secara sistematis pada variabel X, Y, maupun keduanya, dengan proporsi 10%, 20%, dan 30%. Hasil analisis menunjukkan bahwa Tukey's Biweight memberikan estimasi parameter yang lebih stabil pada kondisi pencilan ekstrem, terutama saat pencilan terjadi pada variabel Y atau kombinasi X dan Y. Sedangkan, Huber Loss cenderung menghasilkan *Mean Squared Error* (MSE) yang lebih rendah dalam beberapa kondisi, mencerminkan adanya *trade-off* antara bias dan variansi. Dengan demikian, Tukey's Biweight lebih cocok untuk pencilan ekstrem, sedangkan Huber Loss lebih efisien dalam kondisi pencilan ringan hingga sedang.
Kata kunci: Huber Loss; regresi robust; simulasi data; *Tukey's Biweight*

Abstract. *Linear regression, commonly estimated using the Ordinary Least Squares (OLS) method, is known for its sensitivity to outliers, which can lead to biased and inefficient parameter estimates. Robust regression was developed to overcome the weaknesses of OLS by reducing sensitivity to outliers. Two commonly used loss functions in robust regression are Huber Loss and Tukey's Biweight Loss. This study aims to compare the performance of these two robust regression methods—Huber Loss and Tukey's Biweight in handling various outlier scenarios. Simulated data were generated with intercept and slope parameters set at 3 and 2, respectively, and outliers were systematically introduced to the X variable, the Y variable, or both, in proportions of 10%, 20%, and 30%. The analysis results indicate that Tukey's Biweight provides more stable parameter estimates under extreme outlier conditions, especially when outliers occur in the Y variable or in both X and Y. Meanwhile, Huber Loss tends to yield lower Mean Squared Error (MSE) in certain conditions, reflecting a classic trade-off between bias and variance. Therefore, Tukey's Biweight is more suitable for extreme outliers, whereas Huber Loss is more efficient under mild to moderate outlier conditions.*
Keywords*: data simulation; robust regression; Huber Loss; Tukey's Biweight;*

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

PENDAHULUAN

Regresi linier adalah salah satu metode statistik yang paling banyak digunakan, baik di bidang ekonomi, kesehatan, teknik, maupun ilmu sosial. Metode ini mempelajari hubungan antara satu variabel yang ingin diprediksi (variabel dependen) dengan satu atau lebih variabel yang dianggap berpengaruh (variabel independen). Regresi linier secara umum menggunakan pendekatan Ordinary Least Squares (OLS) untuk mengestimasi parameter. OLS bekerja dengan meminimalkan jumlah kuadrat selisih antara nilai prediksi dan data sebenarnya (James et al., 2021; Yan & Su, 2019).

Pendekatan OLS banyak digunakan karena kesederhanaannya, namun OLS sangat rentan terhadap pengaruh pencilan yang dapat menyebabkan hasil regresi menjadi tidak akurat (Fox, 2016). Salah satu cara untuk mengatasi hal ini adalah dengan menggunakan metode regresi robust. Metode ini dirancang agar lebih tahan terhadap gangguan dari pencilan. Pendekatan paling populer dalam regresi robust adalah M-estimator, yang mengandalkan penggunaan fungsi kerugian (*loss function*) dalam proses estimasi. Fungsi kerugian ini menentukan bagaimana pengaruh tiap observasi terhadap hasil akhir model.

Terdapat dua fungsi kerugian yang sering digunakan dalam regresi robust, yaitu Huber Loss dan Tukey's Biweight Loss. Huber Loss cocok untuk data yang sebagian besar normal tetapi mungkin mengandung beberapa pencilan. Fungsi ini akan memperlakukan data seperti biasa saat nilainya masih dalam batas wajar, namun menjadi lebih hati-hati jika menemukan nilai yang terlalu jauh, sehingga pengaruh pencilan bisa dikurangi (Huber, 1981). Sedangkan Tukey's Biweight Loss bekerja lebih ketat, jika suatu data terlalu jauh dari pola umum, maka fungsinya bisa mengabaikan data tersebut sehingga tidak memberi pengaruh ke model (Rousseeuw & Leroy, 1987). Karena itulah, Tukey sering digunakan ketika data mengandung banyak pencilan ekstrem.

