PENDUGAAN PARAMETER FUNGSI COBB-DOUGLAS GALAT ADITIF DENGAN ALGORITME GENETIKA*

Iqbal Hanif[1], Agus M Soleh[2‡], Aam Alamudi[3]

[1]Telkom, Indonesia, iqbal.hanif.ipb@gmail.com
[2]Dept. of Statistics, Bogor Agricultural University (IPB), Indonesia, agusms@apps.ipb.ac.id
[3]Dept. of Statistics, Bogor Agricultural University (IPB), Indonesia, alamudi@ipb.ac.id
[‡]corresponding author

Indonesian Journal of Statistics and Its Applications
Vol 1 No 1 (2017), 39 - 55

Abstract

Cobb-Douglas function with additive errors is a function which can be used to analyse the relationship between production output and production factors. The method commonly used to estimate the parameter of that function is Nonlinear Least Square (NLS) and a common algorithm for this method is Gauss Newton iteration (NLS-GN). However, NLS-GN method has less-optimum results when analysing multicolinearity data. A possibly better method for this analysis is Genetic Algorithm (NLS-GA). The purpose of this study is to analyse the use of Genetic Algorithm to estimate parameters of Cobb-Douglas function with additive errors. The results show that NLS-GA method could not produce a better parameter estimator than NLS-GN method does but it produced a better parameter estimator in analysing multicolinearity data. NLS-GA method is capable of producing a better model with predictive ability than NLS-GN method does with real data.

Keywords: cobb-douglas function, genetic algorithm, nonlinear least square.

*Received Apr 2017; Accepted Mei 2017; Published online on Oct 2017

1 Pendahuluan

Fungsi Cobb-Douglas dengan galat aditif merupakan salah satu fungsi produksi yang dapat digunakan untuk menganalisis hubungan antara hasil produksi dan faktor-faktor produksi. Salah satu metode yang dapat digunakan untuk menduga parameter fungsi tersebut adalah metode Nonlinear Least Square, yaitu menghitung nilai parameter yang mampu meminimumkan Jumlah Kuadrat Galat (Huet et al., 2003).

Algoritme pengoptimuman dibutuhkan dalam pendugaan parameter dengan metode Nonlinear Least Square. Algoritme yang umum digunakan adalah iterasi Gauss Newton, yaitu algoritme yang menerapkan deret Taylor orde pertama untuk memperoleh titik optimum (Chapra dan Canale, 2002). Metode Nonlinear Least Square dengan iterasi Gauss Newton mampu menghasilkan dugaan parameter yang baik dengan waktu yang relatif cepat. Namun, ketika dihadapkan dengan data yang memiliki masalah multikolinearitas, metode tersebut tidak dapat memberikan hasil pendugaan parameter yang optimal (Erkoç et al., 2010).

Algoritme pengoptimuman alternatif yang dapat digunakan adalah Algoritme Genetika. Algoritme Genetika meniru proses evolusi makhluk hidup untuk menangani masalah pengoptimuman (Sartono, 2010). Algoritme genetika diduga mampu menghindari overfitting, menangani masalah multikolinearitas dengan baik, dan memberikan hasil yang mudah diinterpretasikan (Zhang dan Horvath, 2006). Artikel ini bertujuan untuk mengkaji penerapan Algoritme Genetika sebagai algoritme alternatif pendugaan parameter fungsi Cobb-Douglas dengan galat aditif.

2 Metodologi

2.1 Sumber Data

Penelitian ini menggunakan dua jenis gugus data, yaitu data simulasi dan data produksi pertanian. Data simulasi dibangkitkan berdasarkan skema pembangkitan data yang sama seperti yang dilakukan pada penelitian Goldfeld dan Quandt Goldfeld dan Quandt (1970). Persamaan Cobb-Douglas hasil penelitian Chisasa dan Makina Chisasa dan Makina (2013) digunakan untuk membangkitkan data simulasi, dengan rincian persamaan sebagai berikut:

$$Y = e^{0.1835} X_1^{0.5932} X_2^{-0.0065} X_3^{0.4153} X_4^{-0.0431} \tag{1}$$

Terdapat dua jenis data yang dibangkitkan, yaitu data tanpa masalah multikolinearitas dan data dengan masalah multikolinearitas. Pembangkitan data tanpa masalah multikolinearitas dilakukan dengan cara membangkitkan bilangan acak menyebar seragam(0.1, 100) sebanyak 1.000 bilangan untuk setiap peubah bebas (x_i). Pembangkitan data dengan masalah multikolinearitas dilakukan dengan cara membangkitkan bilangan acak menyebar multivariat normal (1.000 bilangan untuk setiap peubah bebas x_i) dengan nilai rataan = 50 dan matriks kovarian dengan menggunakan matriks korelasi seperti pada Tabel 1. Langkah selanjutnya adalah membangkitkan sisaan (u) menyebar normal(0, 1) sebanyak 1.000 bilangan untuk masing-masing jenis data serta membangkitkan peubah y dengan rumus:

$$y = e^{0.1835} x_1^{0.5932} x_2^{-0.0065} x_3^{0.4153} {x_4}^0.0431 + u \tag{2}$$

Data produksi pertanian yang digunakan di dalam penelitian ini adalah data hasil produksi tanaman yang tergolong tanaman biji-bijian (serealia) beserta faktor-faktor

Tabel 1: Matriks korelasi untuk pembangkitan bilangan acak multivariat normal

	X_1	X_2	X_3	X_4
X_1	1	0.8	0.8	0.8
X_2	0.8	1	0.8	0.8
X_3	0.8	0.8	1	0.8
X_4	0.8	0.8	0.8	1

produksinya, yaitu luas lahan, jumlah alat pertanian yang digunakan, dan konsumsi pupuk. Data produksi pertanian yang digunakan merupakan data *cross-section* yang diperoleh berdasarkan hasil produksi tahun 2004 dan diambil dari 45 negara yang berbeda. Data diunduh dari *database* World Bank yang dapat diakses melalui *website* World Bank (http://data.worldbank.org/indicator).

2.2 Fungsi Cobb-Douglas dengan Galat Aditif

Fungsi Cobb-Douglas dengan galat aditif dikembangkan oleh Goldfeld dan Quandt Goldfeld dan Quandt (1970). Mereka menambahkan komponen stokastik aditif (*additive stochastic term)* ke dalam fungsi produksi yang ditemukan oleh Cobb dan Douglas Cobb dan Douglas (1928), yang kemudian dikenal dengan istilah galat (*error*). Bentuk fungsi produksi Cobb-Douglas dengan galat aditif adalah sebagai berikut:

$$y = e^{\alpha} x_1^{\beta_1} x_2^{\beta 2} ... X_n^{\beta_n} \tag{3}$$

dengan y adalah hasil produksi, x_i adalah faktor produksi ke-i untuk i=1,2,...,n, dan u adalah galat.

