# Pendugaan Data Hilang Menggunakan Metode *Connected* EM-AMMI dalam Bahasa R

Siskha Maulana Basrul<sup>#1</sup>, Atus Amadi Putra<sup>\*2</sup>, Yenni Kurniawati<sup>\*3</sup>

\*Student of Mathematics Department State University of Padang, Indonesia \*Lecturers of Mathematics Department State University of Padang, Indonesia

1siskhamaulana31@gmail.com
2atusamadiputra@gmail.com
3kurniawati.y@gmail.com

Abstract – The Multilocation experiment, complete data are estimated with AMMI (Additive Main Effect and Multiplicative Interaction). But, if data in the research is not complete, then it can not be analysis with AMMI. The method can be use in estimator of unbalance data is "The Method of Connected EM-AMMI with R". The purpose of estimator from unbalance data with connected EM-AMMI in R is for knowing the procedures of estimator unbalance data in R. The source of data is in the research is the harvest's result of rice in four cities's West Sumatra. After that, we get 1.25 % Mean Absolute Percentage Error (MAPE). It shows that the estimator of unbalance data with using connected EM-AMMI's method.

Keywords - AMMI, connected EM-AMMI, unbalance data, MAPE.

Abstrak – Pada percobaan multilokasi, data lengkap dapat diduga dengan AMMI (Additive Main Effect and Multiplicative Interaction). Namun, jika data dalam suatu penelitian tidak lengkap maka tidak dapat dianalisis dengan AMMI. Metode yang digunakan dalam pendugaan data hilang adalah metode Connected EM-AMMI dengan software R. Tujuan pendugaan data hilang menggunakan connected EM-AMMI dalam bahasa R yaitu untuk mengetahui prosedur-prosedur pendugaan data hilang dalam bahasa R. Dalam penelitian ini, data yang diambil sebagai ilustrasi adalah data hasil panen beras di empat Kab/Kota Sumatera Barat. Setelah digunakan prosedur pendugaan data hilang dalam bahasa R maka diperoleh MAPE atau tingkat ketelitian sebesar 1.25%. Ini menunjukkan bahwa pendugaan data hilang menggunakan metode connected EM-AMMI cukup baik untuk digunakan.

Kata kunci – AMMI, connected EM-AMMI, data hilang, MAPE.

## PENDAHULUAN

Data dalam perancangan percobaan merupakan respon atau hasil pengamatan dari perlakuan yang diberikan kepada unit-unit percobaan. Data percobaan terbagi dua yaitu data lengkap dan data hilang. Data lengkap terjadi apabila setiap pengamatan yang sah dilakukan untuk setiap satuan percobaan sedangkan data hilang terjadi apabila pengamatan yang sah tidak dapat dilakukan untuk setiap satuan percobaan[1]. Percobaan memerlukan beberapa faktor atau variabel bebas yang akan diterapkan pada setiap unit percobaan. Percobaan yang dirancang oleh dua faktor dinamakan percobaan multilokasi.

Pada percobaan multilokasi, data lengkap dianalisis dengan AMMI (Additive Main Effect and Multiplicative Interaction). AMMI merupakan gabungan dari Analisis Komponen Utama (AKU) dan Analisis Ragam (ANOVA). AMMI lebih efektif dalam mendeteksi dan menjelaskan interaksi genotip dan lingkungan karena metode ini mempunyai keunggulan dalam menentukan

kestabilan dan daya adaptasi genotip tertentu pada lingkungan yang dicobakan[2].

Konsep penggunaan model multiplikasi dengan metode AMMI memang cukup menarik, namun metode ini masih mempunyai beberapa kekurangan diantaranya adalah tidak bisa menganalisis data yang tidak lengkap. Ada dua cara dalam menganalisis data tidak lengkap yaitu menganalisis data yang teramati saja atau melakukan pendugaan terhadap data hilang dilanjutkan dengan analisis datanya. Jika menganalisis data yang teramati saja maka harus ada baris dan kolom yang dikorbankan. Pengorbanan data bukan merupakan strategi yang baik karena akan menghilangkan informasi dari beberapa perlakuan yang lain[3].