Damayanti dan Susanti (2022) menemukan bahwa Huber Loss lebih unggul dalam memodelkan data tingkat kemiskinan, ditunjukkan oleh nilai *Adjusted* R^2 yang lebih tinggi dibanding Tukey's Biweight Loss. Sebaliknya, Pradewi dan Sudarno (2012) justru merekomendasikan Tukey's Biweight Loss pada data ketahanan pangan karena menghasilkan nilai *MSE* dan determinasi yang lebih baik. Temuan serupa juga disampaikan oleh Latifa (2019), yang menunjukkan bahwa Tukey's Biweight Loss memiliki performa terbaik berdasarkan WRMSE dan MAD dalam memodelkan produksi padi. Penelitian lain oleh Azizah & Wachidah (2022) juga mendukung efisiensi prediksi dan kestabilan estimasi dari Huber Loss pada data pengangguran di Indonesia, sementara Wu & Benkeser (2022) menerapkan Huber loss dalam konteks biaya kesehatan dan menemukan keunggulan pada distribusi data dengan outlier ekstrem. Perbedaan hasil ini menunjukkan bahwa efektivitas fungsi kerugian sangat bergantung pada karakteristik data, khususnya dalam hal pencilan dan distribusi residual. Oleh karena itu, penting untuk melakukan evaluasi lebih lanjut terhadap pemilihan fungsi kerugian, yang menjadi fokus utama dalam penelitian ini.

METODE PENELITIAN

Data

Penelitian ini menggunakan data simulasi agar kondisi dan proporsi pencilan dapat dikendalikan secara sistematis. Simulasi data memungkinkan peneliti mengevaluasi performa metode statistik di berbagai kondisi, termasuk proporsi dan jenis pencilan (Davies & Gather, 1993; Maronna et al., 2019). Variabel dependen (Y) dan independen (X) dibangkitkan secara acak sebanyak 100 dengan skema pencilan sebesar 10%, 20%, dan

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

30%. Terdapat tiga tipe gangguan dalam penelitian ini, yaitu pencilan dalam variabel dependen (Y), pencilan dalam variabel independen (X), dan pencilan dalam keduanya.

Metode Penelitian

Penelitian ini merupakan penelitian kuantitatif eksperimental yang menggunakan data simulasi untuk membandingkan performa fungsi kerugian Huber Loss dan Tukey's Biweight dalam konteks regresi robust (M-estimator).

1. **Regresi Robust dan M-Estimator**

   Regresi robust dikembangkan untuk mengatasi kelemahan metode kuadrat terkecil dengan menurunkan sensitivitas terhadap pencilan. Salah satu pendekatan yang umum digunakan adalah M-estimator. Estimasi ini meminimalkan jumlah dari fungsi kerugian $\rho(.)$ terhadap residual $r_i = y_i - x_i^\top \beta$

   $$\hat{\beta} = arg\min_{\beta} \sum_{i=1}^{n} \rho(r_i) \quad (1)$$

   Fungsi ρ dirancang agar memberikan penalti kecil untuk residual kecil, tetapi membatasi kontribusi residual besar. Pemilihan bentuk fungsi ρ menentukan jenis estimator, seperti Hubner dan Tukey's Biweight. Estimasi parameter diperoleh dengan menyelesaikan persamaan berikut:

   $$\sum_{i=1}^{n} \psi\,(r_i).\, x_i = 0 \quad (2)$$

   dengan $\psi(r_i) = \frac{d\rho(r_i)}{dr_i}$ sebagai fungsi pengaruh (*influence function*). Fungsi pengaruh ini merupakan kunci dalam menentukan seberapa besar kontribusi setiap residual terhadap estimasi parameter. Semakin kecil nilai ψ untuk residual besar, semakin robust metode tersebut terhadap pencilan (Maronna, Martin, & Yohai, 2019).