Nilai parameter β_i di dalam fungsi Cobb-Douglas diinterpretasikan sebagai elastisitas faktor produksi x_i terhadap hasil produksi y. Adanya hukum kenaikan hasil yang semakin berkurang (*law of diminishing returns*) dalam produksi menyebabkan nilai β_i berada pada selang $0 < \beta_i < 1$Soekartawi (2002). Namun, tidak tertutup kemungkinan bagi parameter β_i untuk memiliki nilai negatif. Merujuk pada hasil penelitian Chisasa dan Makina Chisasa dan Makina (2013), parameter β_i memiliki nilai positif apabila penambahan faktor produksi mampu meningkatkan hasil produksi (produktif). Namun sebaliknya, parameter β_i memiliki nilai negatif apabila penambahan faktor produksi justru mengurangi hasil produksi (tidak produktif).

2.3 *Nonlinear Least Square*(NLS)

Erkoç et al. (2010) mendefinisikan bentuk umum dari model nonlinear adalah sebagai berikut:

$$y_i = f(x_i, \beta) + \varepsilon_i \tag{4}$$

dengan y_i adalah nilai peubah respon pada amatan ke-i, x_i adalah nilai peubah bebas pada amatan ke-i, β adalah parameter, dan ε_i adalah galat ke-i yang menyebar normal(0, σ^2I). Metode *Nonlinear Least Square* (NLS) merupakan salah satu metode yang dapat digunakan untuk menduga parameter fungsi nonlinear. Dasar dari metode ini adalah mencari nilai β yang meminimumkan nilai Jumlah Kuadrat Galat (JKG)Huet et al. (2003), yang didefinisikan dengan fungsi objektif berikut:

$$S(\beta) = \sum_{i=1}^{n} [y_i - f(x_i, \beta)]^2 \tag{5}$$

2.4 Iterasi Gauss Newton (GN)

Untuk memperoleh nilai parameter yang mampu meminimumkan JKG, fungsi objektif metode NLS perlu diturunkan terlebih dahulu. Turunan orde pertama atau persamaan normal dari fungsi objektif metode NLS adalah sebagai berikut:

$$\frac{\delta S}{\delta \beta} = \sum_{i=1}^{n} [y_i - f(x_i, \beta)] \left[\frac{\delta f(x_i, \beta)}{\delta \beta_j} \right]_{\beta = \hat{\beta}} = 0 \tag{6}$$

Berbeda dengan persamaan normal pada model linear, persamaan normal diatas cukup sulit untuk diselesaikan, mengingat $f(x_i, \beta)$ adalah fungsi nonlinear Erkoç et al. (2010). Pada tahap ini, pendekatan iterasi Gauss Newton (GN) digunakan untuk melakukan pendugaan parameter dengan metode NLS (NLS-GN). GN menerapkan deret Taylor orde pertama untuk memperoleh nilai dugaan parameter yang meminimumkan JKG Chapra dan Canale (2002), yang dituliskan ke dalam persamaan berikut:

$$f(x_i, \beta) = f(x_i, \beta_0) \tag{7}$$

substitusikan persamaan deret diatas ke dalam bentuk umum model nonlinear:

$$y_i - f(x_i, \beta_0) = \sum_{j=1}^{p} \left[\frac{\delta f(x_i, \beta)}{\delta \beta_j} \right]_{\beta = \beta_0} (\beta_j - \beta_{j0}) + \varepsilon_i \tag{8}$$

jika:

$$D_i = y_i - f(x_i, \beta_0) \tag{9}$$

$$J_{ij}^0 = \left[\frac{\delta f(x_i, \beta)}{\delta \beta_j} \right]_{\beta = \beta_0} \tag{10}$$

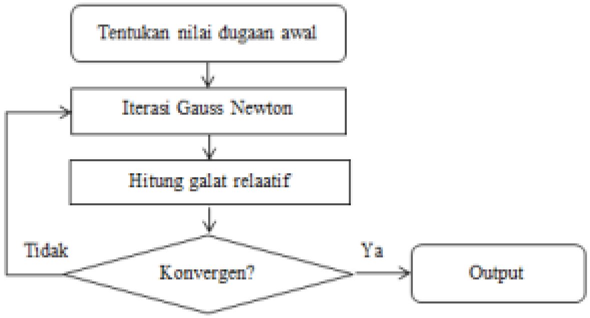
maka persamaan nonlinear dapat dituliskan kembali ke dalam notasi matriks:

$$\underline{D} = \boldsymbol{J^0}(\underline{\beta} - \underline{\beta_0}) + \underline{\varepsilon} \tag{11}$$

Persamaan diatas memiliki bentuk persamaan yang linear, sehingga nilai β dapat diduga melalui pendekatan Metode Kuadrat Terkecil (MKT) secara iteratif dengan persamaan sebagai berikut:

$$\underline{\hat{\beta}_1} = \underline{\beta_0} + (\boldsymbol{J^0}'\boldsymbol{J^0})^{-1}\boldsymbol{J^0}'\underline{D} \tag{12}$$

dengan $\boldsymbol{J^0}$ adalah matriks turunan parsial dari persamaan nonlinear (*jacobian*), $\underline{D}$ adalah vektor nilai selisih antara peubah respon dengan dugaannya, $\underline{\beta_0}$ adalah vektor nilai awal parameter, dan $\underline{\beta_1}$ adalah vektor nilai dugaan parameter hasil iterasi. Bila $\hat{\beta}_1 - \beta_0 = 0$, maka $\boldsymbol{J^0}'\underline{D} = \frac{\delta S}{\delta \beta} = 0$, yang merupakan pemenuhan persyaratan persamaan normal metode NLS-GN. Langkah-langkah pendugaan parameter dengan metode NLS-GN diilustrasikan melalui diagram alir pada Gambar 1.



