

PENDUGAAN DATA HILANG DENGAN METODE YATES DAN ALGORITMA EM PADA RANCANGAN *LATTICE* SEIMBANG

Ni Made Sariayu^{§1}, Made Susilawati², Kartika Sari³

¹Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: sarone00n@gmail.com]

²Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: susilawati.made@gmail.com]

³Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: sari_kaartika@yahoo.co.id]

[§]*Corresponding Author*

ABSTRACT

Missing data often occur in agriculture and animal husbandry experiment. The missing data in experimental design makes the information that we get less complete. In this research, the missing data was estimated with Yates method and Expectation Maximization (EM) algorithm. The basic concept of the Yates method is to minimize sum square error (JKG), meanwhile the basic concept of the EM algorithm is to maximize the likelihood function. This research applied Balanced Lattice Design with 9 treatments, 4 replications and 3 group of each repetition. Missing data estimation results showed that the Yates method was better used for two of missing data in the position on a treatment, a column and random, meanwhile the EM algorithm was better used to estimate one of missing data and two of missing data in the position of a group and a replication. The comparison of the result JKG of ANOVA showed that JKG of incomplete data larger than JKG of incomplete data that has been added with estimator of data. This suggest that we need to estimate the missing data.

Keywords: Missing Data, Yates method, EM algorithm, Balanced Lattice Design.

1. PENDAHULUAN

Setiap rancangan percobaan sangat mungkin mengalami masalah data hilang. Data tersebut hilang disebabkan oleh beberapa hal yang tidak bisa diramalkan dan tidak bisa dihindari. Misalnya percobaan pada bidang pertanian, ketika hasil panen dari beberapa petak percobaan dirusak oleh hewan, hama dan lain-lain, dan juga pada bidang peternakan, ketika hewan mati saat percobaan, serta kejadian lainnya yang bisa saja muncul seperti salahnya pencatatan. Hilangnya data suatu percobaan akan menyebabkan hilangnya keterangan yang diperlukan sehingga informasi yang diperoleh kurang lengkap.

Pendekatan yang biasa digunakan untuk pendugaan data hilang antara lain adalah metode Yates, *hot deck imputation*, *substitution*, *mean imputation*, dan *regression imputation* (Black, *et.al* [1]). Pendekatan yang lebih modern untuk mengatasi data hilang di antaranya adalah algoritma *Expectation Maximization* (EM).

Penelitian pendugaan data hilang telah dilakukan oleh Fatimah [3], Widiarini [11], dan Sriliana [9]. Penelitian yang dilakukan oleh Fatimah [3], menggunakan metode Yates dan algoritma EM pada Rancangan Acak Kelompok (RAK) dan Rancangan Bujur Sangkar Latin (RBSL). Penelitian Widiarini [11], menggunakan metode Yates dan metode Bigger pada Rancangan Acak Kelompok Lengkap (RAKL). Kemudian penelitian yang dilakukan oleh Sriliana [9], menggunakan metode Yates pada Rancangan Bujur Sangkar Latin (RBSL).

Pada percobaan dengan banyak perlakuan, RAL, RAKL ataupun RBSL tidak efisien untuk digunakan karena peneliti sulit mengontrol kehomogenan petak percobaan dengan kelompok besar, dan sulitnya mengontrol galat percobaan. Percobaan tersebut hanya dapat dianalisis dengan Rancangan Kelompok Tidak Lengkap (RKTL) yang salah satunya adalah Rancangan *Lattice* Seimbang (RLS).

Adapun model linier aditif dari RLS (Hinkelman & Kempthorn, [6]) adalah:

$$Y_{ijq} = \mu + \beta_j + \gamma_{ij} + \tau_q + \varepsilon_{ijq} \quad (1)$$

dengan:

- Y_{ijq} = respon ulangan ke- i , kelompok ke- j dan perlakuan ke- q
- μ = rata-rata data keseluruhan
- β_j = pengaruh ulangan ke- i
- γ_{ij} = pengaruh acak pada ulangan ke- i dan kelompok ke- j
- τ_q = pengaruh perlakuan ke- q ;
- ε_{ijq} = galat percobaan pada ulangan ke- i , kelompok ke- j dan perlakuan ke- q

