KONSISTENSI KOEFISIEN DETERMINASI SEBAGAI UKURAN KESESUAIAN MODEL PADA REGRESI ROBUST

Harmi Sugiarti (harmi@ut.ac.id)
Andi Megawarni
Jurusan Statistika FMIPA Universitas Terbuka

ABSTRACT

In statistics, the coefficient of determination can be used to assess the suitability of a model with the data. If there are outliers in the data, the coefficient of determination obtained by the OLS method is not consistent. The purpose of this study was to compare the coefficient of determination of regression lines obtained by the OLS, the M and the LMS methods as a measure of the suitability model. The result showed that when the data contains no-outlier, the LMS method is as consistent as the OLS and the M methods concerning the coefficient of determinations. When the data contain outliers, the LMS method is more consistent than the OLS and the M methods. This result was based on real data with 9.1% outliers.

Keywords: LMS estimator, M estimator, outlier data, robust regression, The Consistency of Coefficient of Determination

ABSTRAK

Dalam statistik, koefisien determinasi dapat digunakan untuk menilai kesesuaian model dengan data. Jika ada outlier pada data, koefisien determinasi yang diperoleh dengan metode OLS tidak konsisten. Tujuan dari penelitian ini adalah untuk membandingkan koefisien determinasi dari garis regresi yang diperoleh melalui metode OLS, M dan metode LMS sebagai ukuran model kesesuaian. Hasil penelitian menunjukkan bahwa ketika data tidak mengandung-outlier, metode LMS adalah konsisten, serupa dengan metode OLS dan metode M terkait dengan koefisien determinasi. Ketika data mengandung outlier, metode LMS lebih konsisten daripada metode OLS dan metode M. Hasil ini berdasarkan ujicoba pada data nyata dengan outlier 9,1%.

Kata kunci: data outlier, konsistensi koefisien determinasi, LMS estimator, M estimator, regresi robust

Sebelum melakukan inferensi parameter model regresi, dianggap perlu untuk mengetahui apakah model yang diperoleh sudah sesuai dengan data yang ada. Ketidaksesuaian model regresi yang dibangun dapat juga disebabkan karena data tidak memenuhi asumsi, misalnya data mengandung pencilan (*outlier),* yaitu pengamatan dengan sisaan yang cukup besar. Penolakan begitu saja suatu pencilan bukanlah prosedur yang bijaksana, karena adakalanya pengamatan pencilan memberikan informasi yang cukup berarti, misalnya karena pencilan muncul dari kombinasi keadaan yang tidak biasa yang mungkin saja sangat penting dan perlu diselidiki lebih lanjut. Pengamatan pencilan dapat merupakan pengamatan yang berpengaruh, artinya pengamatan yang dapat mempengaruhi hasil pendugaan koefisien regresi, sehingga tindakan membuang pengamatan yang berpengaruh akan mengubah secara berarti persamaan regresi serta kesimpulannya (Draper & Smith, 1981).

Menurut Myers (1990), keberadaan pengamatan yang berpengaruh dapat diperiksa dengan melihat perbedaan dugaan peubah tak bebas terbakukan (*DFFITS*) yang dirumuskan sebagai: $(DFFITS)_i = \frac{\hat{y}_i - \hat{y}_{i,-i}}{s_{-i}\ \sqrt{h_{ii}}}$ dengan $\hat{y}_i =$ nilai pendugaan y_i, $\hat{y}_{i,-i} =$ nilai pendugaan y_i tanpa pengamatan ke-i, $s_{-i} =$ dugaan simpangan baku tanpa pengamatan ke-i dan $h_{ii} =$ unsur ke-i dari diagonal matriks topi. Jika p menyatakan banyaknya parameter dan n menyatakan banyaknya pengamatan, maka suatu pengamatan akan merupakan pengamatan berpengaruh dalam persamaan regresi apabila mempunyai nilai $|DFFITS|_i > 2\sqrt{(p/n)}$.

Ada beberapa ukuran yang dapat dipergunakan untuk mengetahui apakah model yang diperoleh sudah sesuai dengan data, diantaranya adalah koefisien determinasi, biasanya dinyatakan dengan R^2 yang menunjukkan proporsi variasi variabel dependen yang dijelaskan oleh variasi variabel independen. Selain memberikan penaksir parameter β_0, β_1, β_2,..., β_p yang bersifat tak bias linear terbaik dari model regresi $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} + \varepsilon_i$, $i = 1, 2, ..., n$, metode OLS (*ordinary least square*) memberikan ukuran R^2 yang sangat diperlukan dalam pemodelan, yakni:

$$R^2_{OLS} = \frac{JKR}{JKT} = 1 - \frac{JKS}{JKT} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

dengan JKR menyatakan jumlah kuadrat regresi, JKS menyatakan jumlah kuadrat sisaan, dan JKT menyatakan jumlah kuadrat total. Koefisien determinasi bernilai $0 < R^2_{OLS} < 1$.