Gambar 1: Diagram alir mekanisme metode NLS-GN

2.5 Masalah Multikolinearitas

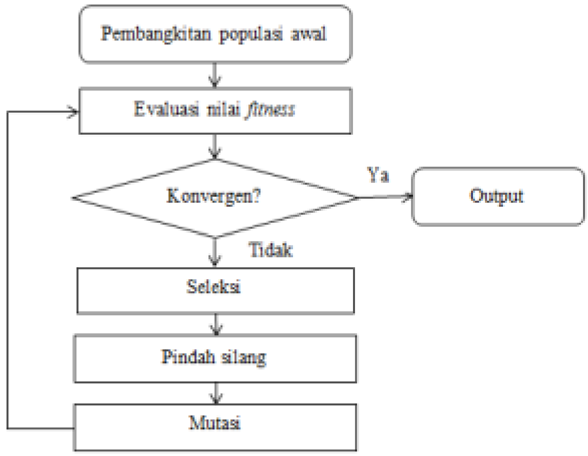
Multikolinearitas adalah sebuah kondisi yang terjadi akibat adanya hubungan linear antar peubah bebas, baik itu interkorelasi sempurna (terpaut linear) atau interkorelasi tidak sempurna tetapi cukup kuat Gujarati dan Porter (2010). Metode NLS-GN tidak dapat memberikan hasil pendugaan parameter yang optimal ketika dihadapkan dengan data yang memiliki masalah multikolinearitas. Misalkan selisih parameter antar iterasi dalam metode NLS-GN ($\beta_1 - \beta_0$) disimbolkan dengan α, maka kita dapat menghitung nilai Kuadrat Tengah Galat (KTG) dari $\hat{\alpha}$ dengan persamaan sebagai berikut:

$$
\begin{aligned}
KTG(\hat{\alpha}) &= E[\hat{\alpha}-\alpha]'[\hat{\alpha}-\alpha] \\
&= \sum_{j=1}^{p} E(\hat{\alpha}_j-\alpha_j)^2 \\
&= \sum_{j=1}^{p} var(\hat{\alpha}_j) \\
&= \sigma^2 Tr(\boldsymbol{J^{0\prime}J^{0}})^{-1} \\
&= \sigma^2 \sum_{j=1}^{p} \frac{1}{\lambda_j} \quad, \lambda_j > 0, \text{ j=1,2...,p}
\end{aligned}
$$

Korelasi yang cukup kuat antar peubah menyebabkan matriks $\boldsymbol{J^{0\prime}J^{0}}$ berada dalam kondisi buruk (*ill-conditioned*), yang menyebabkan minimal satu akar ciri dari matriks $\boldsymbol{J^{0\prime}J^{0}}$ (λ_j) akan bernilai sangat kecil. Akibatnya, var($\hat{\alpha}_j$) dan KTG($\hat{\alpha}$) akan bernilai cukup besar Erkoç et al. (2010).

2.6 Algortime Genetika (*Genetic Algorithm*-GA)

Algoritme Genetika (*Genetic Algorithm*-GA) adalah algoritme optimasi yang bekerja dengan meniru perilaku dalam proses evolusi yang dialami makhluk hidup. GA telah tercatat berhasil memberikan solusi yang sangat baik di berbagai kasus Sartono (2010). Ada berbagai macam skema GA yang dapat digunakan dalam menyelesaikan masalah optimasi, dan pemilihan skema GA dapat bersifat subjektif dan bergantung pada masalah yang dihadapi. Scrucca Scrucca (2013) telah mengembangkan skema GA untuk menangani masalah pengepasan kurva (curve fitting) yang cocok untuk diterapkan pada metode NLS karena memiliki prinsip yang sama, yaitu meminimumkan selisih antara nilai aktual (titik amatan) dengan nilai dugaannya (titik pada kurva). Langkah-langkah pendugaan parameter metode NLS dengan skema GA rancangan Scrucca Scrucca (2013) (NLS-GA) diilustrasikan melalui diagram alir pada Gambar 2.



Gambar 2: Diagram alir mekanisme metode NLS-GA

1. Pembangkitan populasi awal
   Membangkitkan populasi awal sebanyak 1000 individu berupa vektor kromosom $[\hat{\alpha}|\hat{\beta}_1|...|\hat{\beta}_k]$. Nilai dugaan parameter dibangkitkan mengikuti sebaran seragamsesuai dengan batasan yang ditentukan.

2. Evaluasi *fitness* individu
   Kriteria *fitness* yang digunakan adalah nilai JKG. Semakin kecil nilai JKG suatu individu, maka semakin baik individu tersebut. Sebanyak $k\%$ individu terbaik dalam populasi dipertahankan untuk generasi selanjutnya, dan proses ini dikenal dengan istilah elitisme Hopgood (2001).

3. Seleksi individu
   Standardisasi nilai fitness dilakukan terlebih dahulu dengan metode *Fitness Linear Scaling*, yaitu metode standardisasi menggunakan fungsi linear Hopgood (2001) dengan persamaan sebagai berikut:

   $$s_i = af_i + b \tag{13}$$

   jika $f_{min} > 2\bar{f} - f_{max}$ maka:

   $$a = \frac{\bar{f}}{f_{max} - \bar{f}}, b = \frac{\bar{f}(f_{max} - 2\bar{f})}{f_{max} - \bar{f}} \tag{14}$$

   jika $f_{min} \leq 2\bar{f} - f_{max}$ maka:

   $$a = \frac{\bar{f}}{\bar{f} - f_{min}}, b = -f_{min}\frac{\bar{f}}{\bar{f} - f_{min}} \tag{15}$$

   Individu kemudian diseleksi dengan metode *Roulette Wheel*, dengan peluang suatu individu untuk terpilih (P_i) dinyatakan dengan persamaan sebagai berikut:

   $$P_i = \frac{|s_i|}{\sum |s_i|} \tag{16}$$

4. Pindah silang (*cross-over*)
   Metode pindah silang yang digunakan adalah *Local Arithmetic Crossover*, yaitu metode pindah silang dengan cara mengkombinasikan individu induk secara linear Michalewicz (1992). Persamaan yang digunakan adalah sebagai berikut:
   $$s_v^{t+1} = as_w^t + (1-a)s_v^t \tag{17}$$
   $$s_w^{t+1} = as_v^t + (1-a)s_w^t \tag{18}$$
   dengan s_w^t dan s_v^t adalah individu induk, s_v^{t+1} dan s_w^{t+1} adalah individu anak hasil pindah silang dan a adalah bilangan acak seragam(0, 1). Peluang pindah silang yang ditetapkan adalah sebesar p.

5. Mutasi
   Metode mutasi yang digunakan adalah *Uniform Random Mutation*, yaitu mengganti nilai gen yang terpilih dengan bilangan acak seragam. Pemilihan gen yang akan dimutasi dilakukan secara acak, dimana setiap gen memiliki peluang yang sama untuk terpilih Michalewicz (1992). Peluang mutasi yang ditetapkan adalah sebesar q.