2. **Fungsi Kerugian Huber dan Tukey's Biweight**

   Fungsi kerugian Huber bersifat tansisi antara fungsi kuadrat dan linier:

   $$\rho_H(r) = \begin{cases} \frac{1}{2}r^2 & , \text{jika } |r| \le c \\ c\left(|r| - \frac{1}{2}c\right) & , \text{jika } |r| > c \end{cases} \quad (3)$$

   Fungsi pengaruhnya adalah:

   $$\psi_H(r) = \begin{cases} r & , \text{jika } |r| \le c \\ c.\,\text{sign}(r) & , \text{jika } |r| > c \end{cases} \quad (4)$$

   dengan parameter c sebagai titik transisi. Huber loss efisien untuk data normal dan tetap robust untuk pencilan moderat (Li et al., 2021).

   Tukey's Biweight Loss lebih agresif terhadap pencilan dan termauk fungsi *redescending,* yaitu fungsi pengaruhnya kembali ke nol untuk residual yang sangat besar:

   $$\rho_T(r) = \begin{cases} \frac{c^2}{6}\left[1 - \left(1 - \left(\frac{r}{c}\right)^2\right)^3\right] & , \text{jika } |r| \le c \\ \frac{c^2}{6} & , \text{jika } |r| > c \end{cases} \quad (5)$$

   $$\psi_T(r) = \begin{cases} r\left(1 - \left(\frac{r}{c}\right)^2\right)^2 & , \text{jika } |r| \le c \\ 0 & , \text{jika } |r| > c \end{cases} \quad (6)$$

   Tukey sangat cocok digunakan ketika data mengandung pencilan berat, namun kurang efisien pada data yang bersih (Zhou et al., 2017).

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

3. Evaluasi Regresi Robust Melalui Simulasi

Pada penelitian ini, model dievaluasi menggunakan *Mean Squared Error* (MSE), yang merupakan salah satu ukuran kesalahan paling umum dalam analisis regresi. Menurut Ranglani (2024), model dengan nilai MSE yang rendah umumnya menunjukkan trade-off yang optimal antara bias dan varians, terutama pada algoritma ensemble seperti Random Forest dan Gradient Boosting. Secara matematis, MSE didefinikan sebagai berikut:

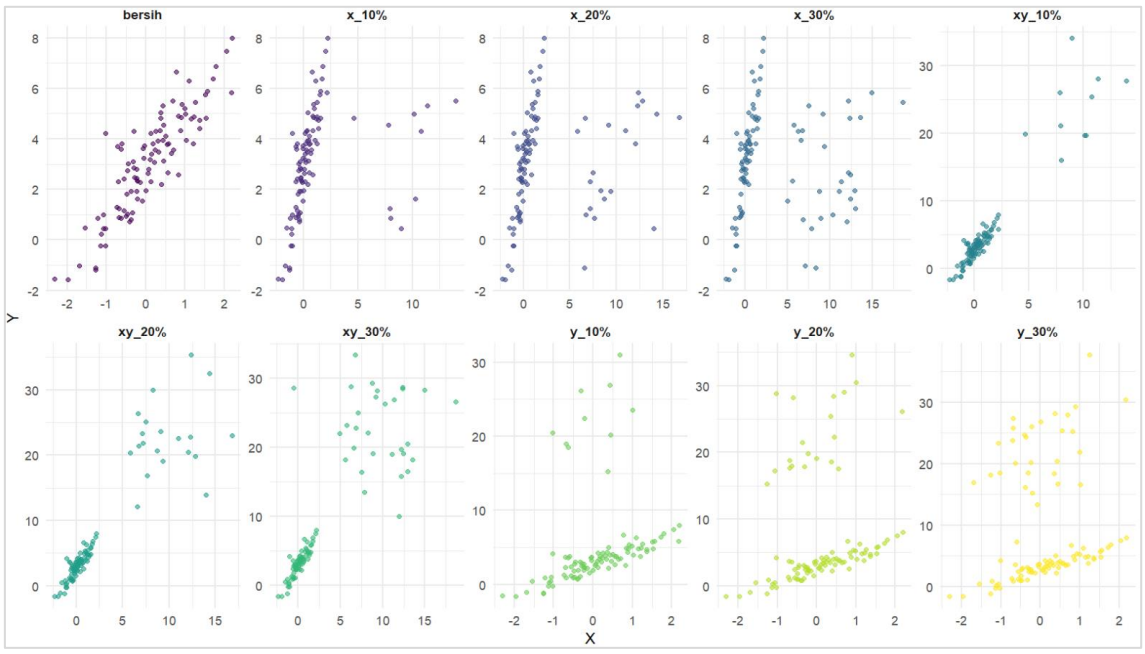
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$