Apabila terdapat asumsi yang tidak dipenuhi, khususnya jika dalam data terdapat pencilan, maka patut dicoba metode yang bersifat tidak sensitif terhadap pelanggaran asumsi-asumsi, yakni regresi *robust*. Beberapa metode pendugaan/penaksiran koefisien garis regresi yang bersifat *robust* telah dikembangkan, diantaranya adalah metode pendugaan parameter regresi berdasarkan pada penduga M *(maximum likelihood estimators)* dan penduga LMS *(least median of square estimators)*.

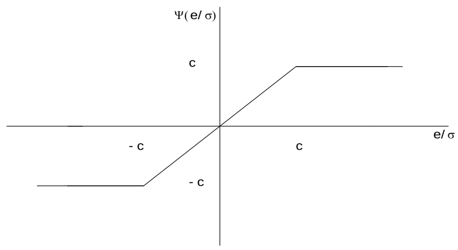
Menurut Staudte dan Sheather (1990), jika hubungan linear antara satu peubah respons dengan peubah-peubah bebasnya dimodelkan sebagai: $y_i = x_i' \beta + \varepsilon_i$ dengan x_i' menyatakan baris ke-i dari matriks rancangan X, β menyatakan parameter model dan ε_i menyatakan suku galat. Nilai $\hat{y}_i$ dan sisaan (e_i) masing-masing didefinisikan sebagai $\hat{y}_i = x_i' \beta$ dan $e_i = y_i - \hat{y}_i = y_i - x_i' \beta$. Penduga M untuk model dengan p parameter $(\hat{\beta}_M)$ diperoleh dengan cara meminimumkan fungsi konveks $\rho(x,e)$ yakni: $\min\sum_i \rho(x_i, e_i) = \min\sum_i \rho\left(x_i, y_i - x_i' \hat{\beta}_M\right)$ atau mencari penyelesaian dari persamaan: $\sum_i x_i\, \Psi\left(x_i, y_i - x_i' \hat{\beta}_M\right) = 0$, dengan $\Psi(x,e) = \rho'(x,e)$ untuk berbagai fungsi konveks $\rho(x,e)$ yang dapat diturunkan dan memenuhi $\Psi(x,0) = 0$. Penduga $\hat{\beta}_M$ yang diperoleh bukan merupakan skala *invariant*, yaitu jika sisaannya $\left(e_i = y_i - x_i' \hat{\beta}_M\right)$ digandakan dengan suatu konstanta akan diperoleh penyelesaian yang tidak sama seperti sebelumnya. Skala *invariant* dapat diperoleh dengan menggunakan nilai $\frac{e_i}{\sigma}$ sebagai pengganti e_i dan σ adalah faktor skala yang juga perlu diduga, sehingga persamaan yang ada menjadi:

$\sum_i x_i \, \Psi\left(x_i , \frac{e_i}{\sigma} \right) = \sum_i x_i \, \Psi\left(x_i , \frac{y_i - x_i' \, \hat{\beta}_M}{\sigma} \right) = \sum_i x_i \left(y_i - x_i' \, \hat{\beta}_M \right) w_i = 0$ dengan fungsi pembobot

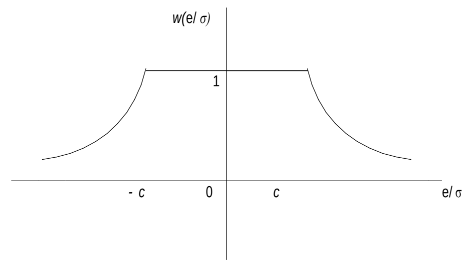
$w_i = w\left(x_i , \frac{y_i - x_i' \hat{\beta}_M}{\sigma} \right) = \dfrac{\Psi\left(x_i , \frac{e_i}{\sigma} \right)}{\frac{e_i}{\sigma}}$ yang bernilai antara 0 dan 1. Secara umum fungsi pembobot dirumuskan sebagai $w_i = w\left(x_i , \frac{y_i - x_i' \hat{\beta}_M}{\sigma} \right) = \frac{\sigma \nu(x_i)}{e_i} \Psi\left(\frac{e_i}{\sigma \nu(x_i)} \right)$ dengan Ψ adalah *influence function*, $\nu(x_i)$ adalah suatu fungsi yang tidak diketahui dan tergantung pada x melalui nilai *leverage*. Nilai pembobot w_i merupakan kombinasi nilai *leverage* dan *studentized residual* melalui $DFFITS$ yang diperoleh dengan memilih fungsi Huber Ψ yang berbentuk:

$$\Psi\left(\frac{e}{\sigma}\right) = \begin{cases} c & , \text{ jika } \frac{e}{\sigma} > c \\ \frac{e}{\sigma} & , \text{ jika } \left| \frac{e}{\sigma} \right| \le c \\ -c & , \text{ jika } \frac{e}{\sigma} < -c \end{cases}$$

dan menentukan nilai $\nu(x_i) = \frac{(1 - h_{ii})}{\sqrt{h_{ii}}}$ serta $\hat{\sigma} = s_{(i)}$. Fungsi Huber Ψ dan fungsi pembobot Huber w masing-masing dapat digambar seperti Gambar 1 dan Gambar 2. Secara singkat nilai pembobot w_i dinyatakan dalam bentuk: $w\left(x_i , \frac{y_i - x_i' \hat{\beta}_M}{\sigma} \right) = w\left(x_i , \frac{e_i}{\sigma} \right) = \min\left(\frac{2\sqrt{p/n}}{|DFFITS|_i} , 1 \right)$. Dengan demikian persamaan $\sum_i \left(y_i - x_i' \hat{\beta}_M \right) w_i x_i = 0$ dapat dituliskan dalam bentuk matriks $X'WX\beta = X'WY$ yang dikenal sebagai persamaan normal kuadrat terkecil tertimbang dengan W adalah matriks diagonal yang berisi pembobot. Solusi persamaan normal tersebut akan memberikan dugaan untuk β yaitu $\hat{\beta}_M = \left(X'WX \right)^{-1} \left(X'WY \right)$ dan penduga-M untuk β diperoleh dengan cara melakukan iterasi sampai diperoleh suatu hasil yang konvergen. Cara ini biasa dikenal sebagai metode kuadrat terkecil tertimbang secara iteratif *(iteratively reweighted least square).*



Gambar 1. Fungsi Huber



Gambar 2. Fungsi Pembobot Huber

Berdasarkan pembobot w_i dan $\hat{\beta}_M$, matriks varians-kovarians untuk $\hat{\beta}_M$ yakni Σ_n dapat didekati dengan persamaan: $\Sigma_n = \frac{1}{n-p}\left(X'D_1X\right)^{-1}\left(X'D_2X\right)\left(X'D_1X\right)^{-1}$ dengan D_1 menyatakan matriks diagonal yang elemen-elemen diagonalnya adalah $\Psi'\left(\frac{e_i}{\sigma v(x_i)}\right)$ dan D_2 menyatakan matriks diagonal dengan elemen-elemen diagonalnya $w_i^2\, e_i^2$ (Staudte & Sheather,1990).

Maronna, dkk (2006) mengusulkan koefisien determinasi dapat dihitung dengan rumus sebagai berikut: $R_M^2 = 1 - \frac{\sum_{i=1}^{n}\rho\left(\frac{y_i - x_i'\hat{\beta}_M}{\hat{\sigma}}\right)}{\sum_{i=1}^{n}\rho\left(\frac{y_i - \hat{\mu}}{\hat{\sigma}}\right)}$ dengan $\hat{\mu}$ merupakan penaksir M untuk $E(y)$ yakni solusi dari $\min \sum_{i=1}^{n}\rho\left(\frac{y_i - \mu}{\hat{\sigma}}\right)$, serta $\hat{\beta}_M$ dan $\hat{\sigma}$ masing-masing adalah penaksir M untuk β dan σ yang diperoleh berdasarkan fungsi ρ.

Metode lain yang bersifat *robust* untuk penaksiran koefisien garis regresi adalah metode LMS *(least median square)*. Metode ini mempunyai keuntungan untuk mengurangi pengaruh dari sisaan. Menurut Rousseeuw dan Leroy (2003), penduga LMS diperoleh dengan mencari model regresi yang meminimumkan median dari *h* kuadrat sisaan (e_i^2) atau didefinisikan sebagai $\hat{\beta}_{LMS} = \arg\min_{\beta}\ \underset{i}{median}\ e_i^2$ dengan $e_i^2 = \left(y_i - x_i'\beta\right)^2$; $i = 1, 2, ..., n$.