Beberapa metode yang digunakan untuk menduga data hilang adalah *connected*, EM-AMMI dan *connected* EM-AMMI. Berdasarkan penelitian sebelumnya, I Made Sumertajaya telah membuktikan bahwa metode *connected* EM-AMMI merupakan metode yang baik dalam menduga data hilang daripada metode *connected* 

atau EM-AMMI saja[3]. Metode EM-AMMI adalah suatu metode yang menggunakan dugaan *connected* sebagai penduga awal dan selanjutnya membentuk model AMMI serta memperbaiki nilai dugaan sampai selisihnya kecil dari 0.0001 maka proses berhenti.

Dalam melakukan pendugaan maka diperlukan suatu ukuran ketelitian nilai dugaan data hilang yang dinamakan dengan *Mean Absolute Percentage Error* (MAPE). Berdasarkan penelitian sebelumnya, I Made Sumertajaya besaran nilai MAPE berada pada selang 0%-10%, semakin kecil nilai MAPE mengindikasikan nilai dugaan semakin teliti[3].

Penelitian sebelumnya, I Made Sumertajaya telah meneliti tentang pendugaan data hilang menggunakan metode *connected* EM-AMMI dalam SAS[3]. Oleh karena itu, Peneliti mencoba menduga data hilang menggunakan metode *connected* EM-AMMI dalam Bahasa R. Penelitian ini bertujuan agar statistikawan dan non statistikawan dapat menggunakan prosedur bahasa R dalam menduga data hilang.

## METODE

Penelitian ini adalah penelitian terapan. Metode yang digunakan adalah analisis teori-teori yang relevan

dengan permasalahan yang dibahas dan berlandaskan kepada kajian kepustakaan. Langkah kerja yang dilakukan adalah mengumpulkan data sekunder (data lengkap), menghilangkan 5% dari data lengkap tersebut, menduga data hilang dengan menggunakan metode connected, menduga data hilang dengan menggunakan metode connected EM-AMMI dan membandingkan nilai data dugaan dengan nilai data lengkap menggunakan MAPE (Mean Absolute Percentage Error).

#### HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data sekunder tentang Hasil Panen Padi di Empat Kab/Kota Sumatera Barat[5]. Data ini merupakan data penelitian eksperimen. Penelitian ini dilaksanakan dari bulan Maret 2012- Desember 2012. Penelitian dilakukan pada empat daerah yang berbeda yaitu Solok, Bukittinggi, Pariaman dan Pesisir Selatan terhadap tujuh varietas yaitu Ciredek, Anak Daro, Randah Putiah, Cantiak Maniah, Mundam, Bakwan dan Sarai Sarumpun. Data lengkap hasil panen padi di empat Kab/Kota Sumatera Barat terdapat pada tabel I.

TABEL I RATA-RATA HASIL PANEN BERAS DI EMPAT KAB/KOTA SUMATERA BARAT (SATUAN BERAT)

	L1	L2	L3	L4	Rata-Rata
G1	18.996	18.427	18.255	15.605	17.82075
G2	18.699	18.249	18.138	15.529	17.65375
G3	18.168	18.316	17.951	15.460	17.47375
G4	14.767	15.197	16.389	15.849	15.5505
G5	18.421	18.670	17.819	15.189	17.52475
G6	17.341	16.053	17.803	17.505	17.1755
G7	12.211	15.712	15.297	17.402	15.1555
Rata-Rata	16.94329	17.232	17.37886	16.077	16.90779

Data tersebut diolah dengan mencari rata-rata Hasil Panen Beras di Empat Kab/Kota Sumatera Barat. Langkah selanjutnya menguji asumsi-asumsi dari analisis ragam yaitu uji keaditifan, uji keacakan, kenormalan dan kehomogenan ragam[2]. Pengujian asumsi diperlukan agar hasil uji-F pada analisis ragam dapat digunakan secara sah. Hasil uji keaditifan menggunakan uji Tukey dihitung menghasilkan nilai F-hitung sebesar 9.36235E - 13.Nilai ini kurang dari  $F_{0.05(1.48)} = 6.302$ , sehingga asumsi keaditifan terpenuhi. Asumsi kebebasan dan keacakan sisaan terpenuhi, hal ini terlihat pada plot sisaan yang tidak membentuk suatu pola atau hubungan tertentu. Asumsi Kehomogenan ragam terpenuhi terbukti dari uji Barlett diperoleh p-value 0.108. Uji Kenormalan juga terpenuhi diperoleh p-value > 0.15. Karena semua asumsi terpenuhi maka dilakukan penghilangan data secara acak.