Adapun langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Membangkitkan data simulasi untuk variabel Y dan variabel X sebanyak $n = 100$.
2. Menambahkan pencilan sebanyak 10%, 20%, dan 30% pada variabel Y, variabel X, dan keduanya.
3. Melakukan analisis regresi robust dengan Huber Loss dan Tukey's Biweight Loss.
4. Menghitung nilai MSE dari tiap skenario model yang diperoleh.
5. Membandingkan nilai MSE yang diperoleh.
6. Menarik kesimpulan.

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data simulasi dengan berbagai skema yang telah dijelaskan sebelumnya. Nilai intersep dan slope pada simulasi ini ditetapkan sebesar 3 dan 2. Gambar 1 merupakan data simulasi yang telah dibangkitkan.



Gambar 1. Eksplorasi Data Simulasi

Gambar 1 menunjukkan hubungan antara variabel X dan Y dari berbagai skema simulasi yang telah dilakukan. Grafik pertama merupakan data simulasi tanpa pencilan yang memperlihatkan hubungan linear antara X dan Y. Titik-titik data tersebar rapat

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

membentuk pola garis lurus, mencerminkan kondisi ideal untuk analisis regresi biasa (OLS). Pada grafik “x_10%”, “x_20%”, dan “x_30%”, pencilan ditambahkan hanya pada variabel X. Artinya, sebagian kecil data "menyimpang" secara horizontal (ke arah sumbu X), sementara nilai Y tetap seperti biasa. Semakin banyak pencilan, maka pola sebaran data semakin tidak jelas. Hubungan linear masih ada, tapi mulai kabur karena titik-titik pencilan mengganggu arah umum data.

Sementara itu, grafik “y_10%”, “y_20%”, dan “y_30%” menunjukkan pencilan hanya pada variabel Y. Berbeda dengan sebelumnya, kali ini titik-titik melonjak secara vertikal (ke arah sumbu Y). Hasilnya, hubungan linear mulai terganggu karena beberapa nilai Y menjadi sangat ekstrem meskipun nilai X-nya masih normal. Semakin banyak pencilan, semakin besar efeknya terhadap bentuk hubungan antar variabel. Hal yang lebih menarik terlihat pada grafik “xy_10%”, “xy_20%”, dan “xy_30%”, di mana pencilan ditambahkan secara bersamaan pada X dan Y. Titik-titik data jadi menyebar secara acak dan jauh dari pola utama. Pencilan seperti ini paling merusak karena menyimpang di dua arah sekaligus, sehingga pola hubungan linear hampir hilang sama sekali, terutama saat proporsinya mencapai 30%.

Secara keseluruhan, Gambar 1 menggambarkan bagaimana pencilan bisa merusak hubungan antara variabel X dan Y. Ketika data bersih, model regresi bisa bekerja dengan baik. Tapi begitu terdapat pencilan apalagi dalam jumlah besar, membuat model menjadi tidak akurat. Oleh karena itu, dalam situasi seperti ini, pendekatan regresi yang lebih tahan terhadap pencilan, seperti regresi robust dengan pendekatan Huber Loss atau Tukey’s Biweight Loss, menjadi solusi yang lebih baik.

Setelah melakukan eksplorasi data simulasi yang telah dibangkitkan, selanjutnya melakukan analisis menggunakan regresi robust dengan pendekatan Huber Loss dan Tukey’s Biweight. Dua pendekatan ini dipilih karena merupakan fungsi kerugian yang secara khusus dirancang untuk mengatasi kelemahan regresi OLS dalam menghadapi pencilan. Pada penelitian ini akan dibandingkan dari dua pendekatan tersebut mana yang lebih baik dalam menghadapi data pencilan dari berbagai skema simulasi yang telah dilakukan. Tabel 1 merupakan hasil estimasi parameter regresi (intercept dan slope) pada berbagai skema pencilan menggunakan metode Huber Loss dan Tukey’s Biweight.