Metode yates merupakan metode untuk analisis data hilang pada rancangan percobaan dengan cara menyisipkan nilai dugaan yang meminimumkan JKG. Misalkan suatu percobaan dengan n sampel pengamatan dan P adalah sebuah data yang hilang pada ulangan ke- a , perlakuan ke- b , dan kelompok terkoreksi ke- c , nilai duga P dirumuskan (Steel & Torrie, [10]):

$$\hat{P} = \left(\frac{r \times R_a + t \times T_b - D}{(t-1)(r-1)} \right) \quad (2)$$

dengan:

- \hat{P} = nilai duga data hilang
- r = banyaknya ulangan
- t = banyaknya perlakuan
- R_a = jumlah nilai ulangan ke- a yang tidak hilang,
- T_b = jumlah nilai perlakuan ke- b yang tidak hilang
- D = jumlah dari data keseluruhan yang tidak hilang

Prosedur pendugaan satu data hilang dengan metode Yates dilakukan dengan menggunakan persamaan (2)

Prosedur pendugaan dua data hilang adalah [5]:

1. Nilai awal dari salah satu data hilang dihitung menggunakan persamaan:

$$x_{ij} = \frac{\bar{t}_i + \bar{b}_j}{2} \quad (3)$$

dengan x_{ij} = nilai dugaan awal, \bar{t}_i = rata-rata untuk data yang tak hilang pada baris ke- i , \bar{b}_j = rata-rata untuk data yang hilang pada kolom ke- j .

2. Nilai dugaan awal yang diperoleh pada langkah 2 digabung dengan data yang tidak hilang. Dengan menggabungkan data yang tidak hilang dan nilai dugaan awal x_{ij} , diduga nilai pengamatan hilang lainnya dengan menggunakan persamaan (2)
3. Selanjutnya x_{ij} dianggap hilang dan diduga kembali dengan menggunakan persamaan (2), langkah 2 dan 3 dilakukan secara bergantian sampai nilai dugaan kedua data konvergen ke suatu nilai tertentu.

Algoritma EM terdiri dari tahap E dan tahap M. Prosedur algoritma EM sebagai penduga data hilang pada RLS (Dempster, *et.al* [2]) adalah:

1. Tahap E

Menghitung nilai duga dari parameter yang terdapat pada model linier rancangan (1) dengan metode kuadrat terkecil [7], yaitu:

$$\hat{\mu} = \bar{G}, \hat{\beta}_i = \bar{R}_i - \bar{G}, \hat{\gamma}_{ij} = \bar{R}_{ij} - \bar{R}_i, \hat{\tau}_q = \bar{T}_q - \bar{G} \quad (4)$$

dengan:

- $\hat{\mu}$ = nilai duga rata-rata keseluruhan
- $\hat{\beta}_i$ = nilai duga parameter pada kelompok ke- j
- $\hat{\gamma}_{ij}$ = nilai duga parameter pada ulangan ke- i dan kelompok ke- j
- $\hat{\tau}_q$ = nilai duga parameter pada perlakuan ke- q
- \bar{G} = nilai rata-rata data
- \bar{R}_i = nilai rata-rata ulangan ke- i
- \bar{R}_{ij} = nilai rata-rata ulangan ke- i dan kelompok ke- j
- \bar{T}_q = nilai rata-rata perlakuan ke- q

2. Tahap M

Substitusi nilai duga parameter yang diperoleh pada tahap E ke dalam model linier RLS (1), agar diperoleh nilai Y_{ijq} yang maksimum [4].

2. METODE PENELITIAN

A. Sumber Data

Data yang dipergunakan dalam penelitian ini adalah data sekunder, yaitu data gabah kering giling (GKG) di Kabupaten Badung tahun 2010 yang terdiri dari 9 perlakuan.

B. Desain Simulasi

Pada desain simulasi ini dibahas evaluasi terhadap kedua metode pendugaan data hilang. Untuk itu dilakukan penghilangan data. Penghilangan data dilakukan sampai dua data hilang pada berbagai pola hilang, yaitu dalam satu perlakuan, dalam satu kelompok, dalam satu kolom, dalam satu ulangan dan secara acak.