Pada proses pengacakan data dilakukan beberapa langkah diantaranya adalah langkah pertama, menomori setiap kolom dimulai dengan angka satu dan seterusnya. Dengan menggunakan software minitab maka dilakukan pengacakan data terhadap 28 respon yang ada, terbukti 2 data respon yang akan dihilangkan yaitu genotip 3, lokasi 2 dan genotip 1, lokasi 3. Langkah selanjutnya mencari nilai dugaan *connected* dan *connected* EM-AMMI seperti berikut ini:

A. Algoritma Bahasa R dalam pendugaan data hilang dengan metode connected EM-AMMI.

###Memasukkan data ke dalam Matriks###

Data\_panen disimbolkan suatu matriks rata-rata yang telah dihilangkan sebesar 5%.

Data\_panen<-matrix(c(data hilang),nrow=banyak kolom)

###Rata-rata hasil genotip dan lokasi### Rataan\_baris <- rowMeans(Data\_panen) Rataan\_kolom<-colMeans(Data\_panen) Rataan\_umum<-mean(Data\_panen)

###Mencari pengaruh genotip dan lokasi### Pengaruh genotip[i]<-rataanbaris[i]-rataanumum Pengaruh\_genotip[i]<-rataanbaris[i]-rataanumum

lokasi[i]<-rataankolom[i]-rataanumum lokasi[j]<-rataankolom[j]-rataanumum

###Mencari dugaan connected### yij<-m1[i,i]+(pengaruhgenotip[j]-pengaruhgenotip[i])</pre>

###Memasukkan dugaan connected kedalam matriks### > data dugaan connected<-matrix (c(data dugaan connected),nrow=banyak kolom)

###Mencari rata-rata genotip dan lokasi### Rataan\_baris <- rowMeans(data\_dugaan\_connected) Rataan\_kolom<-colMeans(data\_dugaan\_connected) Rataan\_umum<-mean(data\_dugaan\_connected)

###Mencari matriks pengaruh Interaksi### Matriks interaksi<-matrix(0,genotip,lokasi)

for (i in 1:genotip){i=0 for (j in 1:lokasi){ interaksi[i,j]<m1[i,j]-rataanbaris[i]-rataankolom[j]+rataanumum}}

jum<-genotipe+lokasi rasiodbKU<-(2\*(jum-1))/(jum\*(jum-2)) nilaieigenvarKU<-svd(varkovKU)\$d vektoreigenvarKU<-svd(varkovKU)\$u

###BIPLOT AMMI-1###

setengah jari2<-

sqrt((akarcirivarKU[2])\*rasiodbKU\*(qf(0.95,2,(jum-2)))) setengah\_jari1<-

sqrt((akarcirivarKU[1])\*rasiodbKU\*(qf(0.95,2,(jum-2))))

jari1<-(setengah\_jari1^2)

jari2<-(setengah\_jari2^2)

kestabilan<-matrix(0,jum,2)

1y1 < -c(0,1:jum)

uy2 < -c(0,1:jum)

for(i in 1:jum){

laktus<-

(((KUI2[i,1])^2)/setengah\_jari1^2)+(((KUI2[i,2])^2)/sete

ngah jari2^2) }

lokasi<-c("lokasi")

genotipe<-c("Genotip")

par(mfrow=c(1,1))

plot(rataanbaris, KUIGEN[,1], xlab='Rataan

Hasil',ylab='KUI1',xlim=range(c(rataankolom,rataanbaris )),ylim=range(KUILING[,1],KUIGEN[,1])),main="Biplo

t AMMI-1",col="red")

text(rataanbaris,KUIGEN[,1],labels=lokasi,col="red")

abline(h=0)

abline(v=rataanumum)

abline(lty=5)

points(rataankolom,KUILING[,1], col="blue",lwd=5) text(rataankolom,KUILING[,1],labels=genotipe,adj=c(0.