Tabel 1. Hasil Estimasi Paramter Regresi

Variabel	Proporsi Pencilan	Huber Loss		Tukey’s Biweight	
		Intersep	Slope	Intersep	Slope
XY	0%	2.84	1.99	2.84	2.01
Y	10%	2.98	1.92	2.80	1.96
Y	20%	3.34	1.90	2.81	1.99
Y	30%	5.68	1.53	2.90	1.99
X	10%	2.90	0.23	2.90	0.22
X	20%	2.87	0.11	2.88	0.11
X	30%	3.01	0.03	3.02	0.03
XY	10%	2.81	2.03	2.80	2.13
XY	20%	2.87	1.91	2.81	1.88
XY	30%	3.09	1.78	2.91	2.11

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Tabel 1 merupakan hasil estimasi parameter menggunakan Huber Loss dan Tukey's Biweight. Pada data bersih, metode Huber maupun Tukey's Biweight memberikan estimasi yang sangat akurat, dengan nilai intersep dan slope yang mendekati parameter sebenarnya, yaitu 3 dan 2. Namun, saat pencilan mulai ditambahkan ke dalam data, performa kedua metode mulai menunjukkan perbedaan. Pada skema pencilan dalam variabel Y, metode Huber menghasilkan estimasi slope yang semakin menurun seiring bertambahnya proporsi pencilan, dari 1.92 pada 10% pencilan menjadi hanya 1.53 pada 30% pencilan. Sebaliknya, metode Tukey menunjukkan kestabilan yang lebih tinggi, dengan nilai slope tetap mendekati 2 bahkan hingga proporsi pencilan sebesar 30%. Hal ini menunjukkan bahwa Tukey lebih tahan terhadap pencilan pada variabel dependen.

Pada skema pencilan dalam variabel X, kedua metode sama-sama mengalami penurunan performa yang sangat drastis. Slope estimasi mendekati nol pada pencilan 30%, baik pada metode Huber maupun Tukey, yang mengindikasikan ketidakmampuan keduanya dalam menangani pencilan ekstrem pada variabel independen. Sementara itu, pada skema pencilan dalam variabel X dan Y, metode Tukey kembali menunjukkan keunggulannya dengan estimasi slope yang tetap mendekati nilai sebenarnya, sedangkan Huber cenderung menghasilkan estimasi yang semakin menjauh dari nilai ideal seiring meningkatnya proporsi pencilan. Secara keseluruhan, hasil ini menunjukkan bahwa Tukey's Biweight lebih stabil dan robust terhadap pencilan, terutama pada variabel Y dan kombinasi pencilan, dibandingkan Huber, yang lebih rentan terhadap deviasi parameter dalam kondisi ekstrem.

Tabel 2 menunjukkan bahwa meskipun metode Tukey's Biweight menunjukkan estimasi parameter yang lebih stabil pada berbagai skema pencilan, nilai MSE yang dihasilkan dalam beberapa kondisi justru sedikit lebih besar dibandingkan Huber. Hal ini mencerminkan adanya *trade-off* klasik antara bias dan variansi. Istilah *trade-off* klasik dalam statistik merujuk pada kompromi antara dua karakteristik model yang saling bertentangan, yaitu bias dan variansi. Secara umum, pengurangan bias dapat menyebabkan peningkatan variansi, sedangkan upaya mengurangi variansi sering kali meningkatkan bias. Kompromi ini menjadi dasar penting dalam pemilihan model yang optimal, terutama dalam konteks regresi robust (James et al., 2021; Hastie, Tibshirani, & Friedman, 2009).

Tabel 2. Nilai MSE

Variabel	Proporsi Pencilan	MSE	
		Huber Loss	Tukey's Biweight
XY	0%	0.93	0.93
Y	10%	39.89	40.59
Y	20%	80.82	84.76
Y	30%	96.72	121.15
X	10%	3.66	3.66
X	20%	3.85	3.85
X	30%	4.03	4.03
XY	10%	4.22	4.12
XY	20%	11.21	11.21
XY	30%	26.95	30.92

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Metode Tukey's Biweight cenderung mengurangi bias dengan cara menghilangkan kontribusi pencilan secara ekstrem, namun hal ini bisa meningkatkan variansi prediksi. Sebaliknya, Huber mempertahankan beberapa informasi dari pencilan dan memberikan estimasi yang cenderung kompromistis, sehingga pada beberapa kasus memiliki MSE lebih rendah (Huber, 1981).