C. Metode Analisis Data

Langkah-langkah analisis dalam penelitian ini adalah:

1. Melakukan pengacakan dengan menggunakan rancangan rencana dasar untuk RLS dan metode bilangan acak, sehingga berdasarkan teori pengacakan diperoleh RLS dengan 9 perlakuan, 4 ulangan, dan 3 kelompok masing-masing ulangan
2. Melakukan analisis ragam dari data lengkap dan data tidak lengkap pada RLS untuk melihat seberapa besar pengaruh dari banyaknya data yang hilang pada RLS.
3. Melakukan pendugaan data hilang dengan langkah-langkah sebagai berikut:
 - a. Melakukan pendugaan data hilang sesuai prosedur pendugaan data hilang dengan Metode Yates
 - b. Melakukan pendugaan data hilang sesuai prosedur algoritma EM
4. Menentukan nilai bias dan ragam dugaan parameter dengan metode Yates dan algoritma EM. Rumus untuk menentukan nilai bias dan ragam parameter dugaan adalah (Musa, [8]):

Nilai bias parameter dugaan

$$= \left| \frac{\sum \hat{\theta}_i}{m} - \theta \right| \quad (5)$$

Nilai ragam parameter dugaan

$$= \frac{\sum (\theta_i - \theta)^2}{m-1} \quad (6)$$

dengan:

θ_i = nilai parameter ke- i

$\hat{\theta}_i$ = nilai dugaan parameter ke- i

m = banyaknya simulasi sampai konvergen

5. Melakukan pengujian asumsi-asumsi analisis ragam pada nilai bias dan ragam dugaan parameter
 - a. Pengujian asumsi keaditifan model dengan uji Non Aditivitas Tukey (NAT).
 - b. Pengujian asumsi kenormalan data dengan uji Anderson-Darling.
 - c. Pengujian asumsi kebebasan galat satu dengan lainnya dengan uji Durbin Watson.
 - d. Pengujian asumsi kehomogenan ragam dengan uji Levene
6. Melakukan analisis ragam dari bias dan parameter dugaan menggunakan analisis ragam RAK faktorial dengan model linier (Mattjik & Sumertajaya, [7]).

$$Y_{ijk} = \mu + A_i + D_j + (AD)_{ij} + \rho_k + \varepsilon_{ijk}$$

dengan bias dan ragam dugaan parameter sebagai respon, posisi data hilang dan metode pendugaan data hilang sebagai faktor dan parameter sebagai kelompok.

7. Melakukan analisis ragam dari data setelah ditambahkan nilai duga dari data hilang.
8. Melakukan perbandingan hasil analisis ragam data tidak lengkap dengan hasil analisis ragam dari data tidak lengkap ditambahkan nilai duga data hilang. Hal ini untuk melihat seberapa besar perbedaan ketika data hilang tidak diduga dengan data hilang dilakukan pendugaan.

3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data produksi gabah kering giling (GKG) dalam ton/ha di Kabupaten Badung. Pada data dilakukan pengacakan sesuai teknis pengacakan dan denah pelaksanaan untuk Rancangan Lattice Seimbang (RLS). Data tersebut merupakan data lengkap RLS dengan 9 perlakuan, 4 ulangan dan

3 kelompok masing-masing ulangan. Kemudian dari data lengkap tersebut dilakukan penghilangan data pada satu data dan dua data pada satu perlakuan, satu kelompok, satu kolom, satu ulangan, dan secara acak. Data yang memuat data hilang selanjutnya disebut data tidak lengkap.

Sebagai langkah awal berikut ini dilakukan analisis ragam terhadap data lengkap dan data tidak lengkap.

A. Analisis Ragam Terhadap Data Lengkap dan Data tidak Lengkap

Analisis ragam terhadap data lengkap dan data tidak lengkap dilakukan untuk melihat pengaruh dari banyaknya data hilang.

Berdasarkan hasil analisis ragam data lengkap dan data tidak lengkap, diberikan nilai JKG seperti tampak pada Tabel 1.

Tabel 1. Nilai JKG Analisis Ragam dari Data Lengkap dan Data Tidak Lengkap

Banyaknya data hilang	Posisi	JKG
Tidak ada	-	11,6409
Satu	-	27,90802
Dua	Perlakuan	56,39846
	Kelompok	32,92388
	Kolom	51,10971
	Ulangan	42,31909
	Acak	28,92939

Berdasarkan Tabel 1 tampak bahwa pada saat data lengkap nilai JKG sebesar 11,6409 sedangkan saat terdapat satu data hilang nilai JKG sebesar 27,90802. Hal ini menunjukkan bahwa hilangnya data menyebabkan nilai JKG lebih besar daripada nilai JKG pada analisis ragam data lengkap. Oleh karena itu, dilakukan pendugaan terhadap data hilang tersebut. Pendugaan data hilang dalam penelitian ini menggunakan metode Yates dan algoritma EM.