6. Langkah 2-5 terus diulangi hingga konvergen, dengan kriteria kekonvergenan adalah m kali iterasi, kemudian dipilih individu terbaik dari populasi terakhir.

2.7 *Minimum Variance Unbiased Estimator* (MVUE)

Suatu dugaan parameter dapat dikatakan sebagai dugaan parameter terbaik apabila memenuhi kriteria *Minimum Variance Unbiased Estimator* (MVUE). Misalkan $Y = u(X_1, X_2, \ldots, X_n)$ adalah dugaan parameter bagi θ, Hogg RV (2005) berpendapat bahwa Y dapat dikatakan sebagai MVUE bagi θ apabila memenuhi kriteria sebagai berikut:

1. $E(Y) = \theta$ (tidak berbias).
2. $Var(Y)$ lebih kecil atau sama dengan ragam dari semua penduga bagi parameter θ (ragam minimum).

2.8 *K-Fold Cross Validation*

Validasi silang adalah metode pemilihan model berdasarkan kemampuan prediksi yang dimiliki oleh model tersebut Good (2006). Kriteria pemilihan model terbaik dengan metode validasi silang adalah memilih model yang mampu meminimumkan sisaan validasi silang, yaitu nilai Kuadrat Tengah Galat (KTG). *K-Fold Cross Validation* adalah salah satu bentuk khusus dari metode validasi silang. *K-Fold Cross Validation* membagi data secara acak menjadi k bagian. Salah satu bagian akan menjadi gugus data tes dan sisanya akan menjadi gugus data latih. Gugus data latih digunakan untuk menentukan model dan gugus data tes digunakan untuk menghitung KTG dari model yang terbentuk dari gugus data latih. Proses ini diulang sebanyak k kali hingga semua bagian digunakan satu kali sebagai gugus data tes. Nilai sisaan validasi silang diduga dengan persamaan berikut:

$$KTG_{PRED} = \frac{1}{k}\sum_{k=1}^{k}\sum_{(x_i,y_i)\epsilon T}(y_i - \hat{y}_{-k}(x_i))^2 \tag{19}$$

dengan y_i adalah peubah respon ke-i pada gugus data tes T dan $\hat{y}_(-k)(x_i)$ adalah dugaan y untuk x_i pada saat bagian ke-k tidak digunakan dalam pembentukan model. Nilai k yang biasa digunakan adalah lima atau sepuluh Izenman (2008).

2.9 Metode

Proses analisis yang digunakan di dalam penelitian ini secara umum terdiri atas tiga poin utama:

1. Penerapan metode NLS-GA untuk pendugaan parameter fungsi Cobb-Douglas dengan galat aditif pada data simulasi tanpa masalah multikolinearitas, serta membandingkan hasilnya dengan dugaan parameter hasil metode NLS-GN.

2. Penerapan metode NLS-GA untuk pendugaan parameter fungsi Cobb-Douglas dengan galat aditif pada data simulasi dengan masalah multikolinearitas, serta membandingkan hasilnya dengan dugaan parameter hasil metode NLS-GN.

3. Penerapan metode NLS-GA pada kasus riil, serta mebandingkan kemampuan prediksi antara model hasil pendugaan dengan metode NLS-GA dengan model hasil pendugaan dengan metode NLS-GN.

Penelitian dilakukan dengan menggunakan komputer dengan spesifikasi CPU Intel Core-i3 1.8 GHz dan RAM sebesar 2 GB. Perangkat lunak (*software*) yang digunakan adalah R 3.1.3 dan RStudio. Fungsi-fungsi R yang digunakan di dalam penelitian antara lain fungsi "nls" untuk menerapkan metode NLS-GN, fungsi "GA" pada paket "GA" yang dikembangkan oleh Scrucca (2013) untuk menerapkan metode NLS-GA, dan fungsi "mvrnorm" pada paket "lestat" untuk pembangkitan bilangan acak menyebar multivariat normal. Beberapa ketetapan untuk masing-masing metode yang digunakan di dalam penelitian adalah sebagai berikut:

1. Untuk metode NLS-GN, nilai awal untuk semua parameter $(\beta_0) = 1$ dan toleransi galat (ε) sebesar 10^-5.

2. Untuk metode NLS-GA, nilai dugaan parameter (gen) pada setiap individu dibangkitkan mengikuti sebaran seragam(-10, 10) untuk parameter intersep dan sebaran seragam(-1, 1) untuk parameter selain intersep, total populasi (N) sebanyak 1000 individu, elitisme (k) sebesar 5%, peluang pindah silang (p) sebesar 0.8, peluang mutasi (q) sebesar 0.05, dan jumlah iterasi (m) yaitu 1000 iterasi.

Prosedur Analisis pada Data Simulasi

Prosedur analisis untuk data simulasi di dalam penelitian ini digambarkan pada diagram alir pada Gambar 3 dengan rincian sebagai berikut:

1. Melakukan resampling tanpa pengembalian sebanyak n data dari data populasi.

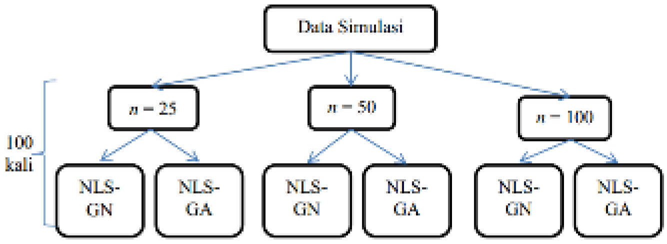
2. Melakukan pendugaan parameter untuk setiap metode.

3. Menyimpan hasil pendugaan parameter ke dalam matriks.

4. Mengulangi langkah 1-3 sebanyak 100 kali, kemudian menghitung rataan dan simpangan baku dari parameter hasil pendugaan.

5. Melakukan langkah 1-4 dengan ukuran contoh yang berbeda-beda ($n = 25$, $n = 50$, dan $n4 = 100$).

Prosedur diatas dilakukan sebanyak satu kali untuk setiap jenis data (data simulasi tanpa masalah multikolinearitas dan data simulasi dengan masalah multikolinearitas). Metode terbaik untuk data simulasi adalah metode yang menghasilkan dugaan yang memenuhi kriteria MVUE, yaitu nilai rataan dugaan parameter yang mendekati nilai parameter sebenarnya dan nilai ragam (dalam penelitian ini diwakilkan oleh simpangan baku) dugaan parameter yang paling kecil.