Ukuran sebaran dari galat dapat ditaksir dengan cara menentukan dulu nilai awal $s^0 = 1,4826\left(1 + \frac{5}{(n-p)}\right)\sqrt{\underset{i}{median}\ e_i^2(\hat{\beta}_{LMS})}$. Faktor $1,4826 = \frac{1}{\Phi^{-1}(0,75)}$ digunakan karena $\frac{median_i\ |z_i|}{\Phi^{-1}(0,75)}$ merupakan penaksir konsisten untuk σ jika z_i berdistribusi $N\left(0, \sigma^2\right)$. Selanjutnya nilai awal s^0 digunakan untuk menentukan pembobot w_i untuk setiap pengamatan, yaitu:

$$w_i = \begin{cases} 1 & \text{jika } |e_i/s_0| \le 2,5 \\ 0 & \text{jika } |e_i/s_0| > 2,5 \end{cases}$$

Berdasarkan pembobot w_i, maka nilai akhir taksiran σ untuk regresi LMS didefinisikan sebagai: $\hat{\sigma} = \sqrt{\left(\sum_{i=1}^{n} w_i\ e_i^2\right) \Big/ \left(\sum_{i=1}^{n} w_i - p\right)}$ dan koefisien determinasi untuk metode LMS adalah:

$$R_{LMS}^2 = 1 - \left(\frac{med|\,e_i\,|}{mad(y_i)}\right)^2 = 1 - \left(\frac{med\left|y_i - x_i'\,\hat{\beta}_{LMS}\right|}{\underset{i}{med}\left\{\left|y_i - \underset{j}{med}\,y_j\right|\right\}}\right)^2$$ Rousseeuw dan Leroy (2003)

Pada penelitian pendahuluan diperoleh hasil bahwa untuk data yang tidak mengandung pencilan, metode regresi *robust* dengan penduga LMS kurang efisien dibanding metode M, sedangkan untuk data yang mengandung pencilan, metode regresi *robust* dengan penduga LMS lebih efisien dibanding metode M (Sugiarti & Megawarni, 2010).

Jika koefisien determinasi untuk model dihitung berdasarkan metode OLS, metode M, dan metode LMS maka diharapkan diperoleh informasi yang lebih detil tentang hubungan yang ada di antara variabel independen dan variabel dependen dalam model regresi. Dengan kata lain penentuan koefisien determinasi diharapkan dapat digunakan sebagai indikator untuk mengetahui apakah model yang diperoleh sudah sesuai dengan data. Tulisan ini bertujuan mengkaji konsistensi koefisien determinasi sebagai ukuran kesesuaian model garis regresi yang diperoleh dengan metode OLS, metode M, dan metode LMS untuk data yang mengandung pencilan maupun tidak.

METODE

Ada dua jenis data yang digunakan dalam kajian ini, yaitu data simulasi berupa data bangkitan yang diperoleh dengan bantuan program MINITAB versi 13.1, serta data terapan berupa nilai Tugas Tutorial Online (Tuton), Nilai Partisipasi Tuton, dan nilai Ujian Akhir Semester (UAS) mata kuliah Metode Statistik I masa ujian 2008.1-2010.1. Adapun langkah-langkah yang dilakukan adalah: (1) membangkitkan sebanyak 40 pasang data sebagai peubah bebas $(x_1,\ x_2)$ dan data galat (ε) dengan $\varepsilon \sim NIID(0,\sigma^2)$, (2) menentukan peubah tak bebas (y) dengan asumsi nilai $(\beta_0, \beta_1, \beta_2)$ tertentu untuk model $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$, (3) mendapatkan pengamatan pencilan dengan mengganti sejumlah tertentu pengamatan y dengan nilai ekstrim sedemikian sehingga diperoleh pengamatan pencilan yang berpengaruh, (4) mencari penaksir OLS, M, LMS untuk koefisien garis regresi untuk data simulasi dengan atau tanpa pencilan, (5) menentukan koefisien determinasi R_{OLS}^2, R_M^2, dan R_{LMS}^2 untuk data simulasi dengan atau tanpa pencilan, (6) mencari penaksir OLS, M, LMS untuk koefisien garis regresi untuk data terapan, dan (7) menentukan koefisien determinasi R_{OLS}^2, R_M^2, dan R_{LMS}^2 untuk data terapan.