B. Pendugaan Data Hilang dengan Metode Yates dan Algoritma EM

Metode Yates adalah metode pendugaan data hilang pada rancangan percobaan dengan meminimumkan JKG. Pendugaan satu data hilang dilakukan dengan menggunakan persamaan (2). Untuk pendugaan dua data hilang

pada satu perlakuan, nilai awal x_{ij} diduga menggunakan persamaan (3), data kedua dihitung menggunakan persamaan (2). Kemudian dugaan awal x_{ij} yang diperoleh dari persamaan (3) digabung dengan data yang tidak hilang. Dengan menggabungkan data yang tidak hilang dan nilai dugaan awal x_{ij} , diduga nilai data hilang lainnya dengan menggunakan persamaan (2). Selanjutnya x_{ij} dianggap hilang dan diduga kembali dengan menggunakan persamaan (2). Langkah tersebut dilakukan secara bergantian sampai nilai dugaan kedua data konvergen ke suatu nilai.

Pendugaan data hilang untuk satu data hilang dengan algoritma EM melalui dua tahap yaitu tahap E dengan menggunakan persamaan (4) dan tahap M dengan menggunakan persamaan (1). Untuk dua data hilang, proses pendugaan dengan algoritma EM, nilai awal diperoleh dari persamaan (4) dan (1). Kemudian, nilai awal tersebut digabung dengan data yang tidak hilang dan data hilang lainnya diduga dengan persamaan (4) dan (1). Selanjutnya nilai awal dianggap hilang dan diduga kembali dengan rumus yang sama. Langkah ini dilakukan secara bergantian sampai nilai dugaan kedua data konvergen ke suatu nilai.

Hasil pendugaan nilai data hilang seperti tampak pada Tabel 2.

Tabel 2. Nilai Duga Data Hilang dan Data Aslinya.

Banyaknya data hilang	Posisi	Nilai Duga				Nilai Asli	
		Yates		EM			
Satu	-	6,05	6,81	7,13			
Dua	Perlakuan	7,81	8,25	8,28	5,66	7,80	8,58
	Kelompok	5,93	7,15	4,80	6,86	4,03	6,15
	Kolom	7,44	7,37	7,21	7,10	7,56	7,28
	Ulangan	5,95	7,90	3,73	8,74	4,03	6,46
	Acak	7,05	5,80	6,44	5,60	7,78	7,13

Berdasarkan data yang memuat data hilang ditambahkan data duga pada Tabel 2, dengan persamaan (4), dihitung nilai duga parameter $\mu, \beta_i (i = 1,2,3,4), \tau_q (q = 1,2,...,9)$ dan $\gamma_{ij} (i = 1,2,3,4; j = 1,2,3)$. Selanjutnya

dilakukan perhitungan untuk nilai bias dan ragam dari dugaan parameter dengan menggunakan persamaan (5) dan (6).

Setelah dilakukan pendugaan nilai data hilang dan diperoleh nilai bias dan ragam dugaan setiap parameter, kemudian dilanjutkan dengan melakukan pengujian asumsi analisis ragam terhadap nilai bias dan ragam dugaan parameter.

C. Pengujian Asumsi Analisis Ragam Terhadap Bias dan Ragam Parameter Dugaan

Proses pengujian asumsi analisis ragam terhadap bias dan ragam parameter dugaan dilakukan untuk mengetahui apakah nilai bias dan ragam parameter dugaan telah memenuhi asumsi-asumsi dasar untuk dilakukan proses analisis ragam. Adapun asumsi yang harus dipenuhi adalah keaditifan model, kenormalan data, kebebasan galat dan kehomogenan ragam.

1. Pengujian Asumsi Keaditifan Model

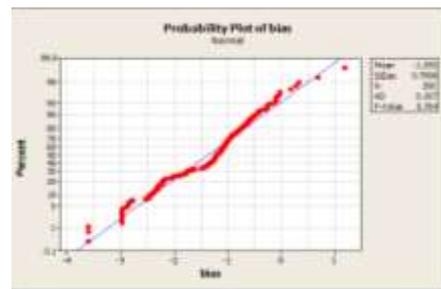
Pengujian asumsi keaditifan model dilakukan dengan uji Non-Aditivitas Tukey. Pada uji Non-Aditivitas Tukey, aditif apabila $F_{hitung} \leq F_{(\alpha,1,db\ galat)}$. Statistik uji NAT untuk nilai bias menghasilkan

$F_{hitung} = 0,00416$ $F_{tabel} = 3,88312$. Oleh karena nilai $F_{hitung} \leq F_{(0,05;1,225)}$ yaitu $0,00416 < 3,88312$, berarti pengaruh utama perlakuan dan kelompok untuk nilai bias dugaan parameter bersifat aditif pada taraf nyata 5%. Dengan cara yang sama Uji NAT untuk nilai ragam menghasilkan $F_{hitung} = 0,06016$ dan $F_{tabel} = 3,90023$. Oleh karena nilai $F_{hitung} \leq F_{(0,05;1;160)}$ yaitu $0,06016 < 3,90023$, berarti pengaruh utama perlakuan dan kelompok untuk nilai ragam dugaan parameter bersifat aditif pada taraf nyata 5%.