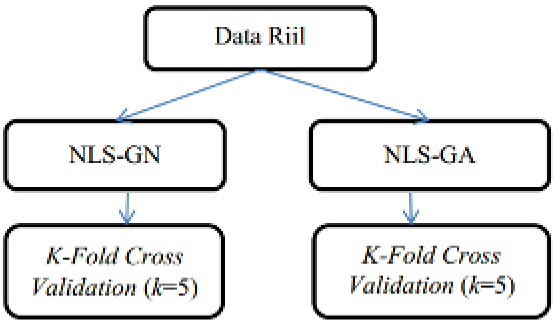
Gambar 3: Prosedur analisis pada data simulasi

Prosedur Analisis pada Data Riil

Prosedur analisis untuk data riil di dalam penelitian ini digambarkan pada diagram alir pada Gambar 4 dengan rincian sebagai berikut:

1. Melakukan pendugaan parameter untuk setiap metode.
2. Melakukan *K-Fold Cross Validation* dengan k=5 (empat bagian terdiri dari sepuluh baris data dan satu bagian sisanya terdiri dari lima baris data)

Metode terbaik untuk data riil adalah metode yang menghasil persamaan dengan kemampuan prediksi yang lebih baik, yaitu persamaan yang menghasilkan nilai sisaan validasi silang yang paling kecil.



Gambar 4: Prosedur analisis pada data riil

3 Hasil dan Pembahasan

3.1 Aplikasi Algoritme Genetika pada Data Simulasi tanpa Masalah Multikolinearitas

Hasil 100 kali pendugaan parameter untuk setiap ukuran contoh ($n = 25$, $n = 50$ dan $n = 100$) menunjukkan bahwa metode NLS-GA mampu melakukan pendugaan parameter dengan cukup baik. Hal ini terlihat dari rataan dugaan parameter hasil metode NLS-GA mendekati nilai parameter ($\alpha = 0.1835$, $\beta_1 = 0.5932$, $\beta_2 = -0.0065$, $\beta_3 = 0.4153$, $\beta_4 = -0.0431$) dan simpangan baku dugaan parameter yang cukup kecil. Namun jika dibandingkan dengan metode NLS-GN, metode NLS-GA belum mampu menghasilkan dugaan parameter yang lebih baik dibanding metode NLS-GN. Hal ini terlihat dari rataan dugaan parameter hasil metode NLS-GN lebih mendekati nilai parameter bila dibandingkan dengan rataan dugaan parameter hasil metode NLS-GA. Metode NLS-GN juga menghasilkan dugaan parameter dengan simpangan baku yang lebih kecil bila dibandingkan dengan simpangan baku dugaan parameter hasil metode NLS-GA. Rincian rataan dan simpangan baku dugaan parameter hasil metode NLS-GN dan NLS-GA dapat dilihat pada Tabel 2 dan Tabel 3.

Tabel 2: Rataan hasil dugaan parameter untuk masing-masing metode di setiap ukuran contoh pada data simulasi tanpa masalah multikolinearitas

Ukuran contoh	Metode	a	b_1	b_2	b_3	b_4
$n = 25$	NLS-GN	0.1925	0.5923	-0.0072	0.4145	-0.0429
	NLS-GA	0.2300	0.5858	-0.0068	0.4105	-0.0418
$n = 50$	NLS-GN	0.1929	0.5926	-0.0068	0.4140	-0.0432
	NLS-GA	0.2295	0.5870	-0.0068	0.4104	-0.0430
$n = 100$	NLS-GN	0.1910	0.5923	-0.0064	0.4140	-0.0428
	NLS-GA	0.2254	0.5870	-0.0063	0.4105	-0.0424

Tabel 3: Simpangan baku hasil dugaan parameter untuk masing-masing metode di setiap ukuran contoh pada data simulasi tanpa masalah multikolinearitas

Ukuran contoh	Metode	a	b_1	b_2	b_3	b_4
$n = 25$	NLS-GN	0.8886	0.4607	0.4105	0.4298	0.3888
	NLS-GA	0.7247	0.2651	0.2280	0.2653	0.2002
$n = 50$	NLS-GN	0.6125	0.3067	0.3158	0.2746	0.3212
	NLS-GA	0.5145	0.1944	0.1745	0.1742	0.1729
$n = 100$	NLS-GN	0.4115	0.1880	0.1835	0.2067	0.1866
	NLS-GA	0.3900	0.1553	0.1190	0.1684	0.1161

Hasil pendugaan parameter juga disajikan dalam bentuk boxplot. Boxplot hasil pendugaan parameter dari masing-masing metode untuk semua ukuran contoh dapat dilihat pada Lampiran 1. Pada Lampiran 1, terlihat bahwa garis tengah boxplot dugaan parameter yang dihasilkan oleh metode NLS-GN lebih mendekati nilai parameter (garis merah) dibanding boxplot dugaan parameter hasil metode NLS-GA, yang menunjukkan bahwa dugaan parameter hasil metode NLS-GN memiliki bias yang lebih kecil dibanding dugaan parameter hasil metode NLS-GA. Boxplot dugaan parameter yang dihasilkan oleh metode NLS-GN juga memiliki lebar kotak yang lebih kecil dibanding boxplot dugaan parameter hasil metode NLS-GA, yang menunjukkan bahwa dugaan parameter hasil metode NLS-GN memiliki keragaman yang lebih kecil dibanding dugaan parameter hasil metode NLS-GA.

3.2 Aplikasi Algoritme Genetika pada Data Simulasi dengan Masalah Multikulinearitas

Hasil 100 kali pendugaan parameter untuk setiap ukuran contoh ($n = 25$, $n = 50$ dan $n = 100$) menunjukkan bahwa metode NLS-GA menghasilkan dugaan parameter yang kurang baik. Hal ini terlihat dari rataan dugaan parameter hasil metode NLS-GA cukup jauh dari nilai parameter ($\alpha = 0.1835$, $\beta_1 = 0.5932$, $\beta_2 = -0.0065$, $\beta_3 = 0.4153$, $\beta_4 = -0.0431$). Hal yang sama juga terjadi pada metode NLS-GN, tetapi jika dibandingkan dengan metode NLS-GA, hasil metode NLS-GN memiliki bias yang lebih kecil. Hal ini terlihat dari rataan dugaan parameter hasil metode NLS-GN lebih mendekati nilai parameter bila dibandingkan dengan rataan dugaan parameter hasil metode NLS-GA.