5,0.5),col="blue")

points(rataanbaris,KUIGEN[,1], col="red",lwd=5)

text(rataanbaris,KUIGEN[,1],labels=lokasi,adj=c(0.5,0.5)

,col="blue")

for(i in 1:genotipe)

 $\{x21 < -c(0,KUILING[i,1])\}$ 

y21 < -c(0,KUILING[i,2])

points(x21,y21,type='1',col="blue",xlab='Rataan

Hasil', ylab='KUI 1')

text(KUILING[i,1],KUILING[i,2],col="darkblue",labels =genotipe[i])}

## B. Menghitung dugaan connected

Connected yaitu penduga data hilang yang disyaratkan kolom yang berisi harus dihubungkan dalam satu garis lurus (garis horizontal maupun vertikal), yang mengikuti arah dari sel-sel terisi[5]. Sedangkan pada struktur data tidak connected terdapat sel-sel data yang ada terpisah menjadi dua kelompok data, dimana kedua kelompok data tersebut tidak dapat dihubungkan dengan suatu garis[3]. Pada tahap ini pengaruh genotip dan lokasi diasumsikan aditif,. Berikut ini ditampilkan syarat connected.

TABEL II SYARAT CONNECTED

	L1	L2	L3	L4
G1	18.996	18 //27	*	15.605
G2	18.699	18.249	18.138	15.529
G3	18.168	*	17.951	15.460
G4	14.767	15.197	16.389	15.849
G5	18.421	18.670	17.819	15.189
G6	17.341	16.053	17.803	17.505
G7	12 211	15.712	15 297	17.402

Dari Tabel II diperoleh bahwa syarat connected sudah terpenuhi maka langkah selanjutnya adalah menduga data hilang dengan menggunakan metode connected dengan cara menentukan rataan lokasi, rataan genotip, pengaruh lokasi, dan pengaruh genotip.

Penduga data hilang dengan metode connected yaitu sebagai berikut:

$$\hat{Y}_{3,2} = \hat{\mu}_{3,2} = \mu_{2,2} + (\alpha_3 - \alpha_2)$$

$$= 18.249 + (0.346 - 0.80675)$$

$$= 17.78825$$

$$\hat{Y}_{1,3} = \hat{\mu}_{1,3} = \mu_{1,2} + (\beta_3 - \beta_2)$$

$$= 18.427 + (0.385833 - 0.20433)$$

## = 18.608503

Dari program R diperoleh bahwa hasil pendugaan genotip 3 dan lokasi 2 adalah 17.788 satuan berat, sedangkan hasil pendugaan genotip 1 dan lokasi 3 adalah 18.609 satuan berat. Hasil Program R diatas sesuai dengan hasil pada program Excel.

## C. Metode Connected EM-AMMI

Langkah selanjutnya yaitu menentukan hasil dugaan dengan metode EM-AMMI tapi menggunakan dugaan awal connected. Sehingga data rata-rata hasil panen padi menjadi

TABEL III DUGAAN CONNECTED

	Lokasi1	Lokasi2	Lokasi3	Lokasi4
Genotip1	18.996	18.427	18.609	15.605
Genotip2	18.699	18.249	18.138	15.529
Genotip3	18.168	17.788	17.951	15.460
Genotip4	14.767	15.197	16.389	15.849
Genotip5	18.421	18.670	17.819	15.189
Genotip6	17.341	16.053	17.803	17.505
Genotip7	12.211	15.712	15.297	17.402

Dengan menggunakan dugaan connected dan prosedur dari EM-AMMI dalam bahasa R diperoleh kontribusi keragaman pengaruh interaksi yang mampu diterangkan oleh masing-masing komponen adalah 69.78% dan 24.68%. KUI1 dan KUI2 telah mampu menerangkan keragaman total hasil panen beras sebesar 69.78% + 24.68% = 94.46% yaitu suatu tingkat keragaman yang cukup tinggi.

Berdasarkan keberhasilan metode (Postdictive success) yaitu KUI yang nyata pada uji-F analisis ragam adalah dua KUI yaitu KUI1 dan KUI2. Sehingga model AMMI dapat diduga dengan model AMMI2 yang dapat ditulis sebagai berikut:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^{2} \sqrt{\lambda_k} \phi_{ki} \theta_{kj}$$

$$Y_{32} = \mu + \alpha_3 + \beta_2 + \sum_{k=1}^{2} \sqrt{\lambda_k} \phi_{k3} \theta_{k2}$$

$$= 17.60345$$

$$Y_{13} = \mu + \alpha_1 + \beta_3 + \sum_{k=1}^{2} \sqrt{\lambda_k} \phi_{k1} \theta_{k3}$$