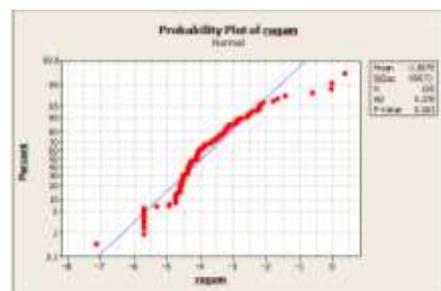
2. Pengujian Asumsi Kenormalan Data

Pengujian asumsi kenormalan data dilakukan uji Anderson-Darling. Hasil uji Anderson-Darling untuk nilai bias dan ragam parameter dugaan dengan bantuan minitab 16

berturut-turut tampak pada Gambar 1 dan Gambar 2.



Gambar 1. Uji Normalitas Nilai Bias



Gambar 2. Uji Normalitas Nilai Ragam

Berdasarkan Gambar 1 dan Gambar 2 terlihat bahwa nilai P uji Anderson-Darling untuk bias dan ragam parameter dugaan secara berturut-turut sebesar 0,764 dan 0,263, yang keduanya lebih besar dari $\alpha = 0,05$ berarti nilai bias dan ragam parameter dugaan mengikuti sebaran normal.

3. Pengujian Kebebasan Galat Percobaan

Pengujian kebebasan galat percobaan dilakukan dengan uji Durbin Watson. Perhitungan uji Durbin Watson untuk nilai bias parameter dugaan menggunakan rumus:

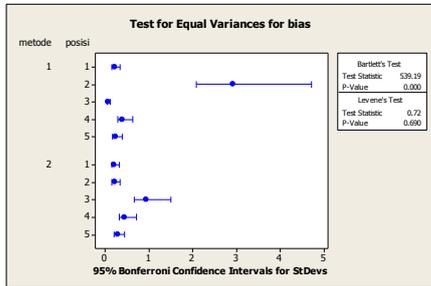
$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Dari sini diperoleh $d = 1,937583$. Hasil uji Durbin Watson untuk bias parameter dugaan $1,937583 > d_L = 1,77344$ Karena nilai $d > d_L$ berarti tidak terdapat korelasi antargalat pada nilai bias parameter dugaan.

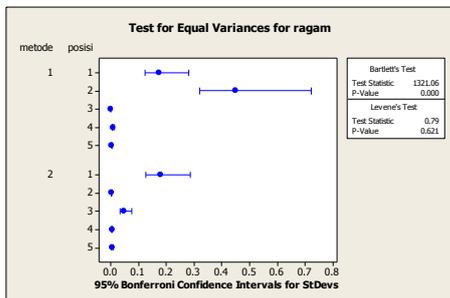
Dengan cara yang sama nilai Durbin Watson dari ragam parameter dugaan diperoleh sebesar $1,999657 > d_L = 1,77344$. Karena nilai $d > d_L$ berarti tidak terdapat korelasi antargalat dari ragam parameter dugaan.

4. Pengujian Kehomogenan Ragam

Pengujian untuk melihat kehomogenan ragam dilakukan dengan uji Levene. Hasil dari uji Levene dengan bantuan program minitab 16 untuk nilai bias dan ragam parameter dugaan secara berturut-turut diberikan pada Gambar 3 dan Gambar 4.



Gambar 3. Uji Kehomogenan Ragam untuk Nilai Bias Parameter Dugaan



Gambar 4. Uji Kehomogenan Ragam dari Ragam Parameter Dugaan

Pada Gambar 3 dan Gambar 4, tampak bahwa P-value untuk uji Levene secara berturut-turut sebesar 0,690 dan 0,621, yang keduanya lebih besar dari $\alpha = 0,05$ berarti bahwa nilai bias dan ragam parameter dugaan memiliki kehomogenan ragam.