Namun, berbeda dengan metode NLS-GN yang menghasilkan dugaan parameter dengan simpangan baku yang besar karena masalah multikolinearitas, metode NLS-GA mampu mengatasi masalah multikolinearitas dalam melakukan pendugaan parameter. Metode NLS-GA menghasilkan dugaan parameter dengan simpangan baku yang jauh lebih kecil bila dibandingkan dengan metode NLS-GN. Rincian rataan dan simpangan baku dugaan parameter hasil metode NLS-GN dan NLS-GA dapat dilihat pada Tabel 4 dan Tabel 5.

Tabel 4: Rataan hasil dugaan parameter untuk masing-masing metode di setiap ukuran contoh pada data simulasi dengan masalah multikolinearitas

Ukuran contoh	Metode	a	b_1	b_2	b_3	b_4
$n = 25$	NLS-GN	0.3982	0.5747	0.0360	0.4172	-0.1242
	NLS-GA	0.4183	0.4022	0.1369	0.3245	0.0349
$n = 50$	NLS-GN	0.3130	0.5880	0.0581	0.3455	-0.0663
	NLS-GA	0.3287	0.4443	0.1259	0.2746	0.0766
$n = 100$	NLS-GN	0.3214	0.5092	0.0923	0.3948	-0.0731
	NLS-GA	0.3288	0.4241	0.1228	0.3363	0.0381

Hasil pendugaan parameter juga disajikan dalam bentuk boxplot. Boxplot hasil pendugaan parameter dari masing-masing metode untuk semua ukuran contoh dapat dilihat pada Lampiran 2. Pada Lampiran 2, terlihat bahwa garis tengah boxplot dugaan parameter yang dihasilkan oleh metode NLS-GN lebih mendekati nilai parameter (garis merah) dibanding boxplot dugaan parameter hasil metode NLS-GA, yang menunjukkan bahwa dugaan parameter hasil metode NLS-GN memiliki bias yang lebih kecil dibanding dugaan parameter hasil metode NLS-GA. Namun dugaan parameter metode NLS-GA mampu menghasilkan boxplot dengan lebar kotak yang lebih kecil dibanding boxplot dugaan parameter hasil metode NLS-GN, yang menunjukkan bahwa dugaan parameter hasil metode NLS-GA memiliki keragaman yang lebih kecil dibanding dugaan parameter hasil metode NLS-GN.

Tabel 5: Simpangan baku hasil dugaan parameter untuk masing-masing metode di setiap ukuran contoh pada data simulasi dengan masalah multikolinearitas

Ukuran contoh	Metode	a	b_1	b_2	b_3	b_4
$n = 25$	NLS-GN	0.8886	0.4607	0.4105	0.4298	0.3888
	NLS-GA	0.7247	0.2651	0.2280	0.2653	0.2002
$n = 50$	NLS-GN	0.6125	0.3067	0.3158	0.2746	0.3212
	NLS-GA	0.5145	0.1944	0.1745	0.1742	0.1729
$n = 100$	NLS-GN	0.4115	0.1880	0.1835	0.2067	0.1866
	NLS-GA	0.4115	0.1880	0.1835	0.2067	0.1866

3.3 Data Riil

Baik metode NLS-GN maupun NLS-GA menghasilkan dugaan parameter yang tidak jauh berbeda pada data riil. Namun, hasil *K-Fold Cross Validation* menunjukkan bahwa nilai sisaan validasi silang dari model yang dihasilkan metode NLS-GA lebih kecil bila dibandingkan dengan model yang dihasilkan metode NLS-GN, sehingga disimpulkan bahwa metode NLS-GA mampu menghasilkan model dengan kemampuan prediksi lebih baik dibanding metode NLS-GN. Rincian hasil pendugaan parameter dan *K-Fold Cross Validation* untuk masing-masing metode dapat dilihat pada Tabel 6 dan Tabel 7.

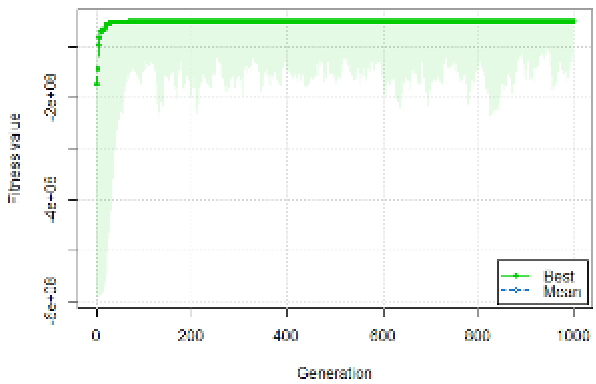
Tabel 6: Hasil dugaan parameter untuk masing-masing metode

Metode	a	b_1	b_2	b_3
NLS-GN	6.69072	0.04271	0.29299	-0.02379
NLS-GA	6.70112	0.04276	0.29031	-0.02361

Tabel 7: Nilai sisaan validasi silang untuk masing-masing metode di setiap ukuran contoh

Metode	*Fold-1*	*Fold-2*	*Fold-3*	*Fold-4*	*Fold-5*
NLS-GN	1643482	786202	1176481	1784237	658039
NLS-GA	1470589	698843	1063334	1285209	617492

Metode NLS-GA membutuhkan waktu yang lebih lama dibanding metode NLS-GN. Metode NLS-GN hanya membutuhkan waktu 0.02667 detik untuk melakukan proses pendugaan parameter, sedangkan metode NLS-GA membutuhkan waktu 1.98653 menit untuk melakukan proses pendugaan parameter. Grafik iterasi proses pendugaan dengan metode NLS-GA dapat dilihat pada Gambar 5.



Gambar 5: Grafik iterasi metode NLS-GA

4 Simpulan dan Saran

Metode NLS-GA mampu menghasilkan dugaan parameter yang cukup baik pada data simulasi tanpa masalah multikolinearitas, namun tidak lebih baik bila dibandingkan dengan metode NLS-GN pada data yang sama. Metode NLS-GA menghasilkan dugaan parameter yang kurang baik pada data simulasi dengan masalah multikolinearitas, namun mampu mengatasi problem yang dihadapi metode NLS-GN ketika dihadapkan pada data dengan masalah multikolinearitas, dengan menghasilkan dugaan parameter dengan simpangan baku yang jauh lebih kecil dibanding metode NLS-GN. Metode NLS-GN maupun NLS-GA menghasilkan dugaan parameter yang tidak jauh berbeda pada data riil, namun hasil *K-Fold Cross Validation* menunjukkan bahwa bahwa metode NLS-GA mampu menghasilkan model dengan kemampuan prediksi lebih baik dibanding metode NLS-GN.