$$= 18.13587$$

Tahap selanjutnya adalah memperbaiki nilai dugaan dengan membandingkan selisih dari nilai dugaan connected dengan nilai dugaan connected EM-AMMI iterasi 1. Dugaan connected adalah sebagai berikut

$$\hat{Y}_{3.2} = 17.78825$$

 $\hat{Y}_{1.3} = 18.608503$ 

Sedangkan dugaan connected EM-AMMI terbaru adalah  $\hat{Y}_{3,2} = 17.60345$ 

$$\hat{Y}_{1.3} = 18.13587$$

Maka selisih dari dugaan connected dengan dugaan connected EM-AMMI adalah

$$\hat{Y}_{3,2} = 17.78825 - 17.60345 = 0.1845498$$

$$\hat{Y}_{1,3} = 18.609 - 18.136 = 0.4731316$$

Karena selisih nilai dugaan lebih besar dari 0.00001 maka dilakukan metode connected EM-AMMI iterasi ke-2. Dalam bentuk matriks datanya menjadi

$$\bar{Y} = \begin{bmatrix} 18,996 & 18.427 & 18.136 & 15.605 \\ 18.699 & 18.249 & 18.138 & 15.529 \\ 18.168 & 17.603 & 17.951 & 15.460 \\ 14.767 & 15.197 & 16.389 & 15.849 \\ 16.443 & 18.670 & 17.819 & 15.189 \\ 17.341 & 18.994 & 17.803 & 17.505 \\ 12.211 & 15.712 & 15.297 & 17.402 \end{bmatrix}$$

Proses memperbaiki nilai dugaan dilakukan sampai iterasi ke-n dimana selisih yang dihasilkan dari dugaan tersebut kecil dari 0.001. Setelah proses iterasi ke-5, selisihnya 0.001 maka

$$\hat{Y}_{3,2} = 17.438$$
  
 $\hat{Y}_{1,3} = 18.792$ 

$$\hat{Y}_{1,3} = 18.792$$

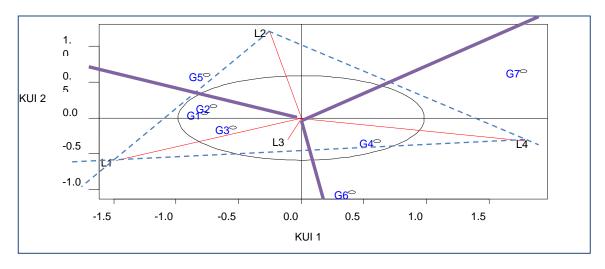
Setelah dilakukan pendugaan dengan metode connected EM-AMMI maka digunakan MAPE untuk mengetahui berapa tingkat ketelitian dari pendugaan data hilang menggunkana metode connected EM-AMMI dengan cara mencari selisih dari nilai dugaan dengan nilai sebenarnya (data lengkap). MAPE connected EM-AMMI dapat terlihat pada tabel IV.

TABEL IV
MAPE CONNECTED EM-AMMI

Genotip	Lokasi	Data Sebenarnya	Dugaan Connected EM-AMMI	MAPE
G1	L3	17.95	17.604	1.9 %
G3	L2	18.25	18.136	0.6 %
Rata-Rata MAPE Connected EM-AMMI				1.25%

Dari tabel IV dapat disimpulkan bahwa metode connected EM-AMMI cukup baik dalam pendugaan data hilang. Ini dibuktikan nilai MAPE yang dihasilkan mendekati 0% dan berada diantara selang 0% - 10%. Untuk menginterpretasikan hasil dugaan connected EM-AMMI maka diperlukan biplot AMMI1(plot antara pengaruh utama dengan skor KUI1), Jika KUI2 nyata maka dapat dilanjutkan dengan Biplot AMMI2 (plot antara KUI1 dengan KUI2). Jarak titik amatan pada

sumbu tegak menggambarkan adanya perbedaan dari pengaruh interaksi. Pengaruh interaksi genotip dan lokasi digambarkan oleh Biplot AMMI2. Untuk mengetahui genotip yang memiliki respon tertinggi di tiap-tiap lokasi, digunakan polygon dengan cara menghubungkan genotip terjauh dari titik pusat, lalu dibuat garis tegak lurus terhadap sisi poligon untuk membagi poligon menjadi beberapa kuadran (jarak yang sama)[5] sebagai berikut:



Gambar 1. Biplot AMMI-2 pada data dugaan connected EM-AMMI

Dari Gambar 1 dapat diinterpretasikan bahwa pada data dugaan *connected* EM-AMMI terdapat empat genotip yang berada didalam ellipse yaitu genotip Ciredek(G1), genotip Anak Daro(G2), genotip Randah Putiah(G3) dan genotip Cantik Maniah(G4). Sedangkan genotip-genotip lainnya seperti genotip Mundam(G5), Bakwan(G6) dan Sarai Sarumpun(G7) berada diluar ellipse sehingga dapat dikatakan tidak stabil, artinya memiliki keragaman antar lokasi yang relatif besar. Genotip stabil artinya genotip tersebut cocok tumbuh di empat lokasi yang ada seperti Solok(L1), Pariaman(L2), Bukittinggi(L3) dan Pesisir Selatan(L4). Sebaliknya jika Genotip tersebut tidak stabil berarti Genotip tersebut tidak cocok tumbuh di empat lokasi tersebut dan hanya tumbuh pada lokasi tertentu.

Dengan memperhatikan poligon yang menghubungkan lokasi-lokasi dengan keragaman terbesar serta kuadran yang terbentuk dari sisi-sisi poligon yang tegak lurus dengan pusat. Terdapat genotip-genotip yang spesifik pada lokasi tertentu artinya genotip tersebut unggul pada

lokasi tertentu yaitu genotip Mundam(G5) spesifik di Pariaman(L2), Genotip Ciredek(G1), Anak Daro(G2) dan Randah putiah(G3) spesifik di Solok(L1) dan Bukittinggi(L3) sedangkan Genotip Cantik maniah(G4), Bakwan(G6) dan Sarai Sarumpun(G7) spesifik di Pesisir Selatan(L4).

## SIMPULAN

Berdasarkan hasil penelitian dan pembahasan menggunakan metode *connected* EM-AMMI dapat di ambil kesimpulan sebagai berikut prosedur bahasa R untuk menduga data hilang menggunakan *connected* EM-AMMI diantaranya adalah mengumpulkan data lengkap, menghilangkan 5% dari data tersebut, memasukkan data ke dalam matriks, mencari rata-rata genotip dan lokasi, mencari dugaan *connected*, memasukkan dugaan *connected* ke dalam matriks yang baru, mencari nilai singular dari matriks dugaan *connected* EM-AMMI, mencari nilai dugaan *connected* EM-AMMI dan memperbaiki nilai dugaan dengan menggunakan MAPE

serta menginterpretasikan hasil dugaan kedalam Biplot AMMI. Dengan menggunakan prosedur *connected* EM-AMMI dalam bahasa R dan diaplikasikan pada data hasil panen beras maka diperoleh hasil pendugaan data hilang menggunakan metode *connected* EM-AMMI dalam bahasa R yaitu rata-rata berat beras Ciredek di Pariaman yaitu 17.59675 satuan berat sedangkan rata-rata berat beras Randah Putiah di Bukittinggi yaitu 18.13587 satuan berat serta MAPE yang dihasilkan metode *connected* EM-AMMI cukup baik. Artinya, jika terdapat data hilang maka metode pendugaan yang baik adalah metode *connected* EM-AMMI.

## REFERENSI

- [1] Basrul, S. M. 2014. Pendugaan Data Hilang Menggunakan Metode *Connected* EM-AMMI dalam Bahasa R, *Skripsi*, Universitas Negeri Padang, Maret 2014.
- [2] Mattjik, A. A., dan Made, Sumertajaya. 2006. Perancangan Percobaan dengan Aplikasi SAS dan MINITAB, Bogor: IPB PRESS.
- [3] Mattjik, A. A. 2011. Pemodelan AMMI, kini dan yang akan datang, Bogor: IPB PRESS.
- [4] Nur, Lastri. 2013. Pengaruh Varietas dan Lokasi Tanam Terhadap Mutu Fisik Beras Padi Sawah Beberapa Varietas Lokal Sumatera Barat, Skripsi, 45 Hal., Universitas Negeri Padang, Padang, Indonesia, Maret 2013.
- [5] Silvianti, P. 2005. Klasifikasi genotip pada data tidak lengkap dengan pendekatan model AMMI, Skripsi, 67 Hal., Institut Pertanian Bogor, Bogor, Indonesia, September 2009.