Karena nilai bias dan ragam parameter dugaan memenuhi asumsi-asumsi dasar untuk dilakukan proses analisis ragam, maka selanjutnya dilakukan analisis ragam terhadap bias dan ragam parameter dugaan.

D. Analisis Ragam Terhadap Nilai Bias dan Ragam Parameter Dugaan

Pengujian analisis ragam terhadap bias dan ragam parameter dugaan dilakukan dengan menggunakan analisis ragam RAK faktorial dengan model linier [7].

$$Y_{abc} = \mu + E_a + H_b + (EH)_{ab} + \rho_c + \varepsilon_{abc}$$

dengan Y_{abc} adalah nilai bias dan ragam dengan data hilang pada posisi a , metode b pendugaan data hilang, dan parameter c , E_a adalah pengaruh data hilang pada posisi a , H_b adalah pengaruh metode b pendugaan data hilang, $(EH)_{ab}$ adalah pengaruh interaksi antara data hilang pada posisi a dan metode b pendugaan data hilang, dan ρ_c adalah pengaruh parameter c serta ε_{abc} adalah pengaruh galat dengan data hilang pada posisi a , metode b pendugaan data hilang, dan parameter c . Dalam hal ini $a =$ satu perlakuan, satu kelompok, satu kolom, satu ulangan dan acak, $b =$ metode Yates, algoritma EM, serta $c = \mu, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{31}, \gamma_{32}, \gamma_{33}, \gamma_{41}, \gamma_{42}, \gamma_{43}, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7, \tau_8, \tau_9$.

Pengujian analisis ragam terhadap bias dan ragam parameter dugaan dilakukan untuk melihat pengaruh metode dan posisi data hilang terhadap bias dan ragam dugaan parameter. Oleh karena itu hipotesis untuk uji analisis ragamnya adalah:

- a. H_0 : Pendugaan parameter tidak berpengaruh nyata terhadap respon
- H_1 : Parameter berpengaruh nyata terhadap respon
- b. H_0 : Posisi tidak berpengaruh nyata terhadap respon
- H_1 : Posisi berpengaruh nyata terhadap respon
- c. H_0 : Metode tidak berpengaruh nyata terhadap respon
- H_1 : Metode berpengaruh nyata terhadap respon
- d. H_0 : Interaksi posisi dengan metode tidak berpengaruh nyata terhadap respon
- H_1 : Interaksi posisi dengan metode berpengaruh nyata terhadap respon

Hasil analisis ragam untuk bias dan ragam parameter dugaan dengan bantuan program minitab 16 secara berturut-turut tampak pada Tabel 3 dan Tabel 4.

Tabel 3. Hasil Analisis Ragam Terhadap Bias Parameter Dugaan

Sumber	DF	JK	KT	F_{hitung}	P	α	Keputusan
Parameter	25	38,9604	1,5584	2,91	0,000	0,05	H_0 ditolak
Posisi	4	4,7091	1,1773	2,19	0,007	0,05	H_0 ditolak
Metode	1	0,2823	0,2823	0,53	0,469	0,05	H_0 diterima
Posisi*Metode	4	0,9550	0,2387	0,45	0,776	0,05	H_0 diterima
Error	225	120,6819	0,5364				
Total	259	165,5887					

Tabel 4. Hasil Analisis Ragam Terhadap Ragam Parameter Dugaan

Sumber	DF	JK	KT	F_{hitung}	P	α	Keputusan
Parameter	25	85,4663	3,5436	6,03	0,000	0,05	H_0 ditolak
Posisi	4	9,0577	2,0938	3,56	0,008	0,05	H_0 ditolak
Metode	1	0,0210	0,0012	0,00	0,964	0,05	H_0 diterima
Posisi*Metode	4	11,1775	2,7944	4,76	0,001	0,05	H_0 ditolak
Error	160	93,9935	0,5875				
Total	194	199,7160					

Berdasarkan Tabel 3 dan Tabel 4 tampak bahwa parameter dan posisi memiliki $p - value < \alpha$. Oleh karena itu H_0 ditolak, yang berarti pendugaan parameter dan posisi data hilang berpengaruh nyata pada taraf 0,05 terhadap nilai bias dan ragam parameter dugaan. Hal ini menunjukkan bahwa ketika hasil pendugaan parameter berbeda maka terdapat perbedaan nilai bias dan ragam dari masing-masing parameter dugaan, demikian juga ketika posisi data hilang berbeda maka terdapat perbedaan nilai bias dan ragam parameter dugaan dari masing-masing posisi data hilang.