Saran untuk penelitian selanjutnya adalah memodifikasi operator Algoritme Genetika pada metode NLS-GA agar mampu menghasilkan dugaan parameter dengan bias yang lebih kecil, baik pada data tanpa masalah multikolinearitas maupun pada data dengan masalah multikolinearitas. Selain itu, disarankan untuk mengujicobakan metode NLS-GA di berbagai kondisi data, diantaranya data dengan masalah heteroskedastisitas atau data dengan masalah autokorelasi. Kendala utama dalam penelitian ini adalah terbatasnya sumber daya, yaitu spesifikasi komputer yang digunakan cukup rendah, sehingga untuk penelitian selanjutnya disarankan untuk menggunakan komputer dengan spesifikasi yang lebih baik.

Daftar Pustaka

Chapra, S. dan Canale, R. (2002). *Numerical Methods for Engineers: With Software and Programming Applications*, 4 edn, McGraw-Hill, New York (US).

Chisasa, J. dan Makina, D. (2013). Bank credit and agricultural output in south africa: A cobb-douglas empirical analysis, *International Business and Economics Research Journal* . 12(4): 387-398.

Cobb, C. dan Douglas, P. (1928). A theory of production, *The American Economic Review* . 18(1): 139-165.

Erkoç, A., Tez, M. dan Akay, K. (2010). On multicolinearity in nonlinear regression models, *Selçuk Journal of Applied Mathematics* . Special Issue: 65-72.

Goldfeld, S. dan Quandt, R. (1970). The estimation of cobb-douglas type functions with multiplicative and additive errors, *International Economic Review* . 11(2): 251-257.

Good, P. (2006). *Resampling Method: A Practical Guide to Data Analysis*, 3 edn, Birkhäuser, Boston (US).

Gujarati, D. dan Porter, D. (2010). *Basic Econometrics*, 5 edn, McGraw-Hill., New York (US).

Hogg RV, McKean JW, C. A. (2005). *Introduction to Mathematical Statistics*, 6 edn, Pearson Prentice Hall, New Jersey (US).

Hopgood, A. (2001). *Intelligent Systems for Engineers and Scientists*, CRC Press, Boca Raton (US).

Huet, S., Bouvier, A., Poursat, M. dan Jolivet, E. (2003). *Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS and R Examples*, 2 edn, Springer, New York (US).

Izenman, A. (2008). *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, Springer, New York (US).

Michalewicz, Z. (1992). *Genetic Algorithms + Data Structures = Evolution Programs*, 2 edn, Springer, New York (US).

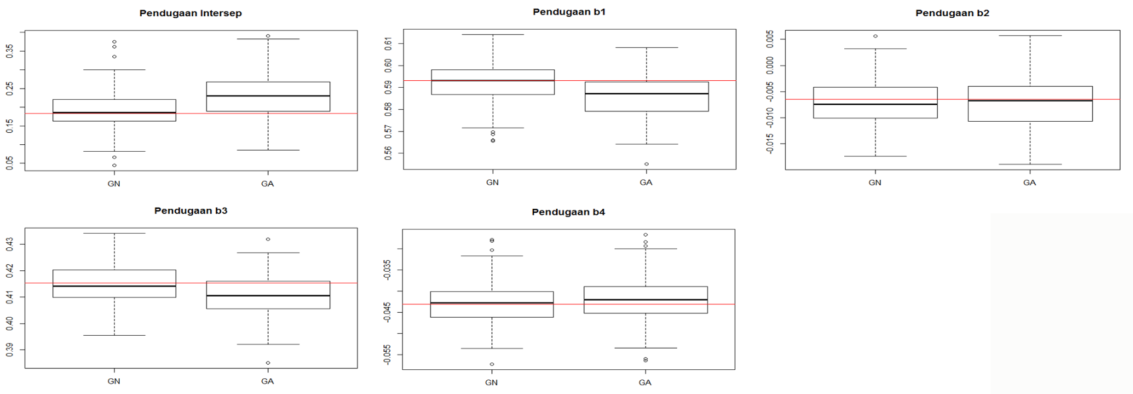
Sartono, B. (2010). Pengenalan algoritme genetik untuk pemilihan peubah penjelas dalam model regresi menggunakan sas/iml, *Forum Statistika dan Komputasi* . 15(2): 10-15.

Scrucca, L. (2013). Ga: A package for genetic algorithms in r, *Journal of Statistics Software* . 53(4): 1-37.

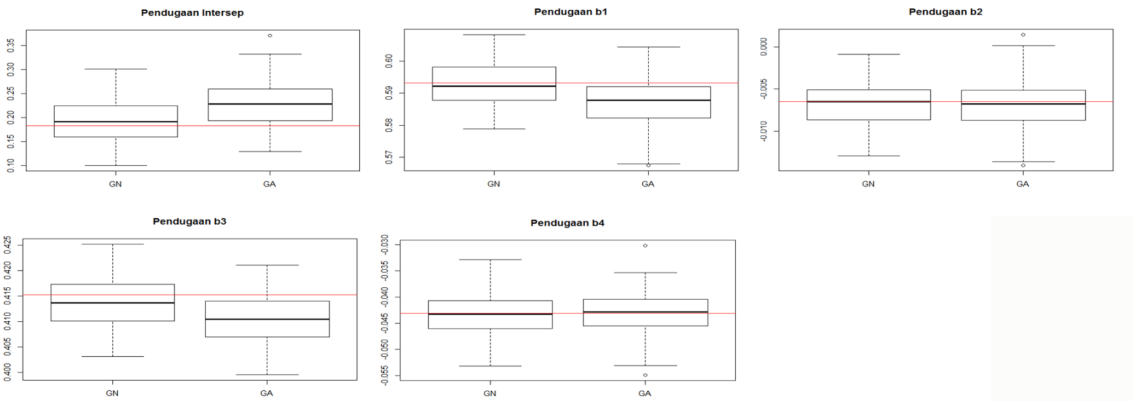
Soekartawi (2002). *Prinsip Dasar Ekonomi Pertanian: Teori dan Aplikasi*, Raja Grafindo Persada, Jakarta (ID).

Zhang, B. dan Horvath, S. (2006). Ridge regression based hybrid genetic algorithms for multi-locus quantitative trait mapping, *Int. J. Bioinformatics Research and Applications* . 1(3): 261-272.