KESIMPULAN DAN SARAN

Hasil penelitian ini memperlihatkan bahwa pencilan dalam data dapat berdampak terhadap hasil estimasi model regresi. Ketika diuji dengan dua metode regresi robust, yaitu Huber Loss dan Tukey's Biweight, terlihat bahwa Tukey's Biweight cenderung memberikan hasil estimasi yang lebih stabil, terutama saat pencilan muncul pada variabel Y atau pada variabel X dan Y. Meski begitu, metode ini menghasilkan nilai MSE yang sedikit lebih besar dibanding Huber Loss. Ini menunjukkan adanya kompromi antara bias dan variansi (*trade-off*), di mana metode Tukey's Biweight lebih fokus mengurangi bias, sementara Huber Loss menjaga variansi tetap rendah. Artinya, pemilihan metode yang tepat sangat bergantung pada kondisi data. Jika data mengandung banyak pencilan ekstrem, Tukey's Biweight lebih bisa diandalkan. Sebaliknya, jika terdapat pencilan tidak ekstrem, Huber Loss bisa menjadi pilihan yang lebih efisien.

DAFTAR PUSTAKA

Azizah, R. J., & Wachidah, L. (2022). Regresi Robust Estimasi-M dengan Pembobot Huber dan Tukey Bisquare pada Data Tingkat Pengangguran di Indonesia Menurut Provinsi Tahun 2020. *Bandung Conference Series: Statistics, 2*(2), 18–26.

Damayanti, R., & Susanti, Y. (2022). Perbandingan Regresi Robust M-Estimator dengan Pembobot Huber dan Tukey pada Data Tingkat Kemiskinan di Indonesia*. Prosiding Seminar Nasional Statistika (Semnas Statistika)*, Universitas Islam Bandung.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd ed.). Sage Publications.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer.

Huber, P. J. (1981). *Robust Statistics*. New York: John Wiley & Sons.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning* (2nd ed.). Springer.

Latifa, N. (2019). *Analisis Regresi Robust Estimasi-M pada Data Produksi Padi dengan Pembobot Huber, Tukey, dan Hampel di Kabupaten Cirebon tahun 2011–2016* (Skripsi, Universitas Jenderal Soedirman). https://repository.unsoed.ac.id/17915

Li, G., Zhong, S., & Zhu, Y. (2021). Robust Regression Estimation Using Adaptive Huber Loss. *Statistics & Probability Letters*, 173, 109073.

Maronna, R. A., Martin, R. D., & Yohai, V. J. (2019). *Robust Statistics: Theory and Methods (with R)* (2nd ed.). Wiley.

Pradewi, E. D., & Sudarno, S. (2012). Kajian Estimasi-M IRLS Menggunakan Fungsi Pembobot Huber dan Bisquaret Tukey pada Data Ketahanan Pangan di Jawa Tengah. Evaluasi Regresi Robust Estimasi-M dengan Fungsi Huber dan Tukey pada Data

Volume 3, No. 2, 2025, 78-85
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Ketahanan Pangan. *Media Statistika*, *5*(1), 1–9. https://doi.org/10.14710/medstat.5.1.1-10

Ranglani, H. (2024). Empirical Analysis of The Bias–Variance Trade off Across Machine Learning Models. *Machine Learning and Applications: An International Journal*, *11*(4), 1–12. 10.5121/mlaij.2024.11401

Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley.

Wu, Z., & Benkeser, D. (2022). *A Huber Loss-Based Super Learner with Applications to Healthcare Expenditures*. arXiv preprint.

Yan, X., & Su, X. G. (2019). *Linear Regression Analysis: Theory and Computing* (2nd ed.). Springer.

Zhou, W., Song, Q., & Wei, Y. (2017). Adaptive Robust Regression for High-Dimensional Data. *Journal of Multivariate Analysis*, 157, 53–66.