Pada Tabel 4 tampak juga untuk interaksi antara posisi dengan metode dengan $p - value = 0,01 < \alpha = 0,05$ sehingga H_0 ditolak, yang berarti interaksi antara posisi data hilang dengan metode pendugaan data hilang berpengaruh nyata pada taraf 0,05 terhadap ragam dari parameter dugaan. Hal ini menunjukkan ketika terjadi interaksi antara posisi data hilang dengan metode pendugaan data hilang maka terdapat perbedaan ragam parameter dugaan ketika metodenya berbeda atau posisinya yang berbeda.

Selanjutnya dilakukan perbandingan untuk melihat metode mana yang lebih baik di antara dua metode yang dipergunakan. Metode

dikatakan lebih baik apabila memiliki nilai bias dan ragam terkecil. Rata-rata nilai bias dan ragam dari nilai duga dan parameter dugaan berturut-turut diberikan pada Tabel 5 dan Tabel 6.

Tabel 5. Bias dan Ragam dari Nilai Duga

Banyaknya data hilang	Posisi	Bias Nilai Duga		Ragam Nilai Duga	
		Yates	EM	Yates	EM
Satu		1,08167	0,32371	1,18182	0,10585
	Satu Perlakuan	0,16953	1,69704	0,02304	4,42193
Dua	Satu Kelompok	1,45658	0,74727	6,76851	4,36448
	Satu Kolom	0,10118	0,26320	0,01124	0,16793
	Satu Ulangan	1,68403	1,28353	2,92450	2,66750
	Acak	1,03088	1,43483	1,16459	2,08868

Tabel 6. Bias dan Ragam dari Parameter Dugaan

Banyaknya data hilang	Posisi	Bias Nilai Duga		Ragam Nilai Duga	
		Yates	EM	Yates	EM
Satu		0,25762	0,22765	0,00018	0,00014
	Satu Perlakuan	0,10938	0,11062	0,03573	0,03646
Dua	Satu Kelompok	0,65083	0,19054	0,09173	0,00081
	Satu Kolom	0,07922	0,28058	0,00010	0,00928
	Satu Ulangan	0,20165	0,18508	0,00191	0,00096
	Acak	0,14885	0,14972	0,00078	0,00097

Berdasarkan Tabel 5 dan Tabel 6 tampak bahwa pendugaan data hilang pada saat terdapat satu data hilang dengan algoritma EM lebih baik daripada metode Yates karena rata-rata nilai bias dan ragam dari nilai duga dan parameter dugaan untuk algoritma EM lebih kecil dari metode Yates.

Dengan demikian metode Yates untuk dua data hilang cenderung lebih baik digunakan pada posisi data hilang pada satu perlakuan, satu kolom dan secara acak. Algoritma EM untuk dua data hilang cenderung lebih baik digunakan pada saat terdapat satu data hilang dan pada posisi data hilang dalam satu kelompok dan ulangan.

Setelah diketahui metode yang lebih baik dalam pendugaan nilai data hilang berdasarkan posisi data hilangnya, kemudian dilakukan analisis ragam terhadap data yang memuat data hilang dan ditambah dengan nilai duga data hilang. Selanjutnya dilakukan perbandingan hasil untuk nilai JKG antara data tidak lengkap dengan data tidak lengkap yang ditambahkan nilai duga.

E. Perbandingan Hasil Nilai JKG dari Data Tidak Lengkap dengan Data Tidak Lengkap Setelah Ditambah Nilai Duga Data Hilang

Sebelumnya telah dilakukan analisis ragam dari data tidak lengkap. Kemudian, untuk melihat seberapa besar perbedaan nilai JKG ketika nilai data hilang tidak dilakukan pendugaan dan nilai data hilang yang dilakukan proses pendugaan, selanjutnya dilakukan analisis ragam terhadap data tidak lengkap dengan ditambahkan nilai duga.

Berdasarkan hasil analisis ragam diberikan nilai JKG dari data tidak lengkap dan data lengkap setelah ditambahkan nilai duga seperti tampak pada Tabel 7.

Tabel 7. Hasil Nilai JKG dari Data Tidak Lengkap dan Data Tidak Lengkap Ditambah Nilai Duga Data Hilang.