HASIL DAN PEMBAHASAN

Sebanyak empat puluh galat berdistribusi Normal dengan mean 0 dan variansi 1 dibangkitkan secara random dengan paket program MINITAB. Jika asumsikan $\beta_0 = \beta_1 = \beta_2 = 1$, simulasi memberikan empat puluh pasang data $(y, x_1,\ x_2)$. Penaksir koefisien garis regresi $(\hat{\beta})$ dan koefisien determinasi (R^2) untuk data simulasi tanpa pencilan dapat dilihat pada Tabel 1.

Tabel 1. Penaksir Koefisien Garis Regresi dan Koefisien Determinasi untuk Data Tanpa Pencilan

Koefisien	OLS	M	LMS
β_0	0,510	0,526	1,107*
β_1	0,973*	0,972*	0,969*
β_2	1,130*	1,129*	0,969*
R^2	0,961	0,994	0,982

$*$ Signifikan pada $\alpha = 5\%$

Pada data tanpa pencilan, ketiga metode memberikan nilai koefisien determinasi yang tidak jauh berbeda, yaitu metode OLS memberikan koefisien determinasi $R_{OLS}^2 = 0,961$ artinya 96,1% variabilitas dalam y dapat dijelaskan oleh x_1 dan x_2. Metode M memberikan koefisien determinasi $R_M^2 = 0,994$ artinya 99,4% variabilitas dalam y dapat dijelaskan oleh x_1 dan x_2 serta metode LMS memberikan koefisien determinasi $R_M^2 = 0,982$ artinya 98,2% variabilitas dalam y dapat dijelaskan oleh x_1 dan x_2. Hal ini menunjukkan bahwa metode OLS, metode M, dan metode LMS menyatakan model regresi linear sesuai untuk data, sehingga inferensi tentang koefisien garis regresi dapat dilakukan.

Penaksir koefisien garis regresi dan koefisien determinasi untuk data simulasi yang mengandung 5% pencilan dapat dilihat pada Tabel 2. Metode OLS memberikan koefisien determinasi $R_{OLS}^2 = 0,616$, metode M memberikan koefisien determinasi $R_M^2 = 0,846$ dan metode LMS memberikan koefisien determinasi $R_{LMS}^2 = 0,983$. Nilai koefisien determinasi yang diperoleh dengan metode OLS dan metode M menjadi lebih kecil, sedangkan nilai koefisien determinasi yang diperoleh dengan metode LMS tidak berubah. Hal ini menunjukkan bahwa metode LMS masih konsisten dalam memberikan nilai koefisien determinasi dibanding metode OLS dan metode M. Pada Tabel 2 juga dapat dilihat adanya perubahan signifikansi koefisien β_0 untuk metode M, yakni menjadi signifikan pada data dengan 5% pencilan.

Tabel 2. Penaksir Koefisien Garis Regresi dan Koefisien Determinasi untuk Data dengan 5% Pencilan

Koefisien	OLS	M	LMS
β_0	2,659	1,252*	0,994*
β_1	0,805*	0,905*	0,925*
β_2	0,969*	1,094*	1,017*
R^2	0,616	0,846	0,983

$*$ Signifikan pada $\alpha = 5\%$

Penaksir koefisien garis regresi dan koefisien determinasi untuk data simulasi dengan 10% pencilan dapat dilihat pada Tabel 3. Metode OLS memberikan koefisien determinasi $R_{OLS}^2 = 0,445$, metode M memberikan koefisien determinasi $R_M^2 = 0,588$ dan metode LMS memberikan koefisien determinasi $R_{LMS}^2 = 0,979$.

Tabel 3. Penaksir Koefisien Garis Regresi dan Koefisien Determinasi untuk Data dengan 10% Pencilan

Koefisien	OLS	M	LMS
β_0	4,015*	2,700*	2,137*
β_1	0,774*	0,862*	0,866*
β_2	0,829*	0,956*	0,885*
R^2	0,445	0,663	0,979

$*$ Signifikan pada $\alpha = 5\%$

Jika dilihat dari konsistensi nilai koefisien determinasi untuk masing-masing metode menunjukkan bahwa R^2_{LMS} lebih konsisten dibanding R^2_{OLS} dan R^2_M. Hal ini menunjukkan bahwa pada data tanpa pencilan, data mengandung 5% pencilan, dan data mengandung 10% pencilan, ukuran koefisien determinasi yang diberikan oleh metode LMS lebih konsisten dibanding metode OLS dan metode M. Demikian juga dengan signifikansi koefisien regresi, metode LMS lebih konsisten dibanding metode OLS dan metode M. Hasil ini sedikit berbeda dengan hasil kajian sebelumnya yang menunjukkan bahwa metode LMS kurang efisien dibanding metode M untuk data yang tidak mengandung pencilan, tetapi metode LMS lebih efisien dibanding metode M dalam menaksir koefisien garis regresi untuk data yang mengandung pencilan (Sugiarti & Megawarni, 2010).