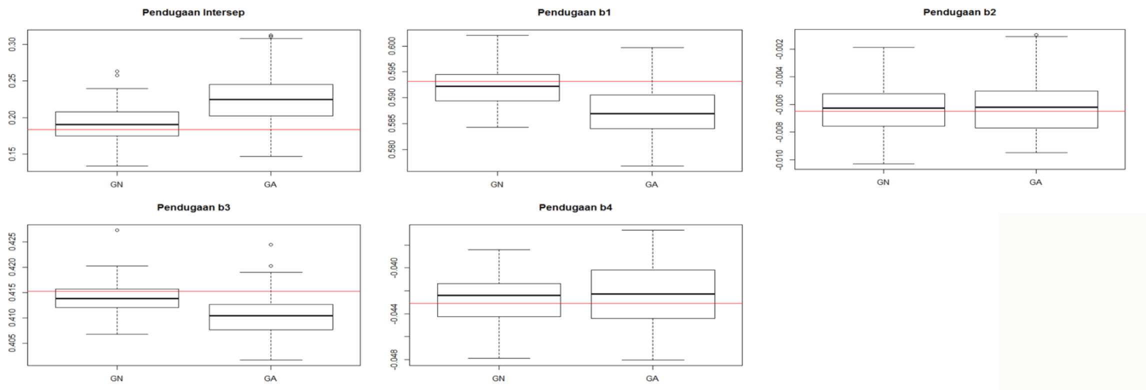
Lampiran 1. Boxplot dari parameter hasil pendugaan melalui metode NLS-GN (GN) dan NLS-AG (GA) pada data simulasi dengan kondisi baik.



Gambar 6: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=25

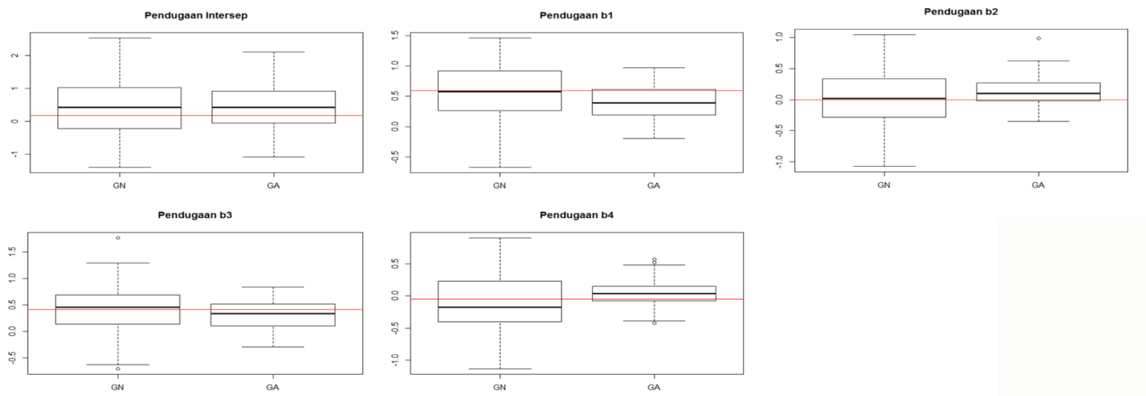


Gambar 7: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=50

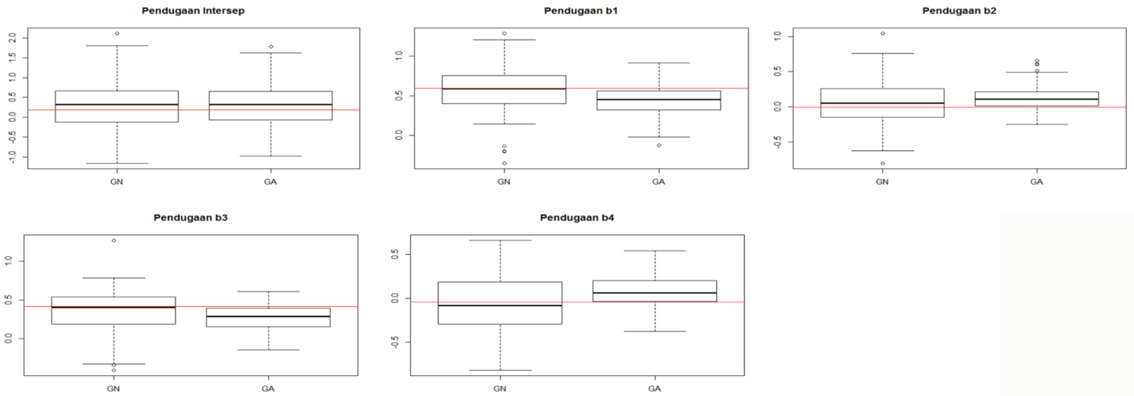


Gambar 8: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=100

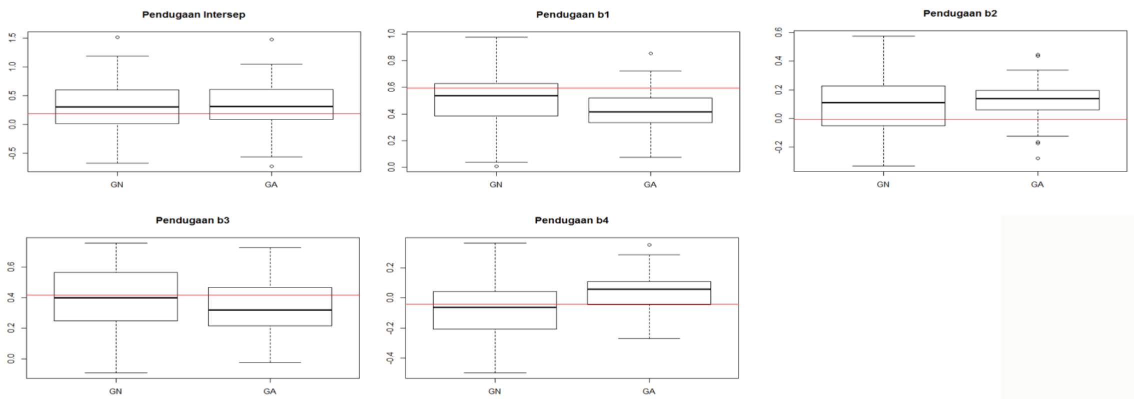
Lampiran 2. Boxplot dari parameter hasil pendugaan melalui metode NLS-GN (GN) dan NLS-AG (GA) pada data simulasi dengan kondisi buruk.



Gambar 9: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=25



Gambar 10: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=50



Gambar 11: Boxplot dari parameter hasil pendugaan dengan ukuran contoh n=100