Banyaknya data hilang	Posisi	Hasil Nilai JKG Sebelum Diduga	Hasil Nilai JKG Setelah Diduga		Hasil Nilai JKG dari Data Lengkap
			Yates	EM	
Satu	-	27,9080	10,8713	11,3048	11,6409
	Satu Perlakuan	56,3985	11,6767	15,7437	
Dua	Satu Kelompok	32,9239	9,6255	10,6682	
	Satu Kolom	28,9294	11,5012	11,3159	
	Satu Ulangan	42,3191	10,7538	14,6987	
	Satu Acak	51,1097	10,9431	11,3876	

Analisis ragam dari data dengan satu data hilang yang belum dilakukan pendugaan data pada Tabel 7 memiliki nilai JKG sebesar 27,9080, sedangkan analisis ragam terhadap data dengan satu data hilang yang telah ditambahkan nilai duga baik diduga dengan metode Yates maupun algoritma EM, memiliki nilai JKG yang lebih kecil dan tidak jauh berbeda dari nilai JKG data lengkap.

Hal yang sama juga berlaku pada analisis ragam dengan dua data hilang yang tidak dilakukan pendugaan data, memiliki nilai JKG lebih besar dibandingkan dengan analisis ragam pada data hilang yang telah ditambahkan nilai duga dan juga nilai JKG dari data hilang setelah ditambah nilai duga tidak jauh berbeda dengan nilai JKG dari data lengkap.

4. KESIMPULAN

Berdasarkan analisis dan pembahasan pendugaan data hilang dengan menggunakan metode Yates dan Algoritma EM pada Rancangan *Lattice* Seimbang maka pendugaan data hilang Metode Yates cenderung lebih baik digunakan pada posisi data hilang dalam satu perlakuan, satu kolom dan secara acak, karena rata-rata nilai bias dan ragamnya lebih kecil. Algoritma EM cenderung lebih baik digunakan pada saat terdapat satu data hilang dan pada saat terdapat dua data hilang pada posisi data hilang dalam satu kelompok dan ulangan.

Hasil analisis ragam dari data dengan data hilang yang tidak dilakukan pendugaan data, memiliki nilai JKG lebih besar dibandingkan dengan analisis ragam dari data dengan ditambahkan nilai data duga dan juga nilai JKG dari data hilang setelah ditambah nilai duga tidak jauh berbeda dengan nilai JKG dari data lengkap, baik dengan metode Yates maupun algoritma EM

DAFTAR PUSTAKA

- [1] Black, W.J. Anderson, R. Hair, J.J, Babin, B. & Tatham, R. 2006. *Multivariate Data Analysis*. 6th ed. New Jersey: Prentice Hall.
- [2] Dempster, A.P. Laird, N.M. & Rubin, D.B., 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), pp.1-38.
- [3] Fatimah, I. 2003. Data Hilang dalam Rancangan Percobaan (Suatu Kajian Algoritma EM dan Metode Yates). *Skripsi*. Bogor: IPB Tidak diterbitkan.
- [4] Goejantoro, R. 2012. Data Tidak Lengkap dan Algoritma Ekspektasi-Maksimisasi. *Jurnal Eksponensial*, 3(2), pp.85 - 88.
- [5] Gomez, K.A. & Gomez, A.A. 1995. *Prosedur Statistik untuk Penelitian Pertanian*. 2nd ed. Terjemahan: Sjamsudin, E & J.S Baharsyah. Jakarta: Universitas Indonesia.
- [6] Hinkelman, K. & O. Kempthorne. 2006. *Design and Analysis of Experiments*. 2nd ed. New York.

- [7] Mattjik, A.A. & Sumertajaya, I.M. 1999. *Analisis Perancangan Percobaan Dengan Aplikasi SAS dan Minitab*. 1st ed. Bogor: IPB Press.
- [8] Musa, M.S. 1998. *Memahami Statistika Sebagai Alat untuk Penelitian : Rancangan Kajian*. 1st ed. Bogor: Jurusan Statistika IPB.
- [9] Sriliana, I. 2013. Penduga Data Hilang Pada Rancangan Bujur Sangkar Latin Dasar. *Kumpulan Makalah Seminar Semirata, F MIPA Universitas Lampung*, pp.275-81.
- [10] Steel, R.G.D. & Torrie, J.H. 1991. *Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik*. 2nd ed. Terjemahan: Bambang Sumantri. Jakarta: PT Gramedia Pustaka Utama.
- [11] Widiharah, T. 2007. Estimasi Data Hilang pada Rancangan Acak Kelompok Lengkap. *Jurnal Matematika*, 10(2), pp.60-65.