Koefisien determinasi untuk data terapan dapat dilihat pada Tabel 4, metode OLS memberikan koefisien determinasi $R^2_{OLS} = 0,042$ artinya 4,2% variabilitas dalam nilai UAS dapat dijelaskan oleh nilai Tugas 1, Tugas 2, Tugas 3, dan nilai partisipasi mahasiswa. Metode M memberikan koefisien determinasi $R^2_M = 0,879$ dan metode LMS memberikan koefisien determinasi $R^2_{LMS} = 0,709$.

Tabel 4. Penaksir Koefisien Garis Regresi dan Koefisien Determinasi untuk Data Terapan

Koefisien	OLS	M	LMS
Konstanta	35,510 *	35,314*	29,678*
Tugas 1	0,092	-0,001	0,064
Tugas 2	0,005	0,004	-0,048
Tugas 3	0,003	-0,054	-0,079*
Partisipasi	-0,057	0,100	0,160*
R^2	0,042	0,879	0,709

$*$ Signifikan pada $\alpha = 5\%$

Berdasarkan hasil simulasi, metode LMS dianggap lebih konsisten memberikan ukuran koefisien determinan dibanding metode OLS dan metode M, sehingga penaksir koefisien garis regresi yang diperoleh dengan metode LMS menunjukkan bahwa nilai Tugas 3 dan Partisipasi cukup signifikan mempunyai pengaruh linear terhadap nilai UAS mahasiswa. Namun, karena penaksir koefisien garis regresi untuk nilai Tugas 3 bernilai negatif $(\hat{\beta}_3 = -0,079)$, materi Tugas 3 perlu ditinjau kembali. Demikian juga untuk Tugas 1 dan Tugas 2, karena penaksir koefisien garis regresi yang diperoleh tidak cukup signifikan mempunyai pengaruh linear terhadap nilai UAS mahasiswa maka materi Tugas 1 dan Tugas 2 perlu revisi sehingga pemberian materi Tugas 1, Tugas 2, Tugas 3, dan partisipasi mahasiswa mempunyai pengaruh linear yang signifikan dalam meningkatkan nilai UAS mahasiswa.

KESIMPULAN

Secara umum dapat disimpulkan bahwa metode OLS, metode M, dan metode LMS memberikan nilai koefisien determinasi yang hampir sama pada data yang tidak mengandung pencilan. Pada data yang mengandung pencilan, metode LMS memberikan nilai koefisien determinasi yang tidak jauh berbeda dengan data yang tidak mengandung pencilan, tetapi metode OLS dan metode M memberikan nilai koefisien determinasi yang jauh lebih kecil. Dengan kata lain, metode LMS lebih konsisten dibanding metode OLS dan metode M dalam memberikan nilai koefisien determinasi. Pada data terapan yang mengandung pencilan, metode LMS lebih konsisten dibanding metode OLS dan metode M dalam memberikan nilai koefisien determinasi.

REFERENSI

Draper, N.R. & Smith, H. (1981). *Applied regression analysis* (*2nd ed*). New York: Wiley.

Maronna, R.A., Martin, R.D., & Yohai, V.J. (2006). *Robust statistics: Theory and Methods*. Chichester, West Sussex, UK: Wiley.

Myers, R.H. (1990). *Classical and modern regression with applications (2nd ed)*. Boston: PWS- Kent.

Rousseeuw,P.J. & Leroy,A.M. (2003). *Robust regression and outlier detection*. New York: Wiley.

Staudte, R.G. & Sheather, S.J. (1990). *Robust estimation and testing*. New York: Wiley.

Sugiarti, H. & Megawarni, A. (2010). Tingkat efisiensi penaksir M terhadap penaksir LMS dalam menaksir koefisien garis regresi. *Jurnal Matematika, Sains, dan Teknologi, 11(2), 90